

FES 524: Natural Resources Data Analysis

Reading 6.1: Repeated measures


Contents

1	Introduction to repeated measures	1
1.1	Repeated measures in time	2
1.2	Repeated measures in space	2
1.3	Necessary study details	3
2	Analysis options	3
2.1	Questions that involve time	3
2.2	Questions that don't involve time	4
3	Assumptions in repeated measures models	6
3.1	Independence of subjects	6
3.2	Within subject dependence	6
3.3	Constant variance	6
4	Correlation	7
4.1	Positive correlation	7
4.2	Negative correlation	8
4.3	Checking for correlation in the residuals	9

1 Introduction to repeated measures

We refer to a study as having a repeated measures design when the investigators repeatedly measured a study unit. Given this definition, you should recognize that we have already been working with some simple repeated measures designs for a couple weeks now. Any time we use sub-sampling, like sampling multiple trees within a stand, we have repeatedly measured a study unit.

Measurements of the same study unit across time is what most people think of when they hear the term *repeated measures*. However, we can and do have repeated measures through space. This week's lab example is a study with repeated measures in space.

The big difference in what we are going to be doing this week compared to what we have been doing since we began working with mixed models in week 3 is that we will now assume some sort of pattern to the correlation among repeated measurements. Usually, with a repeated measures design in space or time, we believe the distance among measurements (in terms of distance in time, space, or space-time ) is related

to the magnitude of correlation between the measurements. We most often think measurements closer in space or time are more alike, but in some studies we are concerned that things closer in space or time are actually less alike. We will see examples later to clarify what this all means.

Having repeated measures can complicate the analysis. We saw the first complication when we started using random effects to deal with simple repeated measures designs like we had last week. Allowing for different types of correlation patterns complicates things even further. So why would we use a repeated measures design?

First, we might have a question involving the repeated measurements, such as how some mean response changes over time. Second, we may use a so-called “within-subject” design to reduce variation. This is common for time series repeated measurements. If a researcher samples different subjects every year they likely have increased variability due to always using different subjects. In addition, variability due to years and variability due to subjects are confounded. Instead, researchers often choose to sample the same subject through time in order to account for subject-to-subject variability and so make more precise estimates. Recognize that using this latter design can limit the scope of inference if one does not also sample many subjects, each sampled through time.

1.1 Repeated measures in time

Below are three examples of studies that involve repeated measures through time. Repeated measures through time can involve short or long time frames. Complicated studies can involve both.

- Example 1: Short time frame
 - Response: Suspended sediment in water samples
 - Measured: After every storm event
 - Time frame: Over one winter
- Example 2: Longer time frame
 - Response: Amount of large wood in a stream reach
 - Measured: Every summer
 - Time frame: 10 years
- Example 3: Short and longer time frames
 - Response: Dissolved oxygen in water samples
 - Measured: Every 5 minutes
 - Time frame: Every summer for 5 years

1.2 Repeated measures in space

Studies involving soil cores are classic examples of repeated measures in space. The study unit, the soil core, is measured multiple times throughout the core for some response variable. This can involve only a few measurements or many measurements.

- Example 1:
 - Response: Carbon-to-nitrogen ratio
 - Measured: At 3 depths
 - Total core length: 60 cm

- Example 2:
 - Response: Amount fine sediment
 - Measured: Every cm
 - Total core length: 7 m

Last week's Lab example was also about repeated measures in space since study units were measured multiple times. Multiple stands were measured in each watershed and multiple locations were measured in each stand. See Week 5 Handout 1 if you need a reminder of that study design.

1.3 Necessary study details

When working with repeated measures designs we need to be able to identify:

1. The subject. This is the study unit that was repeatedly measured.
2. The repeated factor. This is the factor of interest that is measured at the level of the repeated measurements.
3. I will refer to the study units that represents the replicate of the repeated factor as the repeated measurements units. Many times the repeated measurements units are what makes up the total number of observations in the study, but does not necessarily represent a replicate. For me this is often the most difficult thing to identify in a repeated measures study. It helps if you have used explicit naming of all physical/study units, including the units that are at the observation level. This helps keep the factor of interest separate from the physical unit that was measured.
4. Spacing of repeated measurements. To understand what kind of correlation is possible, we need to establish if spacing among measurements is equal/approximately equal or not.

