Sample Size Determination

MacKenzie, Moon, Kakwa System

2017-03-24

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.

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

First is the within-year sampling variability. This measures the standard deviation in CPUE 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. [As will be seen later, process error may be the limiting factor for detecting trends and the actual computations within a year are somewhat moot.]

Second are the year-specific effects (also known as process error). Year-specific effects force the points in year above or below the trend line *en masse*, which is a violation of the key assumption of a regression (trend) analysis. For example, a regression (trend) analysis assumes that the data are always centered about the regression line as shown in Figure 1.

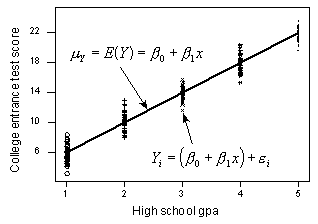


Figure 1. The standard assumption of a regression analysis where at each *X* value, the points are randomly scattered about the trend line.

However, year specific-effects can push the set of points in a year higher or lower around the trend line as shown in Figure 2:

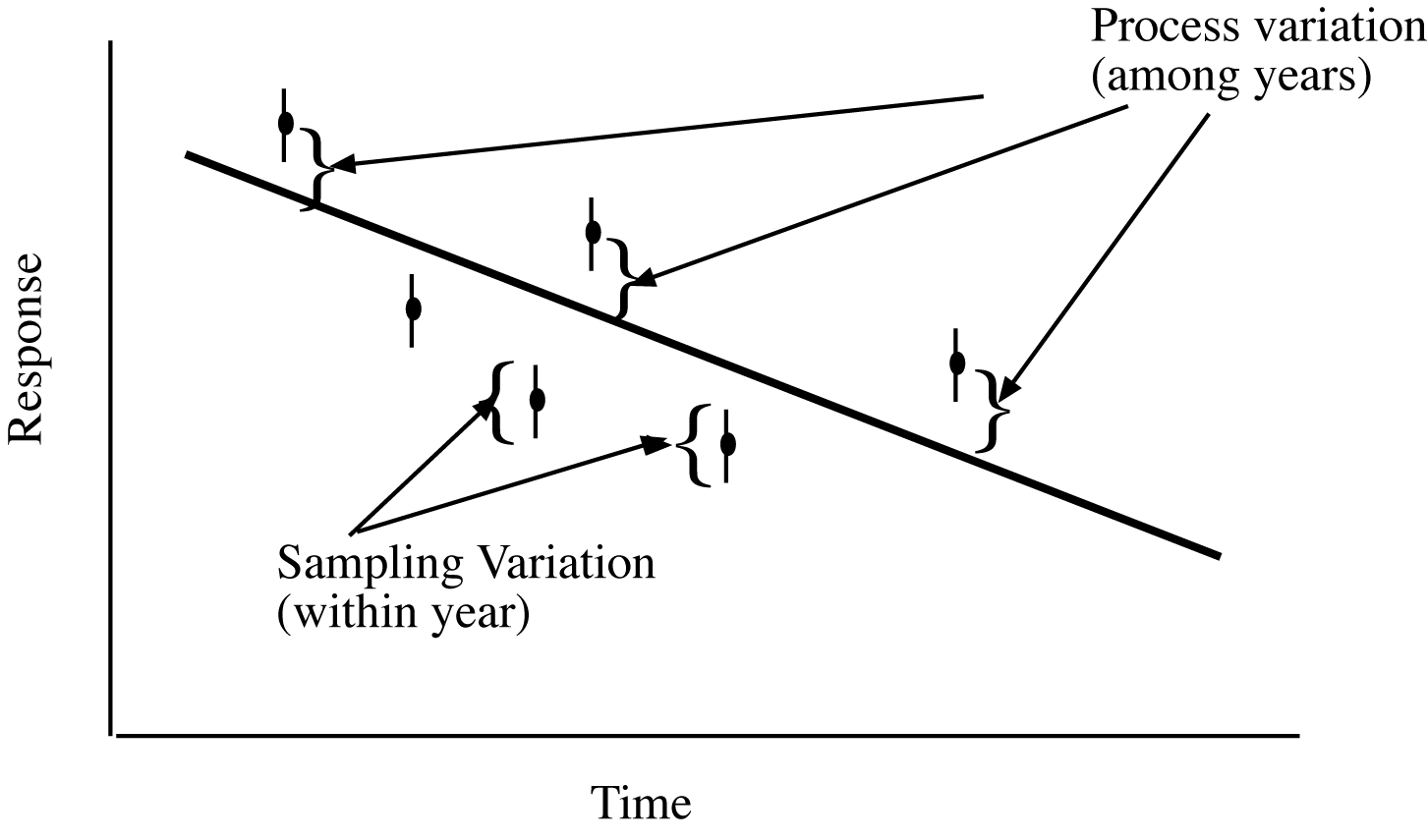


Figure 2. Process error will tend to push all points above or below the underlying trend line.

