

# Spatial autocorrelation: bane or bonus?

Matt. D. M. Pawley<sup>a1</sup>

Brian H. McArdle<sup>b</sup>

<sup>a</sup>Institute of Natural and Mathematical Sciences

Massey University (Albany Campus),

Auckland, New Zealand

<sup>b</sup>Department of Statistics

University of Auckland

Private Bag 92019

Auckland, New Zealand

15<sup>th</sup> December, 2016

---

<sup>1</sup>Corresponding author:

[m.pawley@massey.ac.nz](mailto:m.pawley@massey.ac.nz)

Orcid ID: 0000-0003-4502-3989

## 1 **Abstract**

2           Spatial autocorrelation is a general phenomenon within biogeographical studies.  
3   However, considerable confusion exists about how to analyse spatially autocorrelated data  
4   collected using classical sampling methods (e.g. simple random sampling). We show that  
5   the two common discordant views about autocorrelated data depend upon the desired scale  
6   of inference. Although inferential statistics seeking to generalise to different unsampled  
7   spatial areas need to be adjusted for autocorrelation, if the inference is restricted to the area  
8   from which samples have been taken, then standard tests are applicable. In the latter case,  
9   incorporating autocorrelation into the model may actually improve the precision and power  
10   of the analysis. We found that the scale of spatial inference is rarely discussed, despite  
11   being of central importance to any spatial analysis. We suggest that spatial inference should  
12   rarely be formally generalised to unsampled areas, since it typically requires some  
13   assumption of stationarity, and is thus vulnerable to accusations of “pseudo-replication”.

14

15 **Keywords:** Spatial autocorrelation, inference space, design-based sampling

16

17

## 1 **Introduction**

2           *“everything is related to everything else, but near things are more related*  
3           *than distant things”* (Tobler 1970).

4           Ecologists may argue that they have long been aware of Tobler’s ‘first law of  
5 geography’—the phenomenon of spatial autocorrelation (Student 1914, Moran 1950). A  
6 belief, still commonly held by ecologists, is that classical statistical methods using design-  
7 based inferences (e.g. Simple Random Sampling (*SRS*)), are inappropriate for spatially  
8 autocorrelated data. Following the paradigm introduced in Neyman’s prominent paper  
9 (1934), the standard argument is that in the presence of spatial autocorrelation, the sampled  
10 values are not independent, so you have less information than you think and your estimated  
11 degrees of freedom are too large. Consequently, the variance is underestimated and  
12 confidence intervals are too small (or the type I error is under-reported) (Cliff and Ord  
13 1981, F. Dormann et al. 2007, Kissling and Carl 2008, Bini et al. 2009, Rousset and Ferdy  
14 2014). Contrary to this opinion, classical sampling theory has been demonstrated as quite  
15 applicable to spatial data (de Gruijter and ter Braak 1990, Brus and DeGruijter 1993). These  
16 papers elegantly demonstrated that classical statistics used with design-based methods are  
17 not only valid for autocorrelated data, but actually require fewer assumptions than model-  
18 based (geostatistical) methods that explicitly model the autocorrelation. Although some  
19 ecologists have acknowledged this (Legendre 1993), many continue to state that the  
20 presence of spatial autocorrelation complicates their analysis, and reduces the real power  
21 and precision of statistical tests (Carroll and Pearson 1998, Osborne et al. 2001, Ferguson  
22 and Bester 2002, Fisher et al. 2002, Bester et al. 2002, Diniz-Filho et al. 2003, Schwarz et  
23 al. 2003, F. Dormann et al. 2007, Beale et al. 2010, Marrot et al. 2015).

1           How is it possible to have such disparate views about the same effect? Is one of the  
2 views incorrect? If both views are valid, when can autocorrelation be ignored and when  
3 must it be accounted for? Here, we discuss the mechanism that acts to separate this  
4 dichotomy. We will identify those situations in which classical analytical methods are  
5 inadequate (and why), and we will show that in those circumstances where classical  
6 methods are valid, autocorrelation is actually a bonus, providing an opportunity for  
7 ecologists to *improve* the precision and power of their analyses.

