Assessment of River Phosphorus and Nitrogen from Macroinvertebrate Data using Artificial Intelligence Techniques.



Technical Report

ENVIRONMENT AGENCY

Assessment of River Phosphorus and Nitrogen from Macroinvertebrate Data using Artificial Intelligence Techniques.

Technical Report

W. J. Walley, D. J. Trigg, R. W. Martin, J. Nikhade, M.A. O'Connor and M. Paisley.

Research Contractor :

Centre for Intelligent Environmental Systems,
School of Computing, Staffordshire University

Publishing Organisation:
Environment Agency
Rio House
Waterside Drive
Aztec West
Almondsbury
Bristol BS12 4UD

Tel: 01454 624400 Fax: 01454 624409

Dissemination status
Internal: Released to Regions
External: Released to Public Domain

Statement of use
This document provides information to Environment Agency staff, researchers and external agencies involved in the assessment of biological and nutrient relationships in riverine environments.

Research contractor
This document was produced by:

Centre for Intelligent Environmental System
School of Computing
Staffordshire University
The Octagon
Beaconside
Stafford ST18 0AD

Tel: 01785 353510

Environment Agency's Project Manager
The Environment Agency's Project Manager for this project was:
Dr Rachael Dils - Environment Agency, NCEHS, Wallingford.

CONTENTS

TABLES

Page

FIGURES

APPENDICES

GLOSSARY

The following list gives brief definitions of the technical terms and acronyms that are used throughout the report.

A	Autumn.
AI	Artificial Intelligence.
AMTN	Ammoniacal nitrogen, mg/l.
ASPT	Average score per taxon.
a40Pyx	Code name of an autumn (a) MIR-max model with an input vector of 40 variables, including Presence-only (P) biological data. They input data were divided into two equal sets (x - used for testing, y – used for training).
BBN	Bayesian Belief Network.
BMWP	Biological Monitoring Working Party score system.
BOD	Biochemical Oxygen Demand.
CIES	Centre for Intelligent Environmental Systems at Staffordshire University.
DO	Dissolved Oxygen (% saturation).
EBBN	Eutrophication Bayesian Belief Network.
GQA class	The biological General Quality Assessment class - river quality assessment scheme used by the Environment Agency.
M2R	A validated database of biological GQA survey data consisting of 6039 sites that were surveyed in the spring and autumn of 1995.
MIR-max	Software for the development of models based on pattern recognition.
MLR	Multiple linear regression.
NFAM	Number of BMWP families (See Table 1.2) found in the sample.
NN	Neural Network
N2R	A validated database of matched biological and chemical GQA survey data consisting of 3615 sites, which are a subset of the M2R sites.
N2Rplus	N2R database plus additional data on up to 34 chemical variables.
psab0516 naab1016	Typical code names for neural network models of TRP (p) and TON (n) for spring (s) and autumn (a) based on GQA classes 'a' and 'b' (ab), using threshold concentrations of 0.5 mg/l (05) and 10mg/l (10) for TRP and TON respectively. The final figure (16) is the number of hidden nodes used in the neural network.
r	Correlation coefficient (Pearson).
RIVPACS	River InVertebrate Prediction And Classification System (version III).
RPBBN	River Pollution Bayesian Belief Network (Software developed by CIES).
RPDS	River Pollution Diagnostic System (Software developed by CIES).
S	Spring.
Site Type	A site type classification into one of five types based on alkalinity, altitude and the percentage of sand /silt in the substrate. Defined by Walley *et al.*, 2001.
ST	Site Type.

s86Axy	Code name of a spring (s) MIR-max model with an input vector of 86 variables including Abundance-based (A) biological data. They input data were divided into two equal sets (x - used for training, y – used for testing).
T	Temperature (°C).
TON	Total Oxidised Nitrogen, mg/l.
TRP	Total Reactive Phosphorus, mg/l.

EXECUTIVE SUMMARY

This Technical Report presents the results of a three-month research project financed by the Environment Agency's National Centre for Ecotoxicology and Hazardous Substances. The aim of the project was to answer the following questions, primarily through the development and testing of Artificial Intelligence (AI) models.

- Are there relationships between the composition of the macroinvertebrate communities in running waters and the concentrations of TRP and TON?
- If so, is the TRP relationship stronger in 'cleaner' rivers?
- Are the relationships strong enough to provide reliable predictions of TRP and TON?
- Which taxa are good indicators of TRP and TON?
- Is the occurrence of *Cladophora* related to concentrations of TRP and TON?
- What are the seasonal and geographical distributions of TRP, TON and *Cladophora*?
- Is there a relationship between recorded DO levels and concentrations of TRP and TON?