2 Analysis options

The kind of analysis we can use when working with repeated measures depends on the research question. For many research questions, the analysis will be more complicated than what we have done in the class thus far. However, for certain research questions, we may be able to simplify things.

Let's go through a hypothetical study example and talk about different types of research questions.

Figure 1 a plot of the estimated means from some study with repeated measurements through time. The repeated factor is time. There is a second factor that has three levels, represented by the different shapes. To keep things simple as we talk about research questions, no expression of variability is included on this plot.

2.1 Questions that involve time

If the research question involves the repeated factor, we will need to work with an analysis option where we can account for the correlation caused by repeatedly measuring the same study unit. Below are two examples with research questions that involve the repeated factor.

- Example 1: What are the differences in mean response among groups at each time point? Figure 2 highlights the means that might be used to answer this question for one time point.

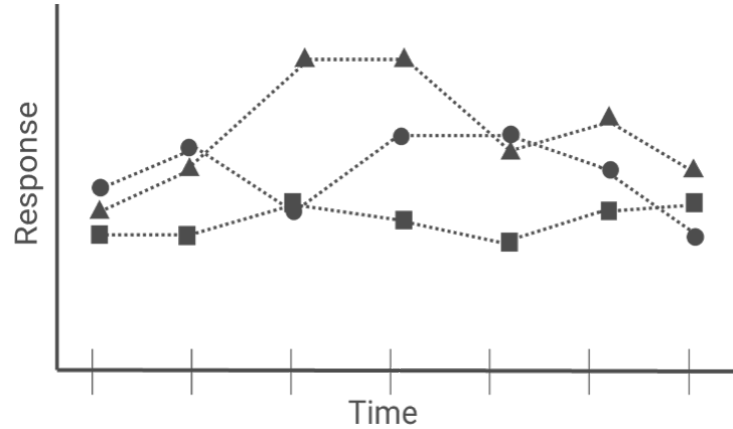


Figure 1: Example dataset with three groups all measured repeatedly through time.

- Example 2: Is the recovery pattern in mean response for the other groups the same as for the control? Comparisons to answer this question will involve estimating differences in mean response across time within a group, such as the component of the figure highlighted in Figure 3. However, the way I've worded this really implies a question about differences in the differences in means. Creating the comparisons of interest is more complicated for such questions.

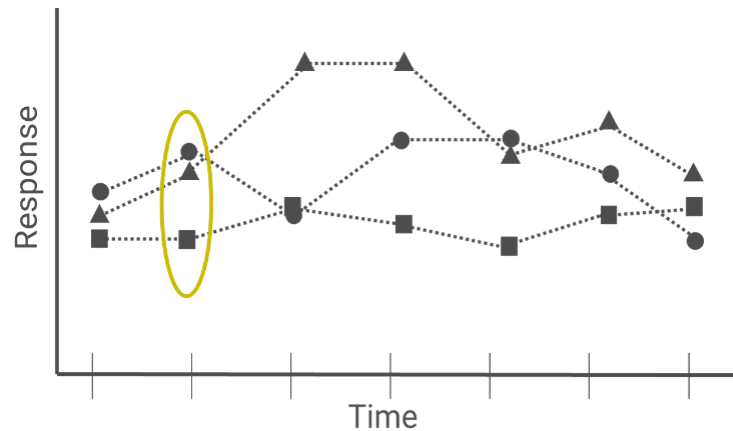


Figure 2: Example dataset with three groups all measured repeatedly through time with the relevant variance in the means for example question 1 circled.

2.2 Questions that don't involve time

If the research question doesn't involve the repeated factor directly, the analysis can likely be simplified. If there is not explicit interest in the repeated factor, it may be that the repeated measurements were really a type of subsampling.