The impact of year-specific effects (process error) are two fold. First, the individual data points (individual measurements of CPUE) 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 3.

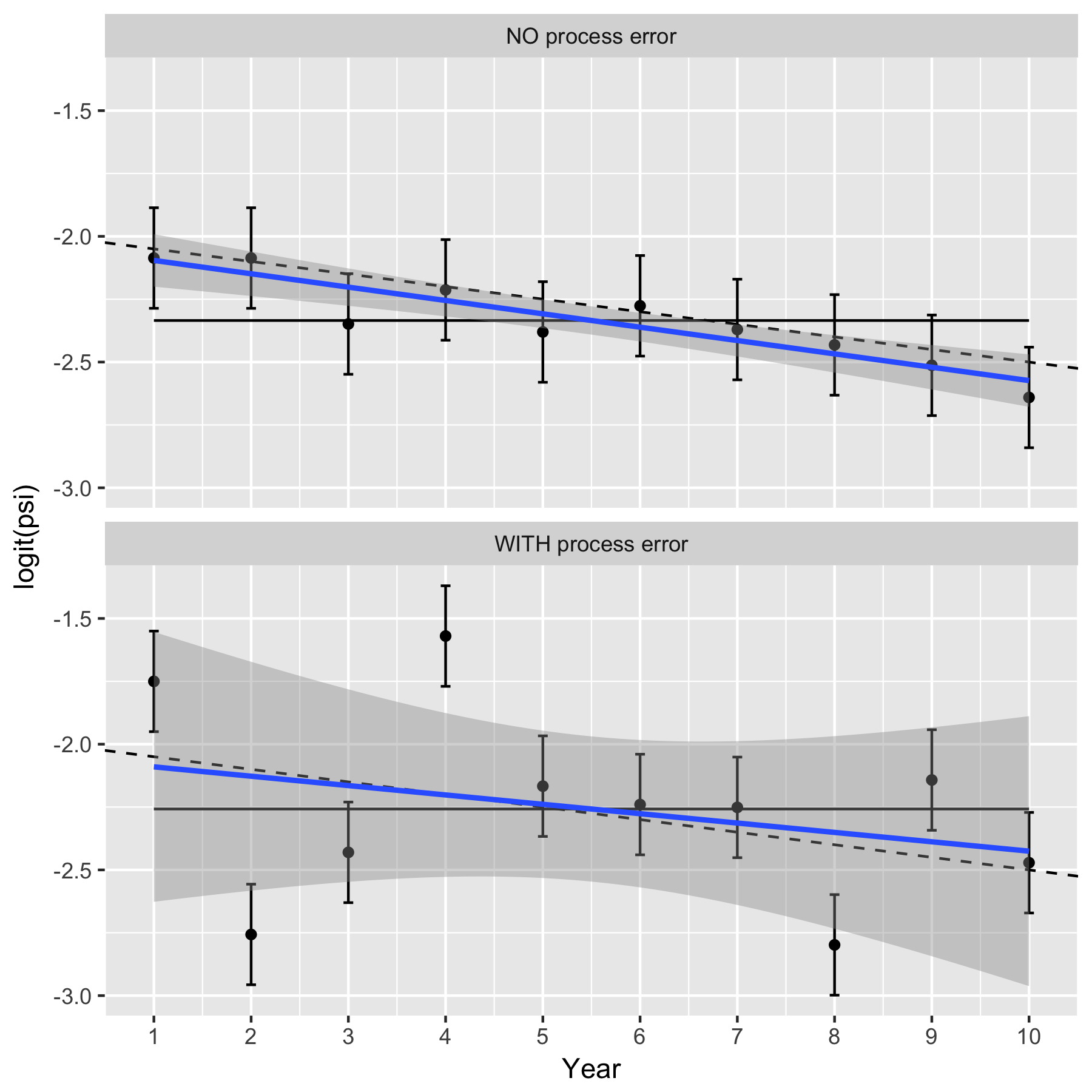


Figure 3. Illustrating the impact of year-specific factors on estimating a trend.

