IDENTIFICATION OF THE KEY BIOLOGICAL INDICATORS OF NUTRIENT ENRICHMENT IN RIVERS FOR USE IN PREDICTIVE/DIAGNOSTIC MODELS

MF Paisley[*1], WJ Walley[1], J Nikhade[1] & R Dils[2]

[1]*Centre for Intelligent Environmental Systems, School of Computing, Staffordshire University, Stafford ST18 0AD, UK.*
[2]*Environment Agency, NCEHS, Evenlode House, Howbery Park, Wallingford OX10 8BD, UK.*

ABSTRACT
This paper reports progress made in elucidating key biological indicators of nutrient enrichment, by phosphorus and nitrogen, in riverine systems. Data from the 1995 survey of river quality in England and Wales, as well as recent data from Northern Ireland, have been analysed using information theory and neural networks. The use of information theory is expected to lead to the identification of indicator taxa of good overall effectiveness, and ecological justification is given for some of the stronger candidates found. However the set of indicators obtained in this way is expected to contain some replication of function. On the other hand, the neural network approach is expected to lead to subsets of indicators with much less replication. The results of the two approaches are compared and are more consistent in the case of phosphorus than nitrogen.

KEYWORDS: Eutrophication, diffuse pollution, biological monitoring, artificial intelligence, modelling.

INTRODUCTION
Eutrophication, the enrichment of waters by nutrients resulting in an array of biological changes, is known to affect many industrialised countries. Symptoms include increased production of algae, enhanced growth of higher aquatic plants, and the deterioration of oxygen supply, which typically lead to a reduction in biodiversity and adverse effects on water use. Recent reports have identified eutrophication as a priority environmental issue for fresh waters in England and Wales (Environment Agency, 1998), as well as Northern Ireland, Scotland and the Irish Republic (McGarrigle, 1998).

Effective eutrophication control measures require an understanding of the relationships between nutrient pressures, nutrient concentrations in the receiving water, and impacts on the ecology (see Withers & Lord, 2002, for example). Implementation of nutrient removal measures at sewage treatment works, to protect sites of high conservation value and under the European Commission's (EC) Urban Waste Water Treatment Directive, are reducing discharges of nutrients from point sources of pollution. Consequently, diffuse sources such as agriculture are likely to represent a relatively more important source of nutrients in the future.

Biological monitoring of river quality has grown in importance over the past few decades due to the recognition of important advantages over chemical monitoring. The recently introduced Water Framework Directive is also founded on a biological approach (EC, 2000). Bio-monitoring requires the development of tools with the capacity to interpret biological and environmental variables in terms of chemistry and vice versa. The response of benthic macroinvertebrate and macrophyte communities in rivers to environmental stresses of various types is acknowledged, and scientifically utilised as a means of assessing water quality. The specific response of macroinvertebrates to nutrient enrichment in rivers has been well studied, although phytoplankton or macrophytes are preferentially used as a means of investigating nutrient-biological response relationships. Despite the amount of work undertaken, it is not clear if a given nutrient regime results in a particular biological community of macroinvertebrates and macrophytes in rivers. The answer is unlikely to be simple as the particular response of the macroinvertebrate and macrophyte community to a particular stressor may be different under various conditions such as geographical location, geology, ambient water chemistry, hydrological regime and season.

Artificial intelligence (AI) techniques have been demonstrated to provide powerful methods of modelling complex relationships in the field of water quality using supervised pattern recognition (Ruck *et al*., 1993; Lek *et al*., 1996; Walley & Fontama, 1998; Gabriels *et al*., 2002), unsupervised pattern recognition (O'Connor & Walley, 2001) and Bayesian belief networks (Trigg *et al*, 2000; Walley *et al*., 2002a). The ultimate objective of the work described here is the construction of belief networks with the ability to diagnose nutrient levels from the biology and characteristics of the local environment, as well as being able to predict changes in the biology from changes in nutrient levels, perhaps as a result of modified agricultural practice.

This paper reports progress made in the prerequisite step of elucidating key biological indicators (macroinvertebrates and macrophytes) of phosphorus and nitrogen in rivers, and builds on earlier work (Walley *et al*., 2002b).

