

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



**Deliverable D6.1-3: WISERBUGS (WISER Bioassessment Uncertainty Guidance Software) tool for assessing confidence of WFD ecological status class**

**User Manual and software: Release 1.1 (Sept 2010)**

Lead contractor: **Bournemouth University (BourneU)**

Contributors: **Ralph Clarke**

Due date of deliverable: **Month 34**

Actual submission date: **Month 19**

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)	



*Deliverable D6.1-3: WISERBUGS (WISER Bioassessment Uncertainty Guidance Software) tool for assessing confidence of WFD ecological status class*

## Content

Content .....	3
1. Overview .....	5
2. Installation .....	11
3. Running the program.....	13
3.1 First run .....	14
3.2 Changing the working directory and your input/output file names.....	14
3.3 Stages of a program run.....	14
4. Input files details .....	17
4.1 Metric specification file (File code name ‘FMETSPEC’).....	19
4.1.(1) Metric names (Column A) .....	20
4.1.(2) Metrics used ?, Observed values or EQR ? (Columns B-D) .....	20
4.1.(3) Ecological status class limits (Columns E-H) .....	21
4.1.(4) Estimates of uncertainty components (Columns I-M).....	22
4.1.(5)-(7) Metric groups, metric weights & Multi-metric rules for combining metrics ...	25
4.1 (5) Codes for multi-metric rule types for combining metrics or their classes .....	25
4.1.(6) Specifying the metric rule type for each top and bottom level group of metrics .....	28
4.1.(7) Specifying the hierarchal grouping of the individual metrics and their weights (Columns N, O, P and Q).....	29
4.1.(8) Number of simulations (cells B15 and B16) .....	31
4.2 Observed values of each metric for each site/waterbody (File code name ‘FOBS’) ...	32
4.2.1 Site/Waterbody names:.....	32
4.2.2 Metric names: .....	32
4.2.3 Observed metric values .....	32
4.2.4 Missing observed metric value indicator.....	33
4.3 EQR metric parameters $E_0$ and $E_1$ (File code names ‘FEZERO’ & ‘FEONE’).....	33
4.3.1 Options for the form and layout of the ‘FEZERO’ and ‘FEONE’ files.....	33
4.3.2 Methods for setting the Reference Condition ( $E_1$ ) value of a metric .....	35
4.4 Correlations between metrics in sampling variation (File code name ‘FCORR’) .....	36
4.5 Direct entry of multi-metric index (MMI) values and their uncertainty .....	37

5. Output file details .....	39
5.1 Log file detailing analysis progress and any problems (file WISERBUGS.LOG) .....	39
5.2 Output EXCEL file of ecological status classes and uncertainty assessment .....	40
HELP .....	44
REFERENCES .....	44
APPENDIX 1. Example input and output files supplied with program WISERBUGS .....	45
APPENDIX 2. Example metric uncertainty standard deviations .....	47

---

Abbreviations:

WB = Water Body, the entity of water management according to the Water Framework Directive (WFD)

BQE = Biological Quality Element, organism groups demanded for assessment and monitoring of water ecological status (fish, benthic macroinvertebrates, aquatic macrophytes, angiosperms, marine macroalgae, benthic freshwater algae, phytoplankton)

EQR = Ecological Quality Ratio, ratio of observed assessment index value to the expected value under reference conditions; EQRs represent normalised index values on a numerical scale between 0 and 1

## 1. Overview

The aim of the software program WISERBUGS is to assist in quantifying uncertainty in the assessment of the ecological status of lakes, rivers, transitional (estuarine) and coastal waters.

WISERBUGS (WISER Bioassessment Uncertainty Guidance Software) is a product of the WISER (Water bodies in Europe: Integrative Systems to assess Ecological status and Recovery) research project (Grant 226273 - WISER) supported by the European Commission under the Seventh Framework Programme (<http://www.wiser.eu/>).

The WISER project aim was to assist in developing methods for calibrating different biological survey results for lakes, transitional and coastal waters against ecological quality classifications to be developed for the Water Framework Directive (WFD) ([http://europa.eu.int/comm/environment/water/water-framework/index\\_en.html](http://europa.eu.int/comm/environment/water/water-framework/index_en.html)). The WFD requires Member States to assess, monitor, and where inadequate, improve the ecological status of water bodies (rivers, lakes, transitional and coastal waters). All such water bodies are to be classified to one of five ecological status classes (high, good, moderate, poor and bad) with the aim of eventually achieving or maintaining good or better status for all water bodies.

The ecological status (i.e. condition) of a water body is often measured using one or more metrics derived from the taxonomic composition and/or abundance obtained from field samples/surveys and/or habitat surveys. The term metric here usually refers to any biological index or other single-valued measure which is designed to measure some aspect of the biological community and its taxonomic composition at a site or water body.

The Articles of the Water Framework Directive (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 can be assigned to an ecological status class.

In addition the WFD requires that the Ecological Status of surface waters of Member States are maintained or improved. However, because of the uncertainties associated with biological monitoring, waterbodies may appear to change Ecological Status over time when, in reality, this is only an artefact due to the uncertainty resulting from the whole bioassessment process and sampling procedures.

A core part of the WISER project was to collect standardised field sample and survey information on 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 and multi-metric classification rules.

The WISERBUGS software program has been written to provide a general means of using simulations to assess uncertainty in estimates of ecological status class for water bodies based on either single metrics or a combination of metrics, multi-metric indices (MMIs) and multi-metric rules. The User provides prior estimates of the relevant sampling uncertainty for each metric and metric value to be involved in the water body assessments, together with metric status class limits and the rules for combining metrics into an overall water body assessment.

WISERBUGS can also 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).

Although initially designed for use with river macroinvertebrate data and metrics, program WISERBUGS is designed to be as generic as possible, so that it can be used with a wide range of metrics derived from field site sampling and survey data for any single or combination of biological quality elements (BQEs, namely phytoplankton, aquatic flora, macroinvertebrates and/or fish) and any type of water body (rivers, lakes, transitional or coastal waters).

The program requires the User to provide a ‘Metric Specification File’ in EXCEL format, in which they specify which metrics are to be used to determine the site or waterbody bioassessments, the individual metric uncertainty estimates and the multi-metric rules for combining information from individual metrics.

The uncertainty in the estimate of the (usually) mean value of a metric for a water body depends on the level of sampling replication on which it was based in terms of replicate sampling, spatial and temporal sampling coverage over the area of the water body to be assessed and the period of time for which the water body assessment is to apply. The estimates of uncertainty in individual metric values can include the sampling standard deviation (SD) due to sampling/sub-sampling variation and (optionally) the SD and bias due to sample sorting and identification.

In practice the uncertainty SD estimates 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.

The ecological status class assessment for individual metrics can be based on just the observed (O) values of metrics or on normalised Ecological Quality Ratios (EQRs) involving the ratio of the observed metric values to the Reference Condition values ( $E_1$ ) of the metric.

More generally, EQRs are determined by:

$$EQR = \frac{O - E_0}{E_1 - E_0} \quad (\text{equation 1})$$

where  $O$  = observed value,

$E_1$  = Reference Condition value (= value of metric for which  $EQR = 1$ )

and  $E_0$  = value of metric for which  $EQR = 0$

Any EQR values calculated from equation (1) which are negative are always reset to zero.

The EQR could be a RIVPACS-type O/E ratio where  $E_1$  is set a RIVPACS model-based site-specific expected value and  $E_0$  is set to zero

When several EQRs are used to create a Multi-Metric Index (MMI) by averaging their values, each EQR is forced into the range 0-1 by setting any EQR values greater than 1 to 1.

If EQRs are used, then the User can provide an estimate of the error SD for the Reference Condition values ( $E_1$ ) of each metric for the group of sites or water bodies to be assessed.

The same User-specified 'Metric Specification File' must also give the ecological status class ('High' to 'Poor') limits for each metric or EQR.

Program WISERBUGS allows the User to specify a wide range of rules for combining individual metrics into multi-metric indices (MMI) or for combining individual metric classes into a metric group class and thus in deriving an overall site/waterbody assessment. For example, it can cope with combining status classes for macro-invertebrate metrics designed to measure one type of stress (e.g. eutrophication, diversity or acidification) and then combine (using the same or a different rule) the classes from these individual stress types into an overall class for macro-invertebrates, and then combine (using the same or a different rule) the overall class for macroinvertebrates with that for one or more other biological quality elements (macrophytes, diatoms or fish).

For each set of sites/waterbodies to be assessed, the program reads the observed values of each metric to be used from a User-specified 'Observed metric values' EXCEL file. The observed values of the metrics must have been calculated previously, outside of program WISERBUGS.

The layout of this input file was designed to provide immediate compatibility with the metric values EXCEL files derived and output from the freshwater macroinvertebrate sample software known as 'AQEMrap' or 'ASTERICS' (obtainable from the EU Fifth Framework Programme river classification project STAR Web site at [www.eu-star.at](http://www.eu-star.at)). The AQEMrap software could be used to calculate and export observed metric values for freshwater macroinvertebrate samples for input to the WISERBUGS program.

The program also needs three other User-specified input files:

'EONE' specifying the Reference Condition ( $E_1$ ) values of each metric for which EQR = 1

'EZERO' specifying the values ( $E_0$ ) of each metric for which EQR = 0

'ECORR' (optional) specifying the correlations between metrics due to sampling variation

The content and format of the User-specified EXCEL input files are explained in detail in Section 4.

WISERBUGS uses the uncertainty estimates for each metric to simulate a large number of other possible observed metric or EQR values which could have been obtained for the site/waterbody. If the sampling SD is considered to be constant on a specified transformed scale, then, in each simulation the observed value ( $x$ ) is first transformed, then a random sampling error ( $z$ ) added and the result back-transformed to obtain a simulated observed value (e.g. for the square-root transformation: simulated observed value =  $(\sqrt{x} + z)^2$ ).

Each simulation leads to a status class based on each individual metric and then groups of metrics in the specified multi-metric rules for site/waterbody bioassessment. From the statistical distribution of simulated values and classes, estimates of the probability of belonging to each status class are obtained.