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 (shaded region), and there is clear evidence that the trend is different from a 0 trend (which is shown by the horizontal line at zero falling outside of the shaded area). 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 (shaded region) much larger and now there is no evidence that the trend line differs from 0 (the horizontal line at 0 is within the shaded area).

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.

Data from the MacKenzie, Kawka and Moon Rivers was provided in a set of Excel workbooks with the BLTR cpue /100m computed from samples in a number of years. A preliminary plot of the data (Figure 4) shows an existing trend (but this could not be detected, see below, except for the Kakwa system). A linear mixed model was used to extract the temporal process error (year-specific effects) and the within-year sampling error as shown in Table 1:

Table 1. Summary of mixed linear regression on the logarithm of CPUE (to estimate % change per year). The sampling and process errors measure the variability within and across years as a proportional value, e.g. the variation of CPUE within a year in Kakwa is about 138% of the mean (on average).

System Measure slope slope.se slope.p SD.sampling SD.process

1 Kakwa BLTR\_100m 0.06 0.02 0.02 1.38 0.39

2 MacKenzie BLTR\_100m 0.11 0.04 0.09 1.78 0.40

3 Moon BLTR\_100m 0.05 0.04 0.27 1.35 0.00

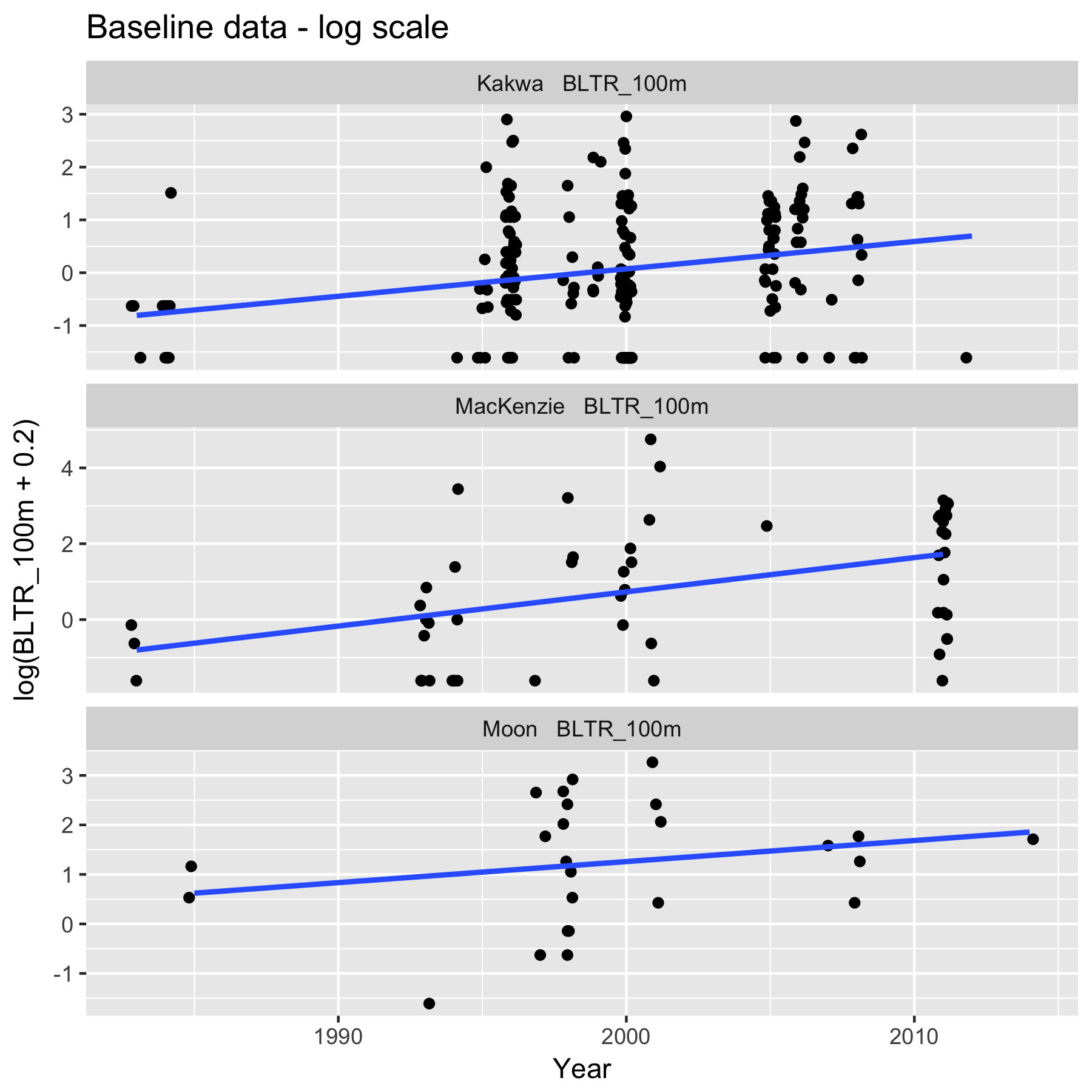


Figure 4. Preliminary plot (on the log-scale with 0.2 fish/100 m added to avoid taking log(0)).