METHODOLOGY
The data used to develop the systems were derived from the Environment Agency's (EA) 1995 biological, environmental and chemical surveys of rivers in England and Wales, as well as recent data from Northern Ireland. The project databases derived from the EA survey data had 6695 records, 3255 for spring and 3440 for autumn. The biological data consisted of the abundance levels of the 76 BMWP (Biological Monitoring Working Party) families that were recorded in the spring

and autumn macroinvertebrate surveys of 1995. There are about 4,000 species of aquatic macroinvertebrates in the British Isles, so for simplicity biomonitoring methods use family level identification. The environmental data were the averages of spring and autumn values. The chemical data recorded alongside each biological/environmental record consisted of the concentrations of up to 34 chemical variables, each averaged over three months preceding the date of the biological sample. The two chemical variables of interest in this research were TRP (Total reactive phosphorus) and TON (Total organic nitrogen).

It is possible that the variations in nutrient concentration and community composition merely reflect the site and catchment characteristics, and that there is little association between nutrients and the macroinvertebrate community. To justify the search for macroinvertebrate taxa with strong associations, a preliminary investigation was undertaken in order to quantify the amount of extra information provided by including the 76 BMWP taxa in addition to the 11 environmental variables as input to the models. Data from the 1995 survey of rivers in England and Wales were used to train and test a neural network with phosphorus as the output variable; firstly with all variables included in the input, and secondly with just the eleven environmental variables. The correlation coefficients obtained between the predicted levels of phosphorus and actual levels in the two cases were 0.810 and 0.596, which correspond to explaining around 65.6% and 35.5% of the variation respectively. Thus, by including the macroinvertebrate taxa the effectiveness of the network was almost doubled. This provided convincing justification of the underlying approach, and confirmed that the search for key indicator taxa was viable.

When characterising indicators it is useful to distinguish between two different types. Firstly, one can identify a set of indicators that all display a strong association with the effect in question. Using a sporting analogy, this can be termed the *squad* approach. However, such a set may contain indicators which replicate the role of others, resulting in a certain amount of overlap and redundancy. Since a useful predictive or diagnostic model needs economy as well as accuracy, further subsets could be identified containing indicators where each has a different role. There may be more than one such subset, and continuing the sporting analogy this might be termed a *team* approach. This paper describes two techniques, one based on information theory and the other on neural networks. These produce indicators belonging to *squad* and *team* sets respectively.

The first approach is based on mutual information (MI) and is designed to identify indicators of TRP and TON belonging to a *squad* set. Mutual information is a measure of the amount of information one random variable contains about another, and is interpreted as the reduction in uncertainty of one random variable due to knowledge of the other. This technique has been applied to the 1995 GQA data of rivers in England and Wales. In an attempt to remove the effects of pollution from high organic load the data analysed included sites of GQA classes A (very good biological quality) and B only (good biological quality). The data was then categorised according to five phosphorus bands, and five types of site (ranging from fast-flowing upland streams to slow-flowing lowland rivers) and each sample allocated to one of these twenty-five classes. This then allowed a distribution of the data to be defined according to the 76 BMWP taxa and the twenty-five classes, from which the mutual information can be calculated.

The second approach adopted was the use of a neural network. Commercial software (NeuralWare Professional II) was used to implement a multi-layer perceptron – a neural network based on supervised learning. Each network was trained and tested on the same data, and although this means that the performance tests were based on dependent data and therefore not true tests of performance, the technique at least allows the more important taxa to be identified. The data analysed was the 1991-2000 survey data of rivers in Northern Ireland. In all cases the network was a multi-layered perceptron with a single output (TRP or TON) and 144 input variables (9 environmental variables and presence/absence data for 76 BMWP taxa and 59 macrophyte species). The networks had one hidden layer of 6 nodes, and the training schedule was Save Best (100,000/1,000,10). This means that during training the network was tested (on the training data) every 1,000 cycles and if a test showed improvement the network was saved. Training continued until 10 successive tests showed no improvement in performance, and the total number of cycles reached 100,000.

A procedure known as impact analysis was used to determine the relative contribution of each input variable to the final prediction. Each input variable is temporarily disabled in turn and the performance of the model determined each time. The percentage reduction in the correlation coefficient below its baseline value (with no inputs disabled) is recorded each time a variable is disabled. At the end of the process the variables can be ranked according to impact, with the most important input variables having the highest percentage impacts. The input vector can then be pruned of variables with negative or very low impacts and the training procedure entered once more. Several cycles of the impact analysis procedure can be used to optimise the input vector, so that by the end the input vector may contain only 40 or 50 of the original 144 variables. Under this approach, variables which replicate the function of others may become labelled as low impact, because the network can function just as well whether they are disabled or not. Such variables may be expected to be pruned from the *team* regardless of the merit of their inclusion in the *squad*.

