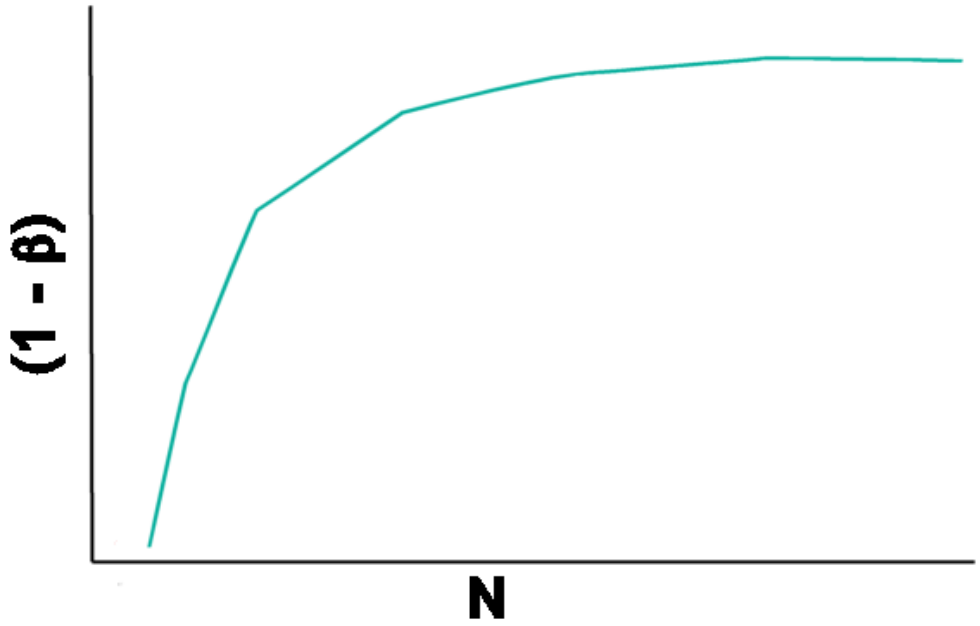
Evaluation of Sample Size and Power Analyses for Trophic Level Data Collected by KTOI Ecosystem Project1



BPA PROJECT 1994-049-00
CONTRACT # 39133
Submitted January 2010

1 Author: **Bahman Shafii, SCS**

Table of Contents

Tables

Figures

Introduction

Modern statistics may be defined as a theory of information with inferences as its objective. Hence, the ultimate objective of any statistical investigation is to make inferences about a population based on information contained in a sample. Since information costs money, the experimenter or manager must determine how much information she/he is willing and able to purchase. Too little information prevents the researcher from making sound estimates about the target parameter(s), whereas too much information results in a waste of resources and money. The quantity of information obtained in the sample depends on the number of items sampled and the amount of variability in the data. Variability can be controlled by the method of selecting the samples, called sampling design or sampling protocol. This, along with sample size, determines the quantity of information in the sample pertinent to the population parameter.

In most research studies, efficient use of resources is mandatory. This is especially true when the study encompasses multiple types of observations taken over space and time. Such is the case with the multiple trophic levels of the Ecosystem project, where sampling of mulitrophic data, i.e. water quality, algae, macroinvertebrate, and fish responses, is carried out over multiple times at multiple sites. Consistent and adequate sampling, both spatial and temporal, is essential to carrying out broader analyses including the investigation of relationships between various trophic levels and their environment. Sample size estimation techniques provide an efficient means of assessing the adequacy and effectiveness of the current sampling protocol. Such assessment can, in turn, help ensure that necessary information is obtained without wasting limited resources.

A related topic, power analysis, assesses the reliability of statistical tests to measure a given hypotheses. Specifically, the power of a statistical test is defined as the probability of rejecting that hypothesis when it is false. That is, it measures the probability of not making a mistake. Sample size can influence this process in two ways: 1) with an inadequate sample size, i.e. limited information, valid conclusions may be missed; or 2) with an excessive sample size, correct conclusions may be gained, but at the cost of wasted resources. Hence, carrying out sample size and power computations are important for answering research questions accurately and efficiently.

Methods

Sample Size

Three pieces of information are required for sample size determination. The first is an estimate of the population variance. This can be estimated from sample statistics such as the data at hand or prior knowledge. The second piece needed is the desired precision (bound on the error of estimation) for future estimates of the population mean. This should be determined by

the researcher(s) and may be specified as either a relative value, such as a percentage of the mean, or an absolute value measured in the units of the specified variables. The final piece of information necessary is the level of confidence (confidence coefficient) desired in the final estimate. Typically, this value is set to a number between 0.9 to 0.99 (e.g. 90 to 99% confidence coefficient). As might be expected, higher levels of confidence will result in larger estimated sample sizes.

The formulation for calculating sample size can be derived from a confidence interval constructed for the population mean and is given by (Cochran, 1977):

$$n = (z*s/d)^2 \quad (1)$$

where n is the estimated sample size, s is the sample standard deviation, d is the desired precision, and z is a tabulated critical value related to the level of confidence and is specified as a quantile of the standard Normal distribution. Given the nature of this value, it should be noted that the above procedure assumes normally distributed data. Failure to meet this assumption will result in inflated (biased) variance estimates, which will in turn produce sample size estimates that may be inaccurate. To meet the assumption of Normality, a transformation of the data may be required prior to sample size estimation.

It should be noted that for all sample size calculations, the resulting sample size values are preliminary, as the calculations are based on available data. Furthermore, the actual precision of mean estimates may vary by the timing or location of sampling. Therefore, sample size estimates should be used cautiously for setting policy regarding future sampling protocols.