- Example 3: How did groups differ in mean response by the end of the study? In this case, while many measurements were taken only the last measurement is of interest. The other measurements are not needed and don't need to be part of the analysis. Similarly, if the question is worded as one that involves only overall differences among groups, the repeated measures could be averaged over like we did in week 2.

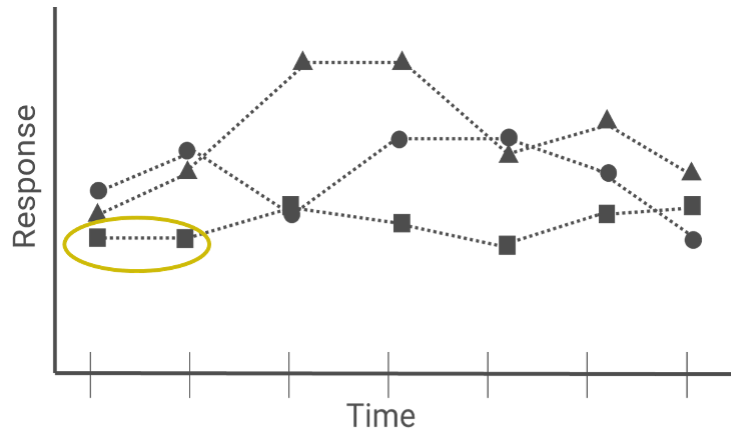


Figure 3: Example dataset with three groups all measured repeatedly through time with the relevant variance in the means for example question 2 circled.

- Example 4: How did the overall change from the beginning to end of the study differ among groups? To answer this question, we could calculate the difference between the first and last measurement for every observation and analyze that as the response variable. Then we would be back to a simple analysis. A second option would be to only work with the first and last repeated measures. While the analysis would still need to account for repeated measures, things are simpler when there are only two repeated measures.

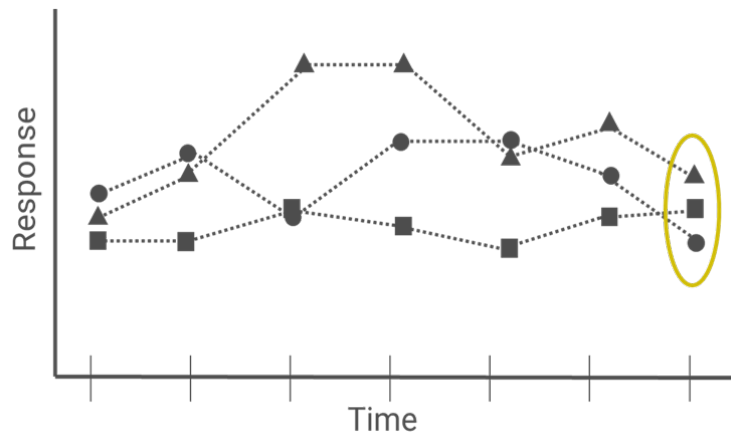


Figure 4: Example dataset with three groups all measured repeatedly through time with the relevant variance in the means for example question 3 circled.

While the question I wrote above is about the change over the duration of the study, sometimes we will have research questions about differences in means among groups at the end of the study *after accounting for the initial measurement*. The initial measurement would be considered the “baseline”. This baseline variable should be included as a covariate in a model using the final measurement as the response variable rather than taking a difference between the first and last measurement and using that as the response variable. The difference is a “baseline as covariate” model as compared to a “change from baseline” study. If relevant to you, a place to start reading about problems with change from baseline is at <https://hbiostat.org/bbr/md/change.html#whats-wrong-with-change-in-general>.

The take home message from these examples is that if we can collapse over the repeated measurements in

some way and can still answer the research question, you may be able to simplify your analysis. However, simplifying the analysis is not required.

3 Assumptions in repeated measures models

Many of the assumptions of models we can use for analyzing data from repeated measures designs should look familiar.