RESULTS AND DISCUSSION

The annual and seasonal MI values of invertebrate taxa for TRP and TON for the 1995 data for rivers in England and Wales are given in Table 1. The strongest 38 out of the 76 taxa are ranked in descending order of annual average. It is clearly not possible to justify the position of every taxon, but an attempt is made to provide ecological justification for some of the top-ranked taxa.

Table 1 Annual and Seasonal Mutual Information (MI) Values of BMWP taxa for TRP and TON based on English and Welsh river data (1995 survey, GQA class A & B only divided into 5 Site Types and 5 TRP/TON Bands).

	Average TRP					Average TON		
Taxon	**Spring**	**Autumn**	**Ann Avg**	**Rank**	**Taxon**	**Spring**	**Autumn**	**Ann Avg**
Rhyacophilidae	0.0471	0.0385	0.0428	**1**	Hydrobiidae	0.0658	0.0335	0.0497
Elmidae	0.0332	0.0491	0.0412	**2**	Goeridae	0.0516	0.0438	0.0477
Heptageniidae	0.0388	0.0321	0.0355	**3**	Gammaridae	0.0452	0.0383	0.0417
Asellidae	0.0378	0.0329	0.0353	**4**	Oligochaeta	0.0564	0.0241	0.0403
Calopterygidae	0.0265	0.0399	0.0332	**5**	Heptageniidae	0.0527	0.0223	0.0375
Ephemerellidae	0.0451	0.0205	0.0328	**6**	Chironomidae	0.0479	0.0238	0.0359
Tipulidae	0.0292	0.0361	0.0327	**7**	Sphaeriidae	0.0444	0.0267	0.0356
Gammaridae	0.0309	0.0318	0.0313	**8**	Hydropsychidae	0.0406	0.0287	0.0346
Erpobdellidae	0.0333	0.0292	0.0313	**9**	Baetidae	0.0428	0.0262	0.0345
Caenidae	0.0309	0.0314	0.0311	**10**	Elmidae	0.0375	0.0304	0.0340
Hydrobiidae	0.0254	0.0368	0.0311	**11**	Leptoceridae	0.0325	0.0258	0.0292
Sericostomatidae	0.0274	0.0346	0.0310	**12**	Ephemerellidae	0.0418	0.0163	0.0291
Leuctridae	0.0306	0.0306	0.0306	**13**	Lymnaeidae	0.0302	0.0254	0.0278
Coenagriidae	0.0190	0.0392	0.0291	**14**	Nemouridae	0.0356	0.0195	0.0276
Nemouridae	0.0199	0.0370	0.0284	**15**	Sericostomatidae	0.0335	0.0204	0.0270
Ancylidae	0.0237	0.0321	0.0279	**16**	Leptophlebiidae	0.0366	0.0156	0.0261
Limnephilidae	0.0215	0.0333	0.0274	**17**	Caenidae	0.0264	0.0257	0.0261
Hydropsychidae	0.0220	0.0325	0.0272	**18**	Limnephilidae	0.0394	0.0127	0.0261
Sphaeriidae	0.0255	0.0289	0.0272	**19**	Ancylidae	0.0290	0.0226	0.0258
Baetidae	0.0328	0.0215	0.0272	**20**	Simuliidae	0.0292	0.0207	0.0250
Planorbidae	0.0207	0.0318	0.0263	**21**	Erpobdellidae	0.0308	0.0190	0.0249
Planariidae	0.0292	0.0230	0.0261	**22**	Hydrometridae	0.0299	0.0190	0.0244
Leptophlebiidae	0.0236	0.0286	0.0261	**23**	Asellidae	0.0258	0.0229	0.0244
Goeridae	0.0237	0.0272	0.0255	**24**	Ephemeridae	0.0274	0.0214	0.0244
Lepidostomatidae	0.0246	0.0250	0.0248	**25**	Tipulidae	0.0284	0.0202	0.0243
Perlodidae	0.0280	0.0211	0.0245	**26**	Planariidae	0.0242	0.0233	0.0237
Leptoceridae	0.0218	0.0270	0.0244	**27**	Lepidostomatidae	0.0275	0.0194	0.0234
Simuliidae	0.0223	0.0260	0.0241	**28**	Hydroptilidae	0.0305	0.0157	0.0231
Oligochaeta	0.0229	0.0248	0.0239	**29**	Neritidae	0.0222	0.0240	0.0231
Chironomidae	0.0245	0.0223	0.0234	**30**	Glossiphoniidae	0.0317	0.0143	0.0230
Ephemeridae	0.0196	0.0267	0.0232	**31**	Rhyacophilidae	0.0224	0.0218	0.0221
Brachycentridae	0.0169	0.0291	0.0230	**32**	Physidae	0.0225	0.0214	0.0220
Gyrinidae	0.0194	0.0259	0.0227	**33**	Brachycentridae	0.0209	0.0227	0.0218
Aphelocheiridae	0.0135	0.0303	0.0219	**34**	Taeniopterygidae	0.0334	0.0101	0.0217
Psychomyiidae	0.0182	0.0229	0.0205	**35**	Leuctridae	0.0221	0.0213	0.0217
Neritidae	0.0170	0.0240	0.0205	**36**	Perlodidae	0.0305	0.0126	0.0216
Hydroptilidae	0.0221	0.0188	0.0205	**37**	Dytiscidae	0.0204	0.0220	0.0212
Glossiphoniidae	0.0226	0.0171	0.0198	**38**	Aphelocheiridae	0.0222	0.0189	0.0206

