PERGAMON
www.elsevier.com/locate/watres

PII: S0043-1354(01)00162-2

Wat. Res. Vol. 35, No. 18, pp. 4285–4292, 2001

Printed in Great Britain
0043-1354/01/$ - see front matter

A REAPPRAISAL OF SAPROBIC VALUES AND INDICATOR WEIGHTS BASED ON SLOVENIAN RIVER QUALITY DATA

WILLIAM J. WALLEY[1]*, JASNA GRBOVIĆ[2] and SAŠO DŽEROSKI[3]

[1]School of Computing, Staffordshire University, Beaconside, Stafford, ST18 0DG, UK; [2]Hydrometeorological Institute, Vojkova 1b, SI-1000 Ljubljana, Slovenia and [3]Jožef Stefan Institute, Jamova 39, SI-1111 Ljubljana, Slovenia

(First received 4 July 2000; accepted in revised form 19 March 2001)

Abstract—The saprobic values and indicator weights used in the Slovenian saprobic system are reappraised using data from the 1990 to 95 river quality surveys of Slovenia. The conceptual basis of the reappraisal is described and then formulated mathematically. The analysis is based on 1106 biological samples and covers 300 taxa. The results are expressed in terms of revised saprobic values and indicator weights that mirror the ones previously assigned by ecological experts. The most significant differences between original and revised values are highlighted and discussed. It is concluded that: (a) the revised values and weights are more representative of their 'true' values than are the original values and weights, but that it would be premature to consider them definitive; (b) the analytical method provides a sound data-based approach to the revision of saprobic values and indicator weights; and (c) the method could help to improve and harmonise the various saprobic systems currently in use across Europe. © 2001 Elsevier Science Ltd. All rights reserved

Key words—river, quality, pollution, classification, monitoring, biology, saprobic system

INTRODUCTION

Biological monitoring methods are playing an increasingly important role in river quality monitoring, mainly due to the fact that the biota are continuous witnesses of the river's state of health and are collectively sensitive to the whole range of potential pollutants. Within the European Union, the Council Directive that established a framework for action in the field of water policy (European Union, 2000) recognised the potential of biomonitoring by adopting it as the key method to be used for monitoring river quality across the Union. However, several different biomonitoring systems are presently in use throughout Europe, all of which are based upon processes and/or numerical values that were subjectively derived by experts. Comprehensive reviews of several European methods of biomonitoring were given by De Pauw and Hawkes (1993) and Grbović (1994).

If the full potential of biomonitoring is to be realised, much work needs to be done to improve existing methods and to develop new methods based on advanced data interpretation methods. The authors have published several papers on the improvement of existing methods (Walley and Hawkes, 1996, 1997; Džeroski *et al*., 1997b) and the use of Artificial Intelligence (AI) in biomonitoring (Walley and Fontama, 2000, 1998; Walley *et al*., 1998; Džeroski *et al*., 1998, 1997a,c; Walley and Džeroski, 1995).

This paper is concerned solely with the saprobic system and the development of a method by which its saprobic values can be objectively reappraised using field data. The saprobic system is used in several European countries, including Austria, Germany, Czech Republic, Slovakia and Slovenia, but the saprobic values used were all derived subjectively and differ from country to country. Comprehensive revisions of saprobic values have been carried out recently in Austria (Moog, 1995) and Germany (DIN38 410, 1990) based upon field experience. However, these revisions were not based upon data analysis, but upon the pooled opinions of expert field biologists. In this paper, we develop and apply a statistical method of analysing field data to produce revised saprobic indices and their associated weights. The method is based on the same basic principle as that used by Walley and Hawkes (1996, 1997) for their reappraisal of BMWP scores. That is, it assumes that the saprobic values originally assigned to the taxa provide a fair first estimate of their true values and that the saprobic index of a monitoring site, being a weighted average of several saprobic values, provides an even better estimate of the site's saprobic

*Author to whom all correspondence should be addressed.
E-mail: w.j.walley@staffs.ac.uk

status. Thus, it is further assumed that the distribution of a taxon with respect to the saprobic status (i.e. quality class) of the sites at which it is found provides a sound basis for the estimation of its 'true' saprobic value. The mathematical formulation of the analysis differs slightly from that used by Walley and Hawkes (1996), because it has been modified to a form more appropriate to the saprobic system.

THE SLOVENIAN SAPROBIC SYSTEM

Slovenia uses the saprobic index method, as developed by Pantle and Buck (1955) and later modified by Zelinka and Marvan (1961), to map biological data to seven discrete quality classes as defined in Table 1 below.

The saprobic index (SI) for a site is derived as follows:

$$\mathrm{SI} = \frac{\sum_{k=1}^{K} s_k w_k h_k}{\sum_{k=1}^{K} w_k h_k} \qquad (1)$$

where: s_k is the saprobic value of the kth taxon found in the sample, w_k the indicator weight of the kth taxon, h_k the abundance rating of the kth taxon ($h = 1$, 3 or 5 if organisms of the taxon are found incidentally, frequently or abundantly, respectively), and K the total number of taxa found in the sample.

The abundance levels 1, 3 and 5 are of a qualitative nature. As a general rule, they correspond roughly to counts of 1–10, 11–100 and over 100 individuals. However, they are dependent on the taxon in question. For example, 200 individuals of Chironomidae or Oligochaeta would be recorded as abundance level 3, while 10 individuals of the taxon *Perlodes* sp. would be recorded as abundance level 5.