3.1 Independence of subjects

We assume that the subjects, the study units that were repeatedly measured, are independent. This is the same assumption we made about the levels of random factors in previous analyses (e.g., blocks in earlier weeks). This means that, in our previous analyses, we assumed that the physical units that were closer in space were more alike than physical units that were farther apart. We usually use scientific expertise to justify this assumption. For example, if the minimum distance between two blocks is large enough that we can consider them independent, then it may be reasonable to assume that blocks even further apart are also independent. If this isn't a reasonable assumption, then you will need a more complex analysis than what we are learning in this class.

3.2 Within subject dependence

For models we have fit so far this term, we assumed that the errors were independent after accounting for other variables that have a systematic effect on the response variable. In studies we are discussing this week, we now expect that the within subject errors are correlated in some way. This means that after accounting for all other variables in the model, we think there is still some sort of pattern of dependence in the errors.

The models we will learn in this class assume that this within subject dependence is shared across all subjects, regardless of if they are in different groups or had different protocols applied. If this isn't reasonably true you will need more complicated analysis tools than what you will see in this class.

3.2.1 Nested designs and dependence

As discussed earlier, nested designs are a type of repeated measurement. We were assuming that there would be some sort of dependence among observations taken within the larger physical unit. However, we assumed that the effect of the larger unit was systematic, which involves a very particular type of correlation. Using a mixed model for analysis was a way to model that type of correlation so the model errors were independent.

We will formalize the kind of correlation that is implied by a mixed model in the second reading of the week and during lecture.

3.3 Constant variance

Constant variance among errors of the repeated factors is still an assumption when doing an analysis of repeated measures data using some sort of linear model. It is not uncommon to need to relax this assumption.

For example, think about a study where some sort of extreme vegetation control is done at the beginning of the study and then vegetation cover is measured through time. In the early years we would expect cover to be less variable, since we removed vegetation cover. In later years, as things start to recover, we would expect to start seeing more variability in vegetation cover. In a study like this we can pretty much expect that we will need to relax the assumption of constant variance of errors through time.

Table 1: Example of using pairs of errors that are spaced at the same distance to estimate a model-based correlation.

ϵ_1	ϵ_2
ϵ_2	ϵ_3
ϵ_3	ϵ_4
\vdots	\vdots

4 Correlation

Now that we have discussed repeated measures in general, let's start to think about correlations more formally.

Take a look at the schematic below. It shows three observations taken on a single subject, along with their errors. Can we calculate a correlation between the first two errors, ϵ_1 and ϵ_2 ?

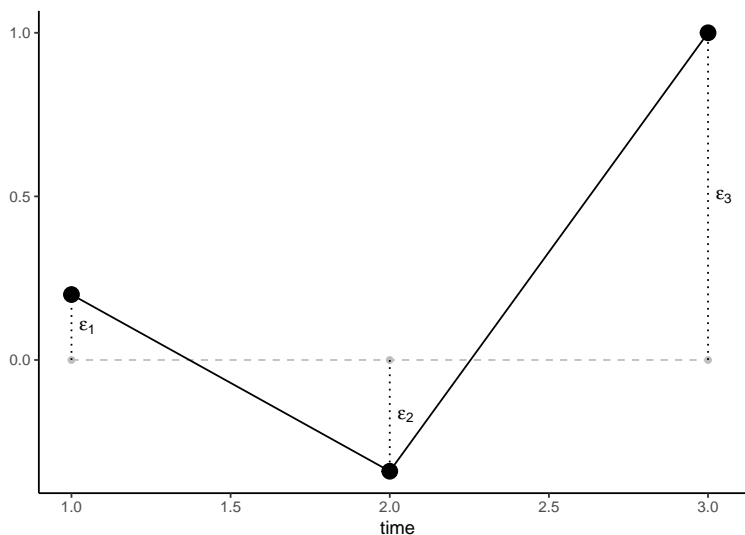


Figure 5: Example of repeated measurements through time on a subject with a mean of 0. The deviations from zero are assume to be random noise.

The answer is no. We need more than just two measurements to calculate a correlation. Recall that a correlation is calculated between two groups. In this example, group one is ϵ_1 and group 2 is ϵ_2 , so we really only have *one* measurement. A variance cannot be calculated for one measurement, and correlation is a standardized co-variance. So how do we estimate correlation in the errors? As usual, that depends.