Power Analysis

As defined above, power refers to the probability of not making an error in a statistical analysis. There are, however, two types of error that can be made during hypotheses testing. The first, or Type I Error, occurs if a null hypothesis is rejected when it is actually true. The probability of this type of error, denoted by α, is controlled by setting the confidence level of the test to $1 - \alpha$. The second type of error, Type II Error, occurs when the null hypothesis is not rejected when it is actually false. The probability of this error, β, is related to the statistical power of the test through the relationship: power = $1 - \beta$ (Hocking, 1985). The goal of experimental and sampling designs should be to minimize both types of error, thereby maximizing the statistical power, $1 - \beta$. Computations for the power of a given test are carried out by examining the significance of the statistical test while varying the assumed sample size and simultaneously holding the variability and confidence level constant. Following these computations, plots of the computed power versus sample size, i.e. power curves, can be made to investigate the effect of sample size on the power of a given hypothesis.

Scope of Analyses

Statistical Consulting Services (SCS) has provided separate written reports on sample size determination to the Kootenai Tribe of Idaho (KTOI) for trophic level data (algae, macroinvertebrate, and fish) and water quality responses since 2002. For the purpose of this report, only certain years and specific trophic level responses are covered. For additional information and detail, please contact the KTOI Ecosystem project manager, Charlie Holderman.

For the bio-monitoring project, two particular years, one prior (2004) and one post (2007) to initiation of nutrient addition were selected for comparative purposes. Site-specific sample size determination results will be provided for the trophic level responses at a given confidence coefficient (95%) and relative precision level (10% of the specified response mean). Response variables included are the TN/TP ratio for water quality, Total Chlorophyll Accrual Rate ($mg/m^2/30d$) for algae, Abundance ($\#/m^2$) for macroinvertebrates and Catch per Effort (#/hr) for fish, respectively*. For the purposes of this report, the macroinvertebrate analyses will concentrate on the aggregated total abundances over all taxa groups, while the fish analyses will utilize data specific to mountain whitefish (Prosopium williamsoni).

These same response variables are used in presenting the power analyses results, pertinent to a specific pre-determined contrast, namely control vs. nutrient addition sites (KR10 vs. KR9 & KR9.1).

As for the fine-scale project, the power analysis results related to water quality (TN/TP) and algal production (Total Chlorophyll Accrual Rate) responses for the year 2007 will be presented. No comparative results for the year 2004 will be presented in this case, as the aforementioned project was not initiated until 2005. Furthermore, due to the spatial arrangement of sites within the applicable river reach (see the related KTOI report) and high sampling intensity, no sample size determination procedures were deemed warranted.

**** Note:*** Interested readers are referred to the SCS's report, entitled "Ecosystem Project Relational Database" (2009), for further information concerning available trophic level data and complete lists of available sites and response variables. That report is currently available on the Ecosystem project BPA Web site. (Reports available on BPA Pisces system for authorized users, or, contact Charlie Holderman, KTOI.)

Empirical Results

Sample Size Estimation

Water Quality

TN/TP sample size estimates for 2004 and 2007 are given in Table 1. In addition, the corresponding mean TN/TP ratios are provided. All except one sample size estimate over both pre and post nutrient addition years are within the specified sampling protocol of six replications per site and sampling date. The lone exception occurs in 2004 at site KR14 (Wardner, B.C.), which is a non-dam impacted (unregulated) control site. Even in this case the estimated sample size of 7 is nominally acceptable. Although some estimates fall below the current protocol level of six samples, the interpretation should be that six replications per site per sampling date are adequate to meet or exceed the specified precision of 10% of the mean value, with a 95% confidence. For TN/TP, the sampling scheme has been sufficient in both the pre and post nutrient addition years.

Algae

Sample size estimates for Total Chlorophyll Accrual Rate are given with the corresponding site means in Table 2. In order to meet the statistical assumptions of normality required for sample size estimation, a natural logarithmic transformation was utilized on the response prior to analysis. All estimates, in both years, are well below the current sampling protocol of six replications per site per sampling date. Hence, the current sampling regime for estimation of total chlorophyll accrual rates has exceeded the specified precision level of 10% of the mean value with a 95% confidence in both years.

Macroinvertebrate

Macroinvertebrate sample size estimates for Total Abundance (aggregated over taxa groups) are given with the associated mean values in Table 3. To meet the statistical assumptions of normality necessary for analysis, the abundance values were transformed using a natural logarithm function prior to estimation. All sample size estimates in both pre and post nutrient addition years are below the current sampling protocol of six replications per site and sampling date. This sampling intensity has met and exceeded the precision of 10% of the mean value with 95% confidence.

Fish

Table 4 gives the sample size estimates and associated means for Catch per Effort for mountain whitefish. A natural logarithmic transformation was utilized on this response prior to analysis to meet the statistical assumptions of the analysis. For the sites

where fish samples are collected, all sample size estimates meet or exceed the current sampling protocol of six replications per site. Site KR4 in 2007 shows a size estimate of 6, however, it is noted that the abundance values for that year and site were relatively low and, in general, mountain whitefish do not occur at high frequencies in the down river sites. Nonetheless, the sampling protocol of six replications was sufficient in both the pre and post nutrient addition years.

Power Analyses

Power analyses were carried out for a specified hypothesis (contrast) on the responses given above, in addition to TN/TP and Total Chlorophyll Accrual for the Fine-scale project. The contrast was with respect to an Analysis of Variance (ANOVA) procedure designed to compare mean differences between sample sites. For the bio-monitoring project, the specified contrast assessed differences between sites potentially affected by nutrient addition (KR9, KR9.1) and a control site (KR10) above the nutrient addition injection point. These power analyses were assessed in both 2004 (pre nutrient addition) and 2007 (post nutrient addition). In the fine-scale project, the contrast compared mean differences between unaffected sites (KRF0, KRF1) and nutrient affected sites (KRF3, KRF5, KRF7, KRF9, and KRF11). The Fine-scale analyses were only carried out for the year 2007.

