Collaborative Project (large-scale integrating project)
Grant Agreement 226273
Theme 6: Environment (including Climate Change)
Duration: March 1st, 2009 – February 29th, 2012

SEVENTH FRAMEWORK PROGRAMME

WISER

Water bodies in Europe: Integrative Systems to assess Ecological status and Recovery

DELIVERABLE

Deliverable D6.2-3: Report from the end-user workshop "Combination of BQEs into a complete water body assessment"

Lead contractor: European Commission Joint Research Centre (EC-JRC)

Contributors: Wouter van de Bund, Rossana Caroni Angel Borja, Richard Johnson, Ralph Clarke

Due date of deliverable: **Month 32**

Actual submission date: **Month 36**

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013) Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Content

Non-technical summary

The main results of WISER Workpackage 6.2 “Combination of BQEs into a complete water body assessment” and related work on uncertainty in ecological assessment were presented to end users during the Final Conference held in Tallinn, Estonia 25-26 January 2012.

Introduction

During the WISER Final Conference the main results of Work package 6.2 "Combination of BQEs into a complete water body assessment" and related WISER work on classification uncertainty were presented to end users. This was done through two presentations and a hands-on demonstration of the WISERBUGS software:

Presentation on combination of BQE results in a complete water body assessment

This presentation, given by Wouter van de Bund (EC-JRC), focused on the main findings of Work Package 6.2, based on analysis of results of data simulations (carried out by EC-JRC), monitoring data from lakes (provided by SLU, analysed using the WISERBUGS software) and monitoring data from coastal waters (provided by AZTI). The full presentation is available on the WISER website (http://www.wiser.eu/download/fin_conf/06_VandeBund_Combination-of-BQEs.pdf).

Introduction

For the assessment of the ecological status of a water body the WFD requires that several biological quality elements (BQEs) are taken into account. Since the ecological status classes are set and intercalibrated at BQE level it is very important for the comparability of the final classification results that all countries apply similar approaches to combine BQE results into a complete water body assessment. According to the CIS guidance document on classification, ecological the lowest class of all relevant BQEs determines the ecological status of a water body (the "one out – all out" principle). There are no specific requirements on how to combine different metrics used within a BQE; this can be done using "one out – all out", but averaging or other approaches are also acceptable.

Differences in approaches between countries

In spite clarity of the guidance, a variety of approaches for combining BQE results is applied in different European countries, varying from a strict application of the "one out – all out" principle to more pragmatic approaches often involving the application of expert judgment, with potentially serious consequences for the comparability of the final classification results. Also, the number of BQEs used varies between countries.

Important factors that influence the final classification

Obviously, the "one out – all out" approach always gives a lower classification than averaging.

Using both simulated data and real monitoring data, it can be clearly demonstrated that the following factors have a strong influence on the final classification:

- Even a single BQE having a high level of uncertainty will strongly affect the reliability of the final classifications using the "one out – all out" approach. This can be remedied by improving the accuracy of the methods, or to exclude methods with high uncertainty.
- The higher the number of BQEs, the larger the negative bias in the classification outcome for the "one out – all out" approach. This can especially be problematic if all BQEs address the same pressure and the effect is proportional to the level of uncertainty associated with the individual BQEs. Averaging may give a more reliable result in such a case
- If the different BQEs address different pressures, averaging results in a positive bias in the classification; here the "one out – all out" approach gives more correct results, provided that the uncertainty associated with the individual BQEs is not too high
- If within the different BQEs metrics are included that are sensitive to different pressures (e.g eutrophication and acidification) it is better to group them by pressure than by BQE. This will improve the reliability of the final assessment (even if this is not consistent with the recommendations of the classification guidance).

Conclusions and recommendations

The "one out – all out" approach only gives acceptable and comparable results if the different BQEs are complementary (showing the effects of different pressures, showing effects on different temporal and/or spatial scales, showing effects on different aspects of ecosystem functioning). Also, the level of uncertainty in the classification should not be too high and not too different between BQEs.

