Quality Assurance Report - Analysis of plant species recording bias in Countryside Survey terrestrial vegetation plots – summary report

Simon Smart[1], Andy Scott[1], Ralph Clarke[2]

1 CEH Lancaster, Library Avenue, Bailrigg, Lancaster LA1 4AP

2 University of Bournemouth, Bournemouth University, Christchurch House
Talbot Campus, Poole, Dorset, BH12 5BB

1. Background

Controlling the level of botanical expertise applied in each Countryside Survey field campaign is a difficult task. The integrity of the time-series relies upon consistent recording effort and skill such that changes in botanical indices are not a function of changes in either the number of species recorded or of increasing or decreasing competence in identification. Various methods have been used to decrease the sensitivity of derived indices to recording error. For example, categorising species into those that are less taxonomically demanding, using cover data sparingly and only by aggregating over growth-forms rather than at the level of individual species, and omitting bryophyte records, since this group are known to be poorly censored.

QA surveys have been a vital tool in assessing and validating the quality of the botanical record in each CS. QA surveys that compared the level of agreement between CS field teams and the same team of independent experts, for a subset of the complete set of CS vegetation plots, have been carried out in 1990, 1998 and, most recently, in 2007. The paired species records from these subsets of plots (the QA plots) have been analysed in a number of ways to measure the consistency of recording effort within each survey. In all three surveys the QA assessors found more species than the CS field teams yet in both the 1990 and 1998 assessments, the results showed that there was no bias in the species-composition of the vegetation recorded, as described by DCA analysis, despite differences in species richness.

2. QA 2007 survey results

The 2007 QA analysis showed a decline in the quality of botanical recording. However this was possibly due to the much less comprehensive recording of common bryophytes than in previous surveys. To clarify the trends in bias over the three surveys, follow-up analyses were commissioned from the same QA assessors. They were asked to examine the 2007 QA versus CS records again but omitting bryophytes from all calculations. They were also asked to undertake an analysis of changes in species richness, species composition (DCA axis scores) and mean unweighted Ellenberg scores for the matched "triplicate dataset", i.e. data from only those QA plots that had been surveyed in all three surveys, 1990, 1998 and 2007. The objective of these further analyses was to determine whether there had been a significant change in the level of bias from survey to survey, excluding bryophytes, and the extent to which this change impacted key botanical indices relied upon for core reporting and as carriers of important environmental signals such as eutrophication, disturbance and succession.

These further analyses are reported in an appendix to the initial QA report, both of which accompany this summary. The most important findings of the follow-up work were as follows:

- Errors attributable to use of the tablet PC software were minor and not deemed significant.
- Results showed an estimated increase in 2007 of the **%** of the total recording error that was due to overlooked species suggesting that the problem was not

mis-identification but **overlooking** species present in plots. Thus change in bias could be related to changes in surveyor effort rather than surveyor skill.

- Analyses of change in the magnitude of the difference between QA and CS records over time were ambiguous. Some test results pointed to a change in bias while others did not. However, the results raised the possibility that a threshold had been crossed whereby the difference in mean species-richness and mean Ellenberg N between CS and QA had themselves increased over time.
- Overall the effects could be interpreted as relatively small yet **they were comparable with some of the magnitudes of change declared significant in previous analyses of 1990 – '98 data**.
- Some patterns seemed to be especially worrying, for example a substantial reduction in CS surveyor recording of sedge species in 2007 compared to 1998
  - For sedges the QA in 2007 recorded 28% more occurrences than the surveyors (17% more in the 1998 QA)
- Grasses were also less well censused in plots by the CS surveyors in 2007
  - For grasses the QA in 2007 recorded 17% more occurrences than the surveyors (6% more in the 1998 QA)

3. Actions triggered by the follow-up work

Results from the additional analyses raised the possibility of a significant bias in the 2007 vegetation plot data. As a consequence a number of new courses of action were put into effect.

