

THE HABITAT CONDITION ASSESSMENT
POWER ANALYSIS TOOL

FOR

NATURAL ENGLAND
CONTRACT NO: 24162

BY

SIMON M.C. POULTON

BIOECOSS LTD.
31ST MARCH 2012



SUMMARY OF MAJOR FINDINGS AND RECOMMENDATIONS

- Σ An interactive Power Analysis Tool was developed in MS Access 2007, based on simulation algorithms run in real-time.
- Σ This tool allows the user to specify values for a number of population, sampling and analysis parameters to generate estimates of power, significance and sample size required to detect change.
- Σ The tool has two main functions; a) to generate theoretical power curves in a graphical format and, b) to calculate sample sizes based on empirical data from the existing SSSI monitoring programme.
- Σ A full user guide to the tool is provided.
- Σ The power tool was used to explore the effects of different parameters on power to detect change. Most importantly, this showed the very strong effects of two population parameters (Initial Proportion and Degree of Change). It also emphasised the greatly increased power obtained by using a dependent (repeat-visit) sampling strategy rather than an independent strategy, but also showed that this was strongly influenced by the amount of turnover in the underlying population.
- Σ The power tool was also used to explore the likely sample sizes required to detect change in the SSSI monitoring programme. It showed that a sample of around 560 sites (0.6% sampling fraction) could provide a significant estimate of overall change, given reasonable values for the parameters. This sample could be reduced to around 100 if dependent sampling was used. However, the analysis also showed that simple one-way and two-way breakdowns, to give change estimates for individual categories, increased the required sample sizes greatly.
- Σ Finally, the lack of intrinsic power in the use of a binary response variable was explored. Two example datasets were compared, identical in all respects except that one had a binary response variable and the other an ordinal 5-point scale. The significance of the change detected in the binary dataset was $p = 0.045$, whereas for the ordinal response variable it was $p \ll 0.001$. This result suggested that a tally of favourable attributes from the CSM methodology would be a much more powerful response variable than just the overall site condition.



CONTENTS

Summary of Major Findings and Recommendations.....	3
Contents.....	4
Tables	5
Figures	5
Introduction.....	1
The Concepts Behind the Power Analysis Tool	2
<i>The Parameters</i>	2
<i>The Simulation Algorithms</i>	4
<i>The Statistical Models</i>	4
User Guide to the Power Analysis Tool	6
<i>Installation & Running the Power Analysis Tool</i>	6
<i>The Main Menu</i>	6
<i>Graphical Analysis of Theoretical Data</i>	7
<i>Tabular Analysis of Empirical Data</i>	10
Using the Power Analysis Tool to Explore the Effects of Parameters on Power.....	14
Introduction.....	14
<i>The Basic Relationship Between Significance and Power</i>	14
<i>The Basic Relationship Between Power and Sample Size</i>	14
<i>The Influence of Significance on this Relationship</i>	15
<i>The Relationship Between Population and Sample Size on Power</i>	15
<i>The Relationship Between Initial Proportion and Sample Size on Power</i>	16
<i>The Relationship Between Change and Sample Size on Power</i>	17
<i>The Effect of Sampling Strategy on Power</i>	17
<i>The Effect of Turnover on Dependent Samples</i>	18
<i>The Effect of Sample Adjustment on Independent Samples</i>	19
<i>The Effect of One-tailed Testing on Power</i>	19
Exploratory Power Analysis for Condition Assessment.....	21
Introduction.....	21
<i>Calculation of a Simple Random Sample</i>	21
<i>One-way Tabulation to Create a Stratified Random Sample</i>	22
<i>Two-way Tabulation</i>	24
<i>Modifying the Fixed Parameters</i>	24
<i>Summary & Conclusions</i>	25
Appendices.....	28



TABLES

Table 1. Proportions of parcels in favourable condition in each of the 19 Priority BAP Habitats.	21
Table 2. Summary of parameter effects on simple random samples. The starting conditions give the basic sample size, followed by the effect on sample size of making one or two changes to various parameters.	22
Table 3. Sample sizes from a one-way tabulation with constant initial proportions for six coastal BAP habitats. a) α -level set to 0.05 to represent six non-independent tests at a significance of 95%. b) α -level set to 0.01 to represent six independent tests at approximately 95% significance.	23
Table 4. Sample sizes from a one-way tabulation with underlying initial proportions for six coastal BAP habitats. The α -level was set to 0.05 to represent six non-independent tests at a significance of 95%.	24
Table 5. Sample sizes from a two-way tabulation by Feature and Qualification for habitat features with sampling fractions less than 1.0.	24
Table 6. The effect of varying the fixed parameters on sample sizes from a one-way tabulation by habitat.	25
Table 7. Mean sample sizes from two filtered samples; a) a simple random sample and b) a one-way tabulation by habitat.	25

FIGURES

Figure 1. Graph definition zone on the Graphical Analysis form.	7
Figure 2. Graph display and management zone on the Graphical Analysis form.	9
Figure 3. Tabulation Definition zone on the Tabulation Analysis form.	10
Figure 4. Multiple categories used to filter the underlying parcel records.	11
Figure 5. Table Display and Management zone on the Tabular Analysis form.	11
Figure 6. Multiple categories used to filter the underlying site records.	12
Figure 7. The relationship between significance and power.	14
Figure 8. The relationship between power and sample size.	15
Figure 9. The influence of significance on the relationship between power and sample size.	15
Figure 10. The relationship between population and sample size on power.	16
Figure 11. The relationship between initial proportion and sample size on power.	16
Figure 12. The relationship between change and sample size on power.	17
Figure 13. The influence of initial proportion on the relationship between change and sample size on power.	17
Figure 14. The effect of sampling strategy on power.	18
Figure 15. The effect of turnover on dependent samples.	18
Figure 16. The effect of sample adjustment on independent samples.	19
Figure 17. The effect of one-tailed testing on power.	19
Figure 18. Means plots (with 95% confidence intervals) from a two-way logistic GLM of a) habitat, b) time and c) their two-way interaction.	26
Figure 19. Frequency histograms of two random samples (N = 79) drawn from populations of a 5-point ordinal scale with means of 1.25 and 2.0.	27



Introduction

The specification for this contract required a power analysis to be carried out, exploring the feasibility of a sample-based approach to habitat condition assessment. Three objectives were defined:

- Σ To better understand and assess our capability to detect change in habitat condition at a national scale and at a more detailed local landscape scale.
- Σ To target in-house staff resource to ensure we carry out sufficient detailed Integrated Site Assessments across SSSI and HLS agreements to allow us to detect change across each priority habitat at both a national and landscape scale.
- Σ To inform external partnerships to ensure they survey sufficient number of sites outside SSSI and HLS to allow us to detect change across each priority habitat at both a national and landscape scale.

The main problem with delivering this type of power analysis is the large number of parameters that have to be incorporated and the huge number of combinations of their values. A “one-off” power analysis can be very useful in indicating the range of powers available given certain sampling parameters. However, it is impossible to incorporate all values of all relevant parameters, so there is always the risk of providing answers for certain scenarios but not others.

To attempt to overcome these problems BioEcoSS has developed an interactive power analysis tool. This gives the user three huge advantages over a passive analysis. Firstly, the user can explore the effects of different parameters on the power of a sampling strategy. For example, by taking a range of sample sizes and rates of change in the underlying population, it is possible to compare the relative effects of these and their interaction. Secondly, certain population parameters, such as the intrinsic rate of change will be unknown. Even if a pilot study has been carried out, this is only likely to give a general indication of the nature of change, so there is always a risk of running a power analysis with incorrect assumptions. Thirdly, it allows the user to undertake “what if” type analyses. This is especially useful for the unknown parameters, but even those under user control can be difficult to assess. So, for example, it would be possible to ask a question such as “How does the power of the analysis change if I set a lower significance level such as 90%, compared to 95%?”.

This report is divided into four main sections:

- Σ **The Concepts Behind the Power Analysis Tool** which explains how the tool works, including the underlying statistical models and the simulation algorithms.
- Σ **A User Guide to the Power Analysis Tool** which guides the user through the use of the tool.
- Σ **Using the Power Analysis Tool to Explore the Effects of Parameters** which addresses the second and fourth of the objectives given above.
- Σ **An Exploratory Power Analysis for Habitat Condition Assessment** which addressed directly the SSSI stock in Wales and attempts to answer the questions posed in the first and third objectives.



The Concepts Behind the Power Analysis Tool

The power analysis tool is designed to explore monitoring schemes based on the Common Standards Monitoring approach. This yields site condition assessments which are essentially binary; Favourable or Unfavourable. In a sample of sites or features, this would be represented as the proportion in favourable condition. Although the method includes “sub-conditions” such as Favourable Declining or Favourable Improving, these have not been built explicitly into the tool. The other main constraint of the tool is that it only allows monitoring data from one time interval, *i.e.* from time 0 (t_0) to time 1 (t_1). Although these may seem like serious constraints at first, careful use of the tool will allow them to be largely circumvented and their inclusion would have increased the complexity of the tool enormously.

The power analysis tool is divided into two main utilities:

- Σ Graphical Analysis of Theoretical Data.
- Σ Tabular Analysis of Empirical Data.

The graphical analysis allows the user a more general approach to power analysis in a number of ways. Firstly, it gives complete control over all of the input parameters and their values. Secondly, it allows a more flexible output by plotting one or more power curves. Thirdly, it gives the user the choice of analysing power or significance, or the relationship between the two.

The tabular analysis gives the user a more empirical approach, firstly, by utilising known values for population size and initial proportions. These data are stored in underlying tables and cannot be modified by the user. In this sense, it is more prescriptive than the graphical analysis, but provides more realistic values. Secondly, it uses known categorisations of sites to allow filtering and breakdowns. Finally, it produces a tabular output of estimated sample sizes, given the parameters specified.

The Parameters

Up to ten parameters can be defined for each analysis. These are divided into three main categories; population, sample and analysis parameters.

There are four population parameters, which define characteristics of the population from which the sample is drawn. All of these are extrinsic parameters, in that they are beyond the control of the surveyor.

- Σ **Population** represents the size of the sampling frame. It is relevant because when the population is small relative to the size of the sample (or, more accurately, the sample is a large proportion of the population) a finite population correction can be applied. This has the effect of reducing standard errors, sometimes by a considerable degree which, in turn, can reduce the sample size required for a given power.
- Σ **Initial Proportion** is the proportion of the underlying population in favourable (or any other binary) condition at t_0 .
- Σ **Change** is the degree of change from the initial proportion of the population at t_0 to a new proportion at t_1 . It is expressed as the proportionate change in the odds. It is very important to understand that it is not the amount of change in absolute percentage terms. So, for example, take an initial proportion of 50% and apply a change of +50%. This does not result in a proportion of 100%. Instead, a proportion of 50% gives odds of 1 (because $0.5 / [1 - 0.5] = 1$). Increasing this by 50% (*i.e.* multiplying by 1.5) gives odds of 1.5. Converting these odds back to a proportion gives 60% (because $1.5 / [1 + 1.5] = 0.6$). This somewhat counter-intuitive result is essential, because it is impossible to express percentage changes in proportions correctly. What would have happened, for example, if we had started with a proportion of 75% and increased this by 50% in absolute terms. If we added 50% to 75% we would achieve 125% and if we multiplied by 1.5 we would achieve 150%, both of which are impossible values – we cannot have



125% of sites in favourable condition. The only way to apply standard rates of change to proportions is to apply the changes to their odds.