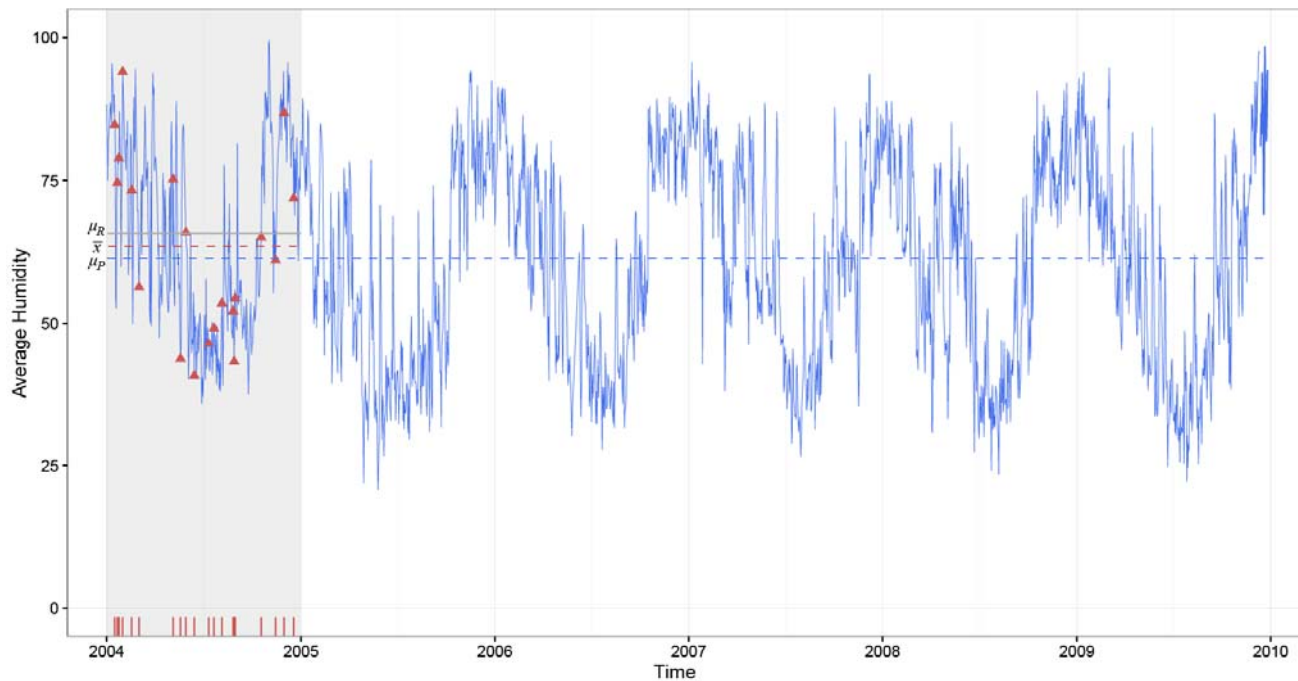
8

## 9 **Temporal vs. Spatial Autocorrelation**

10           Misconceptions about classical methods and spatial autocorrelation almost certainly  
11 arose because autocorrelation has been widely perceived as a problem with measurements  
12 in times series (Chatfield 2004). In fact, autocorrelation in time series data is not a problem  
13 when one can apply the same rules of inference commonly used in design-based methods. If  
14 a time series analysis only required the parameters describing the time-period spanned by  
15 the *sample extent*, i.e. the domain [spatial and/or temporal] within which measurements are  
16 recorded (Wiens 1989, Dungan et al. 2002), then all the statistics derived from a random  
17 sample from that domain would require no correction for autocorrelation.

18           For example, consider Fig. 1, which shows a time series of the mean daily humidity  
19 from the Aranjuez agrometeorological station (Madrid) between 2004 and 2010. Suppose  
20 20 measurements were randomly taken (red triangles) within 2004 (the sample extent, as  
21 depicted by the grey area). This scenario has three different means that may be discussed:  
22 (1) the sample mean,  $\bar{x}$ , of the 20 sample points; (2) the realisation mean,  $\mu_R$ , equivalent to  
23 the mean of all data (sampled and unsampled) within the sample extent, and (3) the process

1 mean,  $\mu_P$ , which equates to the mean of the stochastic process generating the time series  
2 (and is also asymptotically equivalent to the mean of a ‘long-term’ time series. In  
3 probability theory, the observed data are often considered to be only one possible outcome  
4 that the stochastic process *could* (hypothetically) generate—the possibilities that could  
5 happen are known as different ‘realisations’ (Hamilton 1994).



7 Figure 1. A time series showing average daily humidity at Aranjuez station (Spain)  
8 between 2004 and 2010. Twenty random samples (red triangles) were taken within a  
9 particular period of interest (the sample extent [shaded in grey]) and used to  
10 calculate a sample mean,  $\bar{x}$  [red line]. The population mean of the sample extent  
11 ( $\mu_R$ ) is depicted by the horizontal grey line, and the mean of the entire time series  
12 ( $\mu_P$ ) is depicted by the dashed blue line.  
13

1           If the purpose of the analysis in Fig. 1 was to estimate the realisation mean,  $\mu_R$ , then  
2 the estimate of mean humidity would have confidence intervals (and standard errors)  
3 correctly calculated using classical methods that ignore the autocorrelation. That is, if  
4 inferences are limited to the period from which we sampled, then autocorrelation is not a  
5 problem.