- The QA report was given to an independent statistician (Ralph Clarke) to inspect and also to Andy Scott, the Countryside Survey statistician.
- With input from Andy Scott, Ralph Clarke was asked to consider an approach to the derivation of bias-correction factors based on the QA/CS dataset, and to advise on options for their application to the raw CS botanical data.
- Simon Smart requested all data from QA plots from the QA assessors, assembled this into a new database of plant species records, and then calculated all botanical indices in preparation for validation and derivation of bias-correction factors.
- Simon Smart reanalysed the triplicate QA dataset from 1990, 1998 and 2007.
- Andy Scott reanalysed the entire QA dataset (repeat and non-repeat data) using the new modelling approach adopted for the analysis of CS2007 data..
- Simon Smart investigated the extent to which differences between QA2007 and CS2007 could be explained by factors such as earliness of CS survey, botanical skill level of the CS surveyors recording each plot, difference in time interval between CS and QA plot visits, as well as the extent to which these factors varied between the 1990 and 1998 surveys. These analyses were carried out in an attempt to determine whether the source of the bias was a small subset of plots defined by criteria that would enable their exclusion from the analysis phase. This approach relied upon the bias being localised within the plot data.

4. Results

<u>Inspection of the QA reports</u>

- The QA analysis appears sound although sometimes the details presented on the methods used are not sufficient to fully understand their analyses and results.
- A comparison of the bias indicated by the triplicate 90-98-07 dataset with the full 98-07 datasets showed that a much greater difference in species richness had been detected in the former suggesting that the smaller triplicate dataset could itself be a biased subset of the total population of plots in each survey.

Development of bias-correction factors

- A statistical approach was developed to derive estimates of the average bias (CS surveyor minus QA) and its standard error in each survey for each botanical index.
- The approach includes methods which could be used to derive bias-corrected estimates of the mean, standard errors and confidence intervals produced from analysis of change in botanical plot indices

Analysis of QA/CS2007 data – searching for factors that explain the bias

- The only variable that was significantly associated with the magnitude of %accuracy values for individual plots was earliness of their CS survey. Earlier CS plot recording tended to result in lower %accuracy of CS compared to QA records. However, the plots with the worst values tended to be recorded by the best botanists in the survey teams and there was no difference in average botanical skill (i.e. % accuracy) of surveyors of these plots than the rest of the CS plots also subjected to QA resurvey.
- Overall, the CS2007 plots visited for QA had a somewhat higher level of average surveyors' botanical skill applied to them than the total CS2007 plot dataset.
- Examination of the distribution of % accuracy values confirmed that the bias was a diffuse phenomenon spread throughout the dataset and not attributable to a small proportion of outlying values or a few individuals whose values might justifiably be eliminated.

Analysis of the total QA dataset using statistical modelling

- A new database of all QA plots for 1990, 1998 and 2007, excluding bryophytes, has been constructed and all response variables calculated for all plots. This dataset was analysed using a mixed effect statistical model with survey square as a random effect to allow for the possible lack of statistical independence of plots within the same survey squares
- The results show that the differences in the levels of bias of species richness measures across surveys, suggested by the analyses carried out by the QA team, **were not significant.**
- Differences between surveys in the level of bias of derived measures, such as Ellenberg scores, were significant for some response variables but of small

magnitude, rendering the application of bias-correction factors less important and desirable.

- The reasons for the differences in conclusions appear to relate to
  - possible bias in the representativeness of the matched (QA and CS) triplicate dataset of plots
  - dampening of the effect of poorer censusing of sedges and grasses when averaged over many plots
- Further work will be undertaken to give a more precise explanation of these differences in results.
- Bias correction factors have been computed for all survey years and all response variables.

5. Conclusions

The following draft recommendations were made by Scott et al (see below).

1) Bryophytes should be excluded from the analysis of CS2007 data as was done for previous surveys. **Implemented**.
2) No adjustment of CS results to correct for bias between CS and QA results should be made. **Implemented**.
3) For derived measures only (and possibly only Ellenberg scores) an adjustment to the 1990 survey results should be made to correct for the *differences* in bias between this survey and the others. **Implemented**.
4) The adjustment should take the form of a static adjustment to the derived measure values prior to full analysis. **Implemented**
5) If thought necessary, an adjustment to the standard errors of 1990 estimates can be made post-analysis using the correction factors already computed plus their standard errors. **Not implemented**.
6) Lessons for future surveys, in terms of training, recruitment of botanists, quality versus quantity of data recorded, and QC of plot recording during each field campaign need to be clearly defined and taken on board.

Annexes

1. QA Plots 1.
2. QA Plots 2.
3. QA Plots 3.