- Σ **Turnover** is only applicable to sampling strategies that are dependent, *i.e.* they use repeat-visits to the same sites. In independent sampling strategies, this parameter is irrelevant. It represents the proportion of changing sites that exhibit a “reverse change”. For example, imagine a population of 100 sites with an initial proportion of 50%; 50 sites would be in favourable and 50 in unfavourable condition. Applying a +50% change (as explained above) would result in 60 sites being in favourable condition. With 0% turnover, this would be achieved by 10 of the unfavourable sites becoming favourable. However, imagine 11 sites becoming favourable, and one favourable site becoming unfavourable. This would still represent a 60% final proportion in favourable condition, but there would be 10% (1 / 10) turnover. Obviously 100% turnover is the maximum possible, when 20 site become favourable, but 10 sites become unfavourable. Turnover can be thought of as representing the “stability” of a population.

There are three sample parameters that are essentially under the control of the surveyor and so can be considered intrinsic parameters.

- Σ **Strategy** is simply a binary parameter defining either a dependent or independent sampling strategy. Dependent strategies re-visit the same sample of sites at $t1$, whereas independent strategies draw an entirely new sample of sites from the population at $t1$. This has a number of consequences. Firstly, as explained above, turnover only applies to dependent strategies. Secondly however, independent strategies have a different parameter that comes into effect; sample adjustment, which is explained below. Thirdly, different statistical models must be used for analysing these two strategies, which are also explained in more detail below.
- Σ **Sample size** is the parameter that is generally of most interest in a power analysis and is entirely under the control of the surveyor. Clearly, the only constraint on sample size is that it cannot be larger than the population.
- Σ **Sample Adjustment** is only applicable to independent sampling strategies. It represents the adjustment in the size of the sample drawn at $t1$ compared to the sample drawn at $t0$. For example, if the sample at $t0$ was 100 with a -10% adjustment, the sample at $t1$ would be 90. This parameter may be considered either intrinsic or extrinsic, depending on the reason for the two sample sizes being different. The most common reason would be that available resources have changed, so that the second sample has to be smaller (rarely larger!). An extrinsic example would be if the population size had changed greatly, but it was desirable to maintain the same sampling fraction.

The final category has three parameters, all of which are defined explicitly by the surveyor as they are the analysis parameters.

- Σ **Tails** is simply a binary parameter representing one or two-tailed significance testing. If the direction of change is specified as part of the monitoring programme then one-tailed testing, which is intrinsically more powerful, can be used. For example, if the monitoring programme is undertaken to assess the effect of applying management practices designed to increase the proportion of favourable sites, then the one-tailed option can be selected. However, if a surveillance programme is being undertaken, with no *a priori* expectation of either an increase or decrease, then two-tailed testing should be employed. It would be invalid to wait until the direction of change in the sample is detected and then apply a one-tailed test.
- Σ **Power** is the only parameter that is purely predictive, in the sense that it is only of *a priori* value. The power of a statistical test is the probability of rejecting the null hypothesis when it is actually false and so should be rejected. Clearly, the higher the power the better! It is the complement of β (where $\beta = 1 - \text{power}$) which is the probability of committing a so-called Type II error – failing to reject the null hypothesis when it is actually false. Traditionally, this value is often set at 90%, although the power analysis tool allows the user to set any value between 50% and 99% ($\beta = 0.5$ to 0.01).



- Σ **Significance** is the final parameter and is entirely within the control of the survey. Indeed, it must be set prior to any statistical testing of empirical data. It represents the likelihood of correctly accepting a null hypothesis that is in fact true. It is the complement to α (where $\alpha = 1 - \text{significance}$) which is the probability of committing a Type I error – rejecting a null hypothesis when it is actually true. Traditionally, this value is often set at 95%, although the power analysis tool allows users to set any value between 80% and 99.9% ($\alpha = 0.2$ to 0.001)

In the Graphical Analysis utility, all of these parameters can be defined by the user, either as fixed (scalar) values or to be plotted on the graph. However, in the Tabular Analysis utility, two of the parameters (Population and Initial Proportion) are provided by the empirical data and Sample Size is always calculated by the analyses.

The Simulation Algorithms

The tool uses a simulation approach to power analysis. This exploits the immense power of modern desktop computers to undertake thousands of simulations in real-time, from which significance levels can be derived. Both utilities use the same underlying simulation algorithms:

- Σ The three (or four if dependent sampling is defined) population parameters are used to create two populations of binary values in a $2 \times n$ array, where n is the **population size**.
 - α Firstly, the population at t_0 is created as a vector of zeros and ones representing the **initial proportion**. So, if the population is 500 and the proportion is 50%, there will be 250 zeros and 250 ones.
 - α Secondly, the **change** is used to calculate the final proportion at t_1 . This allows the second vector of zeros and ones to be created in the same way. Using the example above, a 50% increase would result in a vector of 200 zeros and 300 ones.
 - α Finally, for dependent samples, the **turnover** is applied to the second vector in the array. Before applying the turnover, the array as a whole would have 200 pairs of zeros, 250 pairs of ones and 50 pairs of zeros and ones, representing the sites that had become favourable. Imagine that the turnover was 20%. This would result in an additional 10 zero/zeros being converted to zero/one, but 10 of the one/ones being converted to one/zero. This results in all four combinations of favourable/unfavourable across the two times, t_0 and t_1 .
- Σ The two (or three if independent sampling is defined) sample parameters are used to draw two samples from the populations. The process is slightly different for the two **sampling strategies**.
 - α Firstly, dependent samples only have one **sample size** – they have no **adjustment**. The sample is drawn by randomly marking the population and then tallying the number in each of the four combinations (zero/zero, zero/one, one/zero and one/one).
 - α Secondly, the independent sampling strategy has two different samples, which may be of different sizes. The sample at t_0 is drawn first by randomly marking the population at t_0 and tallying the number of zeros and ones. Then, if the adjustment is not 0%, a new **sample size** is calculated and this number of items in the population t_1 is marked and tallied. Thus all four combinations of t_0/t_1 and zero/one are stored.
- Σ The sample is then analysed using one of two models depending on the sampling strategy. These are explained further below. The α -value from the analysis is stored.
- Σ The whole process is repeated a large number of times to create a vector of alpha values. The number of simulations is defined by the analysis resolution, which is explained further in the User Guide below, but varies between 500 and 5000.
- Σ Finally, this vector is used to calculate the expected **significance** for a given **power**, or the expected power for a given significance, which will be plotted on the Y-axis of the graphical analyses. A slightly different algorithm is used to find the **sample size** for specific power and significance used in the tabulation analyses.

The Statistical Models

Two different statistical models are used for the two sampling strategies. Firstly, for independent samples, a normal approximation was used to compare two independent proportions (Appendix 1). Although for small



sample sizes this gives optimistic powers, this is completely overwhelmed by the stochastic effects of the random simulations. The minimum sample size allowed by the power tool is 10, and sample sizes less than 25 can be unreliable, especially with small degrees of change.

The statistical test used to analyse dependent samples is a normal approximation to McNemar's Test with a correction for continuity (Appendix 2). This test is subject to similar concerns regarding small sample sizes, although as it is intrinsically more powerful, the effects are smaller.

Both tests are based on the Z-statistic and are subject to a finite population correction (FPC).



User Guide to the Power Analysis Tool

Installation & Running the Power Analysis Tool

The Power Analysis Tool has been developed in MS Access 2007 for use on stand-alone computers. It uses Access 2003 file types, so it will run under this version of Access as well, but it has incompatibilities. There is only one non-standard reference in the VB project, which is to MS Graph 12 (Graph.exe). It will also work with version 11 of this file, although this has not been fully tested. Graph.exe is usually installed with any of the Microsoft Office applications, not just MS Access.

There is no MS-based installation required for the power tool. It is simply a single file, with no external dependencies. You can simply copy the file to a suitable location and double-click it to run it. If you are using a full installation of Access 2007, you should be familiar with the Trust Centre. This allows you to define directories as trusted locations, where Access will not question the security of a file. So, either copy the power tool to an existing trusted location or make its new location trusted.

Using the Access 2007 Runtime

The power tool has been tested successfully under a clean installation of the Access 2007 Runtime. This can be downloaded free from the Microsoft website:

<http://www.microsoft.com/download/en/details.aspx?id=4438>

It takes only a few minutes to download and install, although you will probably need administrator's rights to do so. The runtime also uses trusted locations, but there is no way of adding new locations as there is in the full version of Access 2007. So if you try to run the power tool from any convenient location you will be given a warning every time you open it. To avoid this, the runtime has one trusted location that is automatically created when it is installed:

C:\Program Files\Microsoft Office\Office12\ACCWIZ\

If you move the power tool to this location and double-click it, it will open without the warning dialog. You might like to create a shortcut to the power tool and put it on the desktop.

You can install the runtime alongside earlier versions of Access, but you will need to follow the above procedure, with one modification. To force the power tool to use the 2007 runtime, rather than your previously installed version of Access, you need to modify the target string in the shortcut. You should now copy this string (with the quotes) into the target of the properties of the shortcut:

"C:\Program Files\Microsoft Office\Office12\MSACCESS.EXE" "C:\Program Files\Microsoft Office\Office12\ACCWIZ\Power Tool.mdb"

The Main Menu

When you first open the power tool you will see the main menu form. This has a textbox for a password and three command buttons. Until a valid password is entered, the only button that is enabled is the Exit button – so this is all you can do. The User password is

Urtica

which simply identifies you as a user, and enables the three other buttons. These open the two main forms in the power tool which correspond to the two utilities described in "The Concepts Behind the Power Analysis Tool", plus a Configuration form.

The use of each of these two forms will now be covered in detail, although common functions will only be described in the "Graphical Analysis of Theoretical Data".



Graphical Analysis of Theoretical Data

This form is divided into two main zones. On the left of the form is the graph definition zone and on the right is the graph display and management zone. These contain a large number of controls that allow you to specify values for different parameters. When you close the form, these values are stored so you can return to exactly the same definitions when the form is re-opened.

The Graph Definition Zone

This zone (Figure 1) is further subdivided into a number of groups. On the left-hand side there are ten combo-boxes used for setting the fixed values of the ten parameters plus three display-only text-boxes. On the right-hand side are three groups of controls for setting which parameters, and their range of value, to be plotted on the graph.

Setting the Fixed Parameters

The parameter combo-boxes are grouped into Population, Sample and Analysis as described in “The Parameters” above. Each combo provides a drop-down with a range of valid values that the parameter can take. Two of them (Strategy and Tails) only have two textual values each, but the others give a range of numerical values, both integers and percentages. Three have negative as well as positive values.

Four of the combo-boxes will be disabled (and become greyed out) as shown in Figure 1. Which combos are actually disabled depends on two factors. Firstly, if the parameter has been selected to be plotted on the graph (see “Setting the Plotted Parameters” below) then it will be disabled. Secondly, the Turnover and Adjustment combos are disabled according to the type of Strategy selected; if it’s independent then the Turnover combo will be disabled – if it’s dependent then the Adjustment combo will be disabled. To distinguish this behaviour from the previous type, the word “Irrelevant” is displayed in the disabled combo.

Three of the combo-boxes have additional consequences for the other controls. Firstly, if you select a different Population Size it will alter the range of values that can be selected for the Sample, so that the sample can never be larger than the population.

Secondly, altering the values selected in either the Initial Proportion or Change combos will cause the three display-only text-boxes to be recalculated. These three text-boxes display the stages in the calculation of the Final Proportion, as explained in “The Parameters” above. In the example shown in Figure 1, the initial proportion of 50% equates to odds of 1.

Figure 1. Graph definition zone on the Graphical Analysis form

Increasing this by +100% gives odds of 2 and converting these odds back to the final proportion gives 66.67%.

Setting the Plotted Parameters

On the right-hand side of this zone are three control groups labelled Y-axis, X-axis and Lines. This is where you specify the parameters to be plotted on the graph. Each control group has a set of radio-buttons, which allows exactly one parameter to be selected for each of the three dimensions of the graph. The groups of radio-buttons have a priority from left to right, so that if you select a parameter in one group, its radio-button in the groups to the right become disabled. This is to prevent you from selecting the same parameter for two or more dimensions. (Note that the radio-buttons also become disabled when either the Turnover or Adjustment combos become disabled as described in the previous section.