The estimated trends ranged from 5 to 11% increase per year but there was evidence that the trend differed from 0 only for Kakwa (p=.02). More importantly, the sampling and process standard deviations are extracted. Rather surprising, there was no evidence of process error for Moon watershed.

These standard deviations were used to estimate the power for different number of sites/year to detect a 10%, 30%, 50%, 100% and 200% increase over 5 or 10 years. 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% (compounded) 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 detecting this trend given different numbers of sites/year (assuming a simple random sampling design) and different levels of process error (Figures 5a, 5b, 5c).

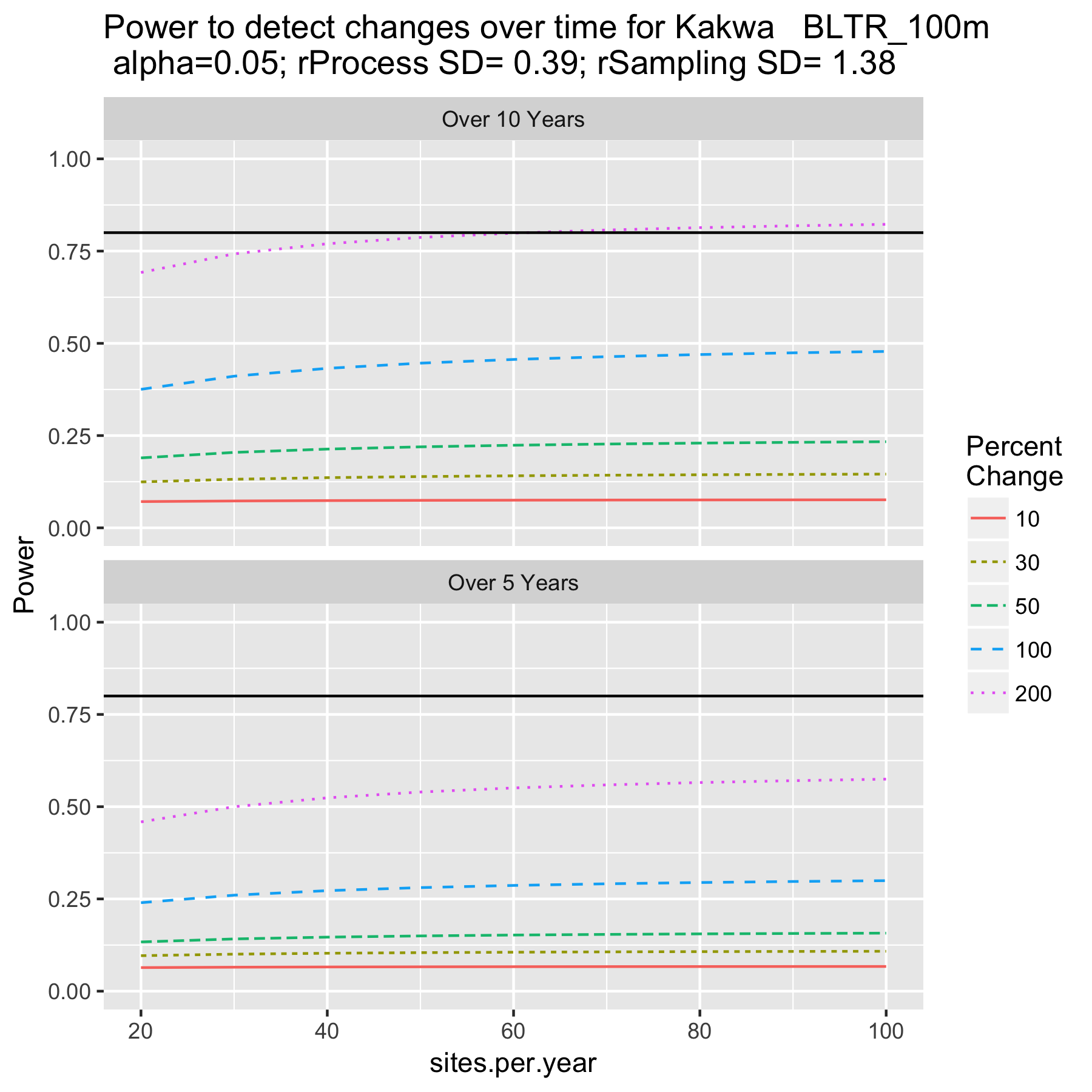


Figure 5a. Power to detect trends at Kakwa.