THE DATA

The field data were supplied by the Hydrometeorological Institute of Slovenia (HMIS), which is part of the Slovenian Ministry of Environment and Spatial Planning. HMIS is responsible for the execution of water quantity and quality monitoring. The data supplied covered the six-year period from 1990 to 1995, and included summer and winter biological samples in which the taxa present were recorded in one of three density bands (incidental, frequent and abundant). Macro-invertebrates were collected from the top 15 cm of the bed using a standard handnet (ISO 7828 (E), 1985). Those living within algae and moss were shaken out and collected in the net, but those firmly attached to the substrate were collected by hand. Periphyton were sampled by scraping them off biotic and abiotic underwater surfaces, whilst the densities of filamentous bacteria, fungi and algae were assessed and recorded on the spot.

The data provided also included the saprobic index derived from each biological sample, plus its corresponding quality class. The database contained 1106 biological samples, in which a total of 839 taxa were recorded, mainly to species or genera level, but a few to family or group level. Some were eliminated from the analysis on the basis that they either had never been assigned a saprobic value or had occurred in fewer than 10 samples. In addition, several genera (i.e. appended with "sp.") were eliminated because it was concluded that they were not clearly representative of the genus, but more representative of those 'other members' of the genus that were not easily identified to species level. The frequency distribution of the remaining 300 taxa is given in Table 2.

THE ANALYTICAL METHOD

The analytical method differs from that used by Walley and Hawkes (1996, 1997) in two respects:

(a) it is based upon one site type and not three, since all sites in the database had been chosen by HMIS to be eroding sites, roughly corresponding to the *Riffle* sites defined by Walley and Hawkes (1996); and
(b) it is based on the saprobic valency method of deriving the saprobic values of individual bioindicators.

The saprobic valency approach was adopted to maintain consistency with the tradition of the saprobic system, but it is worth noting that the resulting procedure for revising saprobic values was mathematically almost identical to that used by Walley and Hawkes (1996) to revise BMWP scores.

Table 1. Definition of quality classes based on derived saprobic index and the distribution of field samples (see "THE DATA" for details) across quality classes

Quality band i	Quality class	Saprobic range	Saprobic index range	Mid-point of range (x_i)	Number of samples in band i	Percentage of samples in band i
1	1	Oligosaprobic	1.0 to 1.5	1.25	55	5.0
2	1–2	Oligosaprobic to β-Mesosaprobic	>1.5 to ≤ 1.8	1.65	286	25.9
3	2	β-Mesosaprobic	>1.8 to ≤ 2.3	2.05	547	49.5
4	2–3	β-Mesosaprobic to α-Mesosaprobic	>2.3 to ≤ 2.7	2.50	150	13.6
5	3	α-Mesosaprobic	>2.7 to ≤ 3.2	2.95	58	5.2
6	3–4	α-Mesosaprobic to Polysaprobic	>3.2 to ≤ 3.5	3.35	7	0.6
7	4	Polysaprobic	>3.5 to ≤ 4.0	3.75	3	0.3

Table 2. Distribution (by number of occurrences) of taxa having saprobic values

No. of occurrences	10–19	20–49	50–99	100–199	200–399	400+
Number of taxa	74	88	58	31	35	14

Revision of saprobic values

The procedure used to revise the saprobic values was as follows.

(1) Analyse the 1106 samples in the database to determine the number of occurrences (n_{ij}) of taxon j, irrespective of density level, in quality band i.
(2) Derive the probabilities of occurrence (p_{ij}) of taxon j in quality band i:

$$p_{ij} = n_{ij}/N_i \quad (2)$$

where N_i is the number of samples taken from rivers in quality band i.
(3) Estimate the saprobic valencies[†] (v_{ij}) of taxon j:

$$v_{ij} = \frac{10p_{ij}}{\sum_{i=1}^{7} p_{ij}} \quad (3)$$

Note that v_{ij} is defined here as a real number, not an integer as has been the convention in the saprobic system. This change was made in the interests of precision.
(4) Derive the raw revised saprobic value ($_rs'_j$) of taxon j:

$$_rs'_j = \frac{1}{10}\sum_{i=1}^{7} v_{ij}\, x_i \quad (4)$$

where x_i is the mid-point saprobic index of quality class i, as defined in Table 1.

At this stage the revised values are valid in relative terms only, because the revision process tends to compress their range, resulting in a reduced standard deviation and slightly modified mean value. In order to ensure that the revised values do not result in an overall shift in saprobic indices, it is necessary to rescale them to preserve their original mean and standard deviation. Walley and Hawkes (1996) derived their rescaling parameters by using *primary lists* of taxa. The purpose in compiling a *primary list* is to minimise the distortions that can arise if some of the $_rs'_j$ values used to estimate the rescaling parameters are unreliable. Since $_rs'_j$ values are most reliable when n_j (the total number of occurrences of taxon j across all quality classes) is large, it is desirable to exclude taxa with small values of n_j from the *primary list*. In this study, a threshold of 100 occurrences was used as the criterion for *primary list* membership. This resulted in a *primary list* of 80 taxa having saprobic values ranging from 1.0 to 3.6 and raw revised values ranging from 1.49 to 3.0.

The next stage was to derive the rescaling parameters and revised saprobic values, as follows.