Bio-Monitoring Project

Water Quality (TN/TP)

Figures 1 and 2 show the power curves for 2004 and 2007, respectively. As seen in Figure 1, the increase in statistical power, e.g. the probability of seeing a true difference, increases slowly with sample size in 2004, while that of 2007 increases much more rapidly. The actual level of sampling (vertical reference line in the figures) indicates only moderate power (0.40 or 40%) in 2004 and nearly 100% in 2007. Mean values (Table 1) would suggest that levels of the TN/TP response were somewhat lower in 2007 at the nutrient addition sites, while in 2004, differences in TN/TP ratios across sites were minimal, thus making it more difficult to discern statistical significance. Such differences may be attributable to changes in the TN/TP ratio due to nutrient addition, as the addition of phosphorous would be expected to decrease the ratio levels.

Algae (Total Chlorophyll Accrual)

Figures 3 and 4 illustrate the power curves for the 2004 and 2007 algae response, respectively. While the power for total chlorophyll accrual is higher in 2004, both years show very high statistical power at more than 95% given the

current level of sampling (vertical reference line in plots).

Macroinvertebrate (Abundance)

Macroinvertebrate power curves are shown in Figures 5 and 6. The pre nutrient addition year (2004) shows minimal power (<10%) at all sample sizes, while that of 2007 is quite high (near 100%). Mean abundances in 2004 were similar across sites, while those for 2007 show substantial increases in the nutrient addition sites, thereby allowing for higher statistical power and increasing the likelihood of observable differences.

Fish (Mountain Whitefish Catch per Effort)

Figures 7 and 8 give the power curves for Mountain Whitefish Catch per Effort. While both 2004 and 2007 show minimal to moderate statistical power, it is noted that power trends across sample size are increasing at a greater rate in 2007, indicating an improvement in the ability to observe differences in this fish metric relative to nutrient addition.

Fine-scale

W***ater Quality (TN/TP)***

Figure 9 gives the power curve for the Fine-scale water quality TN/TP ratio. Good statistical power (>90%) was found at the current level of sampling.

Algae (Total Chlorophyll Accrual)

Statistical power for total chlorophyll accrual is given in Figure 10. Statistical power for this response is high at more than 90%, making it quite likely that the statistical differences between nutrient addition and control sites will be detectable.

Concluding Remarks

Sample size analyses were carried out for the KTOI Ecosystem Project in both the pre and post nutrient addition years (2004 and 2007, respectively). In all cases, the current sampling scheme of six replications was found to either meet or exceed the requirements for a precision level of 10% of the mean and a 95% confidence coefficient. Hence, the Ecosystem project is currently sampling all trophic levels with reasonable effort and should strive to maintain these sampling intensities in the future.

Power analyses (2004 and 2007 for the bio-monitoring project; 2007 for the fine-scale project) demonstrate an increased level of observable response at the current level of sampling for most trophic levels. Those trophic level responses for which statistical power is limited, will need to be monitored closely, however, some of these responses show improvements in detect ability post nutrient addition. Continued sampling at the current levels should ensure detectable response differences due to nutrient addition in future years.

References

Cochran, W. G. (1977), Sampling Techniques, Third Edition, New York, NY: John Wiley & Sons.

Hocking, R.R. (1985), The Analysis of Linear Models, Monterey, CA: Brooks/Cole Publishing Company.

	Year	
Site	2004	2007
KR1	3	3
KR2	3	2
KR3	2	2
KR4	5	3
KR6	2	2
KR7	3	1
KR9	2	1
KR9.1	2	2
KR10	2	2
KR11	2	2
KR12	4	5
KR13	3	1
KR14	7	6

Sample Size

	Year	
Site	2004	2007
KR1	22.29	13.04
KR2	25.49	17.60
KR3	21.62	14.96
KR4	23.16	16.33
KR6	29.55	23.46
KR7	32.57	24.75
KR9	34.75	22.39
KR9.1	31.45	27.36
KR10	34.50	28.89
KR11	32.66	31.84
KR12	33.13	28.50
KR13	39.46	40.33
KR14	13.87	9.18

Means

Table 1. Sample size estimates and associated means for the water quality response, TN/TP, at 13 sample sites in the years 2004 and 2007. Sample sizes determined at a confidence level of 95% and a precision of 10% of the mean value.

	Year	
Site	2004	2007
KR1	1	1
KR2	1	1
KR3	1	1
KR4	1	1
KR6	1	1
KR7	1	1
KR9	1	1
KR9.1	1	1
KR10	1	1
KR11	2	1
KR12	1	1
KR13	1	1
KR14	1	1

Sample Size

	Year	
Site	2004	2007
KR1	23.51	41.62
KR2	13.33	46.02
KR3	22.05	49.10
KR4	26.72	71.64
KR6	2.09	22.09
KR7	4.40	27.06
KR9	3.15	20.56
KR9.1	4.48	26.00
KR10	0.94	34.19
KR11	2.12	31.57
KR12	1.01	22.95
KR13	1.68	17.06
KR14	12.81	43.09

Means

Table 2. Sample size estimates and associated means for the algae response, Total Chlorophyll Accrual ($mg/m^2/30d$), at 13 sample sites in the years 2004 and 2007. Sample sizes determined at a confidence level of 95% and a precision of 10% of the mean value. Sample size estimates are based on a natural logarithmic transformation of the data.