Raised levels of the nutrient phosphorus can lead to enhanced growth of algae, which in turn can cause depleted levels of oxygen, especially at night. The overall expectations might be the absence of invertebrate families intolerant of high phosphorus-low oxygen conditions, and the proliferation of species that are tolerant to these conditions. The relative abundance of such families might be a good indicator of overall phosphorus levels. According to Davies (1974), two families showing intolerance of oxygen depletion are Rhyacophilidae and Gammaridae. Comparison with Table 1 shows that one of these (Rhyacophilidae) heads the list while the other is placed eighth. Furthermore, Asellidae is quite tolerant of oxygen depletion (Davies) so would potentially benefit from conditions of low oxygen that result in the removal of Gammaridae, its principal competitor in oxygenated conditions. Asellidae is ranked fourth. Families tolerant of oxygen depletion are Erpobdellidae and Chironomidae (Davies), placed ninth and thirtieth respectively. Erpobdellidae would be expected to thrive due to the proliferation of macroinvertebrate families on which it preys, such as the highly tolerant Oligochaeta, ranked twenty-ninth.

The same mechanism for oxygen depletion would be expected to occur with raised nitrate levels, but it is complicated by the association of nitrate with ammonia and other nitrogen compounds. High nitrate levels are often associated with the breakdown of organic pollution, represented by ammonia, under conditions of high oxygenation. Although the data only included samples from good quality sites (GQA classes A and B), some of the TON rankings suggest a link with ammonia.

Consider two of the stronger indicators of phosphorus mentioned earlier, Rhyacophilidae and Gammaridae. Both are sensitive to oxygen depletion and Rhyacophilidae is fairly tolerant of ammonia, but Gammaridae is not (Davies). Gammaridae remains a strong indicator of TON (third), but Rhyacophilidae is ranked much lower (thirty-first). Confirmation of the likely influence of organic pollution is provided by the high ranking of Oligochaeta (fourth) and Chironomidae (sixth), both of which would be expected to proliferate in the low oxygen conditions associated with the breakdown of organic matter. Of the two other families mentioned earlier, Erpobdellidae might be expected to thrive with the proliferation of Oligochaeta, and Asellidae decline with the proliferation of Gammaridae. However both of these effects may be masked by only moderate ability of each to tolerate ammonia and the families occur lower down the ranking than before (twenty-first and twenty-third respectively).

While clearly not exhaustive these observations constitute a certain amount of ecological justification for the relative ranking of sensors with regard to TRP and TON obtained by mutual information. Further observations could be made, such as the appearance of four families that are highly ranked for both TRP and TON (Elmidae, Heptagenidae, Ephemerelidae and Hydrobiidae). In addition these rankings can be used as a benchmark to compare with the outcomes from the neural network analysis, which are shown in Table 2. These comparisons must be treated with some caution, given that the datasets used in each case are different and might reflect differences in river ecology between England and Wales on the one hand, and Northern Ireland on the other.

