Modelling of precision and optimisation of sampling design.

Data-file: bqi\_data.xls

Instructions: precision\_optimisation.docx

In this exercise we will use authentic data on benthic quality' index (BQI) from 10 years (2001-2010) at 4 stations in outer Skagerrak to estimate variance components and precision. Each group will use data according to scheme below. Use Excel to solve the tasks below.

Data Data

Group 1 2001 Group 6 2006

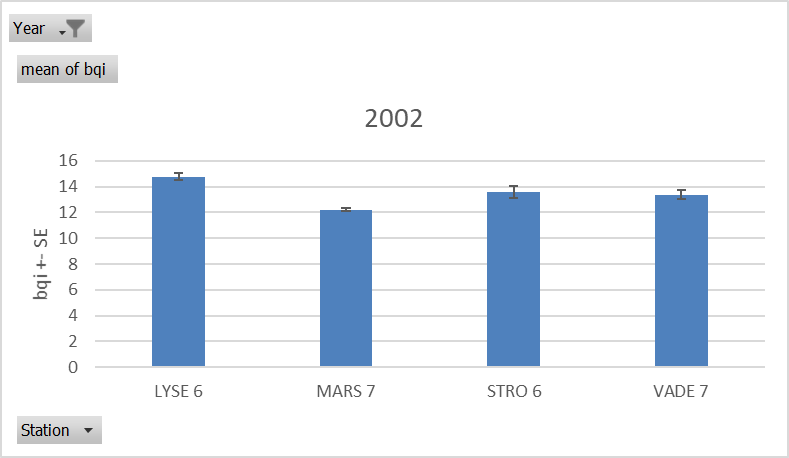
Group 2 2002 Group 7 2007

Group 3 2003 Group 8 2008

Group 4 2004 Group 9 2009

Group 5 2005 Group 10 2010

1. Explore data. Plot means and standard errors for each individual site and year.



2. Calculate variance components for sites and replicate samples for individual years.

Use manual calculations in Excel according to instructions below but if you prefer other statistical software (e.g. r, SPSS, etc.) and know how to estimate components with them it is allowed. Enter results in table below.

3. Calculate overall standard error for outer Skagerrak for individual years using the estimated variance components and existing sampling design. Enter results in table below.

4. Model and plot overall standard error for a range of combinations of sites (b=1-10 and n=1-10).

|  |  |  |  |
| --- | --- | --- | --- |
| SE | 1 | 5 | 10 |
| 1 | 1.177832332 | 0.5267426 | 0.3724633 |
| 2 | 1.085376013 | 0.4853949 | 0.343226 |
| 3 | 1.052754381 | 0.4708061 | 0.3329102 |
| 4 | 1.036058461 | 0.4633394 | 0.3276305 |
| 5 | 1.025910494 | 0.4588011 | 0.3244214 |
| 6 | 1.019089045 | 0.4557505 | 0.3222643 |
| 7 | 1.014188491 | 0.4535589 | 0.3207146 |
| 8 | 1.010497479 | 0.4519082 | 0.3195474 |
| 9 | 1.007617345 | 0.4506202 | 0.3186366 |
| 10 | 1.005307296 | 0.4495871 | 0.3179061 |

5. Use estimates of variability and appropriate formulae for cost-benefit optimisation to calculate the optimal design (i.e. the one that minimises costs\*variance) and its associated standard error. Enter results in table below.

Assume a total budget of 300 000 SEK. The cost for sampling one station is approximated to 15 000 SEK and costs per sample (sampling and sorting) is 3 000 SEK.

*See excel file sheet “task 2” ->left side*

6. Optimise a design for log Biomass (optional if there is time).

*See excel file sheet “task 2” -> right side*

**Instructions for calculation of variance components using Excel:**

Suitable linear model for problem above:

this can be done using routines in r and other statistical software. Because everyone may not have access to such software the variance components may be approximated by the following manual calculations in Excel.

Method 1 (preferred) – based on ANOVA MS estimates

1. Do a one-factor ANOVA using Excel’s “analysis tool pack”.
2. Use residual MS a an estimate of .
3. Calculate by the following formula: =(MSSite-MSResidual)/n

Note that this approximate procedure may result in illogical negative components. If that happens use only estimates from years with positive estimates.

|  |  |  |
| --- | --- | --- |
| Group | Year 1 = 2002 bqi | Year =2002 logbio |
| (from 2) | 0.418 | 0.0605 |
| (2) | 0.969 | 0.0944 |
| Observed SE (3) | 1.1778 | 0.393 |
| nopt  (5) | 1.470 | 1.790 |
| bopt  (5) | 15.457 | 14.728 |
| SEopt  (5) | 0.250 | 0.0800 |