The Y-axis only has two radio-buttons indicating that only Power or Significance can be plotted on the Y-axis of the graphs. However, eight of the ten parameters can be plotted on the X-axis or as Lines on the graphs. Note that two parameters will be unavailable for selection on the X-axis; Power or Significance depending on the Y-axis selection and Turnover or Adjustment depending on the Strategy selection. Similarly, three options will be disabled for the Lines because you cannot select the parameter that you have already selected for the X-axis.

When you select a parameter for either of these two dimensions, the list-box alongside refreshes to display all the valid values that the parameter can take. (This will be identical to the list displayed in the equivalent combo-box if it were enabled.) However, the behaviour of the two list-boxes is slightly different:

- Σ The X-axis listbox must have at least two values selected. You can do this either by clicking with the mouse on a value (which clears all other values) and holding down the <Ctrl> key whilst clicking on a second value, or clicking and dragging to select a block of values. The plotting algorithm only uses the lowest and highest selected values as the range to be plotted on the X-axis.
- Σ The Lines list-box can have between one and five values selected. You can only do this by clicking on individual values – one click selects the value, another click deselects it. Each value will create one line on the graph.

If you select the Change parameter to be plotted either on the X-axis or as Lines, the Final Odds and Final Proportion text-boxes display “Variable”. This is because there will be multiple change values selected, so no single final values can be calculated. Similarly, if you select the Initial Proportion parameter to be plotted on the X-axis or as Lines, the Initial Odds as well will display “Variable”.

The Graph Display and Management Zone

This zone (Figure 2) is where you define how the graphs are actually created and where you manage the set of stored graphs. The zone is dominated by the graph itself, above which there is a text-box that displays the title of the graph. This is created automatically when the graph is first plotted, using the relevant parameters, but you can edit it to help organise saved graphs.

Beneath the graph display are four sets of controls, firstly two sets of radio-buttons:

- Σ **Analysis Resolution**, which allows you to select the number of simulations used in the analysis.
- Σ **Plot Resolution**, which defines how many points will be used to plot each line.

These two values work together, along with the number of lines, to determine the total number of simulations that are carried out. The low analysis resolution option will carry out 500 simulations for each plotted point on the graph. The low plot resolution uses 15 points for each line, so the example in Figure 2 will have 45 points, each of which uses 500 simulations, giving 22,500 in total. This allows the graph to be plotted relatively quickly, giving a somewhat crude graph, but will allow easy exploration of a number of parameter options. In contrast, if high resolutions had been chosen in this example, there would have been a total of 11,250,000 simulations. Clearly, these options will take much longer to plot, so you should use them only when you want to create a final, high quality graph for inclusion in a report.



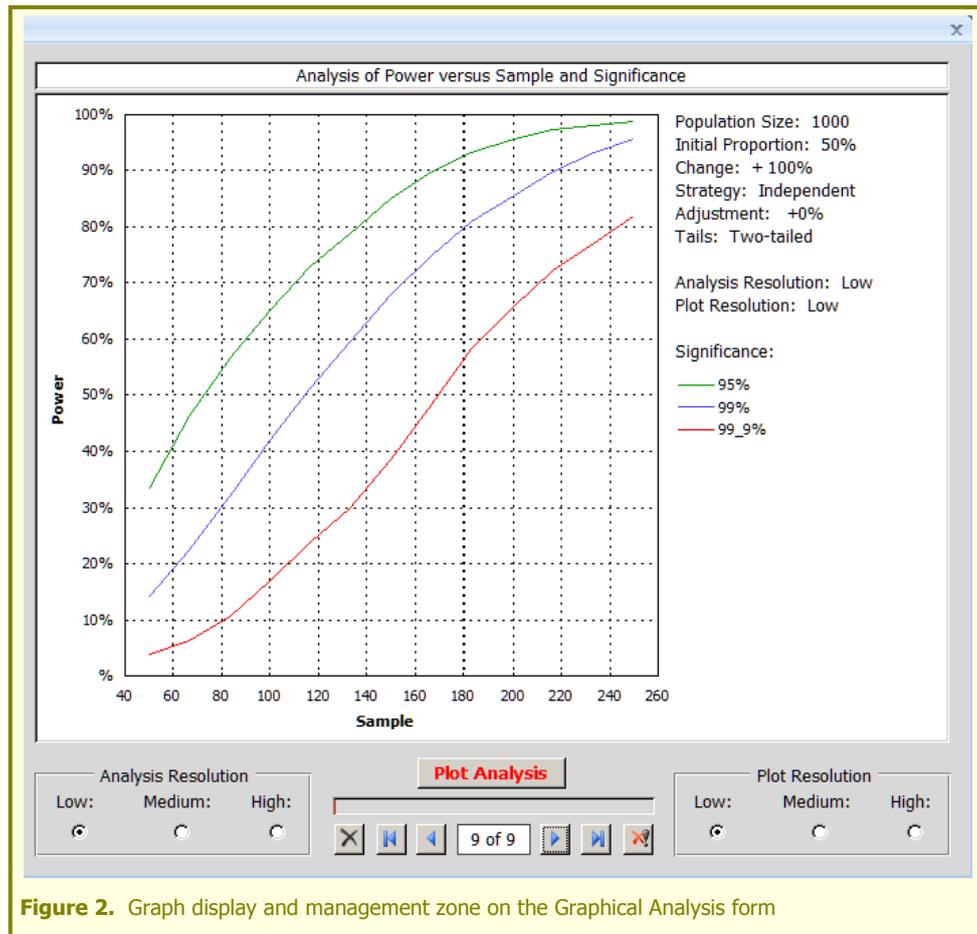


Figure 2. Graph display and management zone on the Graphical Analysis form

- Σ The **Plot Analysis** button works in combination with the progress bar beneath it. Clicking this button will take all the parameter values specified in The Graph Definition Zone, as well as the two resolution options described above to carry out the simulation analysis. Most invalid parameter options are prevented by disabling combos and radio-buttons, but some options cannot be prevented in this way. Instead, these are trapped when the Plot Analysis button is clicked. During the analysis, a Cancel button becomes available which will halt the execution of the analysis. You may find certain occasions when the lines fail to plot, especially if this is the first plot in a series. If so, simply double-click the graph to refresh it and the lines will appear.
- Σ The **Plot Management Buttons** allow you to move between stored graphs. Every time the Plot Analysis button is clicked, a new graph is created and added to the store. This is shown in the counter as, e.g., “9 of 9”. The four inner-most buttons allow navigation to the first, previous, next and last graph in the store. The left-hand button deletes the current graph, and the right-hand button deletes all graphs, after suitable warnings have been given.

Copying Graphs

Although there is no specific function to copy graphs, this can be achieved simply by clicking on the graph and pressing the <Ctrl> + C keys. This is a standard Windows action which copies a graph object, complete with all its data to the clipboard. You can then paste this into a Word document or Excel spreadsheet, where you can edit it or even modify the data.



Tabular Analysis of Empirical Data

This form is similar in basic concept to the Graphical Analysis form, as it is divided into two zones. On the left-hand side is the tabulation definition zone which has certain features in common with the Graphical Analysis form. On the right-hand side, the Table Display and Management zone also has a number of features in common, although its primary tabular display is obviously different from a graph.

The Tabulation Definition Zone

This zone is shown in Figure 3 and is divided into a number of groups. On the left-hand side, there are seven combo-boxes and three text-boxes which are generally identical to those on the Graph Analysis form. However, there are three fundamental differences:

- Σ Three combo-boxes are not visible on this form, although the layout of the remaining combos has been retained to give a familiar appearance. Sample Size and Adjustment are not visible because they are always calculated as the output from the tabulation algorithms. (Adjustment is irrelevant because only a single sample size can be calculated.) Population is not visible because it is always supplied as part of the underlying empirical dataset.
- Σ The Initial Proportion combo-box is disabled by default because, like Population, it is supplied by the underlying dataset. However, these values are often estimated and may not be particularly realistic. To allow some user control over the Initial Proportions, a checkbox allows you to enable the combo and so over-ride the empirical data values with a single user-defined value.
- Σ The Turnover combo-box is still enabled or disabled by the Strategy selection in the same way as on the Graphical Analysis form. However, the other combos remain enabled at all time, as there are no graphical dimensions that need to be specified.

Filters

The underlying dataset has four categorical variables that classify the parcels;

- Σ The **BAP Priority Habitat** to which the parcel belongs.
- Σ Whether or not the parcel is within an **SSSI** designation.
- Σ Whether or not the parcel is within an **HLS** agreement.
- Σ The **NCA** within which the parcel is situated.

The underlying population of parcels can be filtered to exclude certain category values. You can do this with the four list-boxes in the centre of the Tabulation Definition Zone. They are disabled by default, which means that all categories are included in the tabulations. However, by clicking on the adjacent toggle-button with the

Figure 3. Tabulation Definition zone on the Tabulation Analysis form



filter symbol, the list-box can be enabled. Then it is possible to select one or more categories to be included in the tabulation – clicking once selects a value and clicking it a second time deselects it. Disabling a list-box deselects all values, so this is a quick way of clearing the filters.

Filtering uses boolean logic. In other words, multiple categories selected in one list-box are combined with the OR operator, whereas categories selected from different list-boxes are combined with the AND operator. For example, the options shown in Figure 4 would result in the following filter clause:

WHERE ((Habitat = 'Blanket Bog) OR (Habitat = 'Fen')) AND (SSSI = 'Non-SSSI')

Tabulation

The categorical variables used for filtering can also be used for tabulating in one or two dimensions. By default, tabulation is switched off, which means that the analysis will just calculate a single sample size for the whole (or filtered) population.

You can turn on the tabulation options with the two toggle-buttons labelled 1° and 2°. When you click the 1° button, the four radio-buttons beneath become enabled, allowing you to select one category as the Primary Tabulation variable. It will use all of the values (or all filtered values) to undertake a breakdown of the parcels, calculating sample sizes for each category value. Clicking the 1° toggle-button also enables the 2° button – click this to select another category to create a two-way breakdown.

The Table Display and Management Zone

This zone is shown in Figure 5. The title is displayed at the top of the zone and, like the Graphical Analysis form, it is created automatically, but can be edited by the user. Beneath this are six read-only text-boxes that display the scalar values of the parameters. These are stored along with the table that is created by the Tabulate button.

The group of four controls below the table is similar to that found on the Graphical Analysis form.

Σ The Tabulation Resolution corresponds to the Analysis Resolution, although the actual number of simulations is lower because the algorithm for calculating the sample sizes is intrinsi-

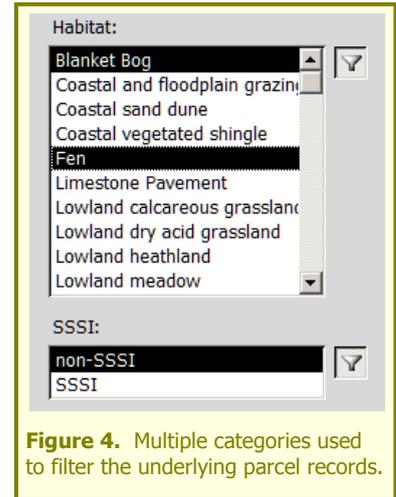


Figure 4. Multiple categories used to filter the underlying parcel records.

