Sample Size Determination

2017-03-23

In the previous document, we examined sample size requirements by looking at the number of years of sampling needed to detect a certain trend. Here we reverse the question and ask for the sample size requirements to detect a specified trend over 5 or 10 years with a given number of samples/year.

As noted in the previous document, there are several sources of variation that need to be specified.

First is the within-year sampling uncertainty. This measures the variation among sites measured that year. At the moment, we will assume that the mean CPUE for that year will be derived using a simple-random sample and so the uncertainty of the mean CPUE for that year is found using SE=sqrt(std dev among sites)/sqrt(number of sites). If a more complex sampling design is used (e.g. a stratified design within each year by HUC or by stream order), then the standard error of the mean for each year needs to be computed in a different fashion.

Second are the year-specific effects (also known as process error). Year-specific effects force the point in year above or below the trend line en masse, which is a violation of the key assumption of a regression analysis. For example, regression analysis assumes that the data are always centered about the regression line as shown in Figure 4.

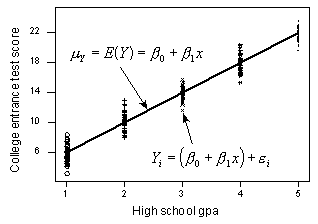


Figure 4. The standard assumption of regression analysis where at each X value, the points are randomly scattered about the trend line.

However, year specific-effects can push the set of points in a year higher or lower around the trend line as shown in Figure 5:

