



Dempster-Shafer reasoning for the biological surveillance of river water quality

M. Boyd, W.J. Walley, H.A. Hawkes

*Department of Civil Engineering, Aston University,
Birmingham B4 7ET, UK*

ABSTRACT

An application of Dempster-Shafer reasoning for the biological surveillance of river water quality is described. A reasoning system is being developed which emulates the expert's ability to determine the quality of river water from samples of benthic macro-invertebrate communities taken from the river bed.

Diagnoses of river water quality are reported in terms of a proposed biological classification scheme. Ecological knowledge on a reference set of indicator organisms is stored in the form of subjective probability distributions, representing the occurrence of the organisms in each of four discrete states of abundance across the quality classes. These distributions provide support for competing propositions in the frame of discernment formed by the biological classes. Combination of this support leads to a classification of river water quality at a sampled site.

The reasoning system has been tested in its ability to match the expert's classification of a set of 53 sampled river sites. The results suggest that Dempster-Shafer reasoning affords a viable procedure for integrating biological evidence in order to classify river water quality.

INTRODUCTION

By determining the biological communities in a river, biologists can assess the impact of pollution on aquatic ecosystems, and thereby monitor the quality of the abiotic environment. Since changes in the quality of river water induce changes to biological community structure, observations of these can be used to indicate trends in river water quality. This is the conceptual basis for using biological surveillance in monitoring river water quality (Hawkes [5]).

In freshwater biological surveys, organisms are sampled, counted and identified to some taxonomic level. Such data can then be processed to report information on the biological quality of water at the sampled site, in summary form suitable for non-biologists



350 Water Pollution

who have responsibility for quality-management programmes. Metcalfe [6] has reviewed the many bio-assessment systems which are based on benthic macro-invertebrates, the taxonomic group most frequently sampled for this purpose. Recently, multivariate statistics have been used to predict benthic communities for unpolluted sites: Wright *et al.* [10] have proposed that this approach could be employed in routine biological surveillance.

The authors are investigating the application of techniques drawn from artificial intelligence in the areas of river water quality monitoring and control. This paper describes the results of a study into the use of an uncertain reasoning scheme, known as Evidential or Dempster-Shafer reasoning, for the biological classification of river water quality. The work forms part of a larger programme to develop an expert system for river pollution (Walley [8]).

UNCERTAIN REASONING FOR RIVER WATER QUALITY ASSESSMENT

The problem of assessing riverine biological quality from benthic data has been reformulated here in terms of a classification decision under uncertainty, a common experience for any expert in the real world. In an earlier paper, Walley *et al.* [9] used Bayesian reasoning to produce a bio-assessment procedure for river water quality. The intention was to improve on current biotic score systems, which, it was argued, have important limitations. The results showed good agreement between the expert's and the model's biological classification of a range of freshwater sites.

Bayesian decision methods remain the primary numerical approach for representing and manipulating uncertainty, and have been employed in expert systems such as PROSPECTOR for mineral exploration. However, alternative numerical reasoning schemes each have their advocates. Of these, the Dempster-Shafer theory of evidence (Shafer [7]) has attracted much attention for its flexibility in the representation and aggregation of evidence. Its advantages include the ability to closely model diagnostic reasoning, and to distinguish between uncertainty, or lack of knowledge, and indifference (Gordon and Shortliffe [4]).

Biological Classification using Dempster-Shafer reasoning

The starting point for the direct interpretation model of river water quality was the adoption of a biological classification scheme, which parallels the chemical classification used by the National Rivers Authority in England and Wales. Thus the scheme uses five classes: B1a, B1b, B2, B3, B4, ranging from unpolluted (B1a) to grossly polluted (B4) waters.