The modelling techniques used were multiple linear regression (to provide a baseline assessment) and three AI-based techniques, namely neural networks (using the multi-layered perceptron), pattern recognition (using the MIR-max system) and plausible reasoning (using Bayesian Belief Networks). In view of the short duration of the project, the majority of the work was based upon neural network models. These offered a quick and effective means of developing and testing a wide range of non-linear TRP and TON predictive models. Some results were derived from two existing AI models, RPDS (a MIR-max model) and RPBBN (a BBN model), but these were not sufficiently specific since they were based on all GQA quality classes. It was only possible to develop and test initial prototypes of more specific MIR-max and BBN models, owing to the limitations of time. All of the TRP and TON predictive models used the Agency's macroinvertebrate and environmental data as input.

The data used to develop the systems were derived from the Agency's 1995 biological, environmental and chemical GQA surveys, and the 1995 survey of perceived environmental stresses. The project database had 6695 records, 3255 for spring and 3440 for autumn.

It was concluded that there are relationships between the composition of macroinvertebrate communities and the concentration of nutrients (TRP and TON), and that these are stronger in the 'cleaner' rivers (GQA classes A and B), and stronger for TON than TRP. The relationships with TRP were considered too weak to provide reliable predictions from macroinvertebrate and environmental data. Models based on present / absent data performed almost as well as those based on abundance data. Lists of indicator taxa for TRP and TON have been drawn up, but owing to poor agreement between the methods used to derive them, they are not considered to be definitive and further analysis is recommended. Increasing concentrations of TRP were found to be associated with lower DO saturation and a smaller number of families (NFAM). Increasing concentrations of TON were found to have no effect on DO, but a very noticeable effect on NFAM, first increasing it and then decreasing it. It is recommended that a more detailed study be carried out using macroinvertebrate and macrophyte data, preferably with some taxa identified to species or genera level.

Keywords: eutrophication, phosphorus, TRP, nitrogen, TON, BMWP, macroinvertebrates, bioindicators, *Cladophora*, artificial intelligence, AI, neural networks, pattern recognition, MIR-max, plausible reasoning, Bayesian belief networks, BBN, multiple linear regression.

1. INTRODUCTION

This Technical Report is the outcome of a three-month preliminary study into the predictability of river phosphorus and total oxidised nitrogen from macroinvertebrate data using Artificial Intelligence (AI) techniques. The project was a logical extension of the work carried out by the Centre for Intelligent Environmental Systems (CIES) in National R&D Project E1-056 which developed two AI-based systems for the diagnosis and prediction of river quality from biological and environmental data (Walley *et al.*, 2002). It was not within the scope of this very small project to set it in the context of existing knowledge and understanding of the relationships between nutrients and invertebrates. Its overall aim was to determine whether analyses of large national databases using AI techniques could shed further light on the relationships. The specific aims were defined by a series of questions that had been drawn up by the client, the Environment Agency.

1.1 Aims

The project attempted to answer the following questions:

a) Are there relationships between the composition of the macroinvertebrate communities in running waters and the concentrations of phosphorus (TRP) and total oxidised nitrogen (TON)?
b) If so, is the TRP relationship stronger in the 'cleaner' rivers (e.g. <0.2 mg P/l)?
c) Are the relationships sufficiently strong to permit predictions of TRP and TON from macroinvertebrate data to be made with an acceptable degree of reliability?
d) Which taxa are good indicators (positively or negatively) of TRP and TON?
e) Is there a relationship between recorded occurrences of *Cladophora* and concentrations of TRP and TON?
f) What are the seasonal and geographical distributions of TRP, TON and *Cladophora*?
g) Is there a relationship between recorded dissolved oxygen (DO) levels and concentrations of TRP and TON?

1.2 Modifications to the stated 'Approach to be Taken'

The approach taken differed slightly from that stated in the original tender submitted by CIES. The main differences were that:

a) 10 environmental variables were used to represent site type effects rather than the five Site Types previously defined by Walley *et al.* (1998);
b) extensive use was made of neural networks to help identify the key characteristics of the data, including the rank order of indicator taxa;
c) Bayesian belief networks have not been used to determine the rank order of taxa, since the neural networks provided a quicker means of doing so; and
d) although maps of recorded TRP and TON have been produced, maps of their predicted values have not, due to the large number of models (and hence predictions) produced.