	Year	
Site	2004	2007
KR1	1	2
KR2	2	3
KR3	3	2
KR4	2	3
KR6	1	1
KR7	1	1
KR9	1	1
KR9.1	1	1
KR10	1	1
KR11	1	1
KR12	1	1
KR13	1	1
KR14	1	1

Sample Size

	Year	
Site	2004	2007
KR1	1169	611.4
KR2	1473	3203
KR3	1688	7959
KR4	767.9	1268
KR6	1695	14959
KR7	1130	12490
KR9	886.5	11609
KR9.1	1295	26194
KR10	1406	2566
KR11	2718	8866
KR12	3476	9087
KR13	10131	28456
KR14	1283	1467

Means

Table 3. Sample size estimates and associated means for the macroinvertebrate response, Abundance (#/m^2), at 13 sample sites in the years 2004 and 2007. Sample sizes determined at a confidence level of 95% and a precision of 10% of the mean value. Sample size estimates are based on a natural logarithmic transformation of the data.

	Year	
Site	2004	2007
KR2	2	1
KR4	2	6
KR6	1	1
KR9	4	2
KR10	2	1
KR14	1	1

Sample Size

	Year	
Site	2004	2007
KR2	14.24	7.633
KR4	25.27	13.92
KR6	131.4	462.8
KR9	63.63	180.1
KR10	59.2	115.2
KR14	301.1	114.8

Means

Table 4. Sample size estimates and associated means for the fish response, Catch per Effort (#/hr), at 6 sample sites in the years 2004 and 2007. Sample sizes determined at a confidence level of 95% and a precision of 10% of the mean value. Sample size estimates are based on a natural logarithmic transformation of the data.

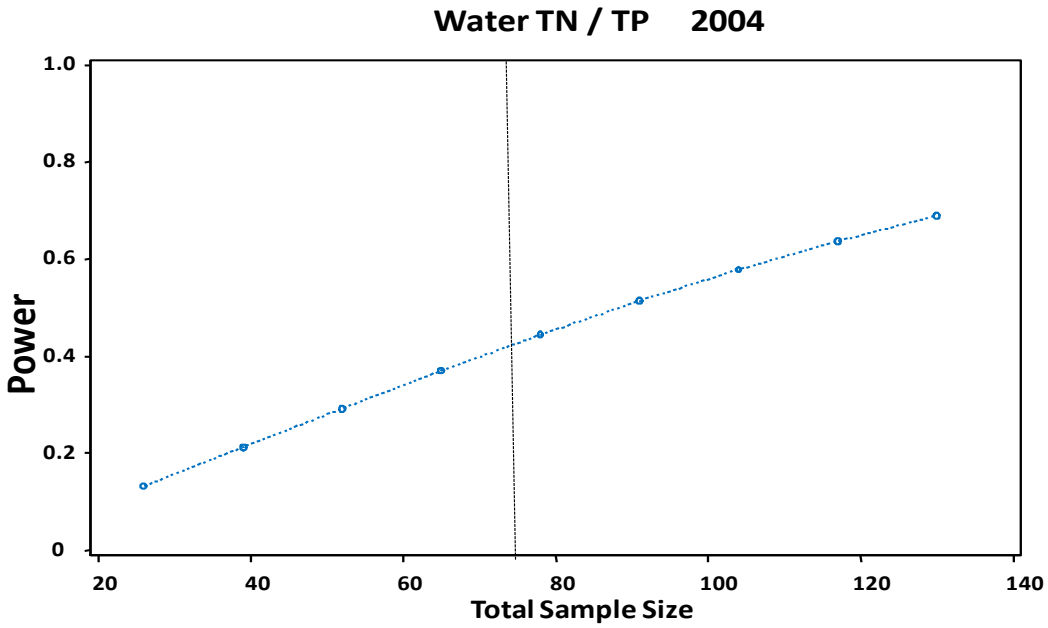


Figure 1. Estimated power curve (blue line) for a contrast of the water TN/TP ratio of the treated zone (KR9, KR9.1) versus the adjacent control site (KR10) in 2004. Black dashed line represents the actual total sample size taken.

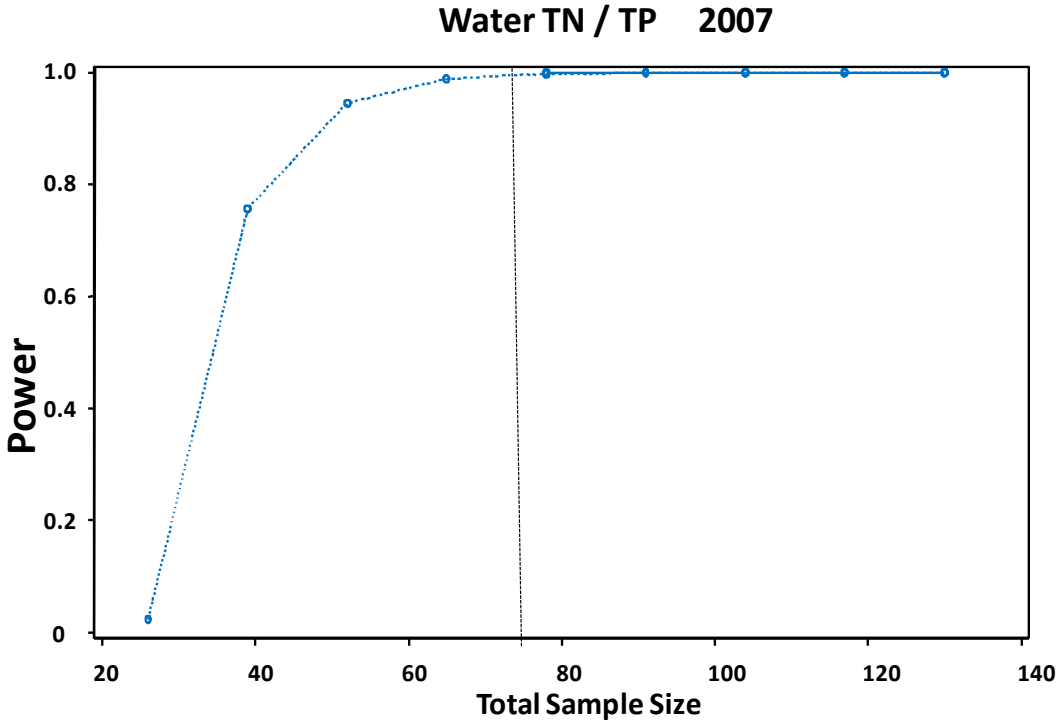


Figure 2. Estimated power curve (blue line) for a contrast of the water TN/TP ratio of the treated zone (KR9, KR9.1) versus the adjacent control site (KR10) in 2007. Black dashed line represents the actual total sample size taken.