(5) Determine the mean (m) and standard deviation (S) of the saprobic values (s_j) of the *primary* taxa.
(6) Determine the mean (m'_r) and standard deviation (S'_r) of the raw revised saprobic values $_rs'_j$ of the *primary* taxa.
(7) Derive the revised saprobic values (s'_j) of all 300 *primary* and *non-primary* taxa by rescaling the raw revised values as follows:

$$s'_j = m + (_rs'_j - m'_r)\frac{S}{S'_r} \quad (5)$$

This ensures that the revised saprobic values of the 80 *primary* taxa have the same mean and standard deviation as their original saprobic values. The corresponding mean and standard deviation of the *non-primary* taxa may differ slightly from their original values, but this is to be expected since their revised values are less reliable than those of the primary taxa, having been estimated from smaller samples.

Revision of the indicator weights

The procedure adopted for the revision of the indicator weights was based upon the same basic principle as that used to derive the original weights, but using less rigid rules with the aim of producing more consistent results. The rules were devised to ensure that the revised weights (w'_j) of the 80 *primary* taxa had the same overall distribution as their original weights (w_j). The revision procedure was as follows.

(1) For each taxon j, examine its saprobic valencies ($v_{1j}\ldots v_{ij}\ldots v_{7j}$) and determine:

$_1V_{\max}$ the maximum value of v_{ij},

$_2V_{\max}$ the maximum value of any two adjacent values of v_{ij} (i.e. the sum $v_{ij}+v_{i+1,j}$),

$_3V_{\max}$ the maximum value of any three consecutive values of v_{ij}, and

$_4V_{\max}$ the maximum value of any four consecutive values of v_{ij}.

Note that $\sum_{i=1}^{7} v'_{ij} = 10$

(2) Apply the following sequence of rules to determine the revised weight w'_j:

IF $_4V_{\max} < 8.50$ THEN $w'_j = 1$, ELSE $w'_j = 2$
IF $_3V_{\max} > 6.25$ THEN $w'_j = 2$
IF $_2V_{\max} > 6.00$ THEN $w'_j = 3$
IF $_2V_{\max} > 7.75$ THEN $w'_j = 4$
IF $_2V_{\max} > 8.50$ THEN $w'_j = 5$
IF $_1V_{\max} > 7.50$ THEN $w'_j = 5$.

Note that these rules were developed for use with the primary taxa (i.e. taxa with more than 100 occurrences, and hence fairly reliable saprobic valencies, $v_{1j}\ldots v_{ij}\ldots v_{7j}$). Their use on taxa with fewer occurrences (esp. <20) tends to result in an overestimation of the weight.

RESULTS

Table 3 shows the mean, standard deviation, maximum and minimum values of the original, raw and revised saprobic values (SVs), and Table 4 shows the distributions of the original and revised SVs with respect to quality class. Distributions of the original and revised indicator weights (IWs) are given in Table 5, whilst the main set of results, showing the original and revised SVs and IWs for each of the 226 taxa having occurrences ≥ 20 are given in Table 6.

[†]Note that saprobic valencies sum to 10, not to unity—which would be computationally more convenient and more in keeping with their function as weights. However, for consistency the convention has been retained, which is why 10 appears in the nominator of Equation (3) and the denominator of Equation (4).

DISCUSSION OF RESULTS

Distributions of revised saprobic values and weights

An important point to note here is that the uneven distribution of samples across quality bands, as indicated in Table 1, did not bias the estimation of revised saprobic values and weights, since the analytical method was based upon probabilities (see Equation (2)) not frequencies of occurrence. However, the shortage of samples from quality bands 6 and 7 did result in less reliable (not biased) revised saprobic values and weights for the taxa that were indicative of these quality classes. In fact, this was true for all taxa having relatively few occurrences and should be taken into account when examining the results given in Table 6.

Table 3 shows that the mean and standard deviation of the raw SVs of the *primary* taxa were noticeably different from their original values, but that after rescaling they were identical to them, as intended. The restoration of the spread in SVs from raw to revised values is also noticeable in the range of the maximum and minimum values, albeit to a lesser degree. The mean and standard deviation of the revised SVs of the 300 taxa having 10 or more occurrences differ slightly from their original values, due to the less reliable estimates derived for infrequently occurring taxa. Note also that the revised minimum of 0.74 and maximum of 4.09 exceeded the original range of 1–4. This was the result of two taxa at the extremes of the original saprobic range being found to be even more extreme, in relation to other taxa, than had previously been thought. These were *Beggiatoa alba* (revised from 4.0 to 4.09) and *Crenobia alpina* (revised from 1.0 to 0.74) which occurred in 15 and 27 samples, respectively. The difference between the original and revised SVs was less than 0.25 for 50.4% of the taxa, and differed by more than 0.75 for only 6.7% of the taxa.

Table 4 indicates that the distribution of original SVs with respect to quality bands was mainly concentrated around quality bands 1 to 2. This is most apparent from the band-width-corrected distribution, which was corrected for the variable band width by dividing frequency by band width. The revised SVs are more smoothly distributed, but spill over at the extremes of the distribution into two new bands (i.e. <1 and >4), for reasons mentioned earlier.