The content and format of the EXCEL output file give the ecological status assessment and the associated uncertainty for each site/waterbody is explained in Section 5

WISERBUGS (and its precursor STARBUGS) was written and produced by Ralph Clarke of Bournemouth University (BourneU) in the UK.



### Disclaimer:

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 or STAR projects 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 underestimate 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.



*Deliverable D6.1-3: WISERBUGS (WISER Bioassessment Uncertainty Guidance Software) tool for assessing confidence of WFD ecological status class*



## **2. Installation**

- Using Windows Explorer, copy the supplied zip file WISERBUGS\_1\_1.zip to the directory from which you wish to install the WISERBUGS program.
- Within the selected directory, click on the zip file and extract all of the files to the same directory.
- To begin the installation, click on the file SETUP.EXE in this installation directory and follow the instructions.
- By default, the WISERBUGS software package, including the supplied example test data files, will be installed to the directory: C:\Program Files\WISERBUGS'.



*Deliverable D6.1-3: WISERBUGS (WISER Bioassessment Uncertainty Guidance Software) tool for assessing confidence of WFD ecological status class*

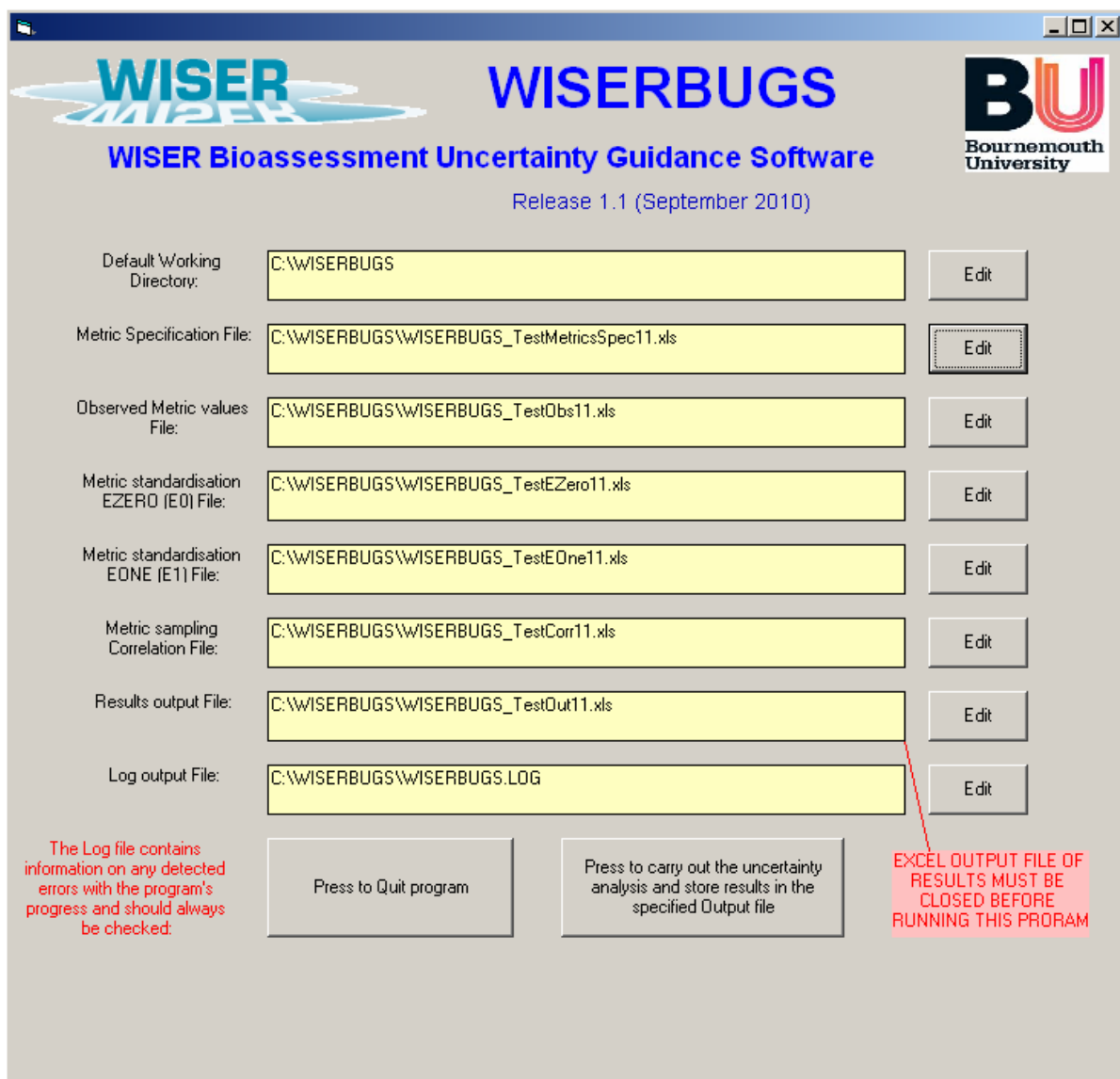
### 3. Running the program

To run the program from the Desktop:

Select the 'Start' menu, 'Programs', 'WISERBUGS', 'WISERBUGS'.

(The WISERBUGS program icon can also be copied to the desktop in the usual way for more immediate access)

The program main menu window is displayed as follows:



The main menu window is where the User specifies all of the input and output files to be used in the analysis. The layout and content of the individual input files are described in section 4 and the output file is explained in Section 5.

### 3.1 First run

On the first run of the software, the default working directory is the start-up directory where the software is stored (e.g. C:\Program Files\ WISERBUGS) and all the input and output files are those supplied with the software package.

On your first run of program WISERBUGS, it is recommended that you simply run the program with the initial default working directory and all of the supplied input and output files, using the instructions below. This will help you familiarise yourself with the way the program works and the type of output you will obtain.

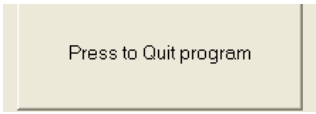
To set up your own analyses on your own data, using your own metrics, multi-metric indices and rules, you must first read Section 4.

### 3.2 Changing the working directory and your input/output file names

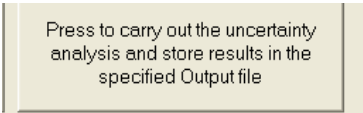
- On start-up, the path of the last used default working directory and the file names (and paths) of the last used input and output files are displayed.
- The default working directory for all of your input and output files can be changed in order to work from any directory of your choice. Simply click on the ‘Edit’ button on the right-hand side of the box and select the required directory (which can be on any directory and drive). This will change the directory of ALL input and output files to this new directory, but will keep the last used filenames.
- Irrespective of whether or not you have just changed the default working directory, you can now click on the ‘Edit’ button for any type of input file to select the filename (and if required independent alternative directory) for your own required previously-created input files of data or parameters. When you change default working directory, it is useful to do this anyway to check that the input files all exist in this new working directory.
- You can click on the ‘Edit’ button for the ‘Results Output file’ or for the ‘Log output file’ to select the filename (and if required independent alternative directory) for these output files. If the output file already exists, you will be asked if you wish to overwrite it. If the output file does not exist or you type the name of a new file in the Windows File Save dialog box, it will be created.

### 3.3 Stages of a program run

- If you want to quit the program, press the button:

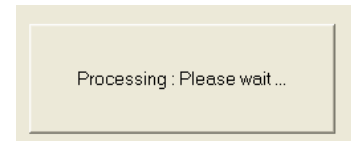


- To begin the analysis using the selected files, press this button



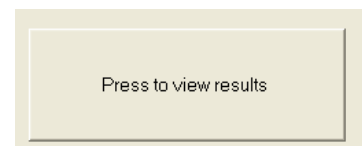
- The program then reads all of the input files and carries out the simulations to assess the uncertainty in your bioassessment of each of the sites/waterbodies in the 'Observed metric values' file.

- Whilst the files are being read and the analysis is being carried out for each site/waterbody in turn, the program displays the message

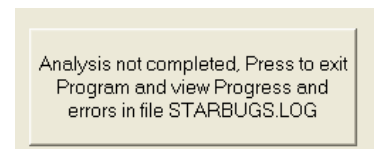


- Progress on reading the input files and working through the analysis is written out to a standard file called 'WISERBUGS.LOG' in the WISERBUGS start-up directory. Any detected errors are also written to the same file – which should always be checked on program completion.

- On completion of the analysis or error detection, this button is displayed. Press the button to continue.

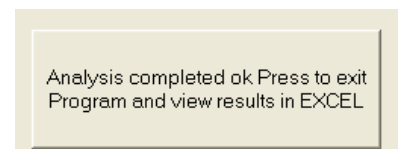


- If the program detected an error in any of the input files, or did not complete the analysis for all, or any, of the sites/waterbodies, then the following message is displayed:



On pressing this button, the Log file WISERBUGS.LOG is displayed using Microsoft Windows' NotePad. Scroll down this file to see how far the reading of the input files and the analysis progressed. The last few lines should usually indicate where the program or analysis failed and what input data and files need to be corrected.

- If the uncertainty analysis was completed ok for all sites/waterbodies then the following message is displayed:



On pressing this button, the results of the uncertainty analysis and multi-metric rules for the assessment of all sites/waterbodies are displayed in the User-specified EXCEL Output file. Details and explanation of the output are given in Section 5 to which you are referred. You could optionally save the WISERBUGS Results Output EXCEL file to another name.

- You must close the Results output file before re-running the software and outputting to the same file, or the program will stop with an appropriate warning.
- On either closing EXCEL or otherwise switching back to the still open WISERBUGS program, you can re-run the program with the same or different input and output files, or press the 'Press to quit program' button to close the program.