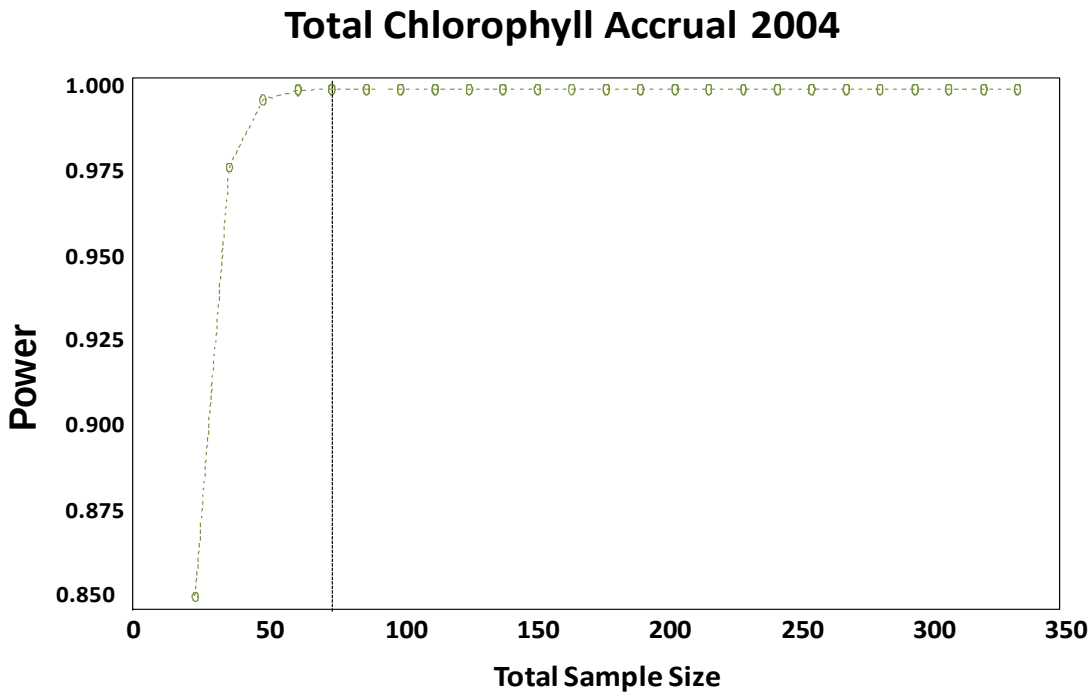


Figure 3. Estimated power curve (green line) for a contrast of the algal total chlorophyll accrual rate of the treated zone (KR9, KR9.1) versus the adjacent control site (KR10) in 2004. Black dashed line represents the actual total sample size taken.

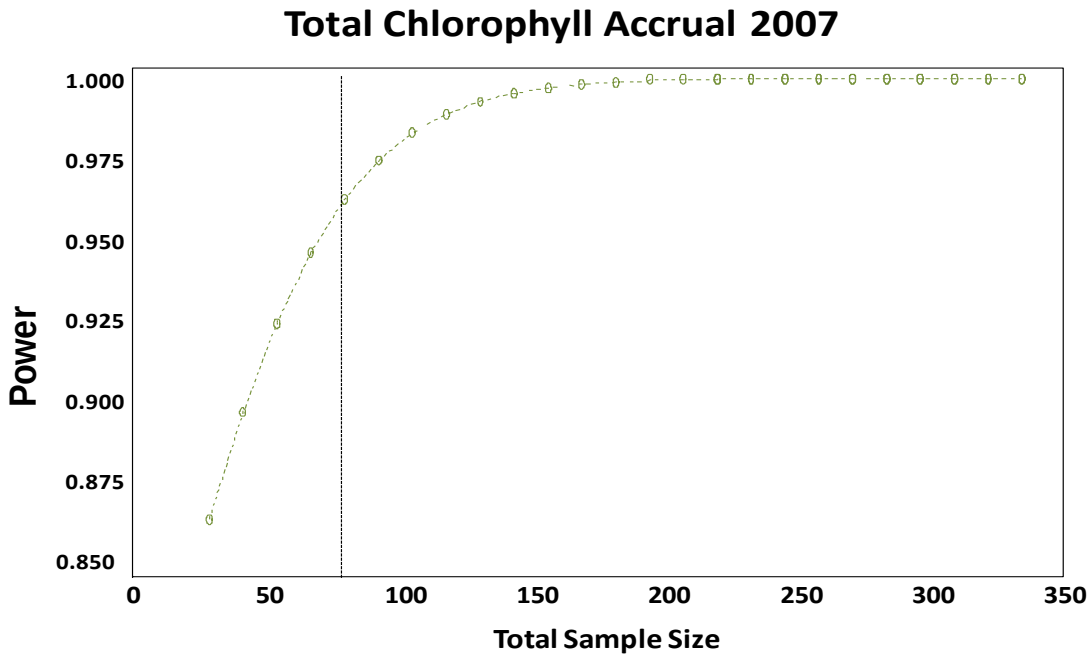


Figure 4. Estimated power curve (green line) for a contrast of the algal total chlorophyll accrual rate of the treated zone (KR9, KR9.1) versus the adjacent control site (KR10) in 2007. Black dashed line represents the actual total sample size taken.