To calculate correlation between any two positions, we need to have multiple observations of those positions. In other words, we need multiple subjects. This is a reason we assume the correlation is the same among subjects. To calculate correlations among errors for position 1 vs position 2, we'd use the errors for all the subjects at time points 1 and 2. Alternatively, we would need a *model* for the correlation in order to use the errors for a single subject to estimate a correlated errors model. For example, in time series analysis, we might assume that any two errors at the same distance from each other in time have the same correlation. Then, we could use the pairs of residuals that are all at a distance of 1 to estimate that correlation.

4.1 Positive correlation

With positive correlation (i.e., correlation > 0), errors are more alike within subjects than between subjects. Positive correlation is also indicated when errors closer in time or space are more alike than errors farther

apart.

We have been assuming positive correlation among observations for blocked and other nested designs. This is equivalent to assuming there is some sort of systematic difference among study units; an “average” effect of the unit. Figures 6 and 7, you can see positive correlation among observations in space and time, respectively. In Figure 6, all the units within the first watershed have large measured response values, while the second watershed shows all small responses. In Figure 7, observations that are close in time tend to collectively be either above or below the mean, but looking at observations farther apart in time, they may be both above, below, or on above and one below with seemingly equal frequency, indicating that the correlation decays through time.

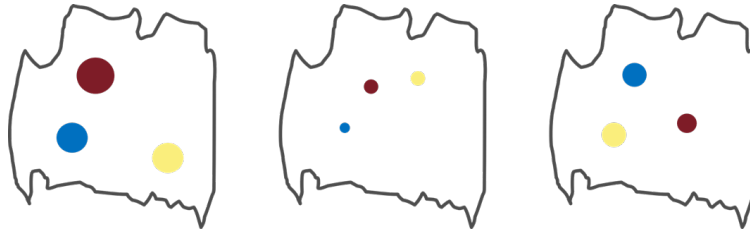


Figure 6: Example of correlation in space. Units measured within a watershed all more similar to one another than they are to units from different watersheds.

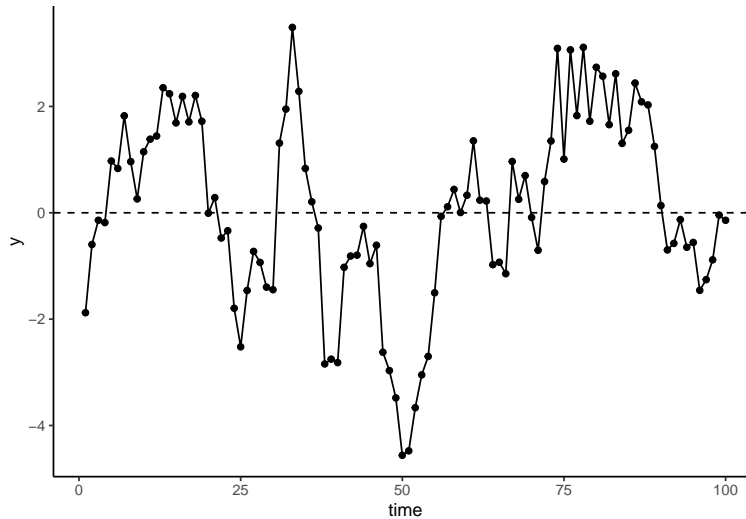


Figure 7: Example of positive correlation in time. Observations that are close together in time tend to be more similar than observations that are far apart in terms of where they are relative to the mean.

4.2 Negative correlation

With negative correlation (i.e., correlation < 0), errors are more alike between subjects than within subjects. Similarly, measurements or errors closer in space or time are less alike than those that are farther apart. Negative correlation is less common than positive correlation in many types of research.

One example of negative correlation is from community ecology. If species fill different niches, then a good year for one may be a bad year for the other, and vice versa. This results in negative correlations between species abundances through time, known as *compensatory dynamics*.