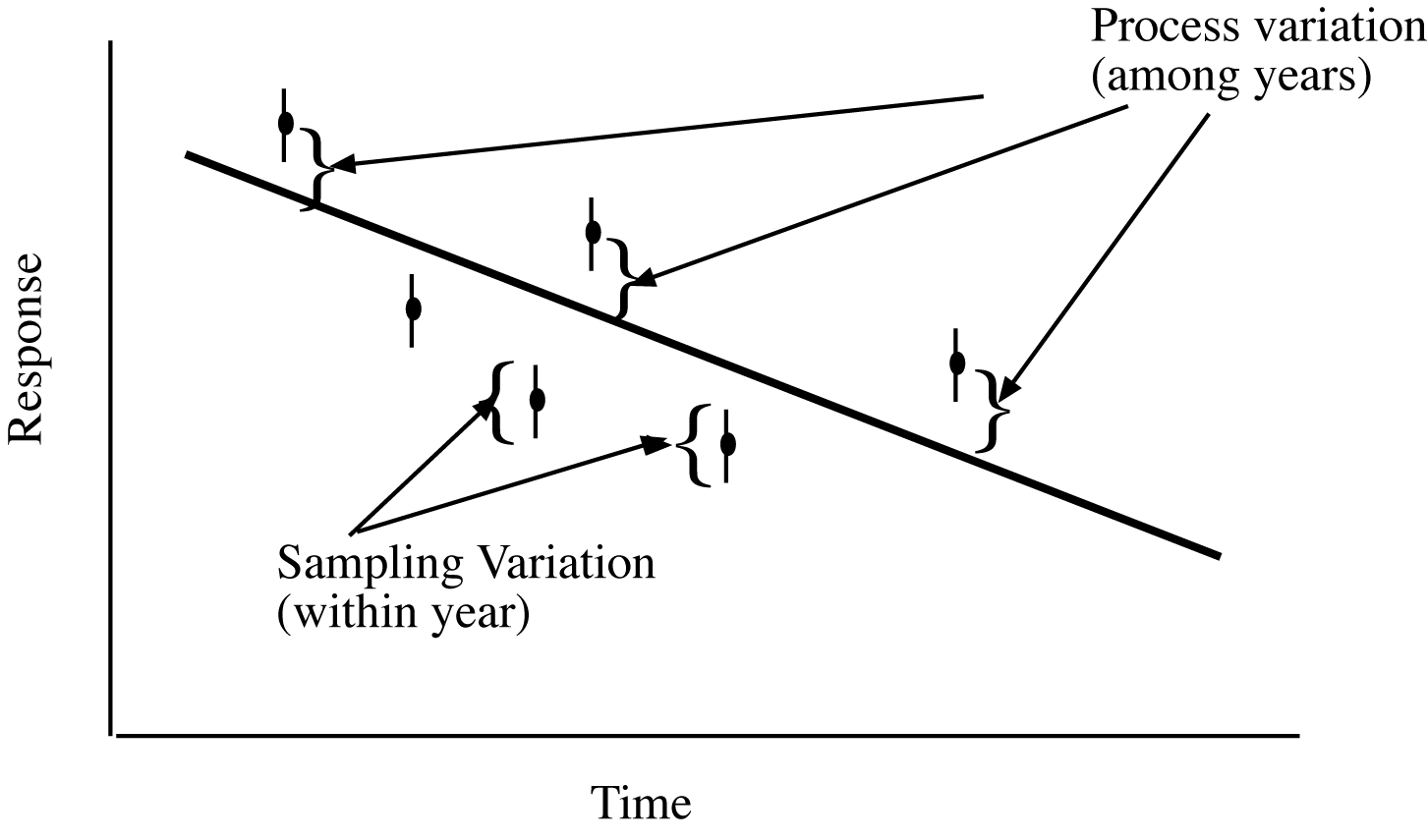


Figure 5. Process error will tend to push all points above or below the underlying trend line.

There is evidence of year-specific effects in the Quirk dataset and in my experience is almost always certainly present in the Nordegg stream, but we are unable to estimate it because we only have 1 year of data. The impact of year-specific effects (process error) are two fold. First, the individual data points become less and less important in determining the trend – only the variation in the AVERGE about the trend line is important. More importantly, to detect trends, the number of years of sampling then becomes the limiting factor as illustrated in Figure 6.