*Deliverable D6.1-3: WISERBUGS (WISER Bioassessment Uncertainty Guidance Software) tool for assessing confidence of WFD ecological status class*



## 4. Input files details

All of the input files are EXCEL (.XLS) files. Five input files are required:

### Input EXCEL files

File Code name	Content
'FMETSPEC'	Metric specification file (uncertainty estimates + multi-metric rules)
'FOBS'	Observed values of each metric for each site/waterbody
'FEONE'	Reference Condition ( $E_1$ ) values of each metric for which EQR = 1
'FEZERO'	Values ( $E_0$ ) of each metric for which EQR = 0
'FCORR'	(optional) Correlations between metrics in their sampling variation

The contents, layout and format of each type of input files are explained in Sections 4.1 – 4.4.

Cells left blank are treated as missing values, zero values or ignored; as appropriate.

Working examples of each type of input file are listed below and in Appendix 1.

### Example Test files

File Code name	Example file
'FMETSPEC'	WISERBUGS_TestMetricsSpec11.xls
'FOBS'	WISERBUGS_TestObs11.xls
'FEONE'	WISERBUGS_TestEZero11.xls
'FEZERO'	WISERBUGS_TestEOne11.xls
'FCORR'	WISERBUGS_TestCorr11.xls - with some sampling correlations
	WISERBUGS_TestCorrNull11.xls - with zero sampling correlations

Two blank input files, listed below, are supplied with the software package.

The blank 'FMETSPEC' file should be used as the starting point to build up your own 'Metric Specification File' detailing the metrics you wish to involve and their class limits and other details as described in Section 4.1.

The contents, layout and purpose of the 'FCORR' file are described in Section 4.4.



Blank input files for modification by the User

File Code name	Example file
'FMETSPEC'	WISERBUGS_BlankMetricsSpec11.xls
'FCORR'	WISERBUGS_BlankCorr11.xls

### 4.1 Metric specification file (File code name 'FMETSPEC')

This EXCEL worksheet file specifies most of the parameters of the analysis, namely the metrics to be used, their uncertainty component estimates and the sets of multi-metric rules to be applied. It is divided into several main sections, as indicated below:

The layout (but not the User-defined parameters) of the spreadsheet and in particular, the rectangular block of cells A1–Q20, is assumed by the program to be fixed and should not be altered or the program will probably crash.

The program is supplied with a working example of this type of file.

To make their own version of a 'Metric Specification file', Users should begin by making a copy of the supplied example file which they then edit, as required.

Codes for each type of Multi-metric rule available      Number of simulations      Multi-metric rule to be used for each named top and bottom level hierarchal grouping of metrics

WISERBUGS : WISER Bioassessment Uncertainty Guidance Software		File specifying sampling variations and other errors, ecological status class limits for metrics and the Multi-metric rules for each metric group		Specification of Metric rule type for each Top level and each bottom level Metric		Bottom level Metric group		Uncertainty in observed (O) value		Uncertainty in reference condition (E) value		Required Hierarchy of metric groups			
Multi-metric rule types:		Base assessment on:		Index is integer		Status class boundaries (for O or OE)		Sampling variation		Sorting/Identification		Top level group		Bottom level group	
1 = Worst class (Default if no type (or zero) specified)		0 (or blank) = No use in error assessment; 1 = Observed values; 2 = EQR values		1 = Yes (or blank) = No		Inclusive class limit		SD constant when Observed values transformed to:		Average bias		Metric Group		Metric Weight (if not equal)	
2 = Average class (rounded better)						high good moderate poor		1 = square root; 2 = fourth root; 3 = arcsine for p; 4 = arcsine for p/s;		bias SD		Metric Group		Metric Weight (if not equal)	
3 = Average class (rounded worse)								0 = untransformed				Metric Group		Metric Weight (if not equal)	
4 = Median class (rounded better)												Metric Group		Metric Weight (if not equal)	
5 = Median class (rounded worse)												Metric Group		Metric Weight (if not equal)	
6 = Average metric values to create Multi-Metric Index												Metric Group		Metric Weight (if not equal)	
Number of simulations required (max 10000) = 10000												Metric Group		Metric Weight (if not equal)	
Integer random number seed (-1 uses current time) = 234												Metric Group		Metric Weight (if not equal)	
Metric name															
Abundance (Indm <sup>3</sup> )															
Number of Tass															
Saprobic Index (Zetinka & Marvan)	1	1				1.7	2	2.3	2.6	0.12	0	0	0	0	2
Average score per Tasson	1	2				0.89	0.78	0.67	0.56	0.249	0	0	0	0.061	2
DSFI	1	1	1			6	5	4	3	0.56	0	0	0	0	2
ISE Aqem															
Diversity (Shannon-Wiener Index)	1	1				2.9	2.4	1.9	1.4	0.23	0	0	0	0	1
Diversity (Margalef Index)															
Acid Index (Hendrikson & Medin)	1	1	1			12	9	6	3	12	0	0	0	0	3
EPT-Tass															
Number of Families	1	2	1			0.85	0.6	0.45	0.3	0.228	1	1.5	1.22	0.53	1
Index A	1	1								0.08					4
Index B	1	1								0.12					4
Index C	1	1								0.14					4
NMI based on Indices A,B and C	6					0.8	0.6	0.4	0.2						4

(1) Metric Names

(2) metric used ? Observed or EQR ?

(3) Ecological status class limits

(4) Estimates of uncertainty components (sampling SD, etc)

(7) multi-metric grouping and weights for metrics

#### 4.1.(1) Metric names (Column A)

Individual metrics can be given any name up to 100 characters providing they are valid as strings within EXCEL cells.

In the examples provided, the names used are often those output from the AQEMrap metric calculation software. However, the program is designed to also be able to use metric values calculated by other software or methods and, if required, for taxonomic groups other than macro-invertebrates.

Program WISERBUGS links the metric specification details in the ‘Metric Specification file’ with the observed and other metric values in the other input files (FOBS, FEZERO , FEONE and FCORR) entirely by matching their metric names. Therefore:

THE NAME OF ANY PARTICULAR METRIC MUST BE EXACTLY THE SAME IN ALL OF THE INPUT FILES.

#### 4.1.(2) Metrics used ?, Observed values or EQR ? (Columns B-D)

Cells left blank are treated as zero (0) values.

*Column B: Use in error assessment ?*

To involve a metric in the bioassessment, enter a ‘1’ in the spreadsheet Column B labelled ‘Use in error assessment’. This highlights the cell in pink.

Where multi-metric indices (MMI) are involved, enter the name for the derived MMI in a new row of column A and the enter a ‘6’ in Column B of the same row. This highlights the cell in light blue (the colour used to denote multi-metric information).

To remove a metric from the bioassessment, either delete its value in Column B or enter ‘0’.

	A	B	C	D	S
18		Use in error assessment:	Base assessment on:	Index is integer	
19		0 (or blank) = No 1 = use as metric 6 = use as MMI	1 = Observed values 2 = EQR values	1 = Yes 0 (or blank) = No	
20	Metric name				
21					
22	Abundance [ind/m <sup>2</sup> ]				
23					
24	Number of Taxa				
25					
26	Saprobic Index (Zelinka & Marvan)	1	1		
27					
28	Average score per Taxon	1	2		
29					
30	DSFI	1	1	1	
31					
32					
33	Number of Families	1	2	1	
34					
35	Index A	1	1		
36	Index B	1	1		
37	Index C	1	1		
38	Multi-metric Index based on Indices A,B and C	6			
39					

**Column C: Base assessment on : Observed values (1) or EQR vales (2) ?**

Enter a ‘1’ if the ecological status class for this metric is to based on its observed values (e.g. as for the Saprobic index in the example below)

Enter a ‘2’ if the ecological status class for this metric is to based on its derived EQR value whereby the observed value is normalised by its expected RC value and range (e.g. as for the ‘Average score per Taxon’ in the example below).

**Column D: Index is integer ?**

Enter a ‘1’ if the metric can only take integer values. This ensures the simulated observed values and the confidence limits are also integers (e.g. as for DFSI in the example below).

**4.1.(3) Ecological status class limits (Columns E-H)**

The program complies with the WFD in using five ecological status classes :

‘high’ ‘good’ ‘moderate’ ‘poor’ ‘bad’.

For each individual metric, or MMI, to be used the User must provide the ecological status class limits for the ‘high’, ‘good’, ‘moderate’ and ‘poor’ classes (*Columns E, F, G and H*).

If the metric is to be used in its EQR form, the limits must be for the EQR (e.g. as for the ‘Average score per Taxon’ in the example below).

	A	E	F	G	H
17					
18		Status class boundaries (for O or O/E)			
19		Inclusive class limit			
20	Metric name	high	good	moderate	poor
25					
26	Saprobic Index (Zelinka & Marvan)	1.7	2	2.3	2.6
27					
28	Average score per Taxon	0.89	0.78	0.67	0.56
29					
30	Index A				
31	Index B				
32	Index C				
33	Multi-metric Index based on Indices A,B and C	0.8	0.6	0.4	0.2
34					

If the observed values of the metric are to be classified, then the limits should be in terms of the observed values (e.g. as for the ‘Saprobic Index’ in the example below).

If higher values of a metric or its EQR indicate poorer status, then the limits entered should be the upper inclusive value for each class (e.g. in the Saprobic Index example below,

sites/waterbodies with observed values greater than 1.7 but less than or equal to 2.0 are classed as ‘good’; all values > 2.6 are classed as ‘bad’).

If lower values of a metric or its EQR indicate poorer status, then the limits entered should be the lower inclusive value for each class (e.g. in the ‘Average score per Taxon’ example, sites/waterbodies with EQR values <0.89 but greater than or equal to 0.78 are classed as ‘good’; all EQR values <0.56 are classed as ‘bad’).

Metrics which are part of a multi-metric index (MMI), but are not individually assigned to status classes as part of the bioassessment do not need any class limits specified. However, the derived MMI, usually scaled to lie in the range 0-1, whose values are to be classified into status classes, does need its class limits defined. (See the example above where metrics ‘Index A’, ‘Index B’ and ‘Index C’ are to be individually normalised to EQRs in the range 0-1 and then their EQR values averaged to form a new MMI, whose name and class limits are specified in the worksheet.).