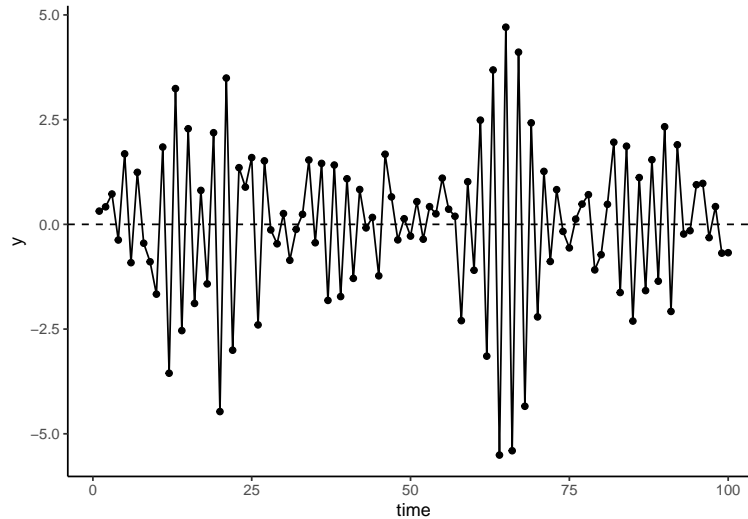


Figure 8: Example of negative correlation in time. Observations that are close together in time tend to be less similar than observations that are far apart in terms of where they are relative to the mean. This results in a pattern of observations that seem alternate above and below the mean, on average.

4.3 Checking for correlation in the residuals

We use residuals to check for dependence among errors within subjects. If we have long time series or many observations in space we have graphical tools to assist with this.

There is no rule for how many constitutes “many” for the number of within subject repeated measures, but generally if we have only very few repeated measurements like the examples in this class (3 or 4), these tools aren’t particularly useful. We will talk about an algorithm to use for when we are in a situation where we can’t graphically check for autocorrelation of errors in the next reading.

The tools we can use for checking for within-subject correlation of the errors depends on if we consider our repeated measures to be equally spaced or not. Time series are often equally spaced (measuring once a week, once a year, etc.) but can be unequally spaced (measuring after every storm event). Measurements in space are most often unequally spaced, but measuring along a soil core could involve equal spacing.

When using tools to check for residual autocorrelation, we must ensure calculations are done only within subjects and not between subjects. This can involve more complicated coding, which we will not cover in this class. If you are in this situation, I recommend doing a lot of reading before conducting analyses. One place to start is with this blog post, <https://aosmith.rbind.io/2018/06/27/uneven-grouped-autocorrelation/>.

4.3.1 Equal spacing

Autocorrelation function (ACF) plots can be used to check for correlation patterns among equally spaced repeated measures. To use these plots, we need a variable that represents the spacing of measurements within each subject so R understands the order of values within each subject. For example, if you have one measurement per subject per year, a year variable would represent the spacing of observations within subjects.

Figure 9 shows the ACF plot for the time series in Figure 7 compared to a *white noise* series (one without correlation).