One has to avoid the blind application of the "one out – al out " rule if those conditions are not met. It is recommended to avoid redundancy between BQEs with regard to the pressures they are responding to, and to use the possibility given by the WFD to exclude BQEs that are too variable wherever necessary. There are good examples where expert judgment is used to avoid the pitfalls mentioned, to ensure that classification is based on the most reliable information available – but the disadvantage of such approaches is that there is always a certain level of subjectivity in the assessment. Such subjectivity can be reduced by decreasing the redundancy and increasing the accuracy of the methods for the BQEs.

Presentation WISERBUGS (WISER Bioassessment Uncertainty Guidance Software) tool for assessing sampling confidence of estimated WFD ecological status of water bodies

In his presentation, given by Ralph T. Clarke, School of Applied Sciences, Bournemouth University, Talbot Campus, Poole, Dorset, BH12 5BB, UK . the WISERBUGS software was introduced to the end users

Overall Purpose

The aim of the software program WISERBUGS is to assist Users in quantifying the effects of (previously- estimated) sampling and other methodological uncertainties on the confidence of estimates of the ecological status of individual water bodies (lakes, rivers stretches, transitional (estuarine) or coastal waters), as required of Member States by the European Water Framework Directive (WFD, 2000).

WISERBUGS software and User Manual was produced by me as WISER Deliverable D6.1.3

The Articles of the WFD (Annex V, section 1.3) require that "Estimates of the level of confidence and precision of the results provided by the monitoring programmes shall be given in the (monitoring) Plan". Thus, water body monitoring and management organisations need to have some understanding and estimates of the confidence to which an individual water body (WB) can be assigned to an ecological status class based on their selected field sampling methods, sample processing protocols and choice of metrics, multiple metric and (optionally) multiple biological quality element (BQE) assessment scheme.

A core part of the WISER project was to collect standardised field sample and survey information on each BQE (phytoplankton, aquatic macrophytes, macroinvertebrates, fish and aquatic habitats) at each of a wide range of lake, transitional and coastal water body sites across Europe. One important reason for this was to improve understanding and provide estimates of the sampling uncertainty (replicate, sub-sample, spatial and temporal) associated with specific sampling/surveying methods, individual metrics, multi-metric indices (MMIs) and multi-metric classification rules. When used with WISERBUGS, this and similar sampling information can help assess which metrics, multi-metric

rules and also combination of BQEs provide the most precise (in terms of sampling uncertainty) assessments of water body status class.

User-specified metrics and multi-metrics and multi-BQE Classification rules

WISERBUGS is designed to be as generic as possible. Therefore, the User has almost a completely free choice (and therefore requirement) to specify the:

i. metrics to be used in the water body assessments,

ii. chosen rules for combining metrics into multi-metric indices (MMI) - examples could be the weighted Inter-Calibration common metrics indices (ICMi)

iii. ecological status class limits for individual metrics, EQR or MMI

iv. rules for combining estimated status classes:

- from individual metrics into a higher level class
- from (stressor-specific) metric groups into an over- all class using that BQE - from each BQE into an overall multi-BQE assessment class for the WB (for example using the worst case rule)

- the combination rules allowed are worst case (one- out-all-out), average (rounded up or down) and median (rounded up or down).

WISERBUGS can usefully be used just to test the effect of new status class limits and multi-metric rules on site/ waterbody status assessments, without any uncertainty assessment (by setting all uncertainty components to zero).

Specification of sampling uncertainty estimates for WB metrics

In order for the software to assess confidence of class based on the User-specified metrics and metric combination rules for a list of specific sampled water bodies, the User must input prior estimates of the relevant sam- pling uncertainty for each estimate of observed metric values or derived EQR for each metric to be involved in the WB assessments. In practice, the estimates of the sampling standard error (SD) for each metric for each water body or site to be assessed within WISERBUGS must be based on best-available information from replicated sampling studies on this or environmentally- similar water bodies, such as as those in the WISER extensive field sampling study. Even where reliable estimates of the sampling uncertainty are not avalable, the User can increase understanding of the consequences for confidnece of class by using a range of trial estimates to represent monitoring sampling schemes of different intensity and thus cost.