Items a) and b) were major improvements to the original plan.

1.3 The Data

The study was based upon the N2Rplus database that was developed for use in National R&D Project E1-056. This database was constructed from biological, environmental and chemical data recorded in the 1995 survey of rivers in England and Wales[1]. This involved matching the sites in a database of validated biological and environmental data (M2R, with 6039 sites) with the chemical monitoring sites. Since only about half of the sites matched and a few of the chemical sites had no data, the N2Rplus database was reduced to 3556 sites, each with spring and autumn biological samples (i.e. a total of 7112 records). The chemical data recorded alongside each biological/environmental record consisted of the concentrations of up to 34 chemical variables, each averaged over the three months preceding the date of the biological sample. This database was further enhanced by the addition of data on the occurrence of *Cladophora* at biological monitoring sites that was recorded during the 1995 Stress Survey. The cover of *Cladophora* (a green alga) is perceived to be a useful indicator of eutrophication in rivers.

The biological data consisted of the abundance levels (0 to 4, see Table 1.1) of the 76 BMWP families listed in Table 1.2 that were recorded in the spring and autumn surveys of 1995. The environmental data were the averages of spring and autumn values of the 10 variables listed in Table 1.3.

Table 1.1 Abundance Levels

Abundance Level	Number of Individuals Found
0	0
1	1 to 9
2	10 to 99
3	100 to 999
4	≥ 1000

Since not all of the 34 chemical variables were recorded at each site, the database was compressed by removing sites that did not have average values for TRP and TON. The final project database had 6695 records, 3255 for spring and 3440 for autumn.

Table 1.2 The 76 BMWP families used in the study.

Planariidae	Gammaridae	Calopterygidae	Rhyacophilidae
Dendrocoelidae	Astacidae	Aeshnidae	Philopotamidae
Neritidae	Siphlonuridae	Cordulegasteridae	Polycentropidae
Viviparidae	Baetidae	Libellulidae	Psychomyiidae
Valvatidae	Heptageniidae	Hydrometridae	Hydropsychidae
Hydrobiidae	Leptophlebiidae	Gerridae	Hydroptilidae
Lymnaeidae	Ephemerellidae	Nepidae	Phryganeidae
Physidae	Potamanthidae	Naucoridae	Limnephilidae
Planorbidae	Ephemeridae	Aphelocheiridae	Molannidae
Ancylidae	Caenidae	Notonectidae	Beraeidae
Unionidae	Taeniopterygidae	Corixidae	Odontoceridae
Sphaeriidae	Nemouridae	Haliplidae	Leptoceridae
Oligochaeta	Leuctridae	Dytiscidae	Goeridae
Piscicolidae	Capniidae	Gyrinidae	Lepidostomatidae
Glossiphoniidae	Perlodidae	Hydrophilidae	Brachycentridae
Hirudididae	Perlidae	Scirtidae	Sericostomatidae
Erpobdellidae	Chloroperlidae	Dryopidae	Tipulidae
Asellidae	Platycnemidae	Elmidae	Chironomidae
Corophiidae	Coenagriidae	Sialidae	Simuliidae

[1] Although this study was carried out in 2002, it was not possible to use data from the 2000 survey of rivers in England and Wales, because such data were not available at the time. Even if they had been, it would have taken a few months to thoroughly validate the biological, chemical and environmental databases and integrate them into a single project database of the same quality as N2Rplus.

Table 1.3 List of the 10 environmental variables used in the study.

Variable	Description	Variable	Description
ALT	Altitude (m)	DISCH	Discharge Category
DIST	Distance from Source (km)	BLDS	Boulders (% of substrate)
SLOPE	Slope (m/km)	PBLS	Pebbles (% of substrate)
WIDTH	Average Width of river (m)	SAND	Sand (% of substrate)
DEPTH	Average Depth of river (cm)	SILT	Silt (% of substrate)

1.4 Overview of the Analyses

The following analyses were carried out on the N2Rplus (1995) data:

Distribution Analyses

a) Basic statistical analyses of the distributions of TRP and TON concentrations, carried out separately for spring and autumn.
b) Derivation of the spring and autumn geographical distributions of TRP and TON.
c) Derivation of the distribution of *Cladophora* by geographical region and GQA class.

Relationships between nutrient concentrations and Cladophora and DO

d) Analysis of the relationship between TRP/TON and the occurrence of *Cladophora*.
e) Analysis of the relationship between TRP/TON and dissolved oxygen (DO % saturation).