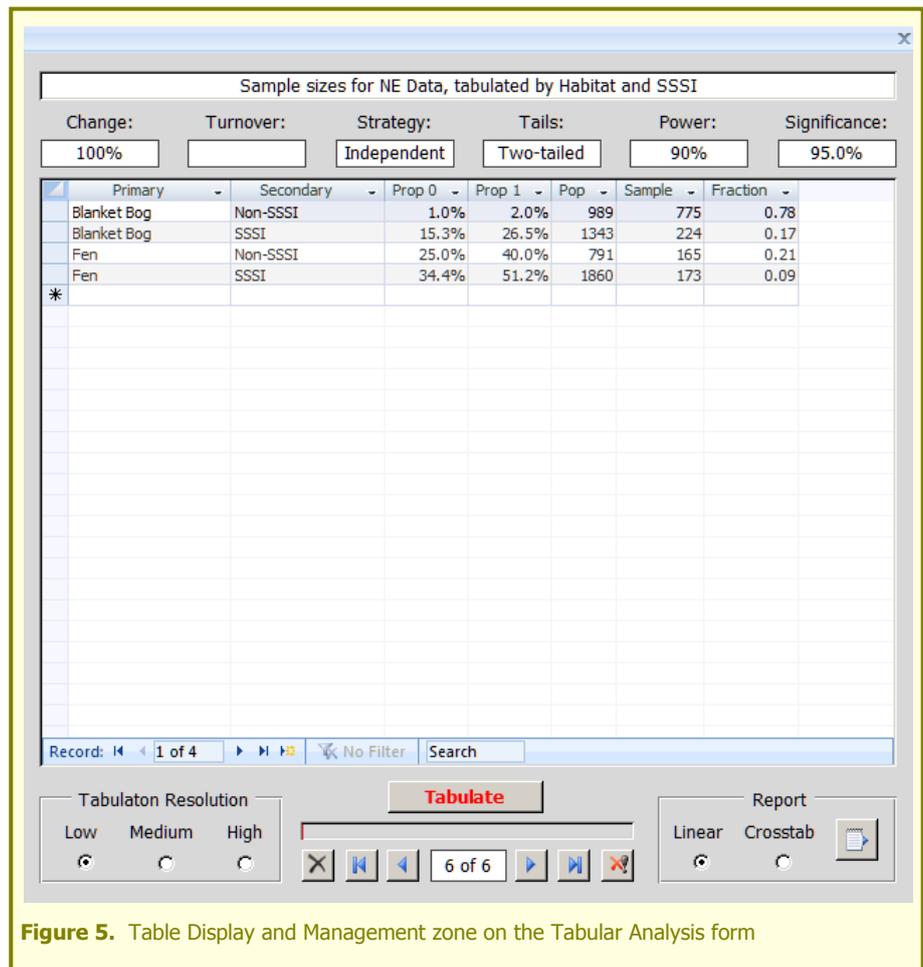


Figure 5. Table Display and Management zone on the Tabular Analysis form



cally less stochastic. However, they are selected in the same way to allow quick exploration with the low resolution option and a more accurate but slower analysis when the high resolution option is chosen.

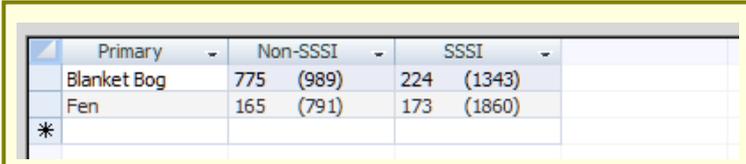
- Σ The Tabulate button and the progress bar work in exactly the same way as the Plot Analysis button, as does the Cancel function.
- Σ The Table Management buttons also work in the same way, enabling you to move between stored tables and delete the current or all tables.
- Σ The main difference between this zone and the Graph zone are the Report Options. Firstly, there are two radio-buttons which allow the table to be displayed as linear or cross-tabulated. The Crosstab button is only enabled when a two-way breakdown has been defined, as described above. There is also a command-button with an “Output to Report” symbol, which sends the specified table to a formatted report.

Linear versus Cross-tabulated Displays

If no tabulation is defined in the Tabulation Definition Zone then the table will be displayed with only a single row and no values in the Primary or Secondary columns. If a one-way breakdown is defined by selecting only a primary tabulation variable, then no categories will be displayed in the Secondary column of the table. However, the example shown in Figure 5 is a two-way breakdown because there are categories shown in the Secondary column as well as the Primary column.

This table is displayed in a linear format, which creates a row for each combination of the primary and secondary categories and displays five columns of data. The first two are the proportions at t_0 and t_1 . This shows that the underlying data have been used, because there are different proportions for the four different combinations of category values. The proportions at t_1 have been calculated from the proportions at t_0 , by applying the user-defined change shown in the text-box above the table. Note that this uses the same algorithm for calculating the change in the odds that is used for the Graphical Analysis. This example gives a good illustration of the effect of applying a change of 100% to the four different initial odds. Following these columns the underlying populations of parcels for each category combination is given. The next column displays the sample size calculated by the simulations. The final column displays the sampling fraction – all populations of ten or less automatically default to a sampling fraction of 1.

However, being a two-way tabulation, this example can also be cross-tabulated as shown in Figure 6. In this case there is only one row for each of the values in the primary tabulation category, whilst the values of the secondary tabulation category become the data columns – in this case there are two.



Primary	Non-SSSI	SSSI		
Blanket Bog	775 (989)	224 (1343)		
Fen	165 (791)	173 (1860)		
*				

Figure 6. Multiple categories used to filter the underlying site records.

Two variables are displayed in a cross-tabulation; the sample size is shown first in each cell and the population size is shown in parentheses. These values are identical to the equivalent columns in the linear display, but they allow you to compare the primary and secondary categories more easily.

Displaying the Table as a Formatted Report

In addition to the spreadsheet style displays on the Tabulation Analysis form, both linear and crosstab tabulations can be displayed as formatted reports. If you click the command-button in the Report options, you will open a print preview of the relevant report. If the display is linear, this will be in portrait format, whereas if the display is crosstab it will be in landscape format, displaying up to eight data columns. You can resize and zoom in on this print preview window in the same way as any Microsoft application.

Copying Tables & Printing Reports

You can copy any table displayed on this form by selecting the part of the table to be copied and pressing <Ctrl> + C. You can select one or more rows with the record selectors (the pale blue squares shown on the



left-hand side of the table in Figure 6), one or more columns by clicking on the column headers, or the whole table by clicking on the white triangle at the top-left of the table. The selection will be highlighted with an orange border and after copying it can be pasted into a Word document or Excel Spreadsheet.

All reports can be printed from the print preview by pressing the <Ctrl> + P keys. This will open the standard Print dialog box, allowing you to select a printer, change the printer properties and alter the page setup options. If you have an Adobe PDF driver, you could print the reports to PDF files.



Using the Power Analysis Tool to Explore the Effects of Parameters on Power

Introduction

This chapter uses the Graphical Analysis form to explore some of the theoretical relationships between the parameters. Clearly, the whole point of the tool is to allow the user to select particular combinations of parameter values from the millions that the tool presents. So this exploration will be just a tiny selection, but its purpose is to illustrate which parameters have the greatest effect on power and sample size..

It is important to emphasise that, as the tool is based on a random simulation process, actual results may vary. However, by selecting the high resolution options, simulations will stabilise to such an extent that the graphs from different analyses will become almost indistinguishable. The only exception to this is when sample sizes, initial proportions and degrees of change are very small.

The Basic Relationship Between Significance and Power

Figure 7 shows the basic relationship between significance and power, whilst holding all of the other parameters constant. This is a negative relationship, so that as the desired power increases, the significance decreases. So, for a population of 1000 and a sample of 200 sites, with a change of 100% (resulting in a final proportion of 66.67%), if a power of 90% was specified, then the resultant significance would be about 98%. This implies that there would be an acceptable probability of detecting this change with this sample size at 98% significance (*i.e.* $\alpha = 0.02$). However, if you needed a higher probability of detecting a change, say 95%, then the likely significance with which you would be able to detect this change declines to just over 95% –still a useful result.

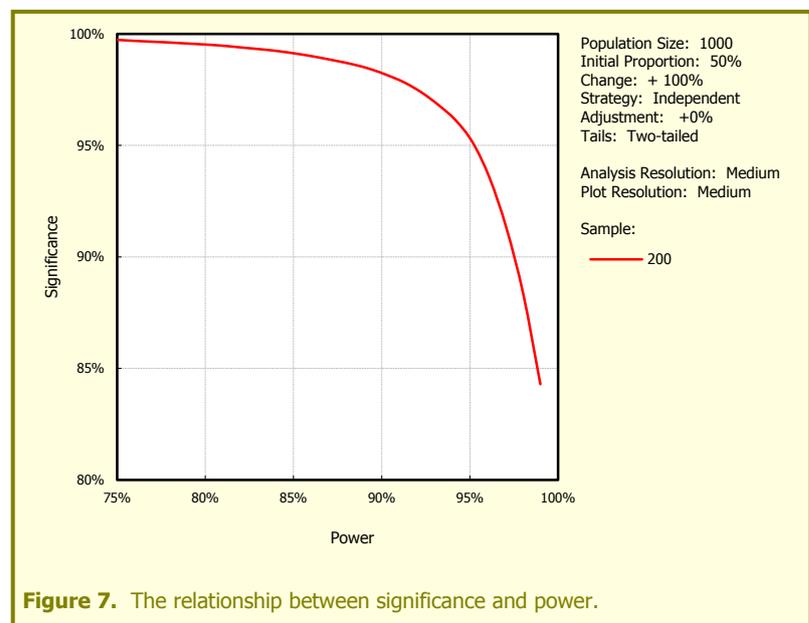


Figure 7. The relationship between significance and power.

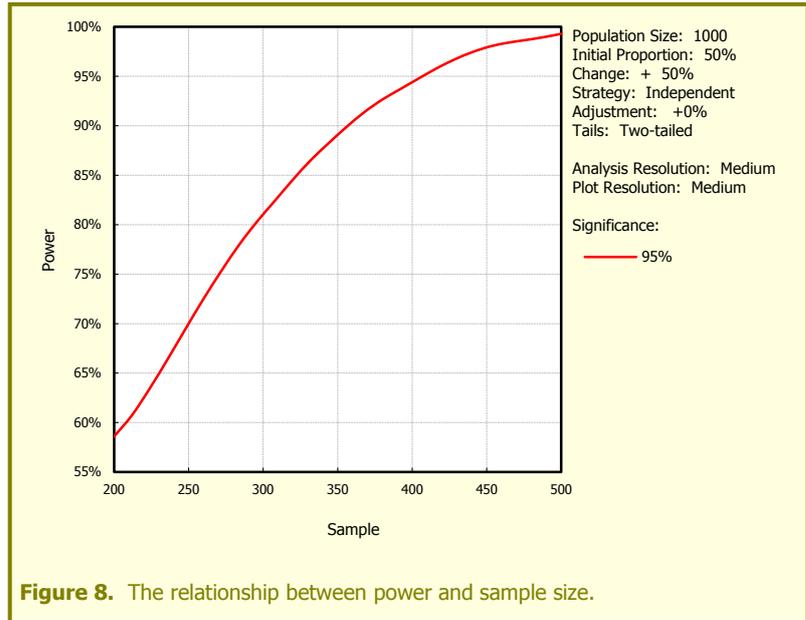
This curve is hyperbolic – meaning that at low powers significance approaches 100%, but never quite reaches it. Similarly, as significance declines, power approaches 100% but does not actually reach it. The consequence of this is that you have to sacrifice a lot of power to achieve relatively little increase in significance.

The Basic Relationship Between Power and Sample Size

This relationship is shown in Figure 8, this time with a change of only 50%. The relationship is positive, with power increasing as sample size increases. In this example, to achieve 90% power, a sample size of approximately 360 would be required. Interestingly, increasing the sample size to 410 doubles the power (95% – because there are only 5 chances in 100 of failing to show a significant change, compared to 10 in 100 at 90% power). Note that as the sample size increases the power continues to rise towards 100% asymptotically, and will in fact reach it when the sample size is the same as the population, because the finite population correction has been applied in all these analyses.