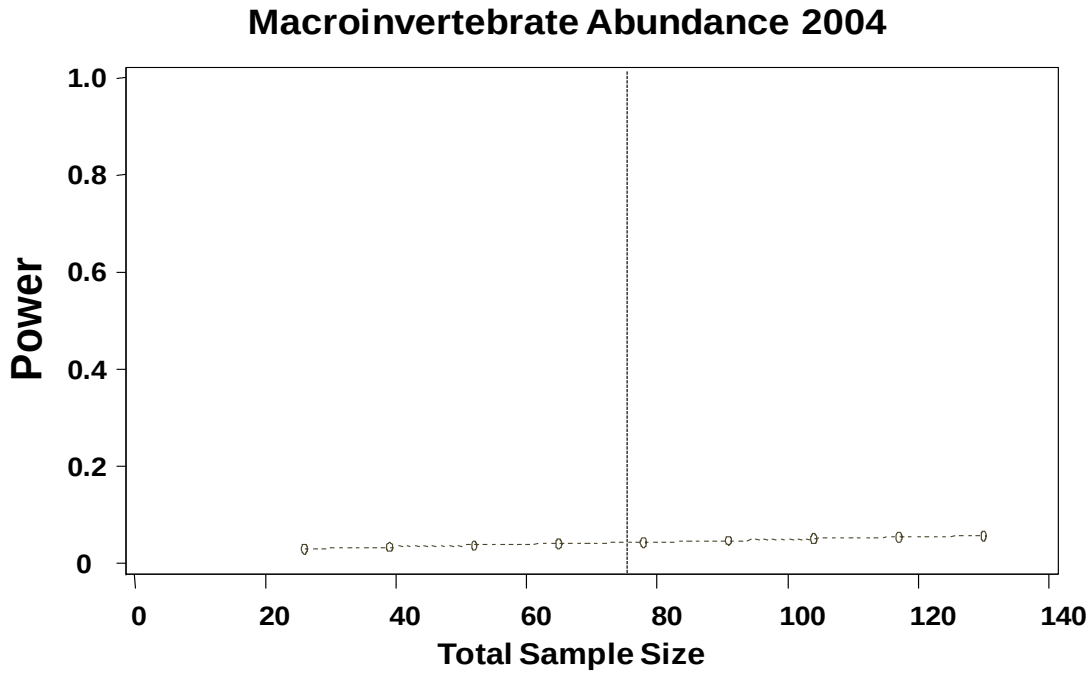


Figure 5. Estimated power curve (brown line) for a contrast of the macroinvertebrate abundance of the treated zone (KR9, KR9.1) versus the adjacent control site (KR10) in 2004. Black dashed line represents the actual total sample size taken.

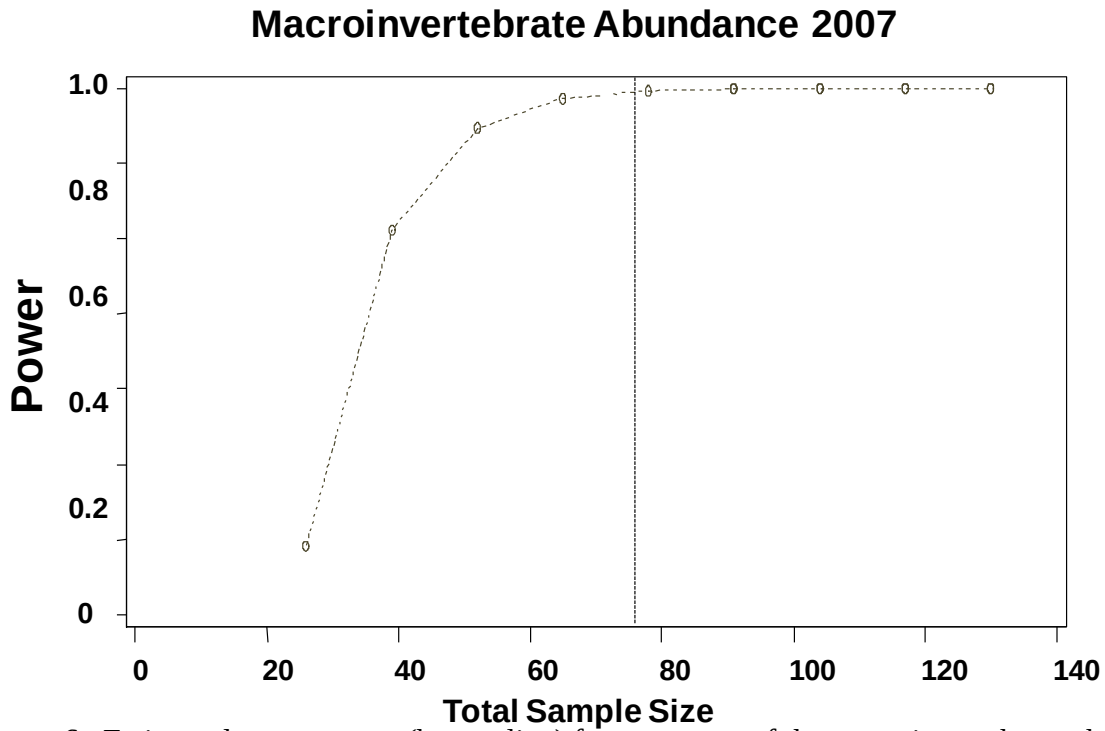


Figure 6. Estimated power curve (brown line) for a contrast of the macroinvertebrate abundance of the treated zone (KR9, KR9.1) versus the adjacent control site (KR10) in 2007. Black dashed line represents the actual total sample size taken.