Predictive Models

f) Construction of various multiple linear regression (MLR) models for the prediction of TRP and TON from biological and environmental data.
g) Construction and testing of various neural network (NN) models for the prediction of TRP and TON from biological and environmental data.
h) Construction and testing of Bayesian belief network (BBN) models for the prediction of TRP and TON from biological and environmental data.
i) Construction and testing of various MIR-max (pattern recognition) models for the prediction of TRP and TON from biological and environmental data.

Identification of the Key Indicator Taxa

j) Analysis of the MLR models to identify and rank the key positive and negative indicators of TRP and TON.
k) Impact analyses to determine the relative importance of each taxon in NN models.

2. DISTRIBUTION ANALYSES

2.1 Spring and Autumn Concentrations of TRP and TON

The distribution of TRP and TON concentrations in the spring and autumn of 1995 are shown in Figures 2.1 and 2.2 respectively.

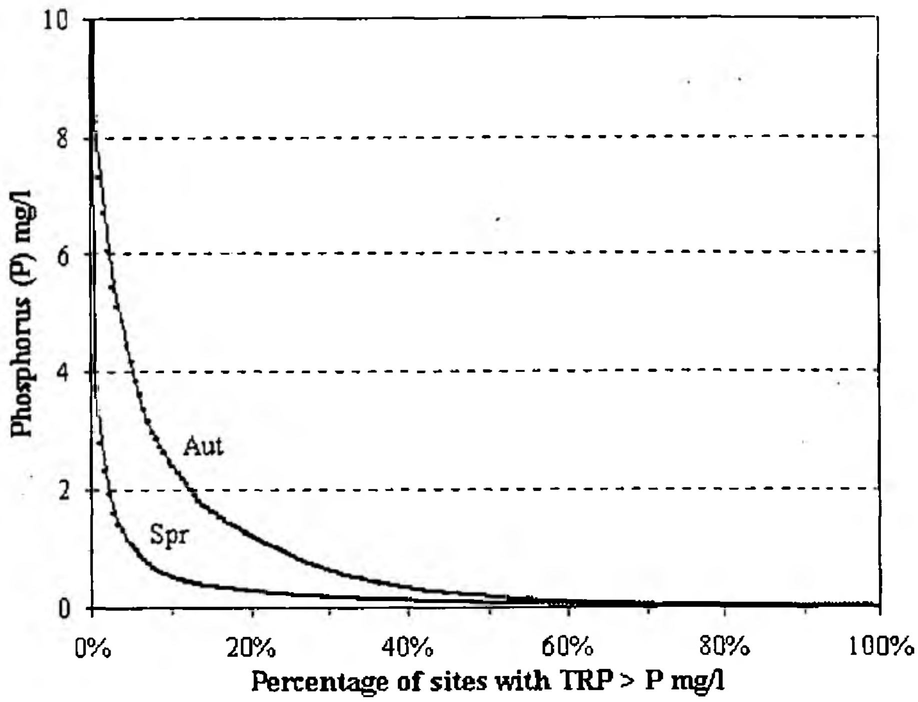

Figure 2.1 Distribution of Phosphorus (TRP) concentrations in spring and autumn 1995.