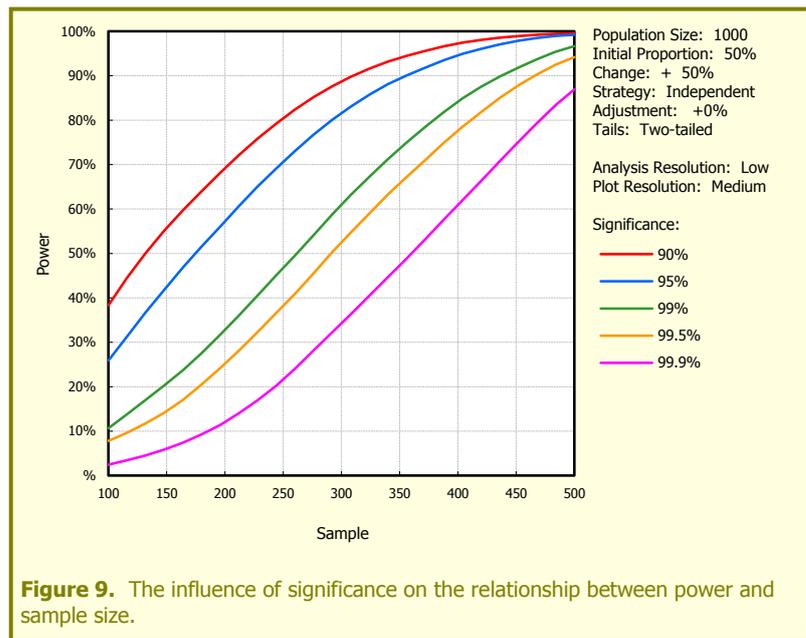


This type of plot is probably the most useful of the power analysis curves, because sample size is the parameter that is most easily under the control of the surveyor and power is the parameter that predicts the likelihood of you obtaining a significant statistical result.



The Influence of Significance on this Relationship

Figure 9 increases the utility of the power tool by including a third dimension. In this case, the influence of significance is included by plotting five different lines, representing significances between 90% and 99.9%. As already demonstrated in Figure 7, a lower significance results in a higher power, so if we were happy with 90% significance, then a sample size of only around 310 would result in a power of 90%. However, if we desired a much higher significance (say 99%), then a sample size of nearly 450 would be needed to achieve a power of 90%.



This graph clearly shows how difficult it is to obtain useful power with small sample sizes, almost regardless of the significance level that is acceptable. Given the values of the fixed parameters, a sample size as small as 100 gives very little power, even for a significance of 90%.

The Relationship Between Population and Sample Size on Power

By definition, the condition assessment of SSSIs in Wales will take place on a known population. This means that any sample drawn from that population will have a known sampling fraction, equal to $\frac{\text{Sample } (n)}{\text{Population}(N)}$. As the sample increases then n/N approaches 1. The finite population correction has already been alluded to,



which equals $1 - n/N$. So as the sample size increases, the FPC decreases. This has the effect of decreasing the standard error in calculating the test statistic, Z and so results in a more powerful test.

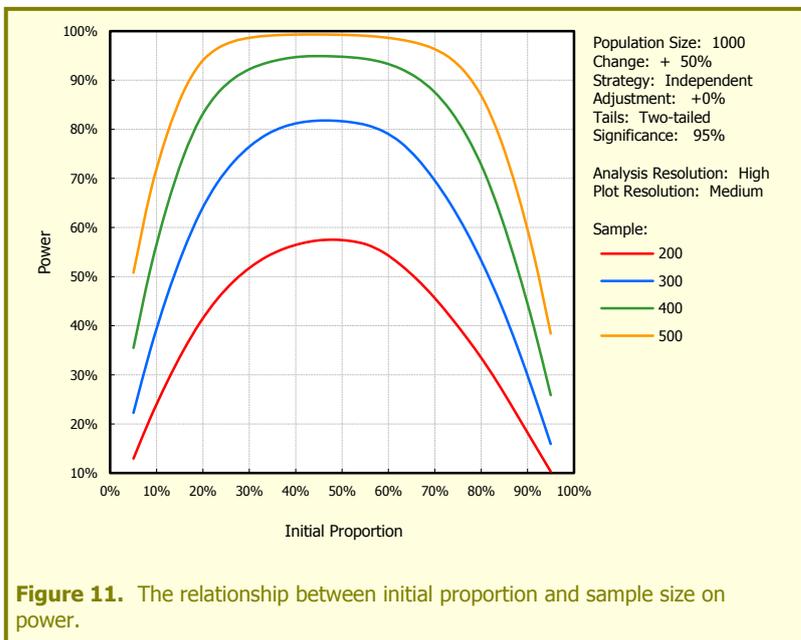
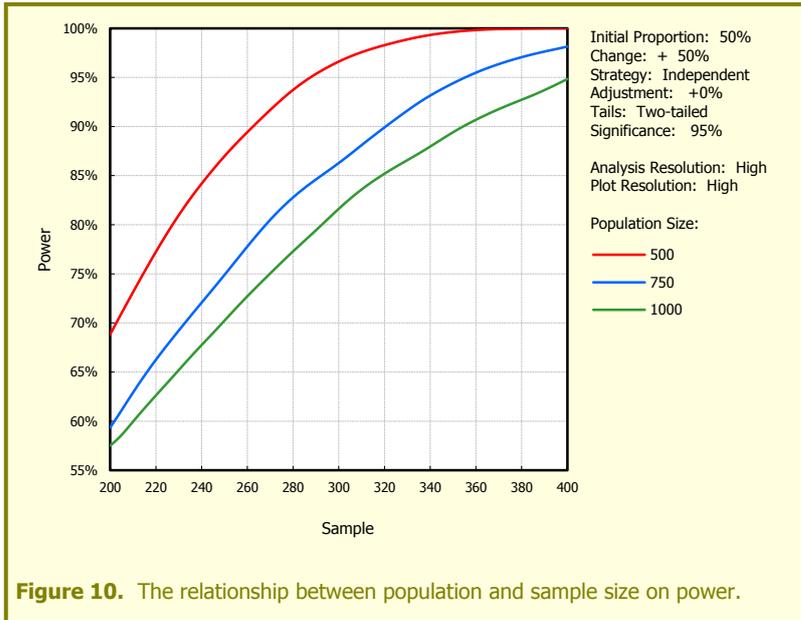
Figure 10 incorporates three different population sizes into the relationship between sample size and power - it clearly shows the effect of the FPC. In a population of 1000, a sample of 300 gives a sampling fraction of 0.3. This sample and population combined with the fixed parameters only gives a power of 84%, implying that this sample would be too small to yield a useful result. However, this sample drawn from a population of only 500 sites

gives a sampling fraction of 0.6. All other things being equal, a sample of 300 gives a power of about 97%. Notice how a sampling fraction of over 75% will give a power that is virtually 100%. Nevertheless, when sampling fraction is below 50% it generally has a smaller effect than absolute sample size.

The Relationship Between Initial Proportion and Sample Size on Power

Initial proportion has a very significant effect on power, primarily because its value is constrained between zero and one. Figure 11 shows this effect very clearly for four different sample sizes. The most important point is that power declines rapidly at the extreme values of proportion. When they are less than 10% or greater than 90% then changes are far less likely to be detectable than when the proportion is nearer 50%. The other interesting thing to note is the flat-topped shape to these curves, especially for larger sample sizes. This implies that the zone of high power becomes wider - *i.e.* a greater range of proportions will give adequate power. Finally, it is worth pointing out the slight asymmetry in these power curves,

with lower powers to the right-hand side of the graph. This is because the change applied in these simulations is positive. So a given percentage increase (in the odds) results in a lower power when the starting proportion is already high, compared to a starting proportion that is very low.



The Relationship Between Change and Sample Size on Power

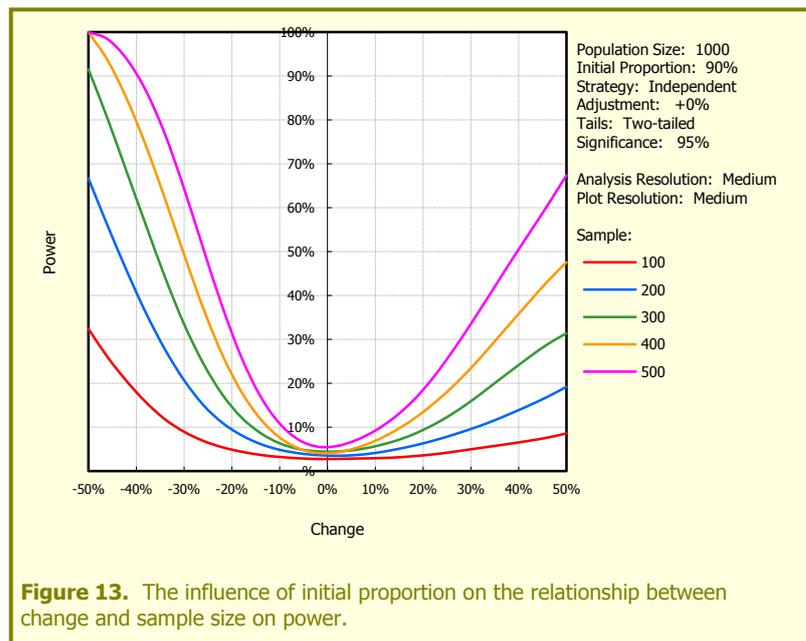
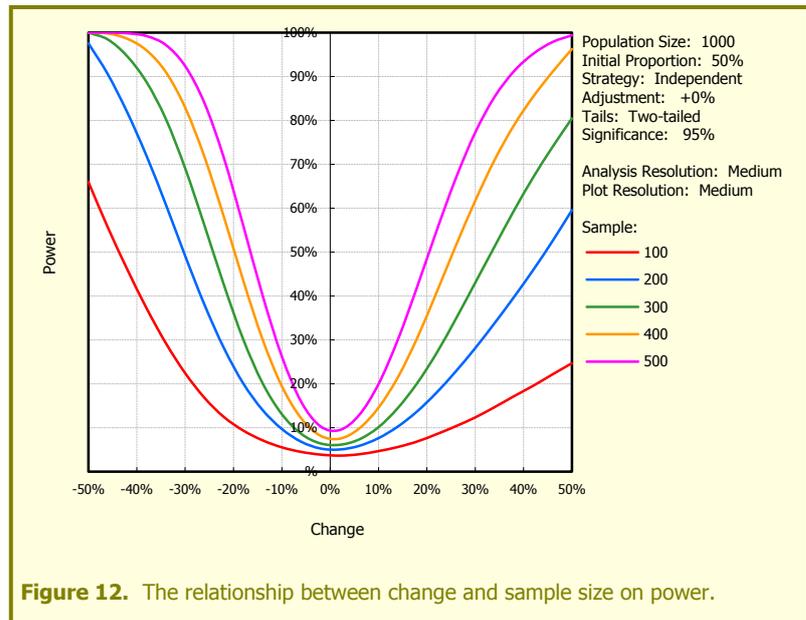
Clearly, the most important parameter in determining the ability to detect change in a monitoring programme is the degree of change itself. Figure 12 shows that, even for large sample sizes, power declines radically as change approaches zero. For example, with a sample size of 500 (large in absolute terms and representing a sampling fraction of 0.5) when the change is +50%, power is over 99%. However, when it is only +5%, the power declines to only about 10%.

Figure 12 also shows the relationship between positive and negative changes. In this case, the larger sample sizes in particular show high degrees of power for both negative and positive changes.