As with probability theory, Dempster-Shafer or Evidential reasoning assumes that there is a fixed set of mutually exclusive and exhaustive hypotheses, the **frame of discernment**, Θ , containing the set of objects (or concepts) of interest, in our case the biological water quality classes. Thus $\Theta = \{B1a, B1b, B2, B3, B4\}$. Each hypothesis in Θ corresponds to a one-element subset (called a *singleton*). However, in D-S reasoning, the

term *hypothesis* is normally used in an enlarged sense. In contrast to the Bayesian approach, one can allocate belief to all the possible subsets of Θ (the "power set"), not only the singletons. The subsets can be viewed as corresponding to various propositions of diagnostic interest in the power set. As an example, the proposition 'Poor Quality water' could be represented by the two-element subset $\{B3, B4\}$, while its complement $\{B3, B4\}^c$ corresponds to the three-element subset $\{B1a, B1b, B2\}$ representing the proposition 'NOT Poor Quality water'.

A number m in the range $[0, 1]$, known as the *basic probability assignment (bpa)* is used to represent the degree to which the evidence supports a proposition. As an example, consider that one piece of evidence from a sampled site results from observing that the freshwater shrimp *Gammarus pulex* is present. Allowing for a certain degree of uncommitted belief, the basic probability assignment suggested by this evidence may be as follows: $m(\{B1a\}) = 0.2$, $m(\{B1b\}) = 0.5$, $m(\{B2\}) = 0.1$. The uncommitted belief $m(\Theta)$ is equal to 0.2 by definition of a *bpa*. Uncertainty can thus be explicitly represented by assigning uncommitted belief to Θ , rather than to any proper subset of it. Thus, if evidence favours a single subset, say $\{B1b, B2\}$, remaining belief can be assigned to Θ rather than to the complement of this subset, as would be required in the Bayesian model.

The benthic taxa may be viewed as sensors which can report their current state and accordingly assign a measure of support to the relevant propositions. The inference task then becomes one of integrating evidence from these disparate, perhaps conflicting sources to produce a report in terms of the adopted classes. This problem of fusing evidential data from multiple knowledge sources using Dempster-Shafer reasoning has been examined by Bogler [1] and Garvey *et al.* [3].

Nature of the evidence

In our application, the evidence is provided by a reference set of benthic organisms, selected by the expert, which are deemed to exist in one of four discrete states: rare, established, abundant, and absent. For each taxon, discrete probability distributions were obtained which depict the likelihood of occurrence across the adopted range of water qualities, for each state. These distributions are not unlike the concept of Saprobic valency (Zelinka and Marvan [11]), but in our case the uncertainty estimates are the expert's personal probabilities of the states of the sensors, given his experience and knowledge of the ecological requirements of the organisms. The reader is referred to Walley *et al.* [9] for a full explanation of the elicitation of this knowledge.

Combination of evidence

Support for the various hypotheses arises from observing multiple items of independent evidence. This support must then be combined in such a way as to represent the aggregation of this evidence. In Evidential reasoning, this combination takes place by Dempster's rule. For two sets of evidence with *bpa's* m_1 and m_2 respectively, Dempster's rule computes a new *bpa* representing their combined effect and referred to as the *orthogonal sum* of m_1 and m_2 . For some proposition Z of Θ , $m_1 \oplus m_2(Z)$ is the sum of all

352 Water Pollution

products of $m_1(X)$ and $m_2(Y)$, where X and Y are subsets of Θ whose intersection is Z . A simple example will illustrate this principle.

Example of combination using Dempster's rule

Suppose that the presence in a sample of the mayfly *Baetis rhodani* supports the proposition $\{B1a, B1b, B2\}$ to degree 0.6 (m_1) while the presence of *LUMBRICULIDAE* supports $\{B1b, B2, B3\}$ to degree 0.7 (m_2). What is the net effect of these two items of evidence? The computation of the orthogonal sum can be shown via the following table:-

Table 1: Illustration of Dempster's rule for two bpa 's

	$m_2(\{B1b, B2, B3\}) = 0.7$	$m_2(\Theta) = 0.3$
$m_1(\{B1a, B1b, B2\}) = 0.6$	$\{B1b, B2\} \quad (0.42)$	$\{B1a, B1b, B2\} \quad (0.18)$
$m_1(\Theta) = 0.4$	$\{B1b, B2, B3\} \quad (0.28)$	$\Theta \quad (0.12)$