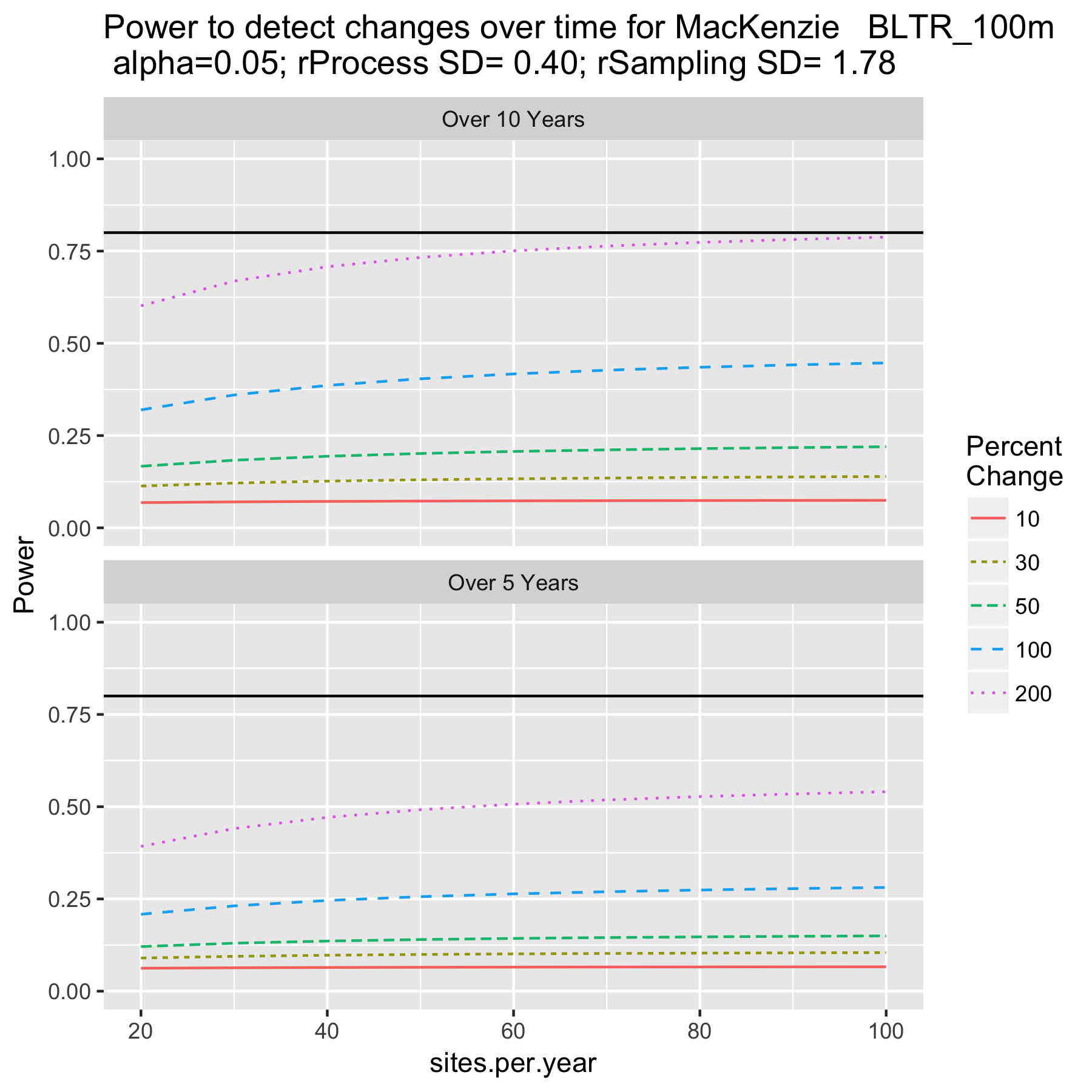


Figure 5b. Power to detect trends at MacKenzie.