#### 4.1.(4) Estimates of uncertainty components (Columns I-M)

For each metric to be used in the bioassessment, best available estimates are required of the various potential sources of uncertainty in the observed metric and EQR values.

	A	I	J	K	L	M
17						
18		Uncertainty in observed (O) value				Uncertainty in reference condition (E) value
19		Sampling variation	SD constant when Observed values transformed to:	Sorting/identification		
20	Metric name	SD	0 = untransformed 1 = square root 2 = fourth root 3 = arcsine for p 4 = arcsine for p%	Average bias	bias SD	SD
25						
26	Saprobic Index (Zelinka & Marvan)	0.12	0	0	0	0
27						
28	Average score per Taxon	0.249	0	0	0	0.081
29						
30	Number of Families	0.228	1	1.5	1.22	0.53
31						
32	Index A	0.08				
33	Index B	0.12				
34	Index C	0.14				
35	MMI based on Indices A,B and C					

The overall joint effect of field sampling variation and subsequent sub-sampling procedures on observed metric values can be estimated by statistical analysis of variance techniques. Within the European 5<sup>th</sup> framework STAR project, replicates samples were taken at a wide range of stream types throughout Europe and this information has been used to provide the example estimates of the sampling standard deviation (SD) of a range of metrics (see Appendix 2). Sampling variance here is the variance between samples from a water body over the assessment

period. Sampling SD is simply the square root of the sampling variance, but is quoted in preference because it is in exactly the same units as the metric values and thus more easily interpreted.

#### *Columns I and J : Sampling variation SD*

The variability between replicate sample values of a metric for a site may tend to increase or vary systematically with their average value for the site. In such cases, statistical analysis may show that the sampling variance is less variable when the metric values are transformed using the square root ( $\sqrt{x}$ ), double square root ( $\sqrt{\sqrt{x}}$ ), arcsine( $\sqrt{x}$ ) or arcsine( $\sqrt{(x/100)}$ ) transformation. As an example, Clarke *et al* (2002) found that sampling SD of the square root of the number of macroinvertebrate families present in RIVPACS samples from UK rivers showed no systematic pattern and was on average estimated as 0.228. To allow for this in program WISERBUGS, the User should enter a '1' in Column J and 0.228 in Column I of the row for 'Number of Families' metric (as illustrated above). Raw abundance metrics are more likely to have constant sampling SD on the double square root scale (which is similar to a logarithmic scale but avoids the need to add one to zero abundances). Metrics which are proportions or percentages (e.g. EPT taxa as a percentage of all individuals) tend to have highest sampling variability at intermediate values and their sampling SD are less variability on a arcsine( $\sqrt{x}$ ) or arcsine( $\sqrt{(x/100)}$ ) transformed scale respectively.

For many metrics, the sampling SD can, at least initially, be assumed to be constant.

#### *Columns K and L : Sorting/identification bias and SD*

This is only applicable where the Reference Condition (RC) metric values ( $E_1$ ) for a site/waterbody were derived from field samples sorted and identified to a higher standard than used for the observed sample metric values. In this case the Observed sample may tend to have more taxa missed or mis-identified than for the RC sites. This means that certain taxonomic richness type metrics will tend to be under-estimated (bias) in the observed samples and underestimate the true site/waterbody quality. For example, in the UK, CEH used a quality audit re-analysing a proportion of the UK government agencies RIVPACS samples to provide annual estimates of the average under-estimation in the number of taxa present; this "bias" and the SD in the number of taxa missed can be incorporated into the site/waterbody bioassessments. As an example above, the 'Number of Families' metric for observed samples is assumed to have an average bias of 1.5 families with a additional SD component of 1.22 ( $= \sqrt{1.5}$ ).

The effect of sample processing errors on other metrics can be more complex is thus harder to synthesise in a simple form.

In the STAR project sampling programme, all of the samples taken from sites within one stream type, including those used to set trial estimates of Reference Condition, were taken and processed to the same standard (usually by the same people). In such cases, there is no bias for

the Observed sample metric values (relative to the RC values) and these components can be ignored or set to zero.

*Column M: Uncertainty in Reference Condition ( $E_I$ ) value*

The Reference Condition value ( $E_I$ ) of a metric for a site/waterbody is usually determined by the observed value of the metric for a set of high quality sites, considered to be in ecological Reference Condition, which are of the same or similar physical/environmental type. The RC value could be either site-specific (and even site and season specific), as in the case of RIVPACS-type predictive models, or stream-type specific, as implied by using the WFD System type A typology for grouping sites within an Ecoregion.

Whenever, the RC metric values are based on a set of sampled sites, they will be prone to estimation error.

For example, in UK RIVPACS, the expected (ie. RC) values of the standardly used metrics 'Number of Families' and 'Average score per taxon' are both based on generally large-sample weighted averages of RIVPACS reference sites; the error SD in the  $E_I$  values were estimated only from the variability due to errors in measuring or estimating the RIVPACS environmental model predictor variables – estimated using multiple personnel, and given for illustration in Column *M* above.

If a stream-type specific target value is used for  $E_I$  and estimated from (say) the mean metric value for a sample of  $m$  Reference Condition sites, then the error SD for the  $E_I$  value for sites in that stream type could be estimated as the standard error (SE) of the mean (i.e. the SD of the  $m$  values divided by  $\sqrt{m}$ ). In the STAR project, most partners used a stream-type specific approach to selecting RC sites and estimating metric values.



#### 4.1.(5)-(7) Metric groups, metric weights & Multi-metric rules for combining metrics

Program WISERBUGS has been designed to be as flexible as possible in allowing the User to specify and assess a wide range of rules for combining information from individual metrics and Ecological Quality Ratios (EQRs) into an overall bioassessment for a site. Indices derived from two or more other metrics or indices are often referred to as ‘multi-metric indices’, denoted here by MMI. EQRs are as defined in Equation (1) in Section 1.

Two intermediate hierarchal levels of grouping the metrics are possible; these are referred to as ‘Top level groups’ and ‘Bottom level groups’, the latter occurring within a subset of a ‘Top level group’ of metrics. Up to 8 groups of metrics are permitted at each level.

*Sequence of (optional) steps in the overall assessment based on all selected metrics:*

- i. convert individual metrics observed values into EQRs
- ii. classify individual metric observed or EQR values into status classes
- iii. combine individual metric EQRs into a Bottom level group MMI and class
- iv. combine individual metric classes into a Bottom level group class
- v. combine individual metric EQRs into a Top level group MMI and class
- vi. combine individual metric classes and/or Bottom level group classes into a Top level group class
- vii. combine individual metric EQRs into an Overall MMI and class for the site
- viii. combine Top level group classes into a Overall status class for the site

Section 4.1.(5) describes the rule types available within program WISERBUGS for combining metric EQRs or their status classes.

Section 4.1.(6) describes how the rule type is specified for combining metrics in each bottom level group, in each top level group, and in the overall site/waterbody bioassessment.

Section 4.1.(7) describes how individual metrics are (optionally) assigned to top or bottom level metric groups, and (optionally) assigned different weights in calculating group averages.

#### 4.1 (5) Codes for multi-metric rule types for combining metrics or their classes

Cells A6:12 specify the codes for the rule types available.

Rules 1-5 are for combining status classes of the metrics within a group

Rule 6 creates a Multi-metric index (MMI)

WISERBUGS	
A	
6	Multi-metric rule types :
7	1 = Worst class (Default if no type (or zero) specified)
8	2 = Average class (rounded better)
9	3 = Average class (rounded worse)
10	4 = Median class (rounded better)
11	5 = Median class (rounded worse)
12	6 = Average metric values to create Multi-Metric Index
13	

Rule 1: Worst class = take the worst of the classes of the individual members of the group

Rules 2 and 3: are both based on the average of the (numerical) classes of the members of the group rounded to the nearest class

(where ‘high’ = 5, ‘good’ = 4, ‘moderate(mod)’ = 3, ‘poor’ = 2, ‘bad’ = 1).

Thus ‘high’ + ‘good’ + ‘poor’ =  $(5 + 4 + 2)/3 = 11/3 = 3.67$  which is rounded to 4 = ‘good’

while ‘good’ + ‘mod’ + ‘mod’ =  $(4 + 3 + 3)/3 = 10/3 = 3.33$  which is rounded to 3 = ‘mod’

The difference between Rules 2 and 3 is when the average is mid-point between two classes.

Thus ‘high’ + ‘good’ =  $(5 + 4)/2 = 9/2 = 4.5$ , for which:

Rule 2 rounds up/better to 5 = ‘high’, while Rule 3 rounds down/worse to 4 = ‘good’

Rules 4 and 5: are both based on the median class of the group of metrics

(median class is always one of the observed classes)

Case (a): Odd number of metrics in group : (same for Rules 4 and 5)

median is middle ranking status class

e.g. median of ‘high’ + ‘good’ + ‘poor’ = ‘good’

median of ‘high’ + ‘high’ + ‘high’ + ‘moderate’ + ‘poor’ = ‘high’

Case (b): Even number of metrics in group :

Rule 4

Rule 5

(round better)

(round worse)

e.g. median of ‘good’ + ‘bad’

: ‘good’

‘bad’

median of ‘high’ + ‘high’ + ‘moderate’ + ‘poor’ :

‘high’

‘moderate’

**Rule 6** : makes a Multi-metric index (MMI) equal to the average value of the Ecological Quality Ratios (EQRs) all of the metrics in the group, and then classifies the MMI into status classes

All metrics involved in the MMI must have been converted into an EQR by specifying option 2='Base assessment on EQR values' in Column C of the 'Metric Specification file'.

This converts the observed (*O*) value into any EQR, based on equation (1), repeated here:

$$EQR = \frac{O - E_0}{E_1 - E_0} \quad (\text{equation 1})$$

where *O* = observed value,

*E*<sub>1</sub> = Reference condition value (= value of metric for which *EQR* = 1)

and *E*<sub>0</sub> = value of metric for which *EQR* = 0

Notes:

When EQRs are used in a MMI, their values are always forced to lie within the range 0 – 1.