Taking the phosphorus (TRP) results first, it can be verified that the strongest macroinvertebrate taxa from the neural network are skewed towards the stronger end of the MI rankings. The best three macroinvertebrate taxa from the neural network tests appear ninth, third and sixth respectively in the MI rankings. The only macroinvertebrate taxon in the top ten MI ranking not to appear in the best twenty from the neural network tests is Asellidae. As mentioned earlier, this does not imply that Asellidae is not a good sensor, but that it probably replicates the function of other macroinvertebrate taxa. Balanced against this good agreement is the occurrence of five of the best twenty macroinvertebrate taxa from the neural network (Dendrocoelidae, Sialidae, Lymnaeidae, Haliplidae and Gerridae) in the weaker half of the MI rankings (not shown in Table 1). However this may just be evidence of the optimisation of the input vector to the neural network which is designed to purge the input vector of taxa which replicate the function of others. Once many replicates have been pruned it might be argued that some of those which remain might well occur lower down the MI rankings. None of the best 20 macroinvertebrate taxa in the neural network appear in the worst twenty of the MI rankings however. An observation of ecological interest is the reliance of the neural network on Aphelocheiridae which is ranked fifth in importance despite being extremely rare. This may reflect the fact that Apheloceiridae breathes via diffusion of oxygen across hairs rather than through gills, implying a very high oxygen saturation requirement and meaning that it would occur only at very low phosphorus levels.

When the comparison is repeated for TON the outcome is less conclusive. Although the highest ranked macroinvertebrate for the neural network (Chironomidae) is placed sixth in the MI rankings, the second-ranked (Siphlonuridae) is seventy-fourth in the MI rankings. In contrast to the results for phosphorus, the best macroinvertebrates for the neural network do not appear skewed towards the stronger end of the MI rankings but rather scattered throughout, with eleven in the stronger half and nine in the weaker half. The same comment applies however, that this may be the natural outcome of the pruning replicate taxa. There may be more fundamental questions however, such as the difficulty of separating out the effects of the association with ammonia. There were also basic differences in the datasets - the Northern Ireland data was based on presence/absence and the English and Welsh data used abundance levels.

CONCLUSIONS

Several conclusions can be drawn from this preliminary study aimed at drawing out biological indicators of TRP and TON in rivers. Firstly, general indicators have been identified using a technique based on mutual information. These have strong association with TRP and TON and in the main are commonly occurring macroinvertebrate taxa. Secondly very specific indicators have been identified using successive impact analyses with neural network models. These indicators are likely to be macroinvertebrate taxa and macrophyte species that are rare and sensitive to very particular conditions, such as very low phosphorus levels. Both kinds of indicators are likely to be useful in the design of diagnostic or predictive models applied to diffuse agricultural pollution.

A very significant finding from analysis of the Northern Irish dataset is that macrophytes did not turn out to be better indicators of TRP or TON than the invertebrates. This challenges the current biological monitoring approach whereby phytoplankton or macrophytes are preferentially used as a means of investigating nutrient-biological response relationships.

Further work is required to complete the analysis, for example to eliminate potential confounding factors such as the association of TON with ammonia, and to fulfil the need for 'like for like' tests on both datasets used. Optimisation of the input vector for the neural network models might be made more efficient using genetic algorithms.

Table 2 Results of impact analyses (macroinvertebrates shown shaded) showing the top 38 indicator variables for TRP and TON based on neural network model of data from Northern Ireland.