6           Time series analyses are often used to predict future values (i.e. forecasting), and  
7 problems with autocorrelation arise when we wish to generalise our descriptions from the  
8 sampled period to other unsampled periods, or to the long term behaviour of the stochastic  
9 process (e.g.  $\mu_p$ ) that generated the sampled year's records. Such analyses require  
10 considerably more restrictive assumptions; in particular, stationarity must be assumed, such  
11 that the properties of the system you have measured are the same as those of the spatial  
12 region, or period of time, you wish to generalise to [see Myers (1989) for a discussion on  
13 how restrictive this assumption must be].

14           Historically, ecologists have proved well aware of the dangers of attempting to  
15 consider a result taken at one place and/or time, and using it to make inferences about other  
16 places and other times (i.e. other realisations). Believing the properties of your single field  
17 to extend beyond your sampled area so as to generalise to other (unsampled) areas is, at  
18 best, optimistic. So in contrast to times series, when sampling an area, the inference space  
19 (the space that we wish to make statistical deductions about) is typically restricted to a  
20 particular finite area, and the standard error of the sample mean refers to the population  
21 contained within the area encompassed by the sample extent, i.e. the realisation mean. If  
22 this is the case, then classical statistical theory will perform as intended. SRS, for instance,  
23 would yield independent sampling units and an unbiased estimate of the population

1 parameters (e.g. mean and variance of the (finite) population of interest) (de Gruijter and ter  
2 Braak 1990).

3         If you wish to generalise your inferences beyond the population bound by your  
4 sample extent, then greater consideration of autocorrelation is required. In the presence of  
5 positive autocorrelation, the variability of data within the sample extent is often less than  
6 that of the process, leading us to possibly overestimate the precision of our sample mean  
7 with respect to the process mean (Pimm and Redfearn 1988, McArdle 1989). In this case, a  
8 correction is required for the fact that your particular realisation (the data from which you  
9 have sampled) may not describe the set of all possible realisations very well. Generalising  
10 inferences beyond the sample extent means that the sample scheme cannot be properly  
11 design-based since inference includes unsampled areas. Consequently, if results are then  
12 generalised beyond the sample extent, then using a design-based scheme such as SRS to  
13 sample a specific (autocorrelated) region will not ensure stochastically independent errors.  
14 In this respect stochastic independence depends not only on the sample design and the  
15 underlying population autocorrelation structure, but also on the inference space that is  
16 chosen. However, the smaller the *range* of autocorrelation (i.e. the distance at which points  
17 became independent) relative to the size of your sample extent, the more representative of  
18 the total process your realisation becomes. The efficacy of a classical 95% confidence  
19 interval using the  $t$ -distribution (i.e.  $\bar{x} \pm t \times se$ ) to estimate the realisation and process  
20 mean, and the effect of the range of autocorrelation is demonstrated via simulation in the  
21 next section.