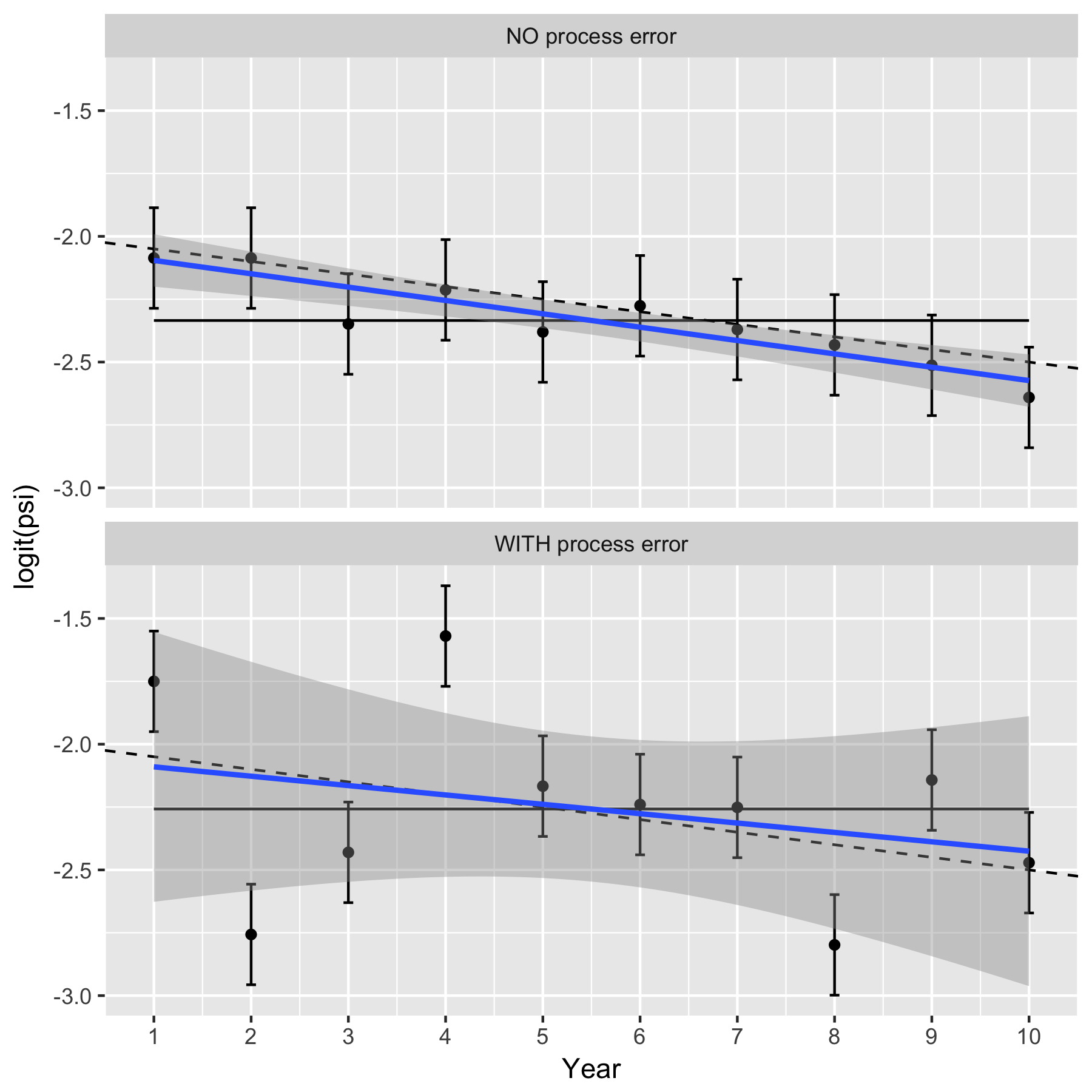


Figure 6. Illustrating the impact of year-specific factors on estimating a trend.

In the top panel, there are no year-specific effects (no process error) and the response changes over time based on the underlying trend line shown as a dashed line. Only sampling variation is present so 95% of the confidence intervals about the observed mean response each year will overlap with the underlying trend line. The fitted line will be close to the true underlying trend line (solid line). The uncertainty is small about the overall trend, and there is clear evidence that the trend is different from a 0 trend (which is shown by the horizontal line). In the bottom panel, the year-specific effects (process error) add extra variation to the underlying response each year due to effects such as weather, food etc. Now, the 95% confidence intervals for mean response still provide valid estimates for the yearly mean response values, but now may not overlap the true underling trend (in dashed). The fitted line will still be unbiased for the true trend (solid line), but the extra variation makes the uncertainty in the fitted line much larger and now there is no evidence that the trend line differs from 0.

If there is substantial year-specific effects (process error), then sampling more sites in a year will NOT be helpful. Sampling more sites in a year will shrink the size of the confidence intervals in the bottom panel above, but has NO impact on the process error and so the variation around the fitted line will only be reduced slightly. In cases of substantial process error, the limiting factor for detecting trends is likely to be the total years of sampling and not the number of sites/year that are sampled.

We did a power analysis/sample size determination to determine how many sites/year are needed to detect a 10%, 30%, 50%, 100% and 200% change over 5 or 10 years for the Nordegg System. Data was extracted from an Excel spreadsheet “Nordegg2016SiteCUESummary” which recorded the number of fish (of 4 species) captured in 39 sites in 2016 through electrofishing. These were converted to the number of fish captured per 100m and per 100s of the survey and the mean and standard deviation were computed over the sites in 2016 with the following results:

Measure n mean SDsampling rSDsampling

1 BKTR\_100m 39 6.16 8.98 1.46

2 BLTR\_100m 39 0.56 1.22 2.20

3 BURB\_100m 39 0.03 0.12 4.61

4 MNWH\_100m 39 0.04 0.14 3.07

5 BKTR\_100s 39 0.80 1.07 1.34

6 BLTR\_100s 39 0.11 0.23 2.02

7 BURB\_100s 39 0.00 0.02 5.20

8 MNWH\_100s 39 0.03 0.11 3.52

The relative sampling sampling standard deviation (last column) is the simply the ratio of the standard deviation to the mean and is a standardized measure of variation (aka the coefficient of variation). The relative sampling standard deviation is quite large – often 2x or 3x the mean.

As noted earlier, we are unable to estimate the process standard deviation. However, from the Quirk system, the relative process standard deviation is on the same order as the relative standard deviation. We examined the impact of a relative process standard deviation of 0, .5, and 1.0 on the power to detect effects.

The change to be detected (e.g. 100% change over 5 years) is converted to the proportional change per year to account for compounding. For example, a 100% increase over 5 years is equivalent to a 19% increase per year. For example, if the system starts at the value of 100 in year 1, it then takes the value of 100\*(1+.19)=119 in year 2; 119\*(1+.19)= 142 in year 3; 142\*(1+.19)=169 in year 4; and 169\*(1+.19)=200 in year 5.

A power was then computed to find the probability of deteting this trend given different numbers of sites/year (assuming a simple random sampling design) and different levels of process error. Sample plots are attached for BKTR \_100m assuming no process error (Figure 1) and a modest process error (Figure 2).

The case of no process error is VERY optimistic – it is quite rare to have cases with no process error in practice. Only very large changes (i.e. 100% or 200% increases over 5 or 10 years) have reasonable power to be detected with 40 or few sites measured/year. Once process error is added (Figure 2), the situation become much more dire. Indeed, the power is essentially independent of the number of sites sampled/year and never reaches the 80% (acceptable) power threshold. In this case, the year-to-year variation simply overwhelms any uncertainty due to sampling.