The distribution of the revised IWs of the *primary* taxa (Table 5) was almost identical to that of the original weights, indicating that the rules used to

Table 3. Comparison between the original, raw and revised saprobic values with respect to their mean, standard deviation, maximum and minimum

	Primary list taxa (80)			All 300 taxa	
	Original SVs	Raw SVs	Revised SVs	Original SVs	Revised SVs
Mean	2.014	2.148	2.014	1.906	1.870
St. deviation	0.563	0.376	0.563	0.589	0.628
Maximum	3.6	3.00	3.29	4.0	4.09
Minimum	1.0	1.49	1.04	1.0	0.74

Table 4. Distribution of original and revised saprobic values of all 300 taxa by quality band

River quality band	<1	1	1–2	2	2–3	3	3–4	4	>4
Saprobic band width	0.4	0.5	0.3	0.5	0.4	0.5	0.3	0.5	0.4
Distribution of original SVs	0	93	58	96	27	17	3	6	0
Band-width corrected distrib. (%)	0.0	26.8	27.8	27.6	9.7	4.9	1.4	1.7	0
Distribution of revised SVs	17	77	49	85	44	20	4	3	1
Band-width corrected distrib. (%)	6.1	21.9	23.3	24.2	15.7	5.7	1.9	0.9	0.4

Table 5. Distribution of original and revised weights by range of occurrence

	Range of occurrences	No. of taxa	Original weights					Revised weights				
			1	2	3	4	5	1	2	3	4	5
Primary taxa	>100	80	8	33	28	9	2	7	33	26	13	1
Non-primary	50–99	58	1	18	21	14	4	1	9	28	10	10
Taxa	20–49	88	4	25	35	16	8	1	8	34	18	27
	10–19	74	2	19	29	15	9	2	7	14	11	40
All taxa (total)	>10	300	15	95	113	54	23	11	57	102	52	78

Table 6. Original and revised saprobic values for taxa with 20 or more occurrences