22

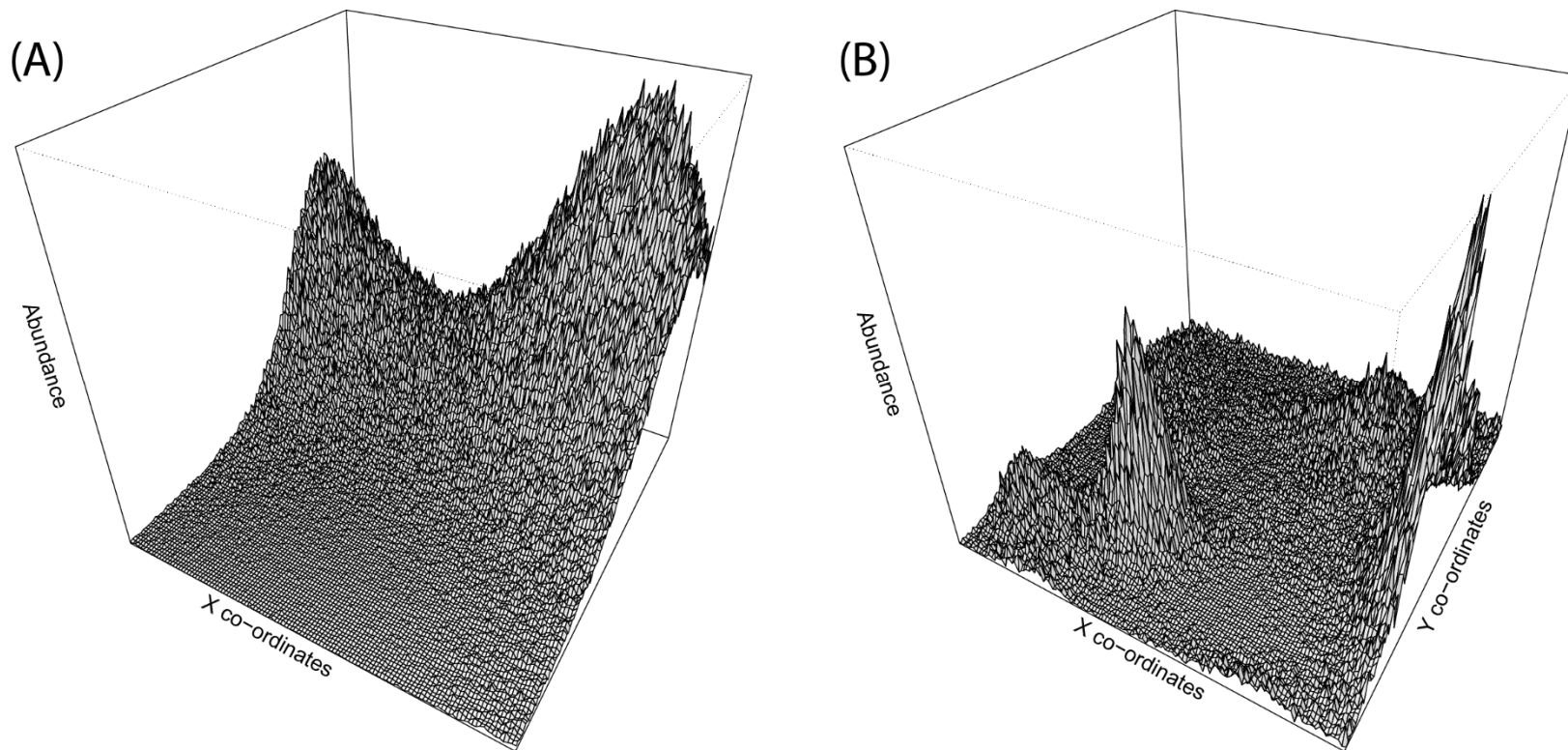
23

## 1 **A simulation study demonstrating the importance of inferences**

2 Fig. 2 shows two surfaces, each representing a simulated spatial distribution of  
3 organisms within the same sized sample extent. These surfaces were created using a  
4 positively autocorrelated random field (with a spherical variogram structure) using the  
5 GaussRF function within the RandomFields library in the ‘R’ software package (Martin  
6 Schlather 2001, R Development Core Team 2015). Since changes in populations are often  
7 considered to be the result of birth and death processes, the processes were simulated on a  
8 multiplicative scale (Legendre and McArdle 1997). Consequently the variogram  
9 (autocorrelation) structure was chosen to represent the spatial distribution of organism  
10 density on the log scale. Further stochasticity was introduced by using the (exponentiated)  
11 random field values as  $\lambda$  when generating from a Poisson distribution. Surface A was  
12 generated with a ‘short-range’ autocorrelation process (the range was around 20% of the  
13 surface) and surface B was generated with a ‘long-range’ autocorrelation process (80% of  
14 the surface), but all other stochastic parameters used to generate these surfaces were  
15 identical.



1



2

3

4

5

Figure 2. Two equal-sized surfaces simulated to represent different spatial distributions of organisms. The range of autocorrelation (relative to the extent of the surface) is relatively 'short' in B and 'long' in A.

1  
2       The distributions shown in Fig. 2 are a single (exemplar) realisation of a simulation  
3 which stochastically generated 10,000 surface realisations each of the ‘long-range’ and  
4 ‘short-range’ autocorrelation structure. For each realisation, the following statistics were  
5 calculated:

- 6       1. the realisation population mean ( $\mu_R$ );
- 7       2. a sample mean ( $\bar{x}$ ) from a single SRS ( $n = 100$ );
- 8       3. a 95% confidence interval (CI) for the population mean using the classical statistical  
9       method (i.e.  $\bar{x} \pm t \times se$ ).

10       In addition, the process means ( $\mu_p$ ) for the ‘long-range’ and ‘short-range’  
11 autocorrelation structure were also calculated. The R-code for the simulation is supplied in  
12 Supplemental 1 (S1).

13       The 95% confidence intervals encompassed their respective realisation means on  
14 approximately 94% of samples from ‘short range’ realisations and 91% of ‘long range’  
15 realisations (Table 1)—the stochastic processes generated highly over-dispersed data, so  
16 coverage for the realization means were slightly lower than 95%. In contrast, coverage of  
17 the process mean was much lower than 95% and depended on the range of autocorrelation  
18 (66% and 33% for the short- and long-range surfaces respectively). Confidence intervals  
19 from the ‘short range’ realisations (e.g. surface B) had a smaller range of autocorrelation  
20 relative to the sample extent, so the sampled data were, on average, more representative of  
21 the total process than data from the ‘long range’ process and consequently confidence  
22 intervals covered the process mean at a higher rate than the ‘long range’ process. Clearly, if

1 you sample a large area (relative to the range of spatial autocorrelation), then your  
2 generalisation is likely to be more reliable than if you only sampled a small one.