How can power be improved. Often process error is the limiting factor. Once way to improve power is to measure covariates that may be related to the process error (e.g. winter temperature) and use to try and “explain” some of the process error. However, if the covariate is also changing over time (e.g. due to climate change), then this is not an advisable process because you would the be correcting for change that will persist over time.

In cases of high process error, there is no need to do much sampling in each year. As Figure 2 showed, the power is essentially independent of the number of sites sampled.

It is not necessary to sample every year. If the same number of sites were sampled every second year (e.g. in years 1, 3, 5) the power is not changed dramatically – what usually is important in cases of process error is the length of time series (i.e. how many years between the start and end of the study).

One important consideration if process error is present, is that comparisons of the mean CPUE between two years is seldom useful. The mean CPUE could up or down simply due to process error and not due to improvements in the system. Only several years of data can separate out the underlying trend from effects of process error.

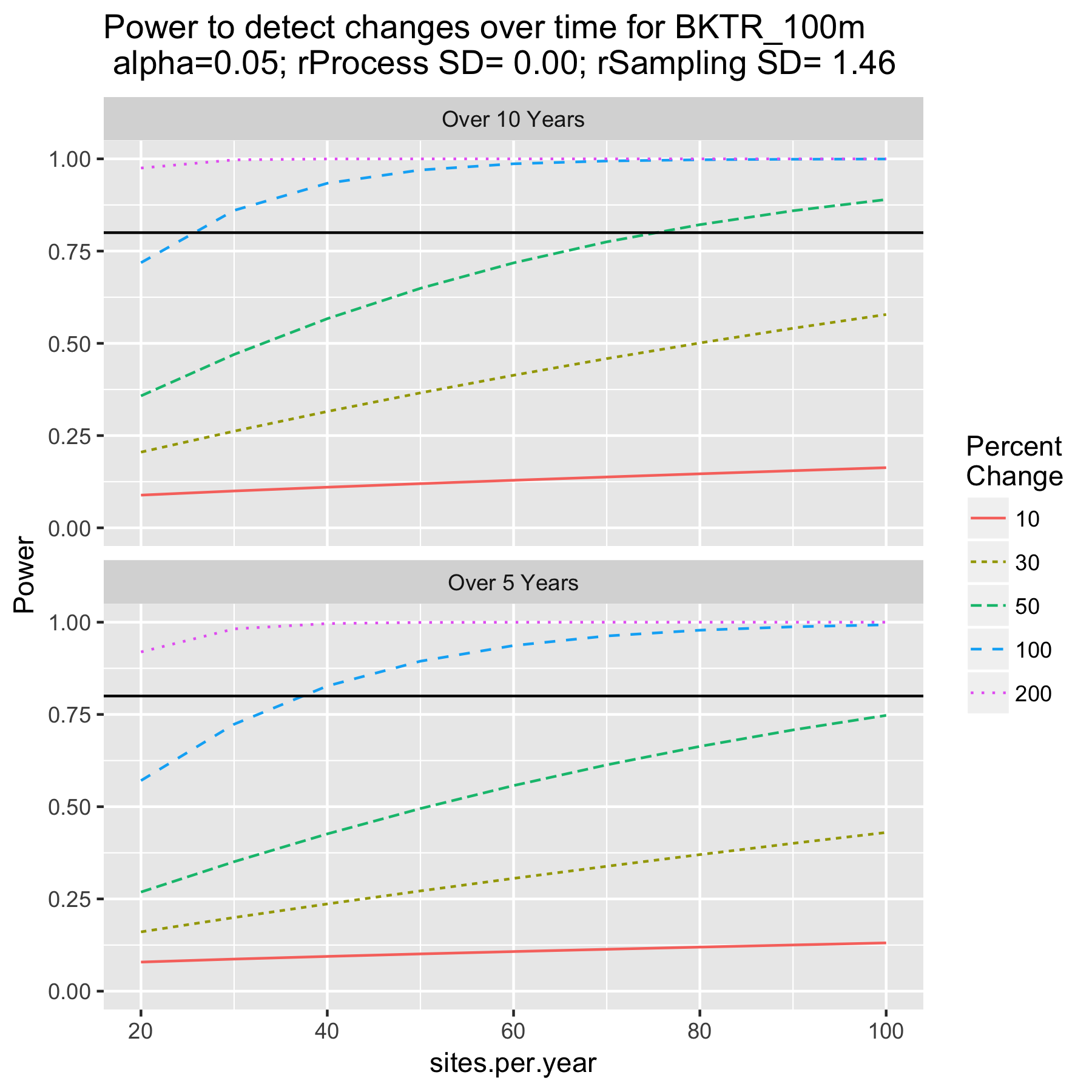


Figure 1. Power to detect various changes over 5 or 10 years with different number of sites sampled/year and no process variation assumed. These will be optimistic because it is quite rare to have no process variation.