Thus any EQR values > 1 are reset to 1, and any EQR values < 0 are reset to 0.

*E*<sub>0</sub> and *E*<sub>1</sub> are specified in separate input files described in Section 4.3

The name for the derived MMI and its status class limits are specified by adding a new row to the 'Metric Specification file' spreadsheet, giving a name for the MMI in Column A, entering a '6' in Column B to indicate it is a derived MMI rather than an ordinary metric, entering the lower limits of each status class of the MMI in Columns E – H. Finally these details are linked to the individual metrics upon which the MMI is based, by assigning the MMI to the same top level metric group (column N) and optionally bottom level group (Column P) as was used for the individual metrics.

Different weights can be attached to each metric involved in determining the average EQR value used as an MMI, as described below under the heading 'Weights in Column O and Q'.

#### 4.1.(6) Specifying the metric rule type for each top and bottom level group of metrics

Cells block E8:N16 of the ‘Metric Specification file’ is used to specify the multi-metric rule to be used for combining the metrics in each top and bottom level group of metrics, Optional relative weights for top level groups can also be specified in Cells N9:N16.

The User must give a name for each top level group of metrics in cells D9:D16.

In the example below, top level group 1 has been named ‘MG1 Taxa/Diversity’ in Cell D9, and the rule for combining classes of metrics in this top level group 1 is coded 2 in cell E9 (i.e. take the Average class (rounded better) as defined in cell A8 and explained above. Similarly, top level group 2 has been named ‘MG2 Eutrophication’ and also uses rule type 2, as does top level group 3 named ‘MG3 Acidification’.

Top level group 4 simply labelled ‘MG4’ in Cell D12, is specified in Cell E12 as using rule code 6 (which indicates from Cell A12 that the EQR values of metrics assigned to this group should be averaged to derive a new multi-metric index (MMI)). The name of the actual MMI, as opposed to its group name, is specified as a new row amongst the individual metrics (see Section 4.1.(7) for further details).

Cell E8 specifies the type of rule to used to combine the top level group metrics into the overall status class for the site.

In this example, the overall class is based on the rule type 1 (‘Worst case’), the default, which means it is the worst status class of all of the top level groups and metrics.

In the example below, this means taking the worst class of the classes based on the top level groups ‘MG1 Taxa/Diversity’, ‘MG2 Eutrophication’, ‘MG3 Acidification’ and the group ‘MG4’ which is the class derived from a MMI.

Cells N9-N16 can be used to specify the relative weights to be assigned to each top level group if an averaging rule (codes 2 or 3) is being used to derive the overall class for a site. In the example below, the relative weights for the top level groups 1-4 are given in Cells N9-N12 as 1, 3, 2 and 1 respectively [Note these weights are not actually used if the overall rule type is ‘Worst case’]

	C	D	E	F	G	H	I	J	K	L	M	N
6	Specification of Metric rule type for each		Bottom level Metric group								Weight for each top	
7	Top level and each bottom level Metric		0	1	2	3	4	5	6	7	8	level metric group in overall assessment (default = 1)
8		Overall	1									
9	Top	MG1 Taxa/Diversity	2									1
10	Level	MG2 Eutrophication	2									3
11	Metric	MG3 Acidification	2									2
12	group	MG4	6									1
13		MG5										
14		MG6										
15		MG7										
16		MG8										
17												

#### 4.1.(7) Specifying the hierarchal grouping of the individual metrics and their weights (Columns N, O, P and Q)

Column N is used to specify the top level grouping each metric

Column P is optionally used to specify the bottom level grouping each metric

By default each metric is given equal weight in determining the group class based on either the average of the individual metric classes or the MMI based on the average of the individual metric EQR values.

However, the User can specify a non-uniform set of relative weights for the metrics in a group. The weights are given in column O for a metric (or even MMI) in top level group or in column Q for a metric in a bottom level group.

	A	N	O	P	Q
18		Required Hierarchy of metric groups			
19		Top level group		Bottom level group	
20	Metric name	Metric Group	Metric Weight (if not equal)	Metric Group	Metric Weight (if not equal)
23					
24	Saprobic Index (Zelinka & Marvan)	2			
25					
26	Average score per Taxon	2			
27					
28	DSFI	2			
29					
30	Diversity (Shannon-Wiener-Index)	1			
31					
32	Diversity (Margalef Index)				
33					
34	Acid Index (Hendrikson & Medin)	3			
35					
36	Number of Families	1			
37					
38	Index A	4	1		
39	Index B	4	2		
40	Index C	4	1		
41	Multi-metric Index based on Indices A,B and C	4			
42					

In the first example above there are only Top level groups: the metrics ‘Diversity (Shannon-Wiener-Index)’ and ‘Number of Families’ are in top group 1, ‘Saprobic Index’, ‘Average Score per Taxon’ and ‘DSFI’ are in group 2, and the metric ‘Acid Index (Hendrikson & Medin)’ is group 3 on its own.

The final top level metric group 4 comprises the three metrics labelled ‘Index A’, ‘Index B’ and ‘Index C’; the row named (by the User) ‘Multi-metric Index based on Indices A, B and C’ in Column A, is also given a ‘4’ in Column N to identify it with the three metrics, whilst the ‘6’ in Column B identifies this row as being a derived MMI (as described above).

*Weights in Column O and Q*

In this example, the MMI is calculated as the average of the EQR values from the three metrics ‘Index A’, Index B’ and ‘Index C’. However, twice as much weight is given to ‘Index B’ as the other two metrics, as specified by the relative weights 1, 2 and 1 specified in Column O above.

The same relative weights (1, 2, 1) could equally have been specified as (0.25, 0.50 and 0.25) or (30, 60, 30).

In general, with  $m$  metrics in a group, and weight  $w_i$  assigned to metric  $i$  with EQR  $X_i$ , the weighted average value for the MMI is:

$$MMI = (w_1X_1 + w_2X_2 + \dots + w_mX_m) / (w_1 + w_2 + \dots + w_m)$$

The same weighting approach can be used to calculate the ‘average’ of the status classes of the individual metrics in a group. For example, with three metrics of class ‘high’(5), ‘good’(4) and ‘poor’(2), and with respective weights 1, 1 and 3, then the weighted average class for the metric group is:

$$\begin{aligned} \text{group class} &= (1 \times \text{‘high’} + 1 \times \text{‘good’} + 3 \times \text{‘poor’}) / (1 + 1 + 3) \\ &= 1 \times 5 + 1 \times 4 + 3 \times 2 / 5 = 15 / 5 = 3.0 = \text{‘moderate’}. \end{aligned}$$

In the second illustrative example below, all the macroinvertebrate metrics of the previous example keep their groups, but these are now treated as bottom level groups and the whole set of macroinvertebrate metrics become one top level group of metrics. In addition, in determining the overall average class for the top level group comprising macroinvertebrates, the bottom level groups 1, 2 3 and 4, are given relative weights 1, 3 1 and 2 as specified in Column O.

	A	N	O	P	Q
18		Required Hierarchy of metric groups			
19		Top level group		Bottom level group	
20	Metric name	Metric Group	Metric Weight (if not equal)	Metric Group	Metric Weight (if not equal)
23					
24	Saprobic Index (Zelinka & Marvan)	1	3	2	
25					
26	Average score per Taxon	1		2	
27					
28	DSFI	1		2	
29					
30	Diversity (Shannon-Wiener-Index)	1	1	1	
31					
32	Diversity (Margalef Index)				
33					
34	Acid Index (Hendrikson & Medin)	1	1	3	
35					
36	Number of Families	1		1	
37					
38	Index A	1		4	1
39	Index B	1		4	2
40	Index C	1		4	1
41	Multi-metric Index based on Indices A,B and C	1	2	4	
42					
43	EFI (European Fish Index)	2			
44					
45	MTR (Mean Tropic rank of Macrophytes)	3			
46					
47	HMI (Habitat Modification Index)	4			
48					

In this example, the overall status class for macroinvertebrates is then combined with the classes of three other top level group metrics, namely fish metric EFI (group 2), macrophytes MTR (group 3) and habitats HMI (group 4) to derive the overall status class for the site/waterbody based on multiple BQEs.

Obviously to derive an uncertainty analysis for this overall bioassessment requires estimates of the sampling and other sources of variation and error for each metric of each BQE involved.

**4.1.(8) Number of simulations (cells B15 and B16)**

The number of simulations of observed and EQR values to be used to assess the uncertainty is specified in cell *B15*.

	A	B
14		
15	Number of simulations required (max=100000) =	100000
16	Integer random number seed (-1 uses current time) =	234
17		

The maximum permitted number of simulations is 100000

A minimum of 10000 simulations is recommended and this is the default..

Using 10000 simulations would ensure that the estimates of the percentage probability of belonging to each status class would vary by at most 2% in at least 95% of repeated runs of the program using different random number seeds (and the low percentage probabilities in unlikely status classes for a water body would vary by very much less).

Using the maximum simulations would give the smoothest probability patterns in graphically plots in studies involving assessing changes in probability of each status class in relation to gradual change in EQR metric values along the quality gradient (e.g. varying EQR from zero to one)

To ensure that exactly the same uncertainty results (in terms of confidence limits and probabilities of assignment to each status class) in repeated runs of the program on the same input data, leave or set the random number seed in cell *B16* to a positive integer. This is recommended and is the initial default.

Setting the random number seed to -1 generates a different set of simulations every run.

## 4.2 Observed values of each metric for each site/waterbody (File code name 'FOBS')

The observed values of each metric for each site/waterbody are stored in a User-specified EXCEL spreadsheet file, called an 'Observed metrics values file', in a standard layout with the site/waterbody names in columns and the metrics as rows, as shown in the example below:

	A	B	C	D	E
1	Metric	U2311073	Cole at Small Heath	U2311173	U2310803 Omore at Bridgend
2					
3	Abundance [ind/m <sup>2</sup> ]	7323	2904	8493	1670.4
4					
5	Number of Taxa	71	39	29	35
6					
7	Saprobic Index (Zelinka & Marvan)	2.008	2.485	2.775	2.025
8					
9	Average score per Taxon	6.500	4.062	3.875	6.188
10					
11	DSFI	6	5	Not Calculated	-9
12					
13	IBE Aqem	11.4	6	8	9
14					

This layout is exactly the same as that output from the AQEM/STAR metric calculation program AQEMrap, so that this output can, if required, be used as the file of observed metrics files for input to program WISERBUGS.