---

Inference	'Long range' autocorrelation (e.g. Surface A)	'Short range' autocorrelation (e.g. Surface B)
<i>Realisation mean</i>	94.4%	91.3%
<i>Process mean</i>	33.1%	66.2%

---

3  
4 One commonly overlooked consideration about autocorrelation is that it is not the  
5 data values that must be independent, but the residuals after the model has been fitted  
6 (ANOVA, regression, t-test etc.) (McArdle 1996). If a hypothesis test refers to parameters  
7 describing the process, then independence of errors may be achieved by the inclusion of  
8 those variables that help determine the spatial distribution. For example, some gregarious  
9 animal might be present at a location because of the presence of the particular habitat type;  
10 including some measure of the habitat in the model may remove much of the spatial  
11 autocorrelation from the residuals. Inclusion of locational covariates (such as longitude and  
12 latitude) can also act as a proxy to some determining covariate variable and help remove  
13 spatial autocorrelation from the residuals (Legendre 1990).

14

## 15 **Regression**

16 The classical model-based technique of regression aims to describe the linear  
17 dependence of a response variable ( $Y$ ), on an explanatory variable ( $X$ ), correcting for the  
18 influence of other  $X$ 's. If the  $X$  values used in the estimation procedure are representative of

1 the sample extent (e.g. from a random sample) then the description (or prediction) identified  
2 will be valid for the population within the sample extent, irrespective of any underlying  
3 autocorrelation. The error variance and inference about the slope parameters (standard  
4 errors and tests) will be correctly estimated, but only for statements about the site in  
5 question. This may be of limited usefulness.

6 Most ecologists would like to generalise any apparent relationships between  
7 variables to a larger class of sites, or at least be sure that any apparent relationship was not  
8 simply an accident of having autocorrelated data. If the range of autocorrelation is large  
9 relatively to the area under consideration then it is quite possible to detect a relationship,  
10 even if no real relationship would be apparent if a larger extent was studied. The problem  
11 arises when a "hill" in the response ( $Y$ ) surface is accidentally aligned with a "hill" (or  
12 "valley") in the explanatory variable ( $X$ ) surface (Legendre 1993). The alignment of  $X$  and  
13  $Y$  may occur due to another variable that is missing in the model ( $Z$ ), which is  
14 autocorrelated and related to both  $X$  and  $Y$ . The impact of autocorrelation on this  $X, Y$   
15 relationship is due to the missing causal variables. In other words, your model was actually  
16 incomplete because you left out an important explanatory variable. Ideally, these  $Z$   
17 variables would be measured and the  $X, Y$  relationship would be clarified by incorporating  
18 them into the model (e.g. by multiple regression, or partial correlation). If you include such  
19  $Z$  variables, the residuals are no longer autocorrelated. As is well known in regression, the  
20 validity of the description and inferences about the parameters depends crucially on the  
21 adequacy of the model. If the model is wrong then the parameter estimates and the tests  
22 upon them may be meaningless (Hawkins 2012). Our ability to generalise is also dependent  
23 on the assumptions of the model. In particular, a crucial assumption is that all the predictor

1 variables upon which  $Y$  depends are in the model; or, if they are not, then the excluded ones  
2 are uncorrelated with the included ones (and can therefore be relegated to the error term  
3 with impunity).

4 We are pseudo-replicated, though somewhat informally, if we use the results from  
5 one study to suggest trends and hypotheses about other areas. We might accept the  
6 generality of the result when many studies in many places and times seem to reproduce it.  
7 The formal generalisation of the result beyond our initial study area relies on there being  
8 enough studies to merit a meta-analysis (Hunter et al. 1982). Ecologists need to ask: what  
9 inferences do we wish to make about our sampled population? If you wish to use a  
10 generalised inference and adjust your test according to the relevant autocorrelation, then  
11 you will be generalising your inference to encompass all those areas with the same  
12 autocorrelation structure. Is this of practical use—will your readers accept the pseudo-  
13 replication?

14

## 15 **Exploiting spatial autocorrelation**

16 Consider the following statement:

17

18 *“For [other] applications, such as estimating the mean of an area, however,*  
19 *samples should be spaced more widely than the range [of autocorrelation], to*  
20 *avoid the effects of spatial autocorrelation.”* (Burrough 1995).

21

22 Such advice is misleading. In fact, if the sampling scheme positions units at distances  
23 much greater than the range of autocorrelation, then the efficiency of the design is actually

1 decreased (towards SRS). As long as inference is restricted to the sample extent, then the  
2 adage that ‘sample points are autocorrelated and are therefore not independent’ can actually  
3 work in your favour. If the underlying data are autocorrelated, then a sample unit also  
4 contains information about its neighbouring points, so you have information about a larger  
5 area of the finite population than the simple sample size implies.