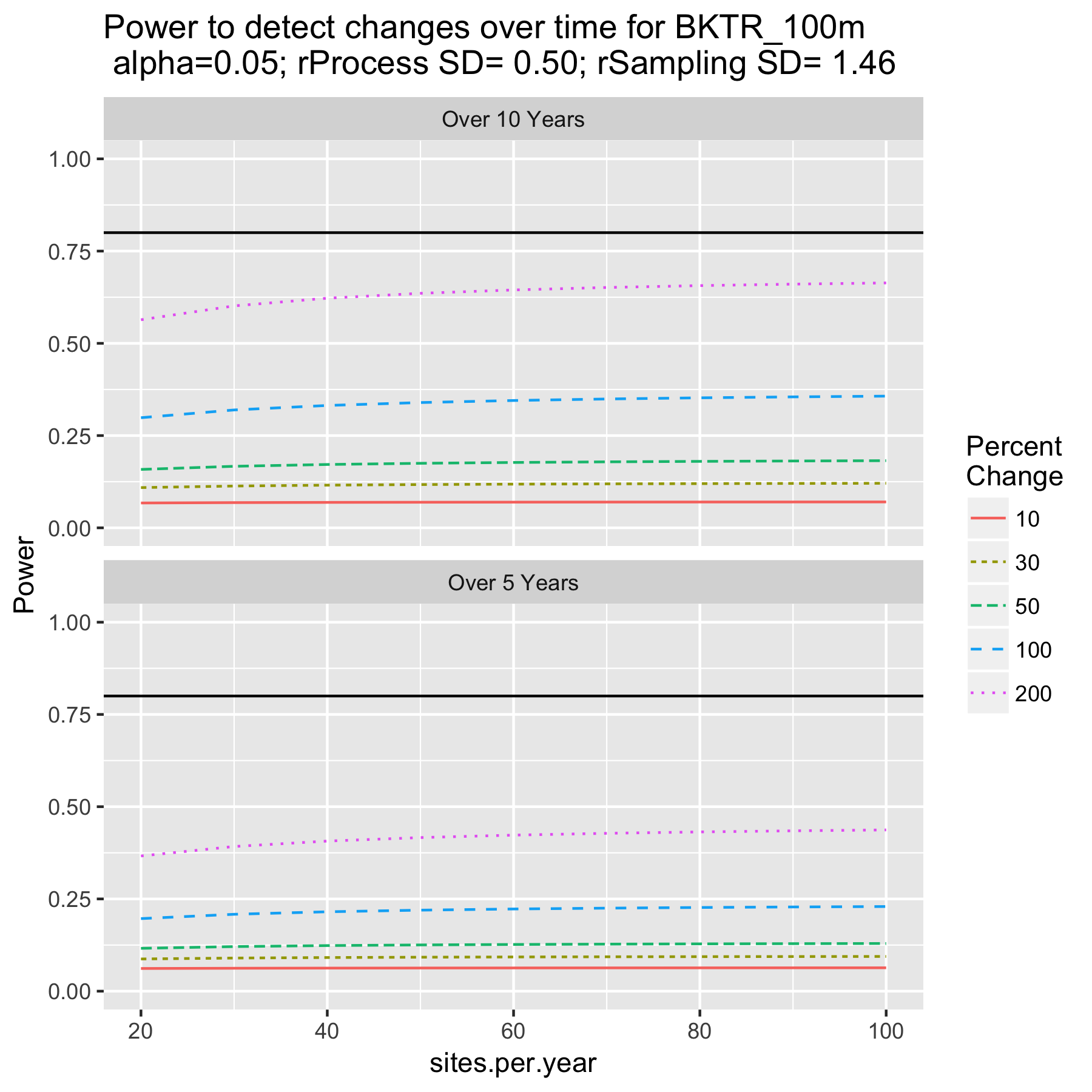


Figure 2. Power to detect various changes over 5 or 10 years with different number of sites sampled/year and a moderate amount of process variation assumed.