### 4.2.1 Site/Waterbody names:

Site or Waterbody names can be any strings of up to 100 characters providing they are valid as strings within EXCEL cells.

### 4.2.2 Metric names:

The names of the metrics must includes all of those specified for use in the site/waterbody bioassessment by the 'Metric Specification file' (see Section 4.1), but it can include other currently unused metrics. The name for a particular metric must be EXACTLY THE SAME in both files. If the name of any of the metrics specified for use is not found in the 'Observed metrics values file', then program WISERBUGS issues an error message in the program log file 'WISERBUGS.LOG' and warns the User that the program failed.

### 4.2.3 Observed metric values

Observed sample values of the metrics to be used need to be numeric.



#### 4.2.4 Missing observed metric value indicator

The missing value indicator is -9. This should be used where the value of a particular metric cannot be, or has not been, calculated for a particular site or waterbody. Empty or blank cells are assumed to be missing values and set to -9. Be careful and check the output.

The AQEMrap program outputs the phrase 'Not Calculated' as the metric value for a site/waterbody sample where it cannot be calculated (usually for reasons related to lack of taxonomic resolution). Program WISERBUGS automatically converts such cells into missing values (-9).

Program WISERBUGS can cope with some metrics within a multi-metric group having missing values, in that it makes the hierarchal bioassessment using the multi-metric rules on the remaining metrics and indicates where one or more metrics values were missing at each stage of the overall assessment (see Section 5 for further details).

### **4.3 EQR metric parameters $E_0$ and $E_1$ (File code names 'FEZERO' & 'FEONE')**

An observed ( $O$ ) metric value is normalised into any EQR using the formula in equation (1), repeated here:

$$EQR = \frac{O - E_0}{E_1 - E_0} \quad (\text{equation 1})$$

where  $O$  = observed value,

$E_1$  = Reference condition value (= value of metric for which  $EQR = 1$ )

and  $E_0$  = value of metric for which  $EQR = 0$

The  $E_0$  values to be used for each site/waterbody for each used metric must be read from a User-specified EXCEL file of any chosen name, but referred to here as the 'FEZERO' file.

The  $E_1$  values to be used for each site/waterbody for each used metric must be read from a User-specified EXCEL file of any chosen name, but referred to here as the 'FEONE' file.

#### 4.3.1 Options for the form and layout of the 'FEZERO' and 'FEONE' files.

One of the three options (1, 2, or 3) MUST be specified in Cell A1 or the 'FEZERO' and 'FEONE' files

Cell A1 = 1:

Individual values of  $E_0$  or  $E_1$  supplied for each site/waterbody of each metric (can be the same or different).

The layout for an option 1 ‘FEONE’ file is exactly the same as for the ‘Observed metrics values file’ detailed in Section 4.2, namely as a matrix of sites/waterbodies in columns and metrics in rows.

Example with

Cell  $AI = 1$

	A	B	C	D	E	F	G
1	1	U2311073	U2311163	U2311173	U2310803	U2310823	
2							
3	Number of Taxa	71	39	29	35	36	
4	Saprobic Index (Zelinka & Marvan)	2.9	2.9	2.9	2.9	2.9	
5	Average score per Taxon	6.23	5.96	6.76	6.27	6.58	
6	Diversity (Shannon-Wiener-Index)	3.5	3.5	3.5	3.5	3.5	
7	EPT-Taxa	18	18	18	16	16	
8	Number of Families	32.5	28.9	32.1	27.6	31.8	

Cell  $AI = 2$ :

Same value of  $E_0$  or  $E_1$  for a particular metric to be used for all sites/waterbodies.

Metric names must be given in Column A, as for the option 1, and the single value of  $E_0$  or  $E_1$  for each metric must be given in Column B.

No samples names are needed in row 1 (or read) because the  $E_0$  or  $E_1$  parameters are the same for all site/waterbodies.

Example with

Cell  $AI = 2$

	A	B	C	D	E	F	G
1	2						
2							
3	Number of Taxa	71					
4	Saprobic Index (Zelinka & Marvan)	2.9					
5	Average score per Taxon	6.23					
6	Diversity (Shannon-Wiener-Index)	3.5					
7	EPT-Taxa	18					
8	Number of Families	32.5					

Cell  $AI = 3$ :

Same ‘null’ value for  $E_0$  or  $E_1$  to be used for all metrics for all sites/waterbodies.

‘null’ value for  $E_0 = 0$

‘null’ value for  $E_1 = 1$

With option 3 set in Cell  $AI$  of the ‘FEZERO’ file, no other cell in the ‘FEZERO’ file needs to be filled, or is read by WISERBUGS. Similarly for the ‘FEONE’ file.

Sample names in the ‘FEZERO’ and ‘FEONE’ files:

Sample names are only required for option 1 (i.e. when Cell  $AI = 1$ )

Sample names in the ‘FEZERO’ and ‘FEONE’ files must be EXACTLY THE SAME and in EXACTLY THE SAME ORDER as in the ‘Observed metrics values file (see Section 4.2).

If there any differences or mis-matches, then program WISERBUGS issues an error message in the program log file ‘WISERBUGS.LOG’ and warns the User that the program failed.

#### *Metric names in the 'FEZERO' and 'FEONE' files:*

Metric names are required in Column A for options 1 and 2 of the 'FEZERO' and 'FEONE' files (i.e. when their Cell AI is set to 1 or 2).

The metric names must include all of those specified for use in the site/waterbody bioassessment by the 'Metric Specification file' (see Section 4.1), but can include other currently unused metrics.

The name for a particular metric must be EXACTLY THE SAME in both files. If the name of any of the metrics specified for use is not found in the 'FEZERO' or 'FEONE' file, then program WISERBUGS issues an error message in the program log file 'WISERBUGS.LOG' and warns the User that the program failed.

#### 4.3.2 Methods for setting the Reference Condition ( $E_I$ ) value of a metric

Various methods of setting the Reference Condition ( $E_I$ ) value of a particular metric for a particular site/waterbody or environmental-similar group of sites/waterbodies can be used depending on the data available. Obviously the reference condition or high quality sites used to determine the  $E_I$  values should be sampled in the same way as the samples for the sites being assessed. The following are several possible options in roughly decreasing order of preference.

- i) If a suitable RIVPACS-type predictive model involving an adequate number of environmental similar reference condition sites is available, then the  $E_I$  values are best based on RIVPACS-type site- and season- specific predictions of the expected fauna and metric values.
- ii) In the absence of a RIVPACS model, if a suitable number of reference condition sites of an environmentally-similar type are available, the  $E_I$  values can be based on the mean or median metric value for these sites.
- iii) If a suitable number of an environmentally-similar type of 'high' quality sites (of uncertain reference condition) are available, the  $E_I$  values can be based on the mean, median or perhaps an upper percentile (75% or 90%) value of the metric for these sites.
- iv) If only a very small number of an environmentally-similar type of 'high' quality sites (of uncertain reference condition) are available, then the  $E_I$  values cannot be reliably estimated and might be based using the maximum of the few values available. However, the maximum value is not a stable measure and increases with the number of sites on which it is based.
- v) If no 'reference condition' or 'high' status sites are available then, some form of hind-casting or extrapolation to reference conditions will be necessary to provide appropriate values of  $E_I$ .

#### 4.4 Correlations between metrics in sampling variation (File code name 'FCORR')

The values of two or more metrics, especially if they are intended to measure the same type of stress, are often correlated across a range of sites. When one such metric gives values indicating relative high quality, the other is also likely to give a similar message. Highly correlated metrics would hopefully give similar status classes for any particular site. A large part of their correlation is often due to the fact that they are responding similarly across one or more stressor gradients. Depending on whether high or low values of each metric indicate high quality, two metrics may be either positively or negatively correlated. For example, the Saprobic Index and ASPT might be expected to be negatively correlated

In uncertainty assessments based on simulations of sampling variation in each metric, the concern is more subtle. A particular site/waterbody at a point/period in time has a certain biological quality. Replicate samples from the same site at one point/period differ only because of sampling, sub-sampling and sample processing variation, collectively referred in WISERBUGS as 'sampling variation'. Metrics which measure similar components of the biological composition, such as giving similar relative weights to taxa, are likely to give correlated values between replicate samples within any one site. Multi-metric indices (MMI) and status classes based on several correlated indices are affected by all such correlations. As an simple example, if two metrics are very highly correlated and their status classes are equivalent, then in any one sample both metrics are likely to indicate the same class. Therefore, a MMI is roughly equivalent to either one of the metrics and a multi-metric class will be the same as the class based on any one metric. Simulations which assumed and generated independent sampling errors would under-estimate the true variability in a MMI average and underestimate the uncertainty in status class estimation.

In practice, such sampling correlations are only important if they are at least 0.5 or so.

Program WISERBUGS has the facility to incorporate sampling correlations between metrics.

The User can supply estimates of the replicated sampling correlations between either all metrics or those pairs of metrics which are thought to be highly correlated.

Program WISERBUGS simulates sampling variation for the metrics (some perhaps on a transformed scale) assuming a multivariate normal distribution for the joint sampling variation, based on the User-supplied estimates of the sampling SD and the sampling correlations. If a particular metric is specified as having constant sampling variance on a square root or double square root scale, then the correlations involving this metric are also assumed to be for when this variable is on its transformed scale

*Layout of the ‘FCORR’ metric correlation file.*

The first row is a header line.

Each subsequent row specifies one practically significant correlation between two metrics.

The metric names are given in Columns A and B, and the correlation is given in Column C.

Correlations can be specified in any order.