6 Ecologists frequently use methods that exploit the spatial correlation present in field  
7 studies, e.g. stratification (for sampling) or restricted randomisation (e.g. blocking, Latin  
8 Squares), for experiments. In both cases, areas of similar values are sampled separately  
9 from one another, and among-strata variation can be removed from the error term by  
10 appropriate modelling methods (e.g., Analysis of Variance). More sophisticated analyses  
11 that explicitly model spatially correlated errors are used in agricultural studies to improve  
12 precision and power (Brownie and Gumpertz 1997).

13 Another common method that exploits autocorrelation is systematic sampling.  
14 Systematic sampling is widely used in practice, due in large part to its simplicity and  
15 convenience, but also because of the intuitive appeal of spreading sampling units evenly  
16 over the population, thus ensuring ‘good coverage’ (Singh and Singh 1977, Iachan 1982).  
17 Standard use of a random-start, systematic sampling design (SYS) equates to randomly  
18 positioning a theoretical grid (with resolution dependent upon the size of the sample) over  
19 the region of interest, and measuring the value of the spatially distributed variable at the  
20 grid intersections. Provided there are no periodic (waveform) patterns in the population’s  
21 surface of values, then, in the presence of positive autocorrelation, SYS is likely to give an  
22 estimate of the population mean that is as good or better than any other sampling scheme  
23 (Hájek 1959).

## 1 **Conclusions**

2           In short, the confusion between the perils of temporal autocorrelation in time series  
3 and spatial autocorrelation in spatial samples arises from the respective analyses typically  
4 having different objectives. Time series methods are designed to allow extrapolation from a  
5 finite sample extent to subsequent unsampled time periods [a *generalised* inference space –  
6 analogous to the ‘super-population’ distribution, a term first coined by Cochran (1946)]. In  
7 contrast, classical methods, as applied to spatial samples and occasionally time series,  
8 usually restrict their description and inference to the sample extent (a *finite* inference space).

9           If you have autocorrelated data, one of the first steps in determining the appropriate  
10 analysis should be examining the scale of inference. This is a step that is rarely mentioned,  
11 and commonly disregarded; leading to a relatively common misconception that correct  
12 analysis of autocorrelated data automatically requires a more sophisticated technique than  
13 the classical methods.

14           Broadly speaking, there are two types of errors that arise from not determining the  
15 appropriate level of inference:

16

17 *1. The appropriate level inference is to generalise to a larger (possibly infinite)*  
18 *space/population, but the investigator chooses an analysis that is appropriate for a finite*  
19 *region (e.g. ignoring autocorrelation).*

20           This error and its consequences have been well researched and documented.  
21 Ignoring autocorrelation in such circumstances commonly leads to underestimation of the  
22 (autocorrelated) parameter variance and confidence intervals/type I error rates that are too  
23 small (Cliff and Ord 1973, Griffith 1978, Odland 1988, Dutilleul et al. 1993, Rousset and

1 Ferdy 2014). If inference is to be generalised, there have been a number of possible  
2 solutions that have been discussed (see Augustin, Muggleston, & Buckland, 1998; Dale,  
3 Blundon, MacIsaac, & Thomas, 1991; F. Dormann et al., 2007; Pierre Legendre, 1990;  
4 Pierre Legendre & Legendre, 2012; Tavaré & Altham, 1983).

5  
6 *2. The appropriate level of inference is restricted to a specific surveyed area/finite*  
7 *population, but the investigator chooses an analysis that generalises to a broader inference*  
8 *space (a process method).*

9 The use of some method that incorporates spatial autocorrelation does not  
10 necessarily mean that such a method is appropriate for generalisation to a broader inference  
11 space. For example, when using model-based geostatistical methods, the standard inference  
12 space is restricted to the sample extent. Using a technique designed to generalise inferences  
13 will give spuriously large confidence intervals (or decrease the power of the test) and will  
14 also give non-optimal parameter estimation by your chosen criterion (e.g. least squares)  
15 since you are estimating the parameter of the process rather than the finite population. For  
16 instance, incorporating non-independent errors by using, say, a generalised least squares  
17 will not only increase the standard error of the results, but will also give a different point  
18 estimate from the independent errors model.

19

20 In summary, when making inferences around a finite population:

21 1. The use of SRS means that inference is unaffected by spatial autocorrelation.



1        2. The use of an efficient sample design (e.g. systematic sampling, stratification) that  
2            can exploit any autocorrelation in the data can improve the power of hypothesis tests  
3            and decrease confidence intervals.

4        3. Advantages may also be gained through the use of model-based analyses (e.g.  
5            geostatistics) that incorporate the autocorrelation structure of the finite realisation  
6            into the inference calculations.

7    If generalising to a ‘super-population’:

8        1. Autocorrelation may be problematic depending upon the extent of the generalisation  
9            relative to the sampling extent (i.e. how far you want to extrapolate your result).

10        Design-based sampling designs, such as SRS, are no longer applicable—so inference  
11        needs to be adjusted for any spatial autocorrelation left in the residuals after fitting  
12        the model.