Taxon		n_j	s_j	s'_j	w_j	w'_j
BACTERIA						
Sphaerotilus natans	[a]	457	3.6	3.29	3	2
Zooglea ramigera		65	4.0	3.78	5	3
MYCOPHYTA						
Asterothrix rhaphidioides		39	2.3	1.88	2	3
Tetracladium marchalianum		49	2.3	2.28	2	3
CYANOPHYTA	[b]					
Lyngbya martensiana		84	1.5	1.48	3	3
Lyngbya sp.	*	136	2.0	1.96	2	2
Merismopedia glauca	*	137	1.8	2.04	4	2
Merismopedia punctata		36	1.9	2.22	5	2
Nostoc sp.		27	1.6	1.16	3	4
Oscillatoria limosa		68	3.1	2.29	2	3
Oscillatoria sp.	*	240	2.3	2.35	1	1
Phormidium autumnale		64	2.5	2.35	2	3
Phormidium foreolarum		67	3.0	2.64	4	3
Phormidium sp.	*	374	2.2	1.72	1	2
Plectonema sp.		44	1.4	1.57	3	3
Pleurocapsa sp.		55	1.4	1.37	3	3
Rivularia sp.		78	1.4	1.25	3	3
CHRYSOPHYTA						
Hydrurus foetidus	*	143	1.4	1.12	3	4
XANTHOPHYTA						
Tribonema sp.		97	1.6	1.31	2	4
Vaucheria geminata		27	1.9	1.86	3	3
Vaucheria sp.	*	247	1.8	2.39	3	1
BACILLARIOPHYTA						
Achnanthes lanceolata		69	2.0	1.19	3	3
Achnanthes minutissima	*	244	2.0	1.62	3	2
Achnanthes sp.	*	376	2.0	1.80	2	2
Amphora ovalis	*	200	1.7	2.10	2	2
Amphora ovalis v. pediculus		53	1.5	1.61	3	2
Asterionella formosa		32	1.4	1.63	3	5
Ceratoneis arcus	*	254	1.4	1.42	3	3
Cocconeis pediculus	*	368	1.7	1.82	2	3
Cocconeis placentula	*	454	1.6	1.62	2	2
Cyclotella comta	*	104	1.2	1.59	4	4
Cyclotella meneghiniana	*	178	2.6	2.62	3	3
Cymatapleura elliptica		43	1.8	1.85	3	3
Cymatopleura librilis		24	2.5	2.42	3	3
Cymatopleura solea	*	171	2.2	2.47	3	2
Cymbella affinis	*	189	1.3	1.52	4	2
Cymbella lanceolata		29	1.5	1.32	3	3
Cymbella minuta	*	157	2.0	1.60	3	2
Cymbella prostrata		76	1.8	2.70	3	2
Cymbella sinnuata	*	186	1.5	1.79	3	2
Cymbella ventricosa	*	494	2.0	2.04	3	1
Denticula elegans		69	1.0	1.08	5	5
Denticula tenuis		33	1.2	1.00	4	5
Diatoma elongatum		85	1.6	1.88	3	2
Diatoma hiemale	*	167	1.0	1.04	5	5
Diatoma vulgare	*	680	2.2	1.84	2	2
Fragilaria capucina		81	1.5	1.53	3	3
Fragilaria construens		58	1.3	1.41	4	3
Fragilaria crotonensis		91	1.4	2.58	2	1
Fragilaria vaucheriae		22	1.8	2.43	2	2
Frustulia vulgaris		27	1.8	2.01	2	4
Gomphonema acuminatum		35	1.7	1.24	4	3
Gomphonema angustatum		64	2.1	1.09	2	4
Gomphonema constrictum		30	1.9	1.70	3	5
Gomphonema intricatum	*	100	1.2	1.16	4	4
Gomphonema int. v. pumil.		21	1.1	1.20	3	3
Gomphonema olivaceum	*	434	2.0	2.04	2	3
Gomphonema parvulum	*	179	2.1	2.29	1	3
Gyrosigma acuminatum		30	2.2	1.94	3	3
Gyrosigma attenuatum	*	110	2.2	1.89	3	4
BACILLARIOPHYTA cont						
Gyrosigma scalproides		99	2.2	1.85	3	4
Hantzschia amphioxys		32	2.4	2.43	2	3
Melosira granulata		38	1.8	2.26	4	3
Melosira varians	*	527	1.7	2.03	2	2
Meridion circulare	*	268	1.1	1.23	5	4
Navicula avenacea	*	119	2.0	2.89	2	2
Navicula bacillum		25	1.5	2.01	3	3
Navicula cryptocephala	*	352	2.4	3.05	2	2
Navicula crypto. v. veneta		27	3.1	2.39	3	3
Navicula cuspidata		38	2.5	2.82	3	5
Navicula gracilis	*	296	1.7	2.04	2	3
Navicula hungarica		51	2.5	2.36	3	3
Navicula pupula		66	1.9	1.92	2	2
Navicula radiosa		64	1.5	1.56	2	2
Navicula rhynchocephala	*	120	2.7	2.82	4	2
Nitzschia acicularis	*	426	2.7	2.40	4	2
Nitzschia dissipata	*	233	1.3	1.76	3	3
Nitzschia fonticola		38	1.3	1.12	3	3
Nitzschia linearis	*	196	1.7	1.91	3	2
Nitzschia pales	*	623	2.6	2.89	1	2
Nitzschia sigmoidea	*	179	2.5	2.06	3	3
Rhoicosphaenia curvata	*	236	1.8	2.06	3	3
Stephanodiscus hantzschii		33	2.7	2.71	4	4
Surirella angusta		40	1.7	2.15	2	3
Surirella ovata	*	370	1.6	2.38	2	2
Synedra acus		72	1.8	1.88	3	3
Synedra ulna	*	509	2.1	2.18	2	2
Synedra vaucheriae		69	2.2	2.22	2	4
Tabellaria flocculosa		23	1.0	1.07	5	4
EUGLENOPHYTA						
Euglena sp.		90	3.0	2.91	2	3
Phacus sp.		30	2.3	2.87	2	5
CHLOROPHYTA						
Ankistrodesmus convolutus		20	2.0	2.59	5	3
Ankistrodesmus falcatus	*	104	2.1	2.29	2	2
Chlamydomonas sp.		49	2.8	2.66	1	4
Chlorella vulgaris		61	3.1	2.59	2	4
Cladophora fracta		32	2.3	2.24	2	2
Cladophora glomerata		95	2.0	2.02	2	3
Cladophora sp.	*	474	2.3	2.08	2	3
Closterium acerosum		21	2.6	2.81	3	5
Closterium ehrenbergii		47	2.0	2.61	3	1
Closterium leibleinii		73	2.7	2.59	4	3
Closterium lunula		28	1.3	1.99	4	3
Closterium moniliferum		26	2.2	2.06	3	2
Closterium sp.	*	222	2.2	2.75	2	1
Cosmarium botrytis		30	2.3	1.96	2	2
Cosmarium sp.	*	150	1.8	3.10	2	1
Gongrosira incrustans	*	294	2.0	2.34	2	2
Microspora amoena		27	1.2	1.42	4	5
Microspora quadrata		60	2.1	1.77	2	2
Mougeotia sp.		99	1.4	1.41	3	3
Oedogonium sp.	*	322	1.4	2.07	3	3
Pandorina morum		53	2.1	2.47	2	3
Pediastrum boryanum		78	1.9	2.40	3	3
Pediastrum duplex		44	2.2	2.62	3	3
Pediastrum tetras		40	1.8	2.59	3	3
Scenedesmus acuminatus		98	2.2	2.34	4	3
Scenedesmus acutus		73	2.0	2.32	4	2
Scenedesmus bijugatus		20	2.0	2.72	5	3
Scenedesmus ecornis		31	1.7	2.19	4	5
Scenedesmus obliquus	*	130	2.8	2.75	3	2
Scenedesmus quadricaud	*	254	2.1	2.87	2	2
Selenastrum sp.		23	1.9	3.28	2	3
Spirogyra sp.	*	225	2.2	1.70	2	3
Stigeoclonium tenue	*	233	2.8	2.70	3	4
Tetraedron sp.		35	1.9	2.24	2	3
Ulothrix tenuissima		82	1.2	1.36	4	3

[a] An asterisk denotes member of the *primary list*.
[b] CYANOPHYTA are now recognised as a form of bacteria and are called CYANOBACTERIA in recognition of this fact.