The name for a particular metric must be EXACTLY THE SAME as in the ‘Metric Specification file’ (see Section 4.1). If a metric name in the ‘FCORR’ file does not match up exactly with the name of a metric specified for use in the ‘Metric Specification file’, then the correlation will be ignored.

Any sampling correlations not specified are treated as zero.

A simple example of an ‘FCORR’ file is given below.

	A	B	C
1	Metric i	Metric j	Sampling correlation ij
2	Saprobic Index (Zelinka & Marvan)	Average score per Taxon	0.6
3	Average score per Taxon	Biological Monitoring Working Party	0.8
4	Average score per Taxon	Number of Families	0.5
5			

**4.5 Direct entry of multi-metric index (MMI) values and their uncertainty**

If you have previously decided on the metrics and combination rules to be used for your multi-metric index (MMI), then you may have already calculated your MMI values for each sample from a prior sampling variability study (such as those carried out in the STAR or WISER projects).

In such cases, it may be possible to calculate an appropriate estimate of the sampling variability and uncertainty in the MMI values directly. The estimates of the observed values of the MMI and estimates of its sampling SD can be supplied directly to the WISERBUGS input files. This avoids the need to derive and supply the sampling correlations between the individual metrics involved in the MMI, as the prior estimate of the sampling SD of the MMI will automated take account of the natural sampling correlations of the component metrics.



*Deliverable D6.1-3: WISERBUGS (WISER Bioassessment Uncertainty Guidance Software) tool for assessing confidence of WFD ecological status class*

## 5. Output file details

All of the output is written to a single EXCEL file. An example output file (WISERBUGS\_TestOut11.XLS) is provided with the software package, as described in Appendix 1.

The contents, layout and format of the output files are explained below.

### 5.1 Log file detailing analysis progress and any problems (file WISERBUGS.LOG)

Text file WISERBUGS.LOG, in the same directory as the software, details progress on reading in information from the User input files and carrying out the uncertainty bioassessment on each sample in turn. This provides very useful information on whether the analysis has proceeded as intended and should be checked.

If program WISERBUGS fails to complete its uncertainty simulation analysis for all sites/waterbodies in the User-specified 'Observed metrics values file', then the program informs the User who is ask to check through the program log text file 'WISERBUGS.LOG' to (hopefully) find information or clues to the data-entry mistake, which should then be corrected and the program re-run.

In particular, the name of each site/waterbody is written to the WISERBUGS.LOG' file as it is processed. Therefore, if some of the sites/waterbodies have been processed, the last 'site/waterbody' name listed should indicate where the problem has occurred.

Note:

Because of the generality and flexibility for User input and analysis provided in program WISERBUGS, it is very difficult to be able to anticipate all possible type of typing and other errors related to the User's input files. Therefore the program will probably occasionally fail without providing any or clear information on the reason. The User should always check the log file 'WISERBUGS.LOG' for any clues.

If required for later reference, the log file 'WISERBUGS.LOG' for a particular run could be saved to a separate directory, as it is overwritten in each run'

If program WISERBUGS completes ok, then the uncertainty analysis results for all of the samples are written to an Output EXCEL file, as described in Section 5.2.

## 5.2 Output EXCEL file of ecological status classes and uncertainty assessment

Rows 1-6 of the Output EXCEL file give the names of each of the input files used in the analysis.

The results for each site/waterbody are all stored in a single row of the Output file, with the site/waterbody name in Column A.

The results for each site/waterbody are given in the following order from left to right:

- i) Overall observed status class and probabilities of belonging to each of the five possible classes
- ii) For each of the top level metric groups in turn:
  - Observed status class for the group and probabilities of belonging to each of the five possible classes
- iii) For each used metric in turn:
  - Observed metric value,  $E_0$ ,  $E_1$ , EQR, standard deviation (SD) of the EQR, lower and upper 95% non-parametric confidence limits of the EQR,
  - Observed status class for the metric and probabilities of belonging to each of the five possible classes
- iv) For each derived multi-metric index (MMI) in turn:
  - Observed MMI EQR value, standard deviation (SD) of the EQR, lower and upper 95% non-parametric confidence limits of the EQR,
  - Observed status class for the MMI and probabilities of belonging to each of the five possible classes

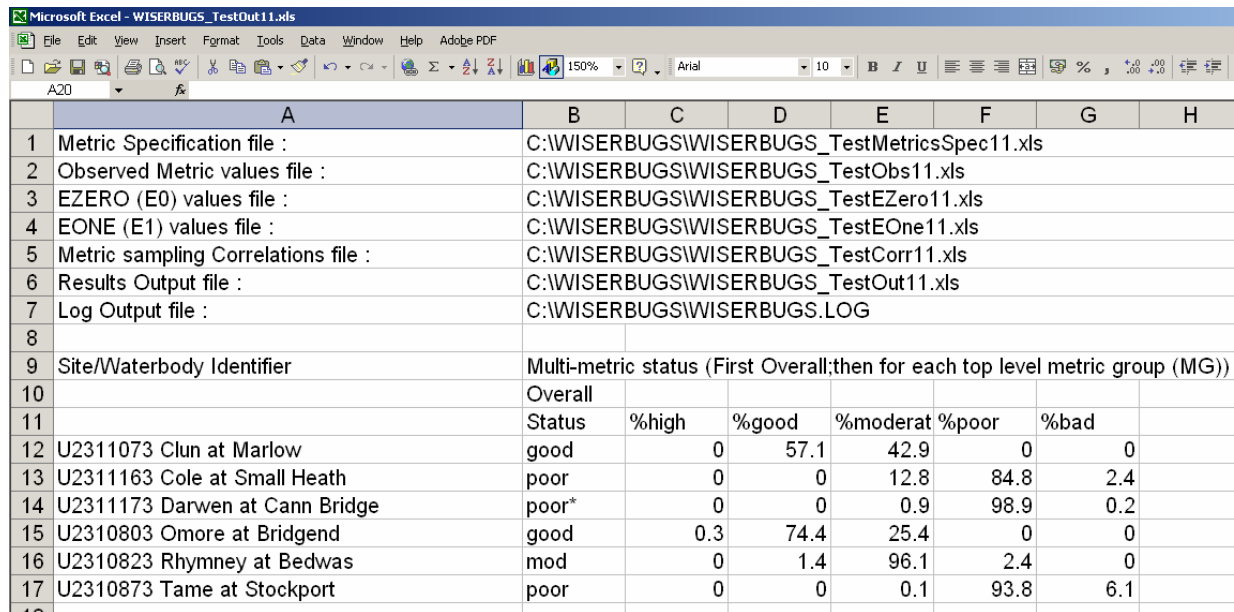
If the status classes for a metric were based on its Observed values (indicated by a '1' in Column C of the 'FMETSPEC' file described in Section 4.1.(2) ), rather than normalised EQR values, then the columns headed 'E0' and 'E1' in the output for that metric are set to values of '-9' (i.e. missing values), the 'EQR' is set equal to the Observed metric values and all information on confidence limits in EQR relates to uncertainty in Observed metric values.



### Overall status class and probabilities of class membership

The example below shows the results on Overall status class for the six example sites.

The first site/waterbody ‘U2301173 Clun at Marlow’ was assigned an overall status class of ‘good’. However, the uncertainty analysis, using the sampling and other variability estimates, suggests that although there is a 57.1% probability the site/waterbody would be classified to ‘good’, there is a substantial 42.9% probability that the site/waterbody would be classified as ‘moderate’ status class.



	A	B	C	D	E	F	G	H
1	Metric Specification file :	C:\WISERBUGS\WISERBUGS_TestMetricsSpec11.xls						
2	Observed Metric values file :	C:\WISERBUGS\WISERBUGS_TestObs11.xls						
3	EZERO (E0) values file :	C:\WISERBUGS\WISERBUGS_TestEZero11.xls						
4	EONE (E1) values file :	C:\WISERBUGS\WISERBUGS_TestEOne11.xls						
5	Metric sampling Correlations file :	C:\WISERBUGS\WISERBUGS_TestCorr11.xls						
6	Results Output file :	C:\WISERBUGS\WISERBUGS_TestOut11.xls						
7	Log Output file :	C:\WISERBUGS\WISERBUGS.LOG						
8								
9	Site/Waterbody Identifier	Multi-metric status (First Overall; then for each top level metric group (MG))						
10		Overall						
11		Status	%high	%good	%moderat	%poor	%bad	
12	U2311073 Clun at Marlow	good	0	57.1	42.9	0	0	
13	U2311163 Cole at Small Heath	poor	0	0	12.8	84.8	2.4	
14	U2311173 Darwen at Cann Bridge	poor*	0	0	0.9	98.9	0.2	
15	U2310803 Omore at Bridgend	good	0.3	74.4	25.4	0	0	
16	U2310823 Rhymney at Bedwas	mod	0	1.4	96.1	2.4	0	
17	U2310873 Tame at Stockport	poor	0	0	0.1	93.8	6.1	

### Metric group status class and probabilities of class membership

Exactly the same information on status class and class probabilities is then given for each of the top level metric groups (see start of statistics for top level group ‘MG1 Taxa/Diversity’ in columns I-J of the example above).

The status class information is also calculated but not currently output for each bottom level metric group (if they are defined). However, all of the hierarchical multi-metric rules are carried to build up to an overall status class for both the observed data and for every simulation. This is how WISERBUGS obtains confidence limits and probabilities of class at all levels of the hierarchical grouping of metrics.

### Individual metric confidence limits, status class and probabilities of class membership

To the right of the observed status class and probabilities of class for the overall site/waterbody assessment and for each top level metric group, the output gives information for each individual metric used in the assessment.

The first example below shows the uncertainty assessment for the metric ‘Saprobic index (Zelinka & Marvan)’. Because the status classes for this metric were based on its observed

values rather than normalised EQR values, no metric conversion parameters  $E_0$  and  $E_1$  were used and these values were set to missing values in the output. The EQR values were set to the observed ‘Saprobic index’ values and the SD and confidence limits represent uncertainty in the observed metric values. The status class is based on User-supplied class limits for the observed metric values. The probabilities of belonging to each class are derived from simulations of observed metric values.