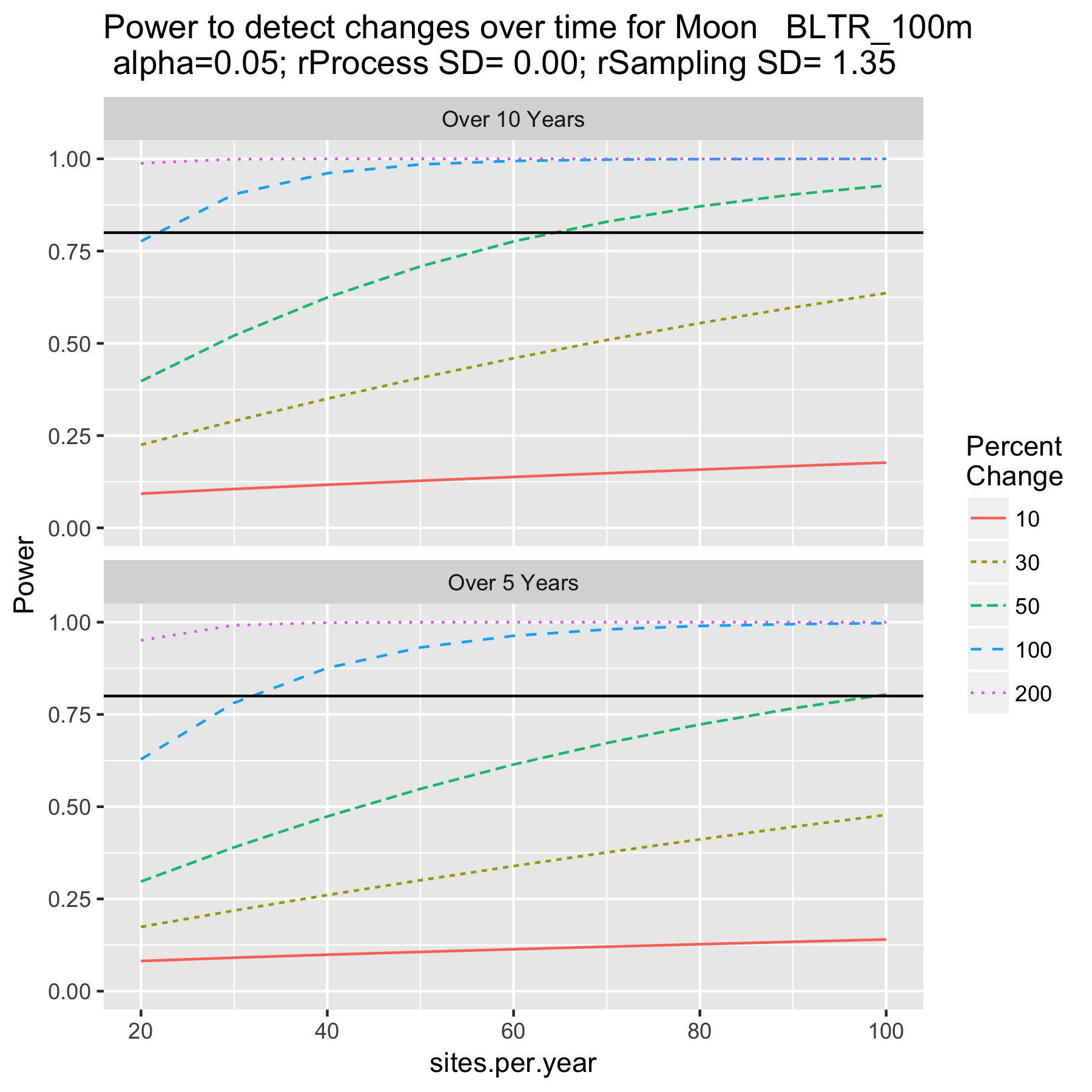


Figure 5c. Power to detect trends in Moon system.

Because of the large process error for the Mackenzie and Kakwa systems, the power is essentially flat as the number of samples increase/year. With large process error, the limiting factor is the number of years of sampling and not the sampling within each year. In no situations for these two systems, could enough sampling be done to detect even the most extreme trends after 5 or 10 years.

Because the estimated process error was 0 for the Moon river system, 30 sites/year would be sufficient to detect the larger trends over 5 or 10 years, but not the smaller trends until well over 100 sites/watershed were sampled.

Once process error is added, the power to detect trends 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.

In order to estimate process error, at least 3 years of sampling are required. IN the absence of this information, a range of process errors from other systems can be used to investigate potential scenarios.

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 5 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.