Table 6. (*Continued*)

Taxon		n_j	s_j	s'_j	w_j	w'_j
CHLOROPHYTA cont						
Ulothrix zonata-o		70	1.1	1.13	5	5
Ulothrix zonata-p		74	2.9	2.65	5	4
Zygnema sp.		62	1.2	1.55	4	3
RHODOPHYTA						
Audouinella chalybea	*	277	2.5	1.77	3	2
Audouinella violacea		31	1.5	1.26	3	4
Bangia atropurpurea		20	1.3	1.42	4	3
Batrachospermum		24	1.2	1.17	4	5
BRYOPHYTA						
Fontinalis antipyretica		90	1.3	1.46	4	3
Fontinalis sp.	*	150	1.3	1.52	4	3
ANTHOPHYTA						
Ceratophyllum demersum		21	2.2	2.29	3	3
Elodea canadensis		20	1.8	1.87	3	5
Myriophyllum sp.		87	1.8	1.97	3	3
Potamogeton sp.		21	1.9	1.54	2	5
RHYZOPODA						
Euglypha alveolata		27	2.0	2.42	3	3
CILIATA						
Carchesium sp.		24	2.9	3.27	2	3
Vorticella sp.		35	3.1	2.71	2	5
HYDROZOA						
Hydra sp.		59	1.8	1.59	3	2
TURBELLARIA						
Crenobia alpina		27	1.0	0.74	5	5
Dendrocoelum lacteum		93	2.4	2.18	3	5
Dugesia lugubris		31	2.1	1.52	3	2
Dugesia sp.		74	1.9	1.49	2	2
Planaria torva		61	2.2	1.94	2	3
Polycelis nigra		21	2.0	1.19	4	4
NEMATODA	*	142	2.8	2.65	4	2
OLIGOCHAETA						
Eiseniella tetraedra	*	254	2.1	2.16	2	2
Lumbriculus variegatus		35	3.1	2.48	3	3
Nais sp.	*	312	2.7	2.56	2	2
Stylaria lacustris		56	2.5	1.99	3	5
Tubifex sp.	*	336	3.6	3.21	3	2
Tubifex tubifex		28	3.6	3.79	3	3
HIRUDINEA						
Erpobdella octoculata	*	319	3.0	2.35	2	3
Glossiphonia complanata	*	106	2.5	2.14	3	4
Helobdella stagnalis	*	114	2.8	2.57	2	4
Piscicola geometra		24	2.1	1.70	2	5
GASTROPODA						
Ancylus fluviatilis	*	244	1.7	1.57	2	3
Lymnaea ovata		22	2.4	2.44	1	3
Lymnaea peregra		27	2.2	1.94	4	4
Lymnaea sp.	*	116	2.3	2.02	1	3
Physa fontinalis		31	1.7	2.42	2	4
Theodoxus danubialis		65	2.0	1.64	4	5
BIVALVIA						
Pisidium sp.		82	2.4	1.83	1	3
AMPHIPODA						
Gammarus fossarum	*	653	1.8	1.74	2	2
Gammarus roeseli		55	2.2	2.05	3	5
Niphargus sp.		48	1.0	1.02	5	5
Synurella ambulans		57	2.0	1.59	4	3
ISOPODA						
Asellus aquaticus	*	223	2.8	2.40	3	4
HYDRACARINA	*	344	1.1	1.44	1	3
EPHEMEROPTERA						
Baetis alpinus	*	128	1.4	1.11	3	4
Baetis fuscatus	*	113	2.1	1.66	3	3
Baetis rhodani	*	351	1.6	1.77	2	2
Caenis sp.		44	1.9	1.90	2	4
Centroptilum luteolum		22	1.9	1.76	3	5
Ecdyonurus sp.	*	209	1.6	1.36	2	3
Ecdyonurus venosus	*	181	1.5	1.12	1	4
Epeorus sylvicola		40	1.1	1.18	5	4
Ephemera danica		28	1.6	1.49	2	5
Ephemerella ignita	*	208	2.1	1.65	2	3
Ephemerella major (belgica)		42	1.5	1.31	3	5
Habrophlebia sp.		46	1.5	1.69	3	3
Heptagenia sulphurea		35	2.2	1.55	3	5
Paraleptophlebia sp.		62	1.6	1.72	2	4
Potamanthus luteus		30	2.2	2.08	3	5
Rhithrogena semicolorata		64	1.2	0.99	4	5
Rhithrogena sp.		74	1.2	1.00	4	5
PLECOPTERA						
Amphinemura sp.		47	1.2	1.02	4	5
Brachyptera sp.		49	1.2	1.20	4	4
Isoperla sp.		56	1.5	1.05	3	5
Leuctra sp.	*	368	1.6	1.23	4	3
Nemoura sp.		59	1.4	1.01	3	4
Perla marginata		29	1.2	1.36	4	3
Perlodes sp.		29	1.2	1.19	4	4
Protonemura sp.		99	1.2	0.91	4	5
ODONATA						
Gomphus sp.		26	2.0	1.88	3	4
HEMIPTERA						
Aphelocherius aestivalis		32	2.0	2.04	3	5
Corixa sp.		21	2.0	2.03	2	5
Sigara sp.		22	2.0	2.44	3	4
TRICHOPTERA						
Hydropsyche sp.	*	493	2.5	1.80	3	3
Limnephilidae	*	226	2.0	1.19	3	4
Odontocerum albicorne		27	1.0	1.07	5	5
Polycentropodidae		43	2.0	1.53	2	3
Polycentropus		21	1.7	1.68	2	5
Rhyacophila sp.	*	342	1.7	1.18	2	4
Sericostoma sp.		45	1.5	1.01	3	5
Silo sp.		24	1.2	1.33	4	3
DIPTERA						
Atherix ibis		30	1.6	1.12	2	4
Atherix sp.		99	1.6	1.14	2	4
Chironomidae (red)	*	228	3.3	2.96	2	3
Chironomidae (green)	*	567	1.7	2.48	1	1
Chironomus thummi		58	3.5	3.68	3	3
Dicranota sp.	*	144	1.9	1.28	2	3
Hexatoma sp.		26	2.0	1.19	1	4
Psychodidae		27	2.4	2.15	1	2
Rheotanytarsus sp.		34	1.8	1.83	2	4
Simulium sp.	*	402	2.0	2.36	2	1
Tabanus sp.		21	2.1	1.84	2	2
Tipula sp.		88	1.9	1.42	2	3
COLEOPTERA						
Elmis sp.	*	372	1.4	1.34	3	3
Esolus sp.		50	1.2	1.11	4	4
Gyrinus sp.		25	1.8	1.77	2	5
Limnius sp.	*	129	1.4	1.29	3	3