13        2. The smaller the range of autocorrelation relative to the extent of your realisation  
14        (sampling extent), the more representative of the process characteristics your data  
15        becomes. Clearly if you sample a relatively large area, your generalisation is likely  
16        to be more reliable than if you only sampled a small area.

17        3. Observational studies can be affected by autocorrelation. Any inferences implying  
18        causation would have had to correct for unknown variables (spatial confounding or  
19        correlated explanatory variables). The validity of the description and inferences  
20        about the parameters depends crucially on the adequacy of the model. If the model is  
21        wrong then the parameter estimates and the tests upon them may be meaningless.

22

1           It is essential to consider exactly what the generalised population (‘super-  
2           population’) is that you wish to make inferences about. We suggest that, in many  
3           instances, ecologists should be restricting their inference to the finite population with a  
4           definite sample frame. Practical reasons may mean that generalising inference to include  
5           non-sampled spatial populations may sometimes be unavoidable. However, it should be  
6           kept in mind that such practice means that your inference is open to the accusation of  
7           pseudo-replication.

8           It is worth remembering that the spatial distribution of ecological phenomena is  
9           sometimes, in itself, the study goal, not a nuisance. The spatial autocorrelation of  
10          ecological phenomena may be a reflection of the underlying processes and functions of  
11          interest and spatial statistics provides a toolbox to be able to discover pattern and model  
12          ecosystems in a spatially explicit manner.

13

## 1 **Acknowledgements**

2 We would like to thank Adam Smith, Jorge David Aguirre, Libby Liggins and David Eme  
3 for their helpful criticisms and suggestions.

4

1 **Literature cited:**

- 2 Augustin, N. H. et al. 1998. The role of simulation in modelling spatially correlated data. -  
3 Environmetrics 9: 175–196.
- 4 Beale, C. M. et al. 2010. Regression analysis of spatial data. - Ecol. Lett. 13: 246–264.
- 5 Bester, M. N. et al. 2002. Population densities of pack ice seals in the Lazarev Sea,  
6 Antarctica. - Antarct. Sci. in press.
- 7 Bini, L. M. et al. 2009. Coefficient shifts in geographical ecology: An empirical evaluation  
8 of spatial and non-spatial regression. - Ecography (Cop.). 32: 193–204.
- 9 Brownie, C. and Gumpertz, M. L. 1997. Validity of Spatial Analyses for Large Field Trials.  
10 - J. Agric. Biol. Environ. Stat. 2: 1--23.
- 11 Brus, D. J. and DeGrujter, J. J. 1993. Design-based versus model-based estimates of spatial  
12 means: Theory and application in environmental soil science. - Environmetrics 4: 123–  
13 152.
- 14 Burrough, P. A. 1995. Spatial aspects of ecological data. - In: R.H. Jongman, C. J. F. T. B.  
15 (ed), Data analysis in community and landscape ecology. Cambridge University Press,  
16 pp. 213–251.
- 17 Carroll, S. S. and Pearson, D. L. 1998. Spatial Modeling of Butterfly Species Richness  
18 Using Tiger Beetles (Cicindelidae) as a Bioindicator Taxon. - Ecol. Appl. 8: 531.
- 19 Chatfield, C. 2004. The analysis of time series: an introduction. - Chapman & Hall/CRC.
- 20 Cliff, A. D. and Ord, J. K. 1973. Spatial Autocorrelation. - Pion Limited.
- 21 Cliff, A. D. and Ord, J. k. 1981. Spatial processes: models & applications . - Pion.
- 22 Cochran, W. G. 1946. Relative Accuracy of Systematic and Stratified Random Samples for  
23 a Certain Class of Populations. - Ann. Math. Stat. 17: 164–177.

- 1 Dale, M. R. T. et al. 1991. Multiple species effects and spatial autocorrelation in detecting  
2 species associations. - *J. Veg. Sci.* 2: 635–642.
- 3 de Gruijter, J. J. and ter Braak, C. J. F. 1990. Model-free estimation from spatial samples: A  
4 reappraisal of classical sampling theory. - *Math. Geol.* 22: 407–415.
- 5 Diniz-Filho, J. A. F. et al. 2003. Spatial autocorrelation and red herrings in geographical  
6 ecology. - *Glob. Ecol. Biogeogr.* 12: 53–64.
- 7 Dungan, J. L. et al. 2002. A balanced view of scale in spatial statistical analysis. -  
8 *Ecography (Cop.)*. 25: 626–640.
- 9 Dutilleul, P. et al. 1993. Modifying the t Test for Assessing the Correlation Between Two  
10 Spatial Processes. - *Biometrics* 49: 305.
- 11 Dormann, C.F. et al. 2007. Methods to account for spatial autocorrelation in the analysis of  
12 species distributional data: A review. - *Ecography (Cop.)*. 30: 609–628.
- 13 Ferguson, J. W. H. and Bester, M. N. 2002. The treatment of spatial autocorrelation in  
14 biological surveys: the case of line transect surveys. - *Antarct. Sci. Antarct. Sci. Ltd*  
15 14: 115–122.
- 16 Fisher, R. N. et al. 2002. Spatial Patterns in the Abundance of the Coastal Horned Lizard. -  
17 *Conserv. Biol.* 16: 205–215.
- 18 Griffith, D. A. 1978. A Spatially Adjusted ANOVA Model. - *Geogr. Anal.* 10: 296–301.
- 19 Hájek, J. 1959. Optimal strategy and other problems in probability sampling. - *Časopis pro*  
20 *pěstování Mat.* 84: 387–423.
- 21 Hamilton, J. D. (James D. 1994. *Time series analysis*. - Princeton University Press.
- 22 Hawkins, B. A. 2012. Eight (and a half) deadly sins of spatial analysis. - *J. Biogeogr.* 39: 1–  
23 9.