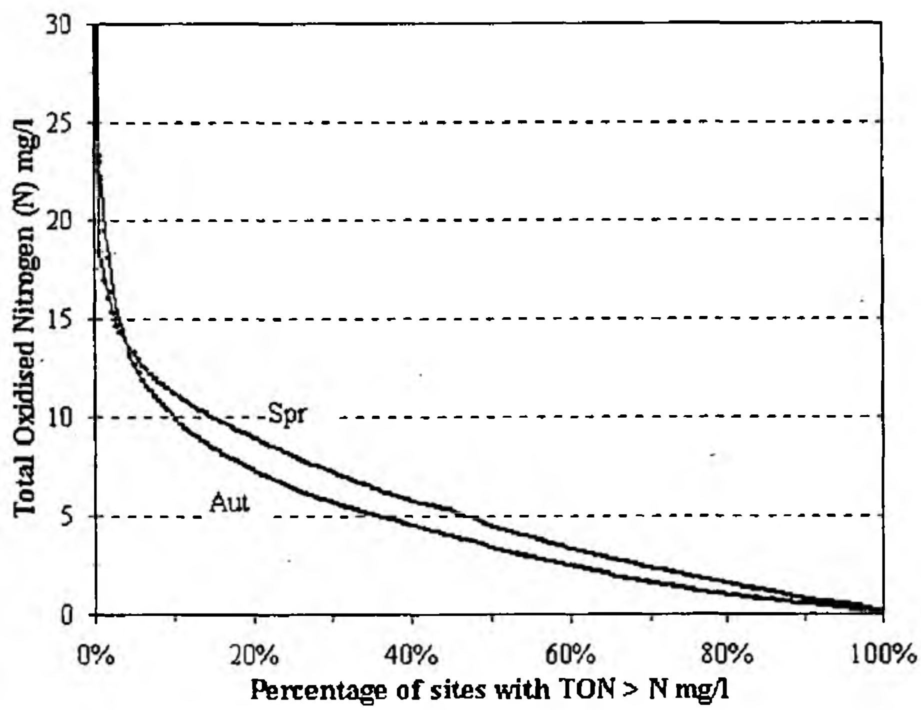

Figure 2.2 Distribution of Nitrogen (TON) concentrations in spring and autumn 1995.

Figure 2.1 reveals that TRP concentrations were noticeably higher on average in the autumn of 1995 than they were in the spring. In fact, the average concentration in autumn was 0.83 mg/l, whereas in spring it was only 0.26 mg/l. However, these high average values were partly due to the fact that a few sites had very high concentrations of TRP (2.0% of sites in spring and 12.5% in autumn had concentrations > 2.0 mg/l).

Figure 2.2 shows that the distributions of TON in the spring and autumn of 1995 were very similar, although in this case the concentrations in autumn were slightly lower than in spring. The average concentrations were 4.59 mg/l in spring and 5.39 mg/l in autumn.

In order to facilitate the MIR-max and BBN analyses, the TRP and TON data were banded into six discrete categories as indicated in Table 2.1.

Table 2.1 Discrete bands used for TRP and TON in the MIR-max and BBN models.

TRP			TON		
Band	Range (mg/l)	No. in Band	Band	Range (mg/l)	No. in Band
1	0 to <0.04	1418	1	0 to <1.5	1611
2	0.04 to <0.08	1394	2	1.5 to <3.0	1159
3	0.08 to <0.20	1263	3	3.0 to <5.0	1169
4	0.20 to <0.50	1104	4	5.0 to <7.5	1163
5	0.50 to <2.00	1018	5	7.5 to <12.0	1150
6	2.00 +	498	6	12.0 +	443

2.2 Geographical Distributions of TRP and TON.

Figures A-1 and A-2 in Appendix A show the spring and autumn distributions of sites in England and Wales which had TRP concentrations of less than 0.2 mg/l in 1995. Both of these figures show high numbers of sites with low TRP concentrations in the North West, South West and Welsh Regions.

Figures A-3 and A-4 in Appendix A show the spring and autumn distributions of sites in England and Wales which had TRP concentrations equal to or exceeding 0.2 mg/l in 1995. These show two main clusters of high TRP concentrations, one in Cheshire and the other in a band from Somerset to Cambridge. However, it is worth noting that many sites on canals, ditches, dykes, drains etc. in Anglian Region were deleted from the database because they were not considered to be running waters, hence the shortage of sites just south of the Wash. Figures A-3 and A-4 also show that there are many more high TRP sites in autumn than in spring, especially in South West, Midlands and North West Regions, and to a lesser degree in Anglian Region. The most noticeable changes from spring to autumn occurred in the North Wessex Area of South West Region and the Southern Area (mainly Cheshire) of North West Region.

Figure A-5 shows the combined spring and autumn distribution of sites at which the concentration of TON was less than 5 mg/l. Figure A-6 shows the distribution of sites at which the concentration of TON equalled or exceeded 5 mg/l. Once again, the sites with low concentration are predominantly in the South West, North West and Welsh Regions, with the exception of the North Wessex Area and Cheshire where high concentrations are common.

2.3 Distribution of Recorded Occurrences of *Cladophora*.

Data from the 1995 Survey of Perceived Stresses were used to derive the distribution of recorded cases of *Cladophora* with respect to: a) biological GQA class; and b) geographical location. However, Technical Report E126 (Martin and Walley, 2000) found that there were considerable inconsistencies in monitoring practices between Regions and even between Areas within Regions. In the case of *Cladophora*, several Areas did not record it, although other Areas within those same Regions did. Thus, the geographic distribution of *Cladophora* recorded by the 1995 Survey of Perceived Stresses (see Figure A-7 in Appendix A) was considered incomplete.