But the asymmetry is also apparent, especially for smaller sample sizes. This is because negative changes, even in the odds, cannot exceed 100% – this represents a change from any initial proportion to a proportion of 0%. A positive change in odds can, however, be any size – but it will eventually approximate to a proportion of 100%.

Finally, this graph shows the very strong interaction between change and sample size. Small sample sizes, such as 100, are incapable of showing a significant change within the range of -50% to +50% shown here. Much larger sample sizes, such as 500, reach a power of 90% within a range of changes of about -28% to +38%.

It is also worth exploring the effect of initial proportion on this relationship. Figure 13 plots the same degree of change on the same sample sizes, but assumes an initial proportion of 90%. As illustrated in Figure 11, this has the effect of reducing the overall power, but it also increases the asymmetry. Now, a change of +50% in a sample of 500 has less than 70% power, whereas -50% still gives almost 100% power.



The Effect of Sampling Strategy on Power

Apart from sample size, the previous explorations have concentrated on population parameters. One of the major factors that can be designed into a sampling strategy is whether the samples will be independent (as illustrated in all previous examples) or dependent, where the same sampling units are revisited. This can result in a major improvement in power because it effectively removes between-site variance.

Figure 14 shows the difference in power for two degrees of change for both independent and dependent samples. Firstly, for a 50% increase with a sample of 200 independent sites, the power would be barely over 20%. However, if the samples were dependent, then the power would increase enormously to over 90%. Similarly, with a larger increase of 100%, a sample of 120 independent sites would only give 40% power, whereas this would be over 90% for a dependent sample. Turning this around, the sample sizes required to achieve 90% power would be nearly 300 for independent sites, but only 120 for dependent sites¹.

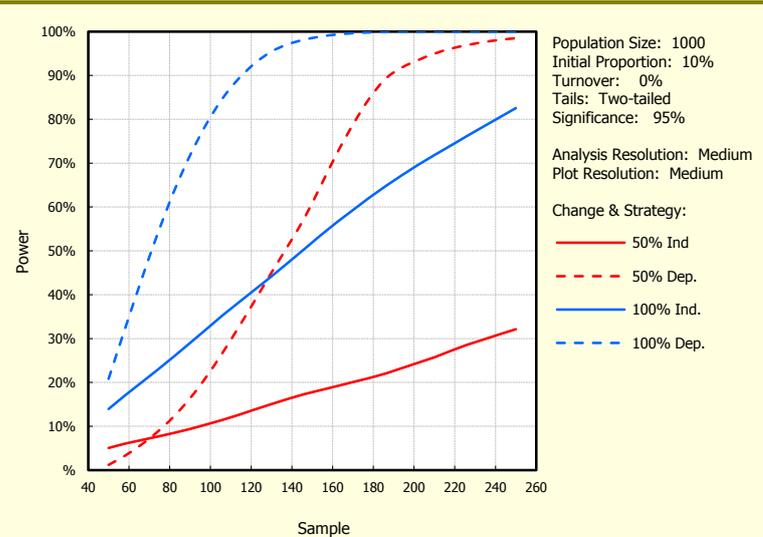


Figure 14. The effect of sampling strategy on power.

The Effect of Turnover on Dependent Samples

Although previous analysis appeared to show that dependent samples are intrinsically much more powerful than independent regimes, it is important to understand the effect of turnover. As explained in “The Concepts Behind the Power Analysis Tool”, turnover is a population parameter that accounts for individual dependent sites changing in the opposite “direction” to the overall change.

Figure 15 shows the relationship between sample size and power for dependent samples with different degrees of turnover. Firstly, with zero turnover, the dependent sampling strategy is very powerful compared to an independent strategy (the grey dotted line). For these fixed parameter values, a sample size of under 100 would be adequate to achieve 90% power. However, just a small turnover of 10% would reduce the power from this sample size to 60% – requiring a sample of over 165 to achieve 90% power. A turnover of 50% has almost identical power, over this range of sample sizes, as an independent sampling strategy. And although not shown on the graph, larger degrees of turnover make a dependent sampling strategy less powerful than an independent design.

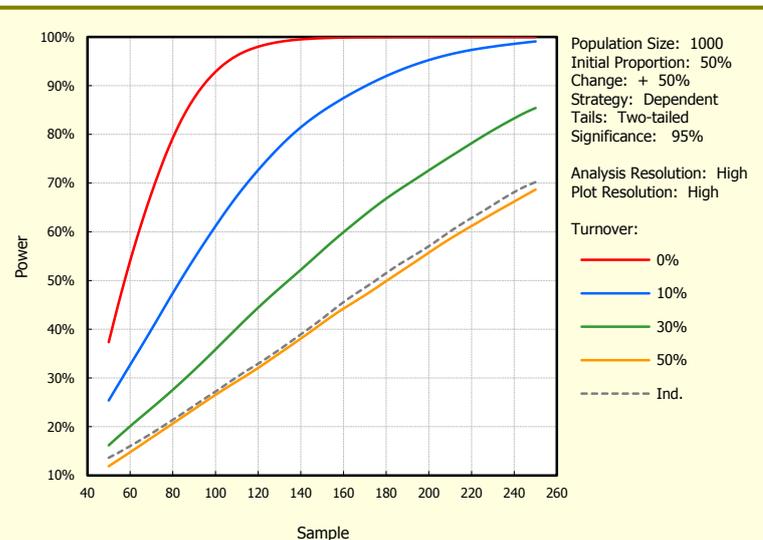


Figure 15. The effect of turnover on dependent samples.

¹ The graph in Figure 14 was created as a compound graph from two runs of the power analysis tool. They were copied to a separate Word document, the data from one graph copied into the other and the line formats and legend then manually edited.



The initial proportion set in Figure 15 is 50%, which makes a large turnover quite unlikely. However, if the starting proportion was only 10% in a sample of 200, then a 50% turnover is easily imagined. In terms of condition assessment, this would equate to 20 of the 180 unfavourable sites becoming favourable, whilst 10 of the 20 favourable sites become unfavourable. The power for this scenario would be only about 25%, whereas if there were no turnover, the power would be over 90%.

The Effect of Sample Adjustment on Independent Samples

An equivalent situation occurs with independent samples, if the sample size on the second visit is different from the visit at t_0 . However, Figure 16 shows that this effect is relatively small. With identical sample sizes, around 165 sites would be sufficient to achieve a power of 90%. However, if the number of sites in the second sample declined by 10%, then it would have been necessary to have started with a sample of approximately 174 (giving a second sample of 157) to achieve 90% power. The equivalent figures for a 50% decline would be about 240 (and 120).

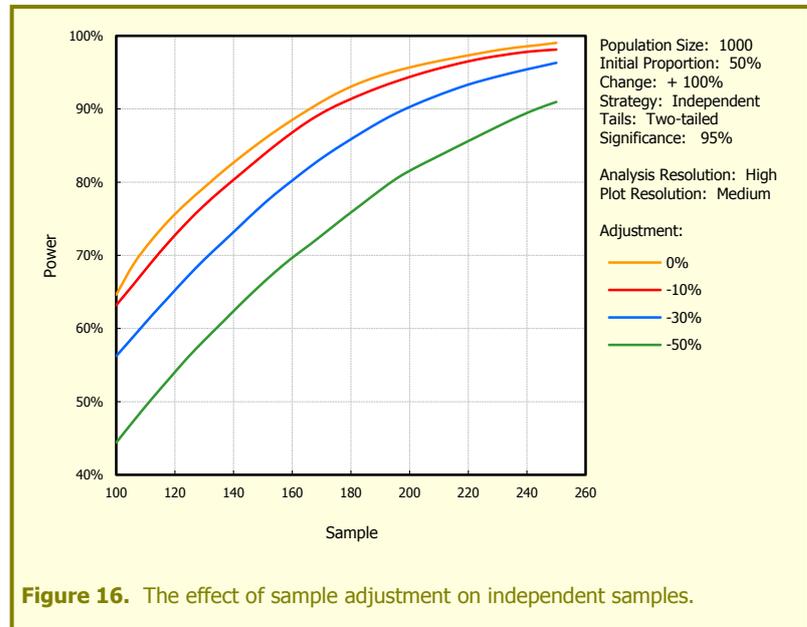


Figure 16. The effect of sample adjustment on independent samples.

The Effect of One-tailed Testing on Power

The final parameter on the power tool not yet addressed is the use of one or two-tailed testing. This has been provided because there may be specific situations where management practices have been introduced to a population of sites which were designed to cause a change in one direction – e.g. to increase the proportion of sites in favourable condition. In this case a one-tailed test will give more power to detect such a change. However, it is important to realise that it will provide no information on the opposite change; these will just be presented as not significant. In surveillance situations where no change is predicted *a priori*, then a two-tailed test should be used.

An example of the increased power of a one-tailed test is given in Figure 17. All parameter values are identical to those in Figure 9, except for the tails. Although the shapes of the power curves are essentially the same, they have all been shifted up the y-axis, indicating greater power. So, for example, with two-tailed testing a sample of 300 would only give a power of 82% for 95% significance, for one-tailed testing,

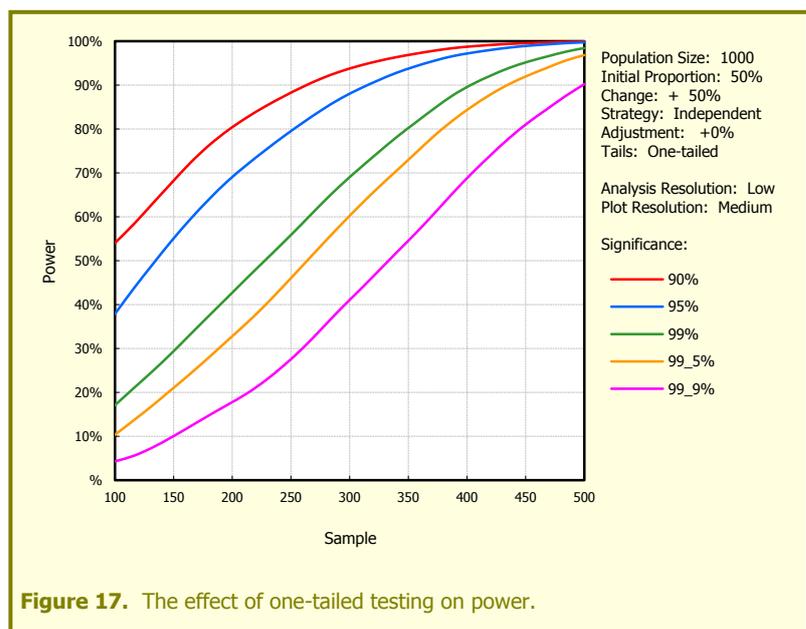


Figure 17. The effect of one-tailed testing on power.

this power would be 88%. Alternatively, the 90% power threshold (for 90% significance) would be achieved with a sample of about 265 with one-tailed testing, but would need to be 310 with two-tailed testing.



Exploratory Power Analysis for Condition Assessment

Introduction

This chapter uses the Tabular Analysis form to explore the sample sizes that will be required, based on the known populations and initial proportions of habitat parcels in different BAP priority habitats, NCAs, etc., as given in the specification document. Overall there are 92,524 parcels 0.25ha or larger, which form the overall statistical population or sampling frame. However, the Configuration form allows the user to define other minimum area thresholds of 0.5, 1.0, 2.0 and 5.0ha to be included in the sampling frame. These give populations of 77,218, 62,354, 46,743 & 27,879 respectively. (To speed up the subsequent examples, they are all based on the latter population.)