allocate the revised IWs preserved the spread of the original values. However, the corresponding distributions for the *non-primary* taxa indicate a tendency for the rules to overestimate the revised IWs as sample size (i.e. number of occurrences) decreases. It may be possible to modify the rules to account for sample size, but this has not been done because the authors feel that it would be better to ensure that any

definitive revisions of IWs are, wherever possible, based upon at least 100 occurrences. The values given in Table 5 for the *non-primary* taxa are presented for guidance only. Overall, 35.0% of revised weights retained their original value, 39.9% differed from the original value by ± 1, 19.9% by ± 2 and 5.3% by ± 3.

Magnitude of revised saprobic values and indicator weights

This section highlights the most significant differences between original and revised saprobic values and indicator weight.

Table 6 lists 131 plant taxa, 91 animal taxa, 2 bacteria and 2 mycophyta arranged in 30 taxonomic groups (mainly Phyla) starting with decomposers, followed by producers and then consumers. Overall, the average difference between the original and revised SVs in Table 6 is negligible (-0.02). However, the average for the animal taxa has decreased by 0.18 whilst that of the plant taxa has increased by 0.1. This may be a consequence of the original values for animals and plants being allocated by different sets of experts, zoologists and botanists. There are also some significant changes in average SVs across the 30 taxonomic groups. Most notably, the average for Chlorophyta has increased by 0.25, suggesting that this group is more tolerant of organic pollution than had originally been thought. Conversely, the averages for Turbellaria and Trichoptera have decreased by 0.42 and 0.35, respectively, suggesting that they are much less tolerant of organic pollution than had originally been thought.

There are 16 *primary* taxa that have differences between the revised and original SVs greater than 0.5. Nine have increased values (indicating poorer quality) and seven, decreased values. In order of decreasing difference they are: *Cosmarium* sp. ($+1.30$), *Navicula avenacea (syn. Navicula lanceolata)* ($+0.89$), Limnephilidae (-0.81), *Surirella ovata* ($+0.78$), Chironomidae green ($+0.78$), *Scenedesmus quadricauda* ($+0.77$), *Audouinella (or Chantransia) chalybea* (-0.73), *Hydropsyche* sp. (-0.70), *Oedogonium* sp. ($+0.67$), *Erpobdella octoculata* (-0.65), *Navicula cryptocephala (v.cryptocephala)* ($+0.65$), *Dicranota* sp. (-0.62), *Vaucheria* sp. ($+0.59$), *Rhyacophila* sp. (-0.52), *Closterium* sp. ($+0.55$) and *Spirogyra* sp. (-0.50).

Large differences between the revised and original IWs (i.e. >1) are observed for 13 *primary* taxa. The largest difference was for *Ecdyonurus venosus,* whose IW increased from 1 to 4. Increases of 2 occurred for *Rhyacophila* sp., Hydracarina, *Lymnaea* sp., *Helobdella stagnalis,* and *Gomphonema parvulum*. These appear to be much better indicators than had originally been thought. Decreases of 2 occurred for *Vaucheria* sp., *Nitzschia acicularis*, *Merismopedia glauca*, *Cymbella affinis*, Nematoda, *Navicula rhynchocephala*, and *Cymbella ventricosa*, implying that these are poorer indicators than had originally been thought.

The way forward

The purpose of this paper has not been to produce a definitive revision of Slovenian saprobic values, but to develop and demonstrate a methodology that could provide a basis for a more comprehensive data-based reappraisal of the saprobic values in use across Europe. However, it will first be necessary for practising river ecologists to be convinced of the validity of the method and the value of its results. Therefore, we invite river ecologists to comment on our proposed revisions in the light of their field experiences. In particular, we would like comments on our revised values at the species level, because valid comparisons at genera or family level are confounded by the fact that the species composition of genera and families vary from region to region. We also hope that similar reappraisals will be undertaken in other countries that use the saprobic system. Consequently, we have fully described our method mathematically in order to facilitate such studies.

CONCLUSION

The principal result of this study is a methodology for the reappraisal of saprobic values based on data analysis, and a set of revised saprobic values and indicator weights for 300 Slovenian taxa, which mirror the original ones assigned by ecological experts.

The study had a few limitations. Firstly, very few samples were available from poor river quality sites, since Slovenian rivers are relatively clean. This may have resulted in unreliable, but not biased, revisions for a few taxa that appear primarily in poor quality waters. Secondly, we disregarded the effect of site type, because Slovenian rivers are mainly torrential streams and all of our samples were collected from sites having eroding substrata. Thus, the vast majority of the sites were 'riffles', as defined by Walley and Hawkes (1996). In general, however, this will not be the case, so that it will be necessary to carry out separate analyses for each site type. Finally, the revised values of taxa that occurred infrequently in the data may be unreliable, so that care must be taken when interpreting the values given in Table 6 if n_j is small.