Table 2.2 gives the distribution of recorded cases of *Cladophora* with respect to biological GQA class. The national distribution of sites by biological GQA classes is also given to facilitate proper comparison of the relative frequencies of occurrence. The table shows that 'Light' and 'Moderate' intensities of *Cladophora* were quite common at Class 'a' sites, but that no 'Severe' cases were found. 'Light' intensities were most frequent in GQA class 'b' rivers, whereas 'Moderate' and 'Severe' intensities were most frequent in GQA class 'c' and 'd' rivers respectively.

Table 2.2 Distribution of occurrences of *Cladophora* by intensity level and biological GQA class.

Biological GQA Class	Intensity of *Cladophora*			Total Recorded	Distribution of GQA Sites Nationally
	Light	Moderate	Severe		
a	36	28	0	64	29.2%
b	55	88	5	148	28.9%
c	33	98	14	145	21.0%
d	9	29	26	64	10.6%
e	0	9	6	15	8.1%
f	0	0	0	0	2.2%
Total	133	252	51	436	100%

3. RELATIONSHIPS BETWEEN NUTRIENT LEVELS, CLADOPHORA AND DISSOLVED OXYGEN

3.1 Nutrient Levels and *Cladophora*

Despite the fact that the recorded distribution of *Cladophora* in 1995 was incomplete, the recorded cases did provide a subset of data that could be used to investigate possible relationships between the intensity level of *Cladophora* and the concentrations of TRP and TON. The results were very disappointing in that: a) only very weak correlations were found with TRP ($r = 0.158$, $p < 0.00005$) and TON ($r = 0.145$, $p < 0.00005$); and b) no threshold value for TRP or TON (i.e. below which *Cladophora* was absent) could be identified. Figure 2.3 (a - d) shows that *Cladophora* occurred over the whole range of TRP and TON concentrations down to 0.01 mg/l of TRP and 0.2 mg/l of TON. However, no 'Severe' cases occurred below 0.06 mg/l of TRP or 1.2 mg/l of TON.