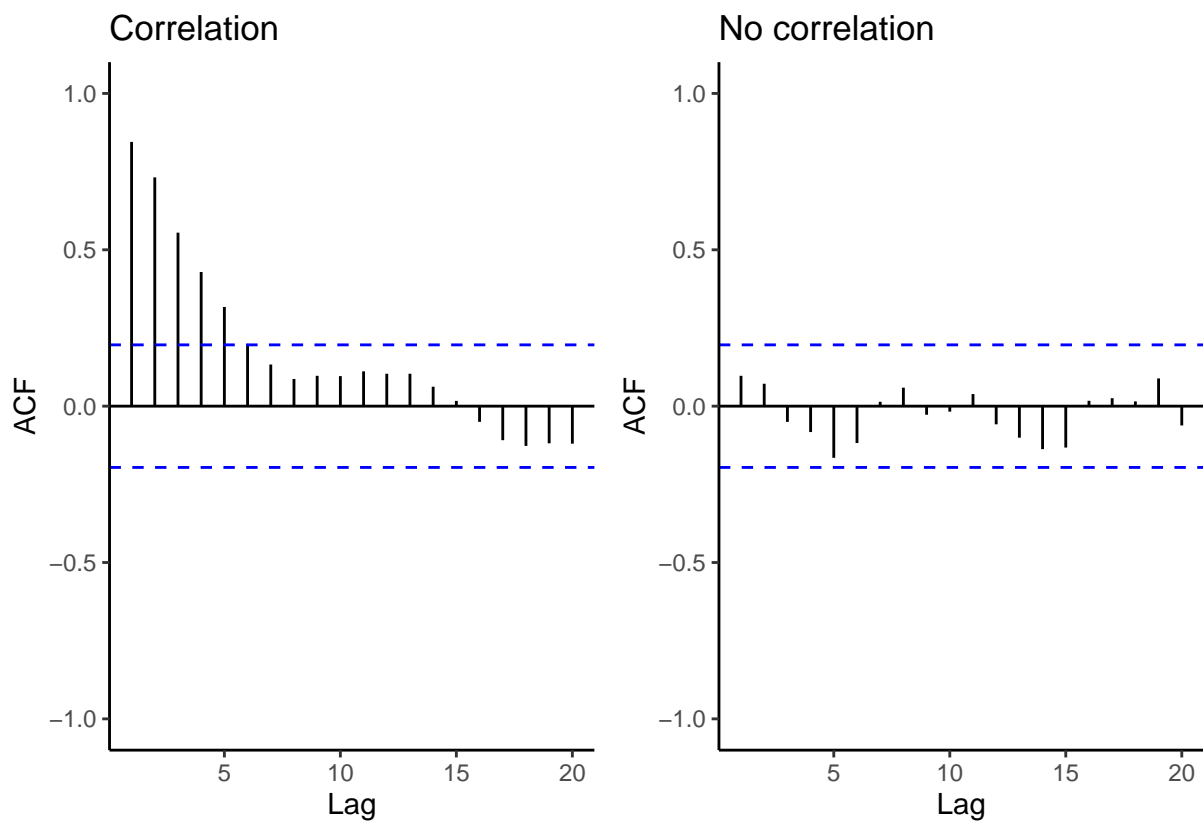


Figure 9: Comparison of two empirical autocorrelation functions, one with correlation in the series and the other without.

4.3.2 Unequal spacing

When observations are unequally spaced we can use semivariograms of the residuals to check for autocorrelation. You will hear about semivariograms most often in spatial statistics, but these are also appropriate to use if you have unequal spacing of measurements through time as well.

In order to use semivariograms, you need one or more continuous variables that represent the distance between observations. This could be Julian day, for example, if you measured the same subject on a variety of different days but the time between days is not constant. Latitude and longitude are common distance variables for measurements in space.

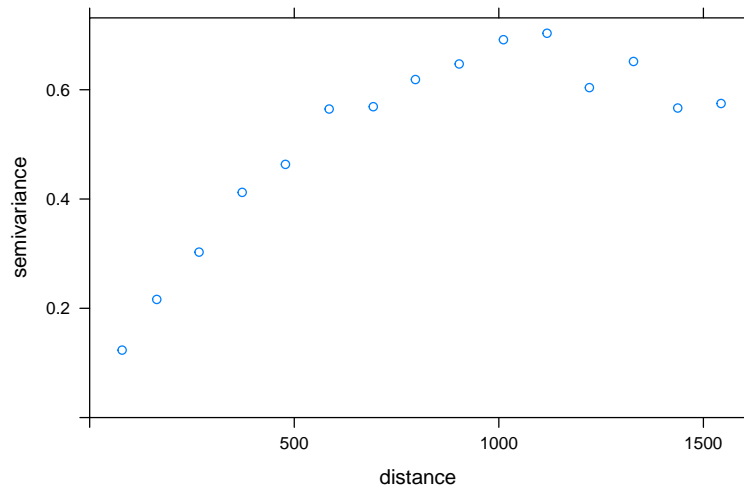


Figure 10: Example semivariogram for spatial data that is correlated in space.