The first row of the table shows the first bpa m_1 in which a belief of 0.4 is left unassigned to the environment Θ . The first column shows the second bpa m_2 . Set intersections are shown in the table alongside the numeric mass product. For instance the set intersection of $\{B1a, B1b, B2\}$ and $\{B1b, B2, B3\}$ is $\{B1b, B2\}$, while the basic probability number assigned to this new proposition is $0.6 \times 0.7 = 0.42$. The orthogonal sum is thus: $m_1 \oplus m_2(\{B1b, B2\}) = 0.42$, $m_1 \oplus m_2(\{B1b, B2, B3\}) = 0.28$, $m_1 \oplus m_2(\{B1a, B1b, B2\}) = 0.18$, $m_1 \oplus m_2(\Theta) = 0.12$.

If two items of evidence do not have any sets in common, the evidence is totally contradictory and the orthogonal sum is not defined. Dempster's rule promotes consensus by only assigning belief to intersecting sets, which represent common elements of evidence. Since the rule is commutative, the evidence may be combined in any order.

Basic probability assignments from the evidence

Probability mass is assigned as a result of a sensor report, i.e. by observation of its state at a sample site. A database of belief measures previously elicited from the expert, H.A. Hawkes, exists in the form of conditional probabilities for each of the reference organisms. The reasoning system allocates these measures to appropriate propositions, according to some pre-determined scheme. These schemes are intended to relieve the user of the potential task of allocating belief measures across the power set.

The allocation schemes derive from the form of the probability distributions, and have been implemented by different versions of the reasoning system. In this paper we confine ourselves to one scheme which we call *singleton support*. For singleton support the scheme allocates for the k th sensor probability masses to each of the singleton subsets $\{B1a\}$, $\{B1b\}$, ... $\{B4\}$, corresponding to the H_i ($i = 1$ to 5) classes. The allocation may be written as: $m(\{B1a\}) = P(H_1|e^k) \cdot (1 - \epsilon)$, $m(\{B1b\}) = P(H_2|e^k) \cdot (1 - \epsilon)$..., $m(\Theta) = \epsilon$, where the subscript j refers to the $j = 1$ to 4 sensor states.

Since the $P(H_i/e^k)$ sum to unity across the five classes, this allocation seems to be the most obvious probability assignment for evidence which can be structured as supporting more than one singleton set. The quantity ϵ is the amount of uncommitted belief, arising perhaps from some uncertainty in the evidence itself, which is assigned to the environment Θ . ϵ is also a measure of *evidential discount*, in which support for each proper subset is reduced by a discount rate (Shafer [7]). Belief which is divided among singleton sets will always entail some amount of contradiction, since the singletons are by definition mutually exclusive states of the world. For such evidence, combination of probability assignments will almost certainly result in conflict, a problem which numerical reasoning schemes must accommodate. Evidential discount is one means of dealing with this phenomenon. Note that if the discount rate $\epsilon = 0$, the mass assignment for the singleton support belief function reduces to a Bayesian belief function.

CLASSIFICATION TESTS

In the Bayesian analysis described in an earlier paper, a conformity index was used to identify potentially contentious evidence whose inclusion would unduly influence the classification, allowing the user to reject the data if necessary or to alert the user to possibly erroneous data (Walley *et al.* [9]). For the purposes of this paper, the analysis was repeated for both Bayesian and Dempster-Shafer systems using evidential discount to manage conflict rather than via the use of conformity indices.

The classifications were carried out for a data-set of 53 freshwater sites in the Yorkshire area. The evidence was allocated as singleton support, in which discount rates were varied for particular states of the evidence over fourteen experiments. If $\epsilon = 1$ for a particular state, that evidence was completely discounted from the combination procedure. Established and Abundant evidence were discounted at a nominal 10% throughout. Table 2 gives details of the 14 tests in evidential discount.