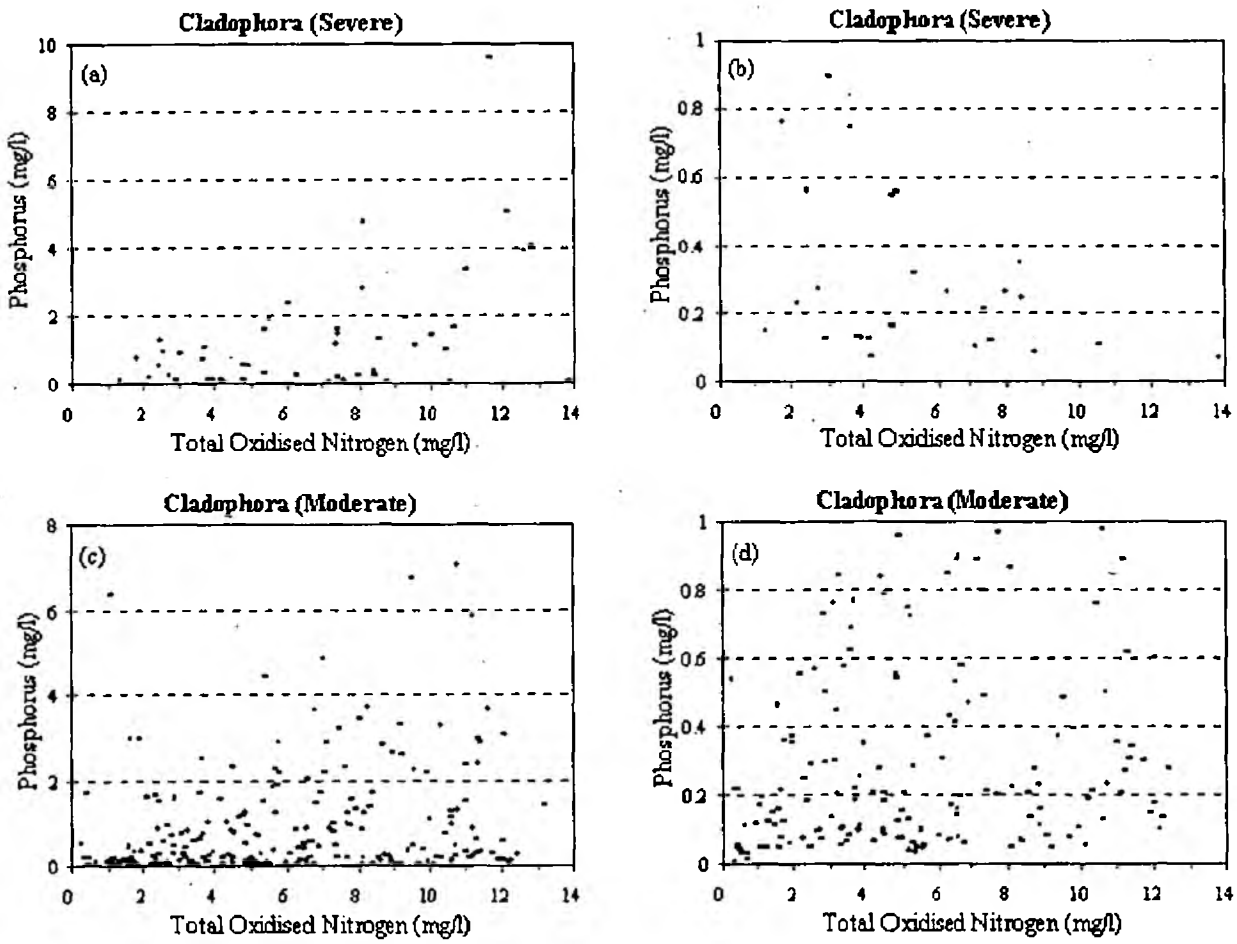

Figure 2.3 Graphs showing the TRP and TON concentrations at which *Cladophora* were recorded as 'Severe intensity' (graphs a and b) and 'Moderate intensity' (graphs c and d). Note that (a) and (b) are identical except for scale of their y-axes, and similarly for (c) and (d).

3.2 Nutrient Levels and Dissolved Oxygen

Several multiple linear regression (MLR) models were developed in which mean DO (% sat) was the dependent variable and TRP, TON, total ammoniacal nitrogen (AMTN), biochemical oxygen demand (BOD), water temperature (T) and Site Type (ST) were the independent variables. Spring, autumn and whole year models were developed using data from all GQA classes and using biological GQA classes 'a' and 'b' only. All analyses were performed using forward step-wise analysis. The resulting models and their correlation coefficients are given below, with the variables listed in the order that they were selected by the step-wise analysis.

Spring (All GQA Classes) $r = 0.4624$ $(p < 0.00005)$

$$DO = 99.60 - 0.95(BOD) - 3.16(TRP) - 1.87(ST) - 2.25(AMTN) + 0.39(TON) + 0.19(T)$$

Autumn (All GQA Classes) $r = 0.4899$ $(p < 0.00005)$

$$DO = 89.84 - 1.18(ST) - 0.95(BOD) - 1.95(AMTN) - 2.10(TRP) + 0.71(T) + 0.24(TON)$$

Whole Year (All GQA Classes) $r = 0.4651$ $(p < 0.00005)$

$$DO = 97.84 - 0.96(BOD) - 2.23(TRP) - 1.47(ST) - 2.13(AMTN) + 0.26(T) + 0.29(TON)$$

Spring (Classes 'a' & 'b' only) $r = 0.3970$ $(p < 0.00005)$

$$DO = 103.48 - 4.59(TRP) + 1.29(BOD) - 8.62(AMTN) - 1.51(ST) + 0.47(TON) - 0.19(T)$$

Autumn (Classes 'a' & 'b' only) $r = 0.4203$ $(p < 0.00005)$

$$DO = 89.5 - 1.78(TRP) + 0.81(T) - 1.44(BOD) - 8.33(AMTN) - 0.99(ST) + 0.32(TON)$$

Whole Year (Classes 'a' & 'b' only) $r = 0.3798$ $(p < 0.00005)$

$$DO = 99.74 - 9.33(AMTN) - 1.32(ST) - 1.28(BOD) - 1.85(TRP) + 0.38(TON) + 0.20(T)$$

In addition to these six-variable models, an equivalent set of four-variable models was developed, using BOD, AMTN, TRP and TON as the independent variables. The results were as follows:

Spring (All GQA Classes) $r = 0.4192$ $(p < 0.00005)$

$$DO = 98.26 - 1.18(BOD) - 2.99(TRP) - 2.39(AMTN) - 0.05(TON)$$

Autumn (All GQA Classes) $r = 0.4372$ $(p < 0.00005)$

$$DO = 98.06 - 1.06(BOD) - 2.04(AMTN) - 1.91(TRP) + 0.08(TON)$$

Whole Year (All GQA Classes) $r = 0.4203$ $(p < 0.00005)$

$$DO = 98.16 - 1.18(BOD) - 1.81(TRP) - 2.26(AMTN)$$

Spring (Classes 'a' & 'b' only) $r = 0.3535$ $(p < 0.00005)$

$$DO = 100.80 - 5.71(TRP) - 1.85(BOD) - 8.53(AMTN) + 0.07(TON)$$

Autumn (Classes 'a' & 'b' only) $r = 0.3397$ ($p < 0.00005$)