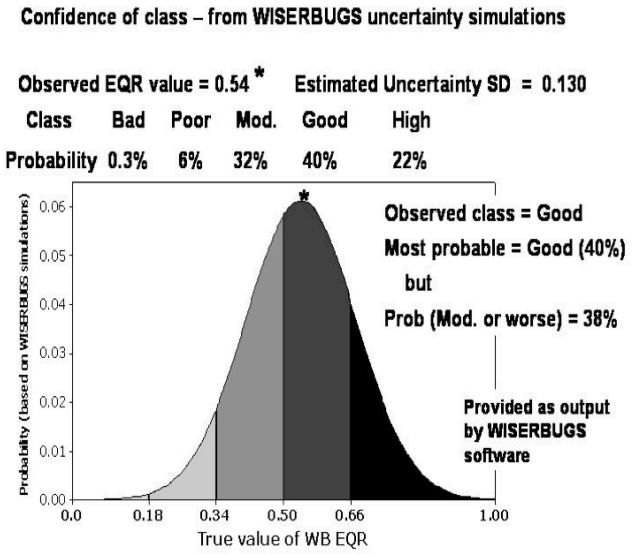

Figure 1: Example of confidence of class estimates from WISER- BUGS simulations.

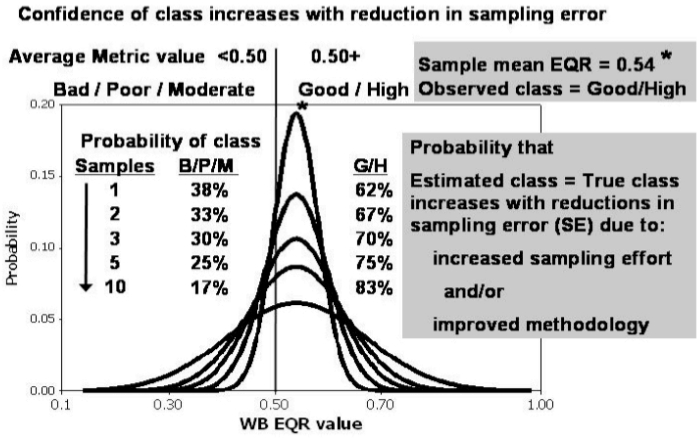

Figure 2: Example showing how confidence of observed class increases with reductions in sampling error.

As an illustrative, (based on WISERBUGS), suppose assessments for a lake are to be based on a single metric EQR for which the status class lower limits are 0.66 (high), 0.50 (good), 0.36 (moderate) and 0.18 (poor). If the lake WB has an estimated observed EQR value of 0.54 based on a single sample and site and the (previously-estimated) sampling uncertainty SD is 0.130, then, although the observed status class is good, with a probability of 40%, there is an estimated 32% chance that the true WB class (based on the average of all possible samples from the lake) is moderate and a 6% prob- ability it is poor (Fig. 1).

However, if the lake mean value of 0.54 was based on more (2,3,5,10) sampling sites around the lake, then our confidence that the true status of this WB is equal to the observed good or better status increases from 62% with one sample to 83% with 10 samples (Fig. 2).

In WISERBUGS, all of the above information on metrics, class limits and rules is supplied by the User in a 'Metric Specification File' in EXCEL format. Detailed help is provided in the WISERBUGS User Manual.

For each set of water bodies to be assessed, the program reads the observed values of each metric or EQR to be used from a User-specified 'Observed metric values' EXCEL file.