Both the Bayesian and Dempster-Shafer systems produce a ranked order of belief in each quality class for a site, but for the Bayesian model, a point value for each quality class is obtained, while for the D-S model, the belief is given as an interval. This support was then mapped to a single Classification Index, this being a weighted average of the values computed for each class. Weights of 0,1,2,3,4 were used for classes B1a to B4. This choice is reminiscent of the Saprobic Index, in which zones of increasing organic enrichment are weighted in a similar fashion.

DISCUSSION

The performance of the classification systems was measured via linear regression analyses, in which the expert's classification is considered to be the independent variable, and the system's classifications the dependent variable. The reasoning systems were tested for their ability to match the expert's classification, obtained by asking the expert to examine the lists of benthic taxa recorded at the sites, and record his opinion on the quality class on this basis. The expert himself used inter-class divisions (such as B2--,

354 Water Pollution

B3+) where appropriate. By converting these to a linear scale bounded by 0 and 4, direct comparison between the system's and the expert's classification index was possible.

Table 2: Discount rates (%) applied to Absent and Rare Evidence

Test	Absent	Rare	Test	Absent	Rare
1	100	10	8	10	60
2	80	10	9	10	40
3	60	10	10	100	100
4	40	10	11	80	100
5	20	10	12	60	100
6	10	100	13	40	100
7	10	80	14	20	100

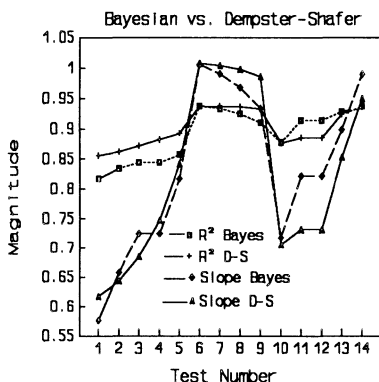


Figure 1: Variation of regression coefficients with discount for Bayesian and Dempster-Shafer systems.

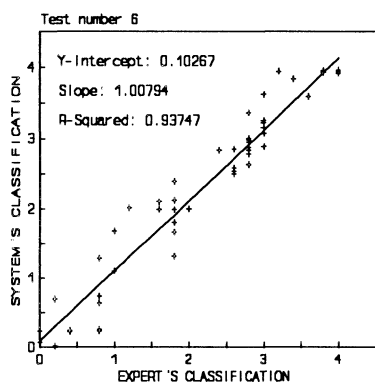


Figure 2 Least-square line for Dempster-Shafer classifier (Test no. 6)

Figure 1 shows a comparison of linear regression coefficients (r^2 and slope) obtained for the tests over the data-set for both Bayesian and Dempster-Shafer reasoning. Initially, absent evidence was completely discounted, resulting in low values of the coefficients. This confirms an observation from the earlier experiments using Bayesian analysis that the inclusion of absent evidence makes a significant difference to the success of the classification. However, varying discount rates for rare evidence had little impact.

Figure 2 shows the linear least-square fit for test 6. In general, the largest errors were caused by the system erring on the poorer-quality side. Note that this device of "erring on the safe side" is adopted by the BMWP biotic score system, in which the most tolerant species within each family is used for allocating points (Department of the Environment

[2]). The two numerical methods were closely matched, with the Dempster-Shafer procedure marginally better for this data set. This may reflect the method's flexibility in dealing with evidential discount.

CONCLUSIONS

In the systems developed so far, probability distributions are required for a reference set of organisms: the choice of these, and the distributions themselves, are those of a single expert in the field. The model could be extended to reflect the consensus of multiple experts, and operate at different taxonomic levels according to the sampling effort undertaken. Although evidential discount is effective in reducing conflict, it is believed that the performance of the reasoning systems can be improved by some form of conformity checks of sample data, in order to highlight evidence which would either be ignored by the expert when assessing river water quality, or which conversely would be deemed to be especially significant. The selective use of evidential discount could weight individual evidence either on the basis of a quantitative conformity index or by heuristic rules.