Although the authors are confident that the revised values, especially those of the *primary* taxa, provide improved estimates of their 'true' values, they are not considered definitive but are offered as a basis for discussion. However, the noticeable difference between the revised saprobic values of plants (up by 0.10 on average) and those of animals (down by 0.18) is thought to be of considerable importance, since it

implies a mismatch between the values allocated by zoologists and those allocated by botanists.

Despite its limitations, the study developed and demonstrated a sound methodology for the objective revision of saprobic values and indicator weights, based on the analysis of field data rather than the opinions of experts. If applied to other national databases, the method has the potential to significantly improve and harmonise the various saprobic systems currently in use across Europe.

Acknowledgements—The data used in the reappraisal were provided by the Hydrometeorological Institute of Slovenia. Thanks are due to Damjan Demšar for performing the preliminary data analysis. Thanks are also due to Dr. Herbert A. Hawkes for his comments on an early draft of the paper.

REFERENCES

De Pauw N. and Hawkes H. A. (1993) Biological monitoring of river water quality. *Proceedings of Freshwater Europe Symposium on River Water Quality Monitoring and Control*, Aston University, Birmingham, pp. 87–111.

DIN 38410-2 (1990) Deutsche Einheitsverfahren zur Wasser-, Abwasser- und Schlammuntersuchung; Biologisch-ökologische Gewässeruntersuchung (Gruppe M); Bestimmung des Saprobienindex (M 2). DIN/Beuth Verlag, Berlin.

Džeroski S., Demšar D., Grbović J. and Walley W. J. (1997a) Learning to infer chemical parameters of river water quality from bioindicator data. *Proceedings of the sixth Electrotechnical and Computer Science Conference*, Vol. B, pp. 129–132. Slovenia Section IEEE, Ljubljana, Slovenia.

Džeroski S., Grbović J., Walley W. J. and Demšar D. (1997b) A computer-based reappraisal of bioindicator zone preferences and weights within the saprobic index using data from river quality surveys in Slovenia. *Proceedings of the fourth International Conference on Water Pollution*, pp. 331–340. Computational Mechanics Publications, Southampton.

Džeroski S., Grbović J., Walley W. J. and Kompare B. (1997c) Using machine learning techniques in the construction of models. Part II: Rule induction. *Ecol. Modelling* **95**, 95–111.

Džeroski S., Grbović J. and Walley W. J. (1998) Machine learning applications in biological classification of river water quality. In *Machine Learning, Data Mining and Knowledge Discovery: Methods and Applications*, eds R. S. Michalski, I. Bratko and M. Kubat, pp. 429–448. Wiley, Chichester.

European Union (2000) Directive of the European Parliament and of the Council (2000/60/EC) Establishing a Framework for Community Action in the Field of Water Policy, Luxembourg, 23 October.

Grbović J. (1994) Applicability of various procedures for the assessment of quality of torrential Streams. Ph.D. Thesis. Biotechnical Faculty, University of Ljubljana, Slovenia (in Slovenian).

ISO 7828 (E) (1985) Water quality: methods of biological sampling: guidance on handnet sampling of aquatic benthic macro-invertebrates. International Standards Organisation.

Moog O., ed. (1995) *Fauna Aquatica Austriaca*. Wasserwirtschaftkataster, Bundesministerium fur Land und Forstwirtschaft, Vienna.

Pantle R. and Buck H. (1955) Die biologische Uberwachung der Gewas und die Darstellung der Ergebnisse. *Gas Wasserfach* **96**, 603.

Walley W. J. and Džeroski S. (1995) Biological monitoring: a comparison between Bayesian, neural and machine learning methods of water quality classification. In: *Proceedings of the International Symposium on Environmental Software Systems*, IFIP Conference Series, eds R. Denzer, G. Schimak and D. Russell. Chapman & Hall, London.

Walley W. J. and Fontama V. N. (1998) Neural network predictors of average score per taxon and number of families at unpolluted river sites in Great Britain. *Water Res.* **32**(3), 613–622.

Walley W. J. and Fontama V. N. (2000) New approaches to river quality classification based upon artificial intelligence. In *Assessing the Biological Quality of Fresh Waters: RIVPACS and Other Techniques*, eds M. T. Furse and J. F. Wright. Freshwater Biological Association, Ambleside, UK.

Walley W. J., Fontama V. N. and Martin R. W. (1998) *Applications of artificial intelligence in river quality surveys. R&D Technical Report E52*. Environment Agency, Bristol.

Walley W. J. and Hawkes H. A. (1996) A computer-based reappraisal of Biological Monitoring Working Party scores using data from the 1990 River Quality Survey of England and Wales. *Water Res.* **30**(9), 2086–2094.

Walley W. J. and Hawkes H. A. (1997) A computer-based development of the Biological Monitoring Working Party score system incorporating abundance rating, site type and indicator value. *Water Res.* **31**(2), 201–210.

Zelinka M. and Marvan P. (1961) Zur Präzisierung der biologischen Klassifikation der Reinheit fliessender Gewässer. *Arch. Hydrobiol.* **57**, 389–407.