The observed metric (or derived EQR) values must have been calculated previously, outside of program WISERBUGS. The layout of this input file (metrics in rows, WB (or samples) in columns was designed to provide immediate compatibility with the metric values EXCEL files derived and output from the freshwater macroinvertebrate sample software known as 'ASTERICS' (obtainable www.eu-star.at).

The ecological status class assessment for individual metrics can be based on:

i. input observed (O) metrics values

ii. input pre-calculated observed (O) EQR values

iii. EQR values derived within the software from the input (O) values as:

$$EQR = (O - E_0) / (E_1 - E_0)$$

where E_1 = Reference Condition value (= value of metric for which EQR = 1), E_0 = anchor value of metric for which EQR = 0 and E_1 and E_0 are supplied by the User (potentially for each WB) in separate input files.

Case (i) and (iii) require uncertainty SD estimates for observed metric values, while case (ii) requires uncertainty SD estimates for pre-calculated EQR. Sampling variability correlations between metrics can also be incorporated to allow for the effect of involving metrics which respond very similarly between samples from the same WB.

Use of simulations to provide estimates of confidence of class

WISERBUGS uses the uncertainty estimates for each metric to simulate a large number of other possible observed metric or observed EQR values which could have been obtained for the same water body using this same sampling monitoring scheme. Non-normally distributed sampling variability of metric values is allowed for by appropriate mathematical transformations in the simulations.

For each simulation, the same User-specified rules that were used for determining single and combined metric status classes from the observed sample WB data values are applied to the simulated observed metric and EQR values to lead to a status class based on each individual metric and combination of metrics in exactly the same way. The resulting frequency distribution of (typically 10000) simulated values and the probabilities of the de- rived classes are used to derive 95% confidence limits for metric values and, most importantly, to provide estimates of the probability of each status class (high, good, moderate, poor and bad) being the true (i.e. complete sampled) class of this WB; based on this user-specified choice of sampling metrics and assessment rules.

The results are all provided in an EXCEL output file, with one line per WB, each with all of the estimates of confidence of each class and EQR confidence limits, for each level of hierarchical grouping of metrics and BQEs, beginning left-most with the overall assessment for the WB and then in increasing detail, all in standard column format. The User can then easily extract the results for their own management or publication purposes.

This provides the type of confidence of status information on single, multimetric and multi-BQE water body assessments required by the WFD.

Disclaimer and Caveat: Assesses precision not accuracy

Uncertainties in estimates of the ecological quality and status class of a site or water body are potentially due to many factors, ranging from the field sampling and sample processing methodology to the choice of high quality sites or metric values to represent the biological Reference Conditions for the site/waterbody.

The approach to assessing 'uncertainty' in program WISERBUGS is simply to estimate the range or variability of estimates of ecological status that could have been obtained using the chosen sampling methods and protocols.

Because the 'true' status class of a site/waterbody is not known, the approach does not try to estimate Type I or Type II errors, but merely to quantify the inherent variability in the methods used to estimate site/waterbody ecological quality. The approach cannot assess whether the metrics used in the bioassessment are good indicators of true ecological quality, but merely whether they give repeatable results. External practical experience with using particular metrics or multi-metric assessments systems must be used to judge their usefulness and reliability to detect the range of biological conditions. Thus the program only assesses aspects of 'precision' rather than 'accuracy'.

The error assessment software must, of necessity, be based on the best available estimates of the various sources of variation and errors in observed metric values and EQRs, as provided by the User (from the WISER project or elsewhere). Sources of variation for which no estimates are currently available are ignored in the error assessment program (and effectively treated as zero). In such cases, the software system will over-estimate the precision and under-estimate the true uncertainty in the assessment of status classes. Any User needs to be made aware of these obvious limitations, especially from the point of view of taking catchment management decisions. However, this software system approach provides a good framework for uncertainty assessment and is a major step forward.

Hands-on session WISERBUGS

Ralph Clarke's presentation on the WISERBUGS software was followed by a hands-on demonstration where end users could explore the possibilities of the software, using the Swedish lake data set used in Work Package 6.2