In this example, the first site/waterbody ‘U2301173 Clun at Marlow’ has an observed sample Saprobic Index value of 2.008, but based on its supplied sampling SD, the 95% confidence limits for possible values are estimated as 1.771 – 2.248. The sample’s status class is ‘moderate’ based on this metric, but there is a 47% probability the site/waterbody could have been classed as ‘good’ and a very small probability it could have been either ‘high’ or ‘poor’ status.

	A	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW
9	Site/Waterbody Identifier													
10		Saprobic Index (Zelinka & Marvan)												
11		Observed	E0	E1	EQR	St Dev	Lower 2.5%	Upper 2.5%	Status	%high	%good	%moderate	%poor	%bad
12	U2311073 Clun at Marlow	2.008	-9	-9	2.008	0.122	1.771	2.248	mod	0.5	47	51.6	0.9	0
13	U2311163 Cole at Small Heath	2.485	-9	-9	2.485	0.118	2.255	2.719	poor	0	0	5.6	77.7	16.7
14	U2311173 Darwen at Cann Bridge	2.775	-9	-9	2.775	0.119	2.539	3.01	bad	0	0	0	7.2	92.8
15	U2310803 Omore at Bridgend	2.025	-9	-9	2.025	0.119	1.794	2.256	mod	0.3	42.1	56.6	1.1	0
16	U2310823 Rhydney at Bedwas	2.358	-9	-9	2.358	0.12	2.119	2.594	poor	0	0.2	31.2	66.5	2.2
17	U2310873 Tame at Stockport	2.493	-9	-9	2.493	0.121	2.253	2.729	poor	0	0	5.5	76.2	18.3

The second example, shown below, is for the metric ‘Average score per Taxon’. This metric, usually referred to as ASPT, is used here in its EQR form.

	A	AY	AZ	BA	BB	BC	BD	BE	BF	BG	BH	BI	BJ	BK
9	Site/Waterbody Identifier													
10		Average score per Taxon												
11		Observed	E0	E1	EQR	St Dev	Lower 2.5%	Upper 2.5%	Status	%high	%good	%moderate	%poor	%bad
12	U2311073 Clun at Marlow	6.5	0	6.23	1.043	0.042	0.961	1.125	high	100	0	0	0	0
13	U2311163 Cole at Small Heath	4.062	0	5.96	0.682	0.043	0.598	0.766	mod	0	1.2	59.2	39.4	0.2
14	U2311173 Darwen at Cann Bridge	3.875	0	6.76	0.573	0.038	0.5	0.648	poor	0	0	0.5	62.9	36.7
15	U2310803 Omore at Bridgend	6.188	0	6.27	0.987	0.042	0.904	1.07	high	99.1	0.9	0	0	0
16	U2310823 Rhydney at Bedwas	5.412	0	6.58	0.822	0.039	0.745	0.9	good	4.5	82	13.5	0	0
17	U2310873 Tame at Stockport	3.2	0	6.93	0.462	0.036	0.39	0.534	bad	0	0	0	0.3	99.7

In WISERBUGS, all EQR values are derived from the formula in equation (1), repeated here for convenience :

$$EQR = \frac{O - E_0}{E_1 - E_0} \quad (\text{equation 1})$$

By setting the  $E_0$  values to zero, and the  $E_1$  values to the RIVPACS-type model expected value under Reference Conditions, the EQR values become RIVPACS-type O/E ratios of the observed (O) to expected (E) values of metrics or biotic indices(e.g. see Clarke *et al.* 1996; Wright, Sutcliffe & Furse 2000, Clarke, Wright & Furse 2003).

In the example above, for the second site/waterbody, the observed ASPT is 4.062, the expected ( $E_1$ ) values is 5.96, give an EQR (O/E) of 0.682. The 95% confidence for the EQR were 0.598 – 0.766, placing it in the ‘moderate’ with a probability of 59.2%. However, there is also a 39.4% that the site/waterbody could be classed as ‘poor’ status.

All EQR values are constrained to be greater than or equal to zero by resetting any negative values to zero. However, unless a metric EQR is being used in a multi-metric index (MMI), EQR values greater than one are left unchanged. This is to permit compatibility with the widely-used RIVPACS-type O/E ratios.

*Multi-metric index (MMI) confidence limits, status class and probabilities of class membership*

Finally, if any multi-metric indices were involved, estimates are provided of their MMI values, together with their confidence limits, status classes and probabilities of class membership for each site/waterbody.

*Reporting status class for a site/waterbody with missing values for one or more metrics*

In the example above, the third site/waterbody 'U2311173 Darwen at Carn Bridge' has an overall status class of 'poor\*'. The '\*' symbol is used to indicate that one or more metrics required by the User to be in the site/waterbody assessments had missing values for this site/waterbody. This could be because, it was just not derived, or more likely that the taxonomic resolution or taxonomic detail and/or the taxa present were not sufficient to calculate a meaningful value for the metric.

In such cases, the output file gives missing values '-9' for the statistics of any metric with missing values. The status class of missing metrics is set to 'none'. Any MMI values are based on the average EQR values of the remaining metrics involved in the MMI. Each status class, which is derived from metrics with one or more missing values is given a '\*' at the end of the class name to indicate it is based on incomplete information.

## HELP

If you have any queries or comments about the software, please contact Ralph Clarke at Bournemouth University ([rclarke@bournemouth.ac.uk](mailto:rclarke@bournemouth.ac.uk)).

This Release of the software and User manual is the first version of the software for trial general use.

It would be helpful if you could report any bugs/errors found with the software or User Manual to Ralph Clarke at Bournemouth University ([rclarke@bournemouth.ac.uk](mailto:rclarke@bournemouth.ac.uk)).

## REFERENCES

- Clarke, R.T., Furse, M.T., Wright J.F. & Moss, D. 1996. Derivation of a biological quality index for river sites : comparison of the observed with the expected fauna. *Journal of Applied Statistics*, 23, 311-332.
- Clarke, R.T. 2000. Uncertainty in estimates of river quality based on RIVPACS. In: Assessing the biological quality of freshwaters: RIVPACS and similar techniques. Wright, J.F., D.W. Sutcliffe and Furse, M.T. (eds). pp 39-54. *Freshwater Biological Association*, Ambleside.
- Clarke, R.T., Furse, M.T., Gunn, R.J.M., Winder, J.M. & Wright, J.F. 2002. Sampling variation in macroinvertebrate data and implications for river quality indices. *Freshwater Biology*, 47, 1735-1751.
- Clarke, R.T. & Hering, D. 2006. Errors and uncertainty in bioassessment methods – major results and conclusions from the STAR project and their application using STARBUGS. *Hydrobiologia*, 566, 433-440.
- Staniszewski, R., Szoszkiewicz K., Zbierska, J., Lesny, J., Jusik, S. & Clarke, R.T. 2006. Assessment of sources of uncertainty in macrophyte surveys and the consequences for river classification. *Hydrobiologia*, 566, 235-246.
- Sundermann, A. , Pauls, S.U., Clarke, R.T. & Haase, P. 2008 Within-stream variability of benthic invertebrate samples and EU Water Framework Directive assessment results. *Fundamental and Applied Limnology (Archiv für Hydrobiologie)*, 173, 21-34
- Wright J.F, Sutcliffe D.W. & Furse M .T. (eds) (2000) *Assessing the biological quality of fresh waters: RIVPACS and other techniques*. Freshwater Biological Association, Ambleside.

## APPENDIX 1. Example input and output files supplied with program WISERBUGS

### Input EXCEL file types

File Code name	Content
'FMETSPEC'	Metric specification file (uncertainty estimates + multi-metric rules)
'FOBS'	Observed values of each metric for each site/waterbody
'FEONE'	Reference Condition ( $E_1$ ) values of each metric for which EQR = 1
'FEZERO'	Values ( $E_0$ ) of each metric for which EQR = 0
'FCORR'	Correlations between metrics in their sampling variation

### Output EXCEL file

File Code name	Content
'FOUT'	Results of uncertainty & status analysis for each site/waterbody in turn

### Example Test files

File Code name	Example file
'FMETSPEC'	WISERBUGS_TestMetricsSpec11.xls
'FOBS'	WISERBUGS_TestObs11.xls
'FEONE'	WISERBUGS_TestEZero11.xls
'FEZERO'	WISERBUGS_TestEOne11.xls
'FCORR'	WISERBUGS_TestCorr11.xls - with some sampling correlations
	WISERBUGS_TestCorrNull11.xls - with zero sampling correlations
'FOUT'	WISERBUGS_TestOut11.xls

Blank input files for modification by the User

File Code name	Example file
‘FMETSPEC’	WISERBUGS_BlankMetricsSpec11.xls
‘FCORR’	WISERBUGS_BlankCorr11.xls

The contents, layout and format of each files is detailed in Section 4.1 – 4.5.

It is highly recommended that the User leave the supplied example input files unchanged.

A User ‘Metric Specification file’ (FMETSPEC) should be built up from the Blank version supplied or from one of the User’s existing ‘Metric Specification files’.

This is because the layout (but not the User-defined parameters) of the spreadsheet and in particular, the rectangular block of cells A1–Q20, is assumed by the program to be fixed and should not be altered or the program will probably crash.

The details on the individual metrics to be used should be added in rows 21 onwards.

## APPENDIX 2. Example metric uncertainty standard deviations

Example estimates of the replicate sampling standard deviation (SD) of a range of macroinvertebrate metrics derived from the STAR replicated sampling programme. All SD based on untransformed metric values.

Metric	Sampling method	Replicate sampling SD
ASPT	AQEM/STAR	0.384
	RIVPACS	0.236
Saprobic Index (Zelinka & Marvan)	AQEM/STAR	0.075
	RIVPACS	0.048
Diversity (Shannon-Weiner)	AQEM/STAR	0.253
	RIVPACS	0.180
Total number of Families	AQEM/STAR	3.35
	RIVPACS	4.23
Number of EPT Taxa	AQEM/STAR	1.53
	RIVPACS	2.14