TRP			TON	
Variable	Impact %	Rank	Variable	Impact %
Silt	6.629	1	*Width*	8.187
Depth	6.602	2	Chironomidae	6.905
Pebbles	5.434	3	Siphlonuridae	5.851
Width	5.378	4	*Sand*	5.326
Erpobdellidae	4.159	5	Lemanea	5.06
Heptageniidae	3.612	6	*Altitude*	4.876
Ephemerellidae	3.595	7	Leuctridae	4.401
Dendrocoelidae	3.566	8	Ephemeridae	4.148
Petasites	3.533	9	*Depth*	3.835
Sparganium	3.238	10	Heracleum	3.623
Sparganium	3.11	11	Solanum	3.121
Aphelocheiridae	3.062	12	Lythrum	2.989
Nuphar	2.76	13	Potamogeton	2.935
Phalaris	2.746	14	Phalaris	2.71
Elmidae	2.293	15	Hydrobiidae	2.667
Ranunculus	2.233	16	Haliplidae	2.613
Sialidae	2.208	17	Oenanthe	2.525
Sphaeriidae	2.08	18	Elmidae	2.108
Glossiphoniidae	2.049	19	Astacidae	1.962
Ancylidae	1.901	20	Polycentropidae	1.83
Planorbidae	1.843	21	Planariidae	1.79
Fontinalis	1.742	22	Ranunculus	1.625
Boulders	1.715	23	Coenagriidae	1.605
Simuliidae	1.458	24	Sparganium	1.547
Gyrinidae	1.444	25	Asellidae	1.458
Caenidae	1.412	26	Juncus	1.103
Vaucheria	1.344	27	Heptageniidae	1.04
Tipulidae	1.311	28	Limnephilidae	0.926
Rhyacophilidae	1.163	29	Cladophora	0.907
Lymnaeidae	1.013	30	Hildenbrandia	0.886
Calopterygidae	0.951	31	Amblystegium	0.854
Diatoms	0.892	32	Gerridae	0.838
Haliplidae	0.881	33	*Pebbles*	0.822
Potamogeton	0.85	34	Fontinalis	0.689
Gerridae	0.827	35	Dendrocoelidae	0.671
Callitriche	0.826	36	Glossiphoniidae	0.614
Astacidae	0.821	37	Veronica	0.495
Asellidae	0.783	38	Callitriche	0.371

ACKNOWLEDGEMENTS

The authors are indebted to Dr H A Hawkes for assistance in interpreting some of the results described here, and to Dr J Murray-Bligh for his comments on the manuscript.

REFERENCES

Davies L. J. PhD Thesis, Aston University UK, 1974.

EC, 2000. European Parliament and Council Directive 2000/60/EC of 23 October 2000 establishing a framework for community action in the field of water policy. Official Journal of the European Communities, L327, 22/12/2000.

Environment Agency (1998). The state of the environment of England and Wales: fresh waters. The Stationary Office, London.

Gabriels W., Goethals P. L. M., & De Pauw N. (2002). Prediction of macroinvertebrate communities in sediments of Flemish watercourses based on artificial neural networks. Verh. Internat. Verein. Limnol. **28**, 777-780.

Lek S., Dimopoulos I. & Fabre A. (1996). Predicting phosphorus concentrations and phosphorus load from watershed characteristics using backpropagation neural networks. Acta OEcologica, **17**, 43-53.

McGarrigle M.L. (1998). Aspects of river eutrophication in Ireland. *Annals of Limnology*, **29**, 1051-63.

O'Connor M. A. & Walley W. J. (2001). River Pollution Diagnostic System (RPDS) – Computer-based analysis and visualisation for bio-monitoring data. Water Science and Technology **46**, 6, 17-23.

Ruck B. M., Walley W. J. & Hawkes H. A., 1993. 'Biological Classification of River Water Quality using Neural Networks' In: G. Rzevski, J. Pastor & R. A. Adey (Eds) *Applications of Artificial Intelligence in Engineering, Vol 2: Applications and Techniques*,. Elsevier/CMP. 361-372.

Trigg D. J., Walley W. J. & Ormerod S. (2000). A prototype Bayesian belief network for the diagnosis of acidification in Welsh rivers. In *Development and Application of Computer Techniques to Environmental Studies*. Eds Brebbia CA, GIbarra-Berastegui and Zanetti P. CMP.

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

Walley W. J., O'Connor M. A., Trigg D. J. & Martin R. W. (2002a) Diagnosing and predicting river health from biological survey data using pattern recognition and plausible reasoning. R&D Technical Report E1-056/TR2. Environment Agency, Bristol. ISBN 1 85705 816 X

Walley W. J., Trigg D. J., Martin R. W., Nikhade J., O'Connor M. A. & Paisley M. F. (2002b). Assessment of river phosphorus and nitrogen from macroinvertebrate data using artificial intelligence techniques. R&D Technical Report, Environment Agency, Bristol.

Withers P. J. & Lord E. I. (2002). Agricultural nutrient inputs to rivers and groundwaters in the UK: policy, environmental management and research needs. *The Science of the Total Environment*, 282-283, 9-24.