- 1 Hunter, J. E. (John E. et al. 1982. *Meta-analysis*: cumulating research findings across  
2 studies. - Sage.
- 3 Iachan, R. 1982. *Systematic Sampling: A Critical Review*. - *Int. Stat. Rev. / Rev. Int. Stat.*  
4 50: 293.
- 5 Kissling, W. D. and Carl, G. 2008. Spatial autocorrelation and the selection of simultaneous  
6 autoregressive models. - *Glob. Ecol. Biogeogr.* 17: 59–71.
- 7 Legendre, P. 1990. *Quantitative Methods and Biogeographic Analysis*. - In: *Evolutionary*  
8 *Biogeography of the Marine Algae of the North Atlantic*. Springer Berlin Heidelberg,  
9 pp. 9–34.
- 10 Legendre, P. 1993. Spatial Autocorrelation: Trouble or New Paradigm? - *Ecology* 74:  
11 1659–1673.
- 12 Legendre, P. and McArdle, B. 1997. Comparison of surfaces. - *Ocean. Acta* 20: 27–41.
- 13 Legendre, P. and Legendre, L. 2012. *Numerical ecology*. - Elsevier.
- 14 Marrot, P. et al. 2015. Spatial autocorrelation in fitness affects the estimation of natural  
15 selection in the wild. - *Methods Ecol. Evol.* 6: 1474–1483.
- 16 Martin Schlather 2001. *Simulation and Analysis of Random Fields*. - *R News* 1: 13–20.
- 17 McArdle, B. H. 1989. Bird population densities. - *Nature* 338: 628–628.
- 18 McArdle, B. H. 1996. Levels of evidence in studies of competition, predation, and disease. -  
19 *N. Z. J. Ecol.* 20: 7–15.
- 20 Moran, P. A. P. 1950. Notes on continuous stochastic phenomena. - *Biometrika* 37: 17–23.
- 21 Myers, D. E. 1989. To be or not to be... stationary? That is the question. - *Math. Geol.* 21:  
22 347–362.
- 23 Neyman, J. 1934. On the Two Different Aspects of the Representative Method: The Method

- 1 of Stratified Sampling and the Method of Purposive Selection. - J. R. Stat. Soc. 97:  
2 558.
- 3 Odland, J. 1988. Spatial autocorrelation. - Sage Publications.
- 4 Osborne, P. E. et al. 2001. Modelling landscape-scale habitat use using GIS and remote  
5 sensing: a case study with great bustards. - J. Appl. Ecol. 38: 458–471.
- 6 Pimm, S. L. and Redfearn, A. 1988. The variability of population densities. - Nature 334:  
7 613–614.
- 8 R Development Core Team 2015. R: A Language and Environment for Statistical  
9 Computing. in press.
- 10 Rousset, F. and Ferdy, J.-B. 2014. Testing environmental and genetic effects in the presence  
11 of spatial autocorrelation. - Ecography (Cop.). 37: 781–790.
- 12 Schwarz, P. A. et al. 2003. FACTORS CONTROLLING SPATIAL VARIATION OF  
13 TREE SPECIES ABUNDANCE IN A FORESTED LANDSCAPE. - Ecology 84:  
14 1862–1878.
- 15 Singh, D. and Singh, P. 1977. New systematic sampling. - J. Stat. Plan. Inference 1: 163–  
16 177.
- 17 Student 1914. The elimination of spurious correlation due to position in time or space. -  
18 Biometrika 10: 179–180.
- 19 Tavaré, S. and Altham, P. M. E. 1983. Serial Dependence of Observations Leading to  
20 Contingency Tables, and Corrections to Chi-Squared Statistics. - Biometrika 70: 139.
- 21 Tobler, W. 1970. A computer movie simulating urban growth in the Detroit region. - Econ.  
22 Geogr. 46: 234–240.
- 23 Wiens, J. A. 1989. Spatial Scaling in Ecology. - Funct. Ecol. 3: 385.