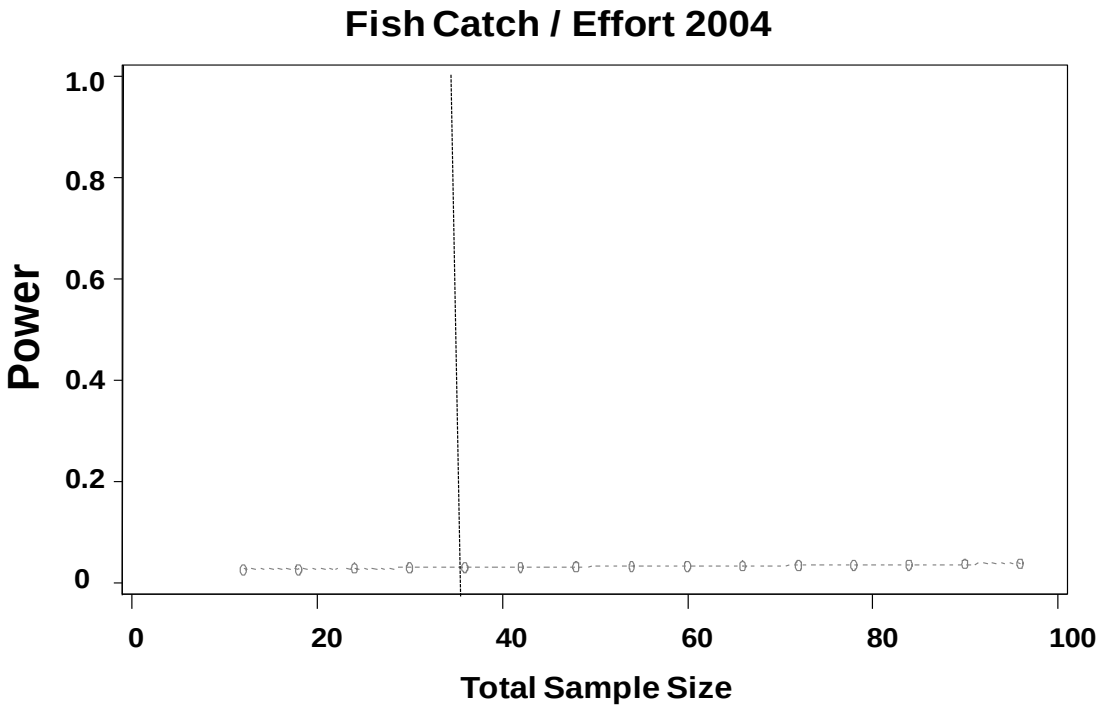


Figure 7. Estimated power curve (gray line) for a contrast of the fish catch per effort of the treated zone (KR9, KR9.1) versus the adjacent control site (KR10) in 2004. Black dashed line represents the actual total sample size taken.

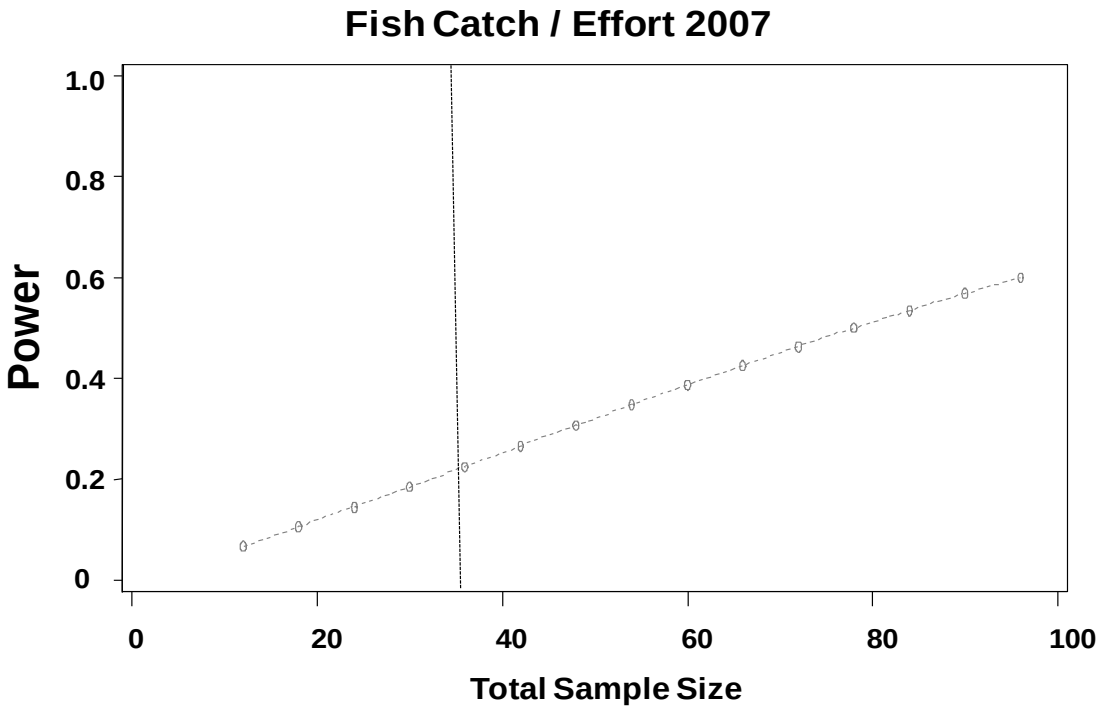


Figure 8. Estimated power curve (gray line) for a contrast of the fish catch per effort of the treated zone (KR9, KR9.1) versus the adjacent control site (KR10) in 2007. Black dashed line represents the actual total sample size taken.