Each of the 92,524 parcels has been classified according to the four categories, as described in “The Tabulation Definition Zone” of the User Guide section. This means that the population of parcels in any combination of values of the four categories can be calculated in real time, during a tabulation. So, for example, if a one-way tabulation by SSSI were undertaken, it would show that there were 65,074 non-SSSI parcels and 27,450 SSSI parcels within the overall population. In a similar way, with one important exception, each of the parcels was classified as being in favourable or unfavourable condition. This means that the average condition of parcels in any given breakdown of categories can also be calculated in real time. In this example, the average proportion of SSSI parcels in favourable condition was 39.1%.

The important exception is that the favourable/unfavourable classification was only available for SSSI parcels. To give a similar functionality for non-SSSI parcels, overall proportions in favourable condition were estimated for a simple, one-way breakdown by Habitat. Eight of the 19 categories had values provided by NE. The remaining values were set arbitrarily at 0.25, or the SSSI value, whichever was lower (Table 1). No further estimates, broken down by the other categories, were made. These figures gave an overall proportion for non-SSSI parcels of 22.9% favourable condition.

Calculation of a Simple Random Sample

The first analysis excludes any filtering or tabulation, so that a single sample size was calculated from the whole population of 27,879 parcels, based on an initial favourable proportion of 29.1%. Change was set at +50%, which resulted in a final proportion of 38.1%. Sampling strategy was set to independent, with a two-tailed analysis, a power of 90% and a significance of 95%. With these parameter settings, the required sample size was about 560, which equates to a sampling fraction of 0.020¹.

Table 1. Proportions of parcels in favourable condition in each of the 19 Priority BAP Habitats. Figures for SSSI parcels are exact values calculated from the raw data. Figures for non-SSSI parcels are a) estimates provided by NE (*), b) a general value of 0.25 where these were not supplied, or c) the SSSI value if lower than 0.25(+).

Habitat	SSSI	Non-SSSI
Blanket Bog	0.150	0.01*
Coastal and floodplain grazing marsh	0.437	0.25
Coastal sand dune	0.494	0.25
Coastal vegetated shingle	0.595	0.25
Fen	0.342	0.25
Limestone Pavement	0.275	0.25
Lowland calcareous grassland	0.436	0.28*
Lowland dry acid grassland	0.318	0.23*
Lowland heathland	0.292	0.05*
Lowland meadow	0.467	0.18*
Lowland raised bog	0.133	0.13+
Maritime cliff and slope	0.688	0.25
Mudflat	0.523	0.25
Purple moor grass and rush pasture	0.354	0.35*
Reedbed	0.533	0.25
Saline lagoon	0.721	0.25
Upland calcareous grassland	0.429	0.01*
Upland hay meadow	0.692	0.07*
Upland heathland	0.154	0.15+

¹ This estimated sample size is based on the mean from a number of runs of the tabulation tool with the resolution set to low. Due to the nature of the algorithm used for calculating sample sizes, this appears to be a better strategy for obtaining an accurate estimate than using the medium or high resolution, although not as simple.

So, what does this analysis tell us? Essentially, it is saying that if we had a population of 27,879 parcels, with 29% in favourable condition at t_0 , and this changed to 38% at t_1 , then we could be 90% certain of detecting this change by drawing two independent samples of 560 parcels and analysing these with a two-tailed test, setting the critical α to 0.05.

Clearly there is one unknown parameter here, and a number that can be varied by the analyst. It would be worthwhile, in this simple analysis, to see what effect altering the parameters has on the predicted sample sizes.

Firstly, let's just see what influence the analysis parameters have, as these are the simplest for the analyst to modify. Assuming that the monitoring programme is only interested in detecting an increase in proportion of favourable features, a one-tailed test could be used. This would give sample sizes of around 460, which reduces the sampling fraction to 0.016. Instead, if a significance of only 90% ($\alpha = 0.1$ – a very unsatisfactory test!) is acceptable, then the sample size required would be about 454 – roughly the same effect as using a one-tailed

Table 2. Summary of parameter effects on simple random samples. The starting conditions give the basic sample size, followed by the effect on sample size of making one or two changes to various parameters.

Parameter	Starting Conditions		Single Change		Two Changes	
	Value	Sample	Value	Resultant Sample	Value	Resultant Sample
Population	27,879	560				
Initial Proportion	29%					
Change	+50%		100%	188	+100% & one-tailed	155
Turnover	-				+100% & dependent strategy	55
Strategy	Independent		Dependent	100	Dependent & 10% turnover	213
Tails	Two		One	460		
Power	90%		80%	423	80% & one-tailed	332
Significance	95%		90%	454		

test. Finally, if we are happy to be only 80% sure that we will obtain a significant result, then we would only need sample sizes of about 423. Accepting both the reduced power and applying a one-tailed test gives an additive improvement, so that the required sample sizes would now be only around 332.

Secondly, changing the strategy to dependent samples has a huge effect. Setting all the other parameters back to their initial values results in a sample size of only around 100 – less than one fifth the sample size of the independent samples and a sampling fraction of 0.004. Adopting a one-tailed test reduces the sample size further to about 85, as did reducing the power to 80%. However, as indicated in the previous chapter, the effect of turnover on dependent samples can be very great. If there is just 10% turnover, the required sample size would increase to about 213. So, quite a plausible degree of turnover more than doubles the sample size. If the turnover is much greater, say 50%, then the required sample size would be around 577 which gives a worse situation than drawing two independent samples.

Finally, there is one other unknown population parameter that should be investigated, even though it is beyond the control of the analyst. The degree of change in the underlying population has already been shown to have a large effect on the sample size required to detect that change. So, for example if the change is +100% (resulting in a proportion at t_1 of 45%), then the required sample would only be about 188. This could be further reduced to about 155 by employing a one-tailed test. An even greater reduction in sample size to around 55 would occur if a change of +100% is combined with a dependent sampling strategy. On the other hand, a much smaller degree of change, say 25% (resulting in a t_1 proportion of 34%) would require a sample of around 1800 to be detected. These parameter effects are summarised in Table 2.

One-way Tabulation to Create a Stratified Random Sample

A stratified random sample can be created in a number of different ways;



- Σ Firstly, the simplest way would be to draw equal sized samples from each stratum, but this ignores the sizes of the populations in each stratum. There is a risk in this strategy of over-sampling the rarer strata and under-sampling the more common strata.
- Σ Proportional sampling incorporates population size by using a constant sampling fraction across strata. This ensures that more effort is expended on the commoner strata but may risk under-sampling rarer strata.
- Σ Proportional sampling with a minimum threshold can be used to overcome the latter problem.
- Σ Finally, the most desirable method for stratification is one that ensures equal power within each stratum. This ensures that the only the amount of effort likely to provide a significant result is expended on each stratum.

To investigate the effect of one-way tabulation, we will use a subset of the data comprising all coastal habitat parcels ≥ 5 ha in area. This gives a population of 10,767 parcels with a mean proportion in favourable condition of 33%. Applying the starting parameters given in Table 2 results in a sample of 520, or a sample fraction of 0.048.

Tabulation with a Constant Initial Proportion

The coastal habitats comprise six priority BAP habitats as shown in Table 3. We can use the checkbox and combo-box under Population Parameters to set a constant initial proportion of 35% and then carry out a one-way tabulation on these six habitats. The change of +50% results in a proportion at $t1$ of c.45% for all habitats. However, the population sizes vary enormously, ranging from 49 for saline lagoons to 8686 for grazing marsh. The latter would require a sample size of around 500 or a sampling fraction of 0.058 – not greatly different from the overall sampling fraction. However, the saline lagoons would require almost a complete census – 47 out of 49 parcels – or a sampling fraction of 0.959. It can be seen clearly in Table 3a that the sampling fractions for the intermediate populations are proportionate to their sizes, ranging from less than 0.4 to over 0.8. When the samples for these six habitats are tallied, the overall fraction becomes 0.138. This means that even a simple stratification that enables change to be detected in six different habitats requires a much larger overall sample – in this case around 1484 compared to 520.

Furthermore, this doesn't take into account the problem of multiple testing, which means that the specified α -level of 0.05 should be reduced to provide **independent** test results. The simplest way to do this is to use the Bonferroni procedure of dividing the desired α -level by the number of tests, in this case six. The effect of this is to increase the required sample sizes to those given in Table 3b; an overall increase of about 350 sampling units. An alternative way of looking at this would be to retain the approximate sample sizes shown in Table 3a and the α -level of 0.01, which would reduce the power from 90% to about 75%.

Table 3. Sample sizes from a one-way tabulation with constant initial proportions for six coastal BAP habitats. a) α -level set to 0.05 to represent six non-independent tests at a significance of 95%. b) α -level set to 0.01 to represent six independent tests at approximately 95% significance.

Category Value	Proportions		Population	a) Non-Independent		b) Independent	
	$t0$	$t1$		Sample	Fraction	Sample	Fraction
Coastal and floodplain grazing marsh	35%	45%	8686	504	0.058	687	0.079
Coastal sand dune	35%	45%	261	183	0.700	201	0.770
Coastal vegetated shingle	35%	45%	120	100	0.833	106	0.883
Maritime cliff and slope	35%	45%	783	322	0.411	393	0.501
Mudflat	35%	45%	868	329	0.379	406	0.467
Saline Lagoon	35%	45%	49	47	0.959	48	0.980
Total			10767	1484	0.138	1841	0.171



Tabulation with Intrinsically Different Initial Proportions

To investigate the effect of different initial proportions on sample size, the tabulation have been re-run using the underlying proportions as explained in the Introduction to this section (Table 4). These range from 28% for grazing marshes to 59% for maritime cliffs and saline lagoons. Those proportions that are closer to 50% than the universal proportion of 35% applied in the previous analysis, result in slight reductions in the sample sizes. However, the one category which has a smaller proportion (grazing marshes) shows an increase in sample size, resulting in an overall increase of about 40 sampling units.

Table 4. Sample sizes from a one-way tabulation with underlying initial proportions for six coastal BAP habitats. The α -level was set to 0.05 to represent six non-independent tests at a significance of 95%

Category Value	Proportions		Population	Sample	Fraction
	t0	t1			
Coastal and floodplain grazing marsh	28%	37%	8686	548	0.063
Coastal sand dune	40%	50%	261	180	0.691
Coastal vegetated shingle	56%	66%	120	104	0.867
Maritime cliff and slope	59%	69%	783	325	0.415
Mudflat	46%	56%	868	320	0.368
Saline Lagoon	59%	68%	49	46	0.939
Total			10767	1523	0.141

Two-way Tabulation

The problems introduced by tabulation are exacerbated when two-way tabulation is employed (Table 5). This shows the six habitat types, cross-tabulated by SSSI designation. The overall sample size remains at 10,767, but now the sampling fraction has increased to 0.206, resulting in a sample of 2216.

Part of the explanation for this is that the initial proportions for the non-SSSI parcels were all set at 0.25 as described in the Introduction. Furthermore, with one exception, these categories all had small populations, which resulted in sampling fractions of 0.8 or greater. The single exception was the grazing marsh category, which had a sample of 7197. However, this still resulted in a required sample size of 570 – larger than the combined sample size of 548 shown in Table 4. When this is combined with the sample of 377 required for the SSSI designated grazing marsh parcels, it is easy to see how the overall sample size increases rapidly.

Table 5. Sample sizes from a two-way tabulation by Feature and Qualification for habitat features with sampling fractions less than 1.0..