The results obtained to date indicate that both the Bayesian and Dempster-Shafer numerical reasoning schemes provide good decision algorithms for rapidly assessing the biological quality of river water in terms of a proposed classification scheme, provided that the data are appropriately weighted. Benthic organisms within the reference set which are absent from the sample provide significant evidence, which, although not as important as that from established or abundant taxa, markedly improve the success of the reasoning systems in classifying water quality. Evidence from taxa occurring in rare numbers seems to be of less significance.

Of the two reasoning methods, the Dempster-Shafer calculus allows a more direct representation of uncertainty in the evidence, but testing on other data sets as well as the use of mass-assignment schemes other than singleton support is required before reaching any firm conclusions on the relative merits of the two approaches. However, the Dempster-Shafer theory of evidence provides an intuitively attractive framework for integrating and interpreting biological sensor data which is noisy, imprecise, and sometimes conflicting. With the increasing importance of the ecological monitoring of surface water quality, uncertain reasoning algorithms could provide a powerful means of improving data interpretation in this area.

ACKNOWLEDGEMENT

The work described in this paper was carried out, in part, under a UK Science and Engineering Research Council studentship.

REFERENCES

1. Bogler, P.L. 'Shafer-Dempster Reasoning with Applications to Multisensor Target Identification Systems' *IEEE Trans. Systems, Man, and Cybernetics*, Vol. SMC-



356 Water Pollution

17, No. 6, November/December 1987.

2. Department of the Environment. *Assessment and presentation of the biological qualities of rivers in Great Britain*. Final report of the Biological Monitoring Working Party of the Standing Technical Advisory Committee on Water Quality, 1980.
3. Garvey, T.D., Lowrence, J.D. and Fischler, M.A. 'An inference technique for integrating knowledge from disparate sources'. In: *Proceedings of the Seventh Int. Joint Conference on Artificial Intelligence*, Vancouver, BC, 319-325, 1979.
4. Gordon, J. and Shortliffe, E.H. 'A method of managing evidential reasoning in a hierarchical hypothesis space' *Artificial Intelligence*, **26**, 323-357, 1985.
5. Hawkes, H.A. 'Conceptual basis for the biological surveillance of river water quality'. In Hughes, J.G. and Hawkes, H.A. (collators) *Biological surveillance of river water quality*. Proceedings of Section K, jointly with Section D, of the British Association for the Advancement of Science, Aston 1977, pp. 1-14, 1978.
6. Metcalfe, J.L. 'Biological Water Quality Assessment of Running Waters Based on Macroinvertebrate Communities: History and Present Status in Europe' *Environmental Pollution*, **60**, 101-139, 1989.
7. Shafer, G. *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, N.J., U.S.A., 1976.
8. Walley, W.J. 'Artificial Intelligence in River Water Quality Monitoring and Control'. In *Freshwater Europe: Proceedings of Symposium on River Water Quality Monitoring and Control in Europe*, Birmingham, England, 1993. (Paper presented).
9. Walley, W.J, Hawkes, H.A., and Boyd, M. 'Application of Bayesian Inference to River Water Quality Surveillance' *Proceedings of the Seventh International Conference on the Applications of Artificial Intelligence in Engineering*, pp.1033-1047, 1992.
10. Wright, J.F., Amritage, P.D., Furse, M.T. and Moss, D. 'A new approach to the biological surveillance of river quality using macroinvertebrates' *Verhandlungen Internationale Vereinigung für Theoretische und Angewandte Limnologie*, **23**, 1548-1552, 1988.
11. Zelinka, M. and Marvan, P. 'Zur präzisierung der biologischen klassifikation der reinheit fliessender gewässer'. *Arch. Hydrobiol.*, **57**, 389-407, 1961.