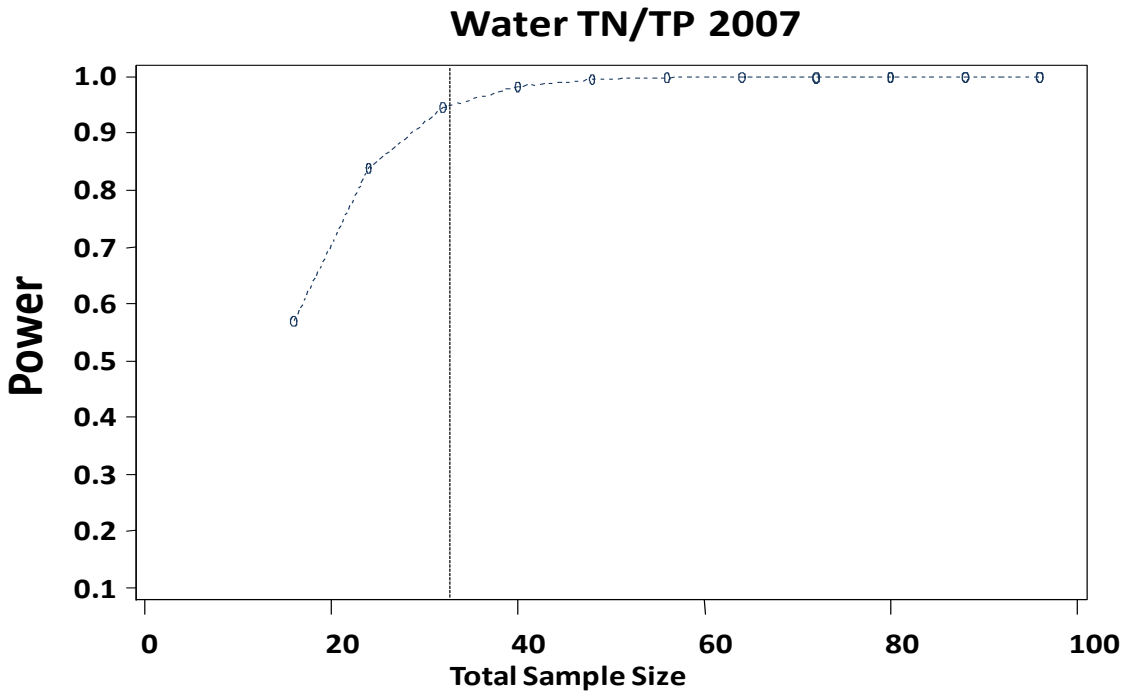


Figure 9. Estimated power curve (blue line) for a contrast measuring the Fine-scale project treatment effect (KRF0, KRF1 versus the remaining sites, KRF3-KRF11) on the TN/TP water quality ratio in 2007. Black dashed line represents the actual total sample size taken.

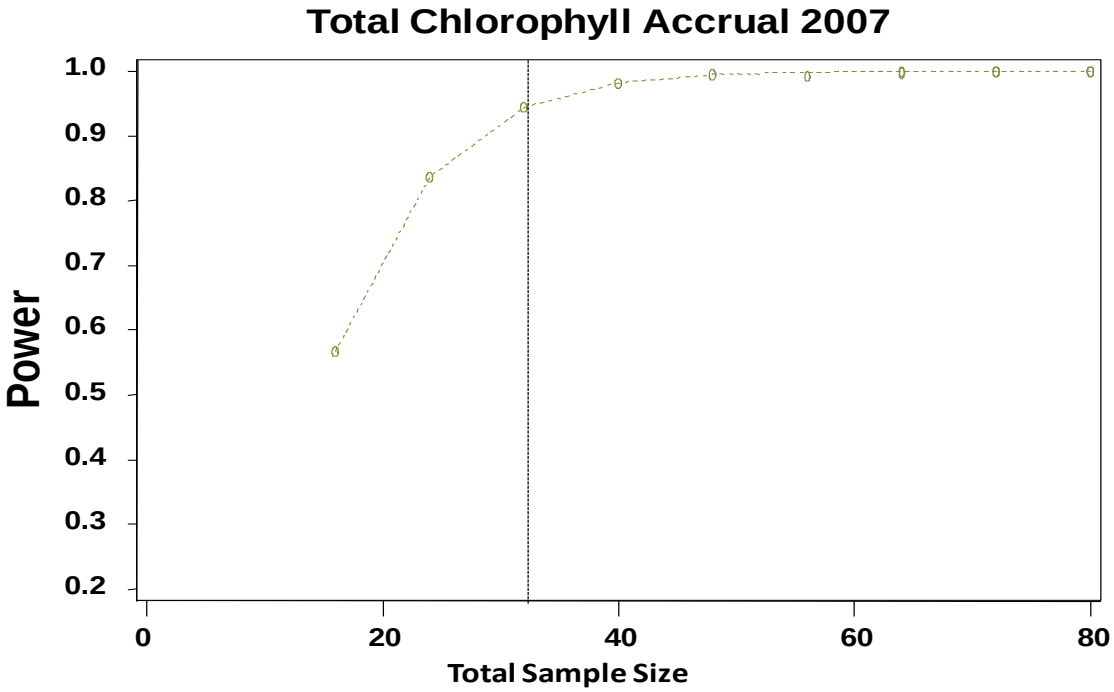


Figure 10. Estimated power curve (green line) for a contrast measuring the Fine-scale project treatment effect (KRF0, KRF1 versus the remaining sites, KRF3-KRF11) on the total chlorophyll accrual rate in 2007. Black dashed line represents the actual total sample size taken.