Primary Category	Secondary Category	Proportions		Population	Sample	Fraction
		t0	t1			
Grazing marsh	Non-SSSI	25%	33%	7197	570	0.079
Grazing marsh	SSSI	45%	55%	1489	377	0.253
Sand dune	Non-SSSI	25%	33%	78	71	0.910
Sand dune	SSSI	47%	57%	183	140	0.765
Vegetated shingle	Non-SSSI	25%	33%	12	12	1.000
Vegetated shingle	SSSI	59%	69%	108	93	0.861
Cliff and slope	Non-SSSI	25%	33%	177	140	0.793
Cliff and slope	SSSI	69%	77%	606	324	0.535
Mudflat	Non-SSSI	25%	33%	176	146	0.828
Mudflat	SSSI	51%	61%	692	294	0.425
Saline lagoon	Non-SSSI	25%	33%	9	9	1.000
Saline lagoon	SSSI	67%	75%	40	40	1.000
Total				10767	2216	0.206

The single exception was the grazing marsh category, which had a sample of 7197. However, this still resulted in a required sample size of 570 – larger than the combined sample size of 548 shown in Table 4. When this is combined with the sample of 377 required for the SSSI designated grazing marsh parcels, it is easy to see how the overall sample size increases rapidly.

Modifying the Fixed Parameters

All the previous tabulations use the same values for the fixed parameters; +50% change, independent sampling strategy, two-tailed testing with 90% power and 95% significance. Even if we accept this change value as a reasonable estimate of an unknown value, there are still four parameters that we have control over, although one of these (significance) should probably not be reduced below 95%.

The easiest parameters to modify are the tails and power of the statistical tests. By employing a one-tailed test, rather than two-tailed, the sample size from the one-way tabulation by habitat, as shown in Table 4, can be reduced from 1523 to 1332 – a reduction in sampling fraction from 0.141 to 0.124 (Table 6). Furthermore, by



reducing the power to 80% the sampling fraction can be reduced further to 0.101. But, as described in the “Calculation of a Simple Random Sample”, the most dramatic effect on sample size can be obtained by using a dependent sampling strategy. Even with a two-tailed test at 90% power, the required sample size falls dramatically to 449, and even further to 344 with one-tailed testing at 80% power. Note, however, that dependent samples are almost as susceptible to turnover when tabulated. With only a 10% turnover in the latter scenario, the sampling size increases from 344 to 546.

Table 6. The effect of varying the fixed parameters on sample sizes from a one-way tabulation by habitat.

Strategy	Tails	Power	Sample	Fraction
Independent	Two-tailed	90%	1523	0.141
	One-tailed	90%	1332	0.124
Dependent	Two-tailed	80%	1091	0.101
	One-tailed	90%	449	0.042
	One-tailed	80%	344	0.032

Summary & Conclusions

As a final stage to this exploratory analysis it would be worthwhile to extract an example of the simulated data from the power tool for two purposes.

- Σ Firstly, it would be possible to confirm the accuracy of the power tool, by analysing a sample using a standard statistics package.
- Σ Secondly, it is worth noting that the power tool derives its sample sizes based on independent assessments within each sub-group. In other words, it assumes that the sample required to detect a change in, for example, the grazing march habitat will be analysed independently from the other sub-groups, using a normal approximation to detect a difference between two proportions. But, in reality, a one-way breakdown by feature is most likely to be analysed as a whole, using a statistical model such as logistic regression. By extracting a sample with a one-way breakdown such as this, it would be possible to undertake these more sophisticated analyses.

The power tool was used to generate two random samples of the six coastal habitats used in the preceding analyses. The first sample was a simple random sample with no tabulation (Table 7a), whilst the second sample included a tabulation by the six habitats (Table 7b). The sample sizes were calculated as the arithmetic means of ten runs of the power tool to reduce stochastic effects. The fixed parameter values were; change = +100%, Strategy = independent, two-tailed test, significance = 95% and, most importantly, power = 50%. The reason for choosing this value for power was that it gives a 50:50 chance of a test carried out on this sample being significant. In other words, the result is likely to hover around the $\alpha = 0.05$ level – it may or may not be significant.

This dataset was analysed in a proprietary statistics package¹ in a number of different ways. Firstly, χ^2 and Yates’s-corrected χ^2 gave test results of $p = 0.044$ and 0.064 respectively. Fisher’s exact test gave $p = 0.064$, all of which were very close to the anticipated value. Secondly, a logistic regression model was constructed with a single binary predictor variable (time), which had a significance of $p = 0.045$ – again just as expected. This model gave predicted values for the proportions at t_0 and t_1 of 32.9% and 49.3%, which were almost exactly the values given in Table 7a.

The logistic regression approach was then used to analyse the tabulated sample (Table 7b).

Table 7. Mean sample sizes from two filtered samples; a) a simple random sample and b) a one-way tabulation by habitat.

Category Value	Proportions		Population	Sample	Fraction
	t_0	t_1			
a)					
All six habitats	32.7%	49.3%	10767	74	0.007
b)					
Grazing marsh	28.3%	44.2%	8686	79	0.009
Sand dune	40.1%	57.3%	261	60	0.229
Shingle	55.9%	71.7%	120	53	0.438
Maritime cliff	59.2%	74.3%	783	75	0.096
Mudflat	45.5%	62.5%	868	67	0.077
Saline lagoon	58.9%	74.1%	49	33	0.682
Total			10767	367	0.034

¹ Statistica v 9.1. © Statsoft Inc. 1984 – 2010.



Firstly, each subset was analysed separately in one-way models. This gave values of $p = 0.037, 0.069, 0.091, 0.045, 0.039$ & 0.203 respectively. The three values that were not significant were those where the sample was drawn from smaller populations, especially the saline lagoons. This was almost certainly due to the lack of a FPC in these simple analyses. Because these were included in the power tool, it would have proposed smaller sample sizes in anticipation of their use, which would then fail to give a significant result.

Finally, a two-way logistic model was built with time and habitat as categorical predictor variables, plus their interaction. This accurately modelled all three effects with their predicted mean proportions closely matching the specified values (Figure 18). The difference in favourable proportions between habitats is shown in the graph (a), which was very highly significant ($p < 0.0001$). Similarly, the difference between the two times (b) was highly significant ($p < 0.0001$), whereas the interaction (c) was highly non-significant ($p > 0.9999$). This shows that there was no difference at all in the rates of change between the six different habitats – exactly as was modelled in the power tool.

This latter point is important. In reality, rates of change between habitats (or any other category) would probably have been quite different. The outcome of this is that the simple random sample from this group of features as a whole would have had a higher variance than the power tool created, so that the one-way test would not have given the significant result it did ($p = 0.045$). However, by using a two-way model, any variance between features (both the main effect and the changes) could be accounted for, which would probably result in a more significant overall change.

In summary, the power tool accurately predicted sample sizes that would yield significant results when using simple, independent test statistics. However, more sophisticated analytical models would probably give greater power to detect change and so the sample sizes predicted by the power tool should be considered conservative.

The final point to note from the exploratory analyses is that, for anything other than a simple random sample, the sample sizes will be quite large. This is primarily due to the lack of information in a binary response variable. Rather than just being “favourable” or “unfavourable”, if the response variable had been a numeric value, even just an ordinal scale, it would have yielded greater power.

To illustrate this, two samples of 74 cases were drawn from such a population, with mean values of 1.67 and 2.5 – representing proportions of 0.33 and 0.49 from a 5-point ordinal scale (Figure 19). These two samples have the same size and means as the binary samples shown in Table 7a, which gave a barely significant result of $p = 0.045$. In contrast, when the ordinal samples were analysed for a change in mean score over time, assuming a normal approximation and using one-way ANOVA, the result was $F_{(1, 146)} = 28.5, p \ll 0.001$.

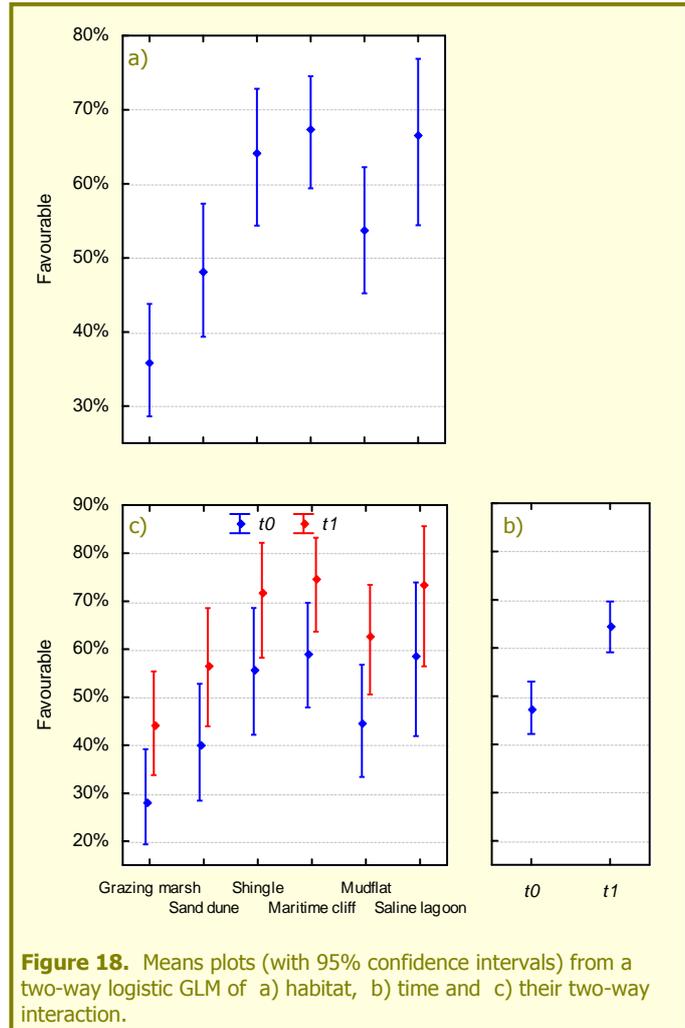
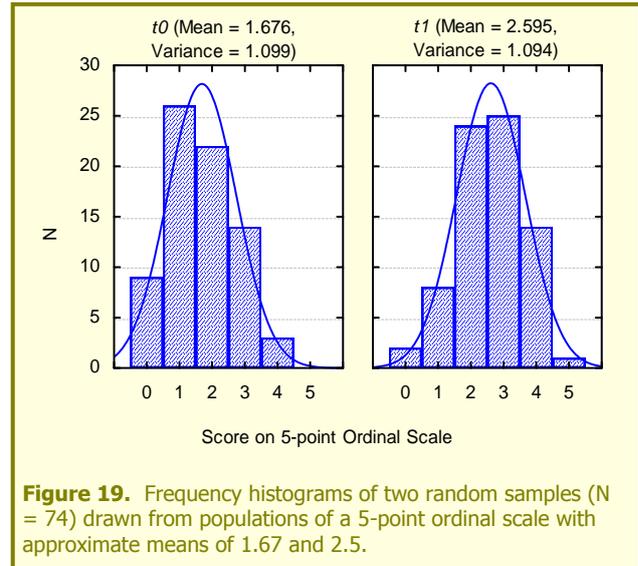


Figure 18. Means plots (with 95% confidence intervals) from a two-way logistic GLM of a) habitat, b) time and c) their two-way interaction.



So, simply changing the response variable from a 2-point scale to the 5-point scale used in this example has an enormous effect on power. Instead of just using the overall condition of the site as a binary response variable, the number of mandatory attributes that were passed could be utilised very easily from the existing CSM methodology.



Appendices

1. Normal Approximation for Difference Between Two Independent Proportions

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}} \times FPC}$$

Where

$$\bar{p} = \frac{X_1 + X_2}{n_1 + n_2}$$

and

$$\bar{q} = 1 - \bar{p}$$

and

$$FPC = 1 - \frac{n}{N}$$

with correction for continuity applied to \bar{p} and \bar{q} .

2. Normal Approximation to McNemar's Test for a Dependent Binary Sample

$$Z = \frac{|f_{12} - f_{21}| - 1}{\sqrt{(f_{12} + f_{21}) \times FPC}}$$

Where

f_{12} = tally of 'zero/one' pairs

and

f_{21} = tally of 'one/zero' pairs