$DO = 100.62 - 1.52(TRP) - 9.05(AMTN) - 1.58(BOD)$

Whole Year (Classes 'a' & 'b' only) $r = 0.3346$ ($p < 0.00005$)

$DO = 100.88 - 10.15(AMTN) - 1.80(BOD) - 1.59(TRP)$

A summary of the TRP, AMTN and TON coefficients in these equations is given in Table 3.1.

Table 3.1 Summary of coefficients of TRP and AMTN in the MLR models of DO.

	Coefficients in MLR Models					
	GQA Classes 'a' & 'b'			All GQA Classes		
	TRP	AMTN	TON	TRP	AMTN	TON
Six-variable models						
Spring	-4.59	-8.62	+0.47	-3.16	-2.25	+0.39
Autumn	-1.78	-8.33	+0.32	-2.10	-1.95	+0.24
Whole Year	-1.85	-9.33	+0.38	-2.23	-2.13	+0.29
Average	*-2.74*	*-8.76*	*+0.39*	*-2.50*	*-2.11*	*+0.31*
Four-variable models						
Spring	-5.71	-8.53	+0.07	-2.99	-2.39	-0.05
Autumn	-1.52	-9.05	-	-1.91	-2.04	+0.08
Whole Year	-1.59	-10.15	-	-1.81	-2.26	-
Average	-2.94	-9.24	-	-2.24	-2.23	-

There are several points worth noting from Table 3.1, and the MLR models.

a) Although decreasing DO levels were associated with increasing TRP concentrations, no such relationship was found between DO and TON.
b) The inverse association between DO and TRP was stronger in spring than in autumn.
c) The models covering all GQA classes indicated that the association between DO and TRP was very similar in magnitude to that between DO and AMTN (i.e. in terms of percentage reduction in DO per mg/l of TRP or AMTN).
d) The models based on data from GQA classes 'a' and 'b' indicate that the association between DO and TRP is noticeably weaker than that between DO and AMTN. One extra mg/l of TRP was associated with a 3% fall in DO(%), whereas an extra mg/l of AMTN was associated with about a 9% fall in DO(%).
e) Overall, the reductions in DO varied from about 4 to 5 percent per mg/l of TRP in spring to about 1.5 to 2.0 percent per mg/l of TRP in autumn.

4. DEVELOPMENT AND TESTING OF TRP AND TON PREDICTIVE MODELS

In order to determine how predictable TRP and TON are from invertebrate and environmental data, various predictive models were developed, tested and evaluated. These included:

a) multiple linear regression models (MLR);
b) neural networks (multi-layered perceptrons);
c) MIR-max pattern recognition models; and
d) Bayesian Belief Networks (BBN).

The MLR models were developed to provide an initial screening of the data and a baseline performance against which to judge the other, non-linear, models. The following sub-sections give details of each model type and their test results.

4.1 Multiple Linear Regression

Multiple linear regression models were developed to predict TRP and TON from the 76 BMWP taxa and 10 environmental variables listed in Tables 1.2 and 1.3 respectively. The first models were based on all of the available data (i.e. whole year and all quality classes), but then separate models were developed for spring and autumn. Then, in order to separate the worst effects of organic pollution from those of eutrophication, a range of models were developed using data from quality classes 'a' and 'b' only. Finally, in order to check the importance of the environmental variables in the above models, spring and autumn models were developed using the 76 BMWP families only. The results of all the tests are given in Table 4.1.

Table 4.1 Results of the Multiple Linear Regression tests.

Dependent Variable	GQA Quality Classes	Season	Correlation Coefficient
Models using 76 taxa and 10 environmental variables as input			
TRP (conc.mg/l)	All ('a'-'f')	Year	0.4204
TRP (conc.mg/l)	All ('a'-'f')	Spring	0.3852
TRP (conc.mg/l)	All ('a'-'f')	Autumn	0.4596
TRP (conc.mg/l)	'a' and 'b'	Year	0.5029
TRP (conc.mg/l)	'a' and 'b'	Spring	0.5081
TRP (conc.mg/l)	'a' and 'b'	Autumn	0.5467
TON (conc.mg/l)	All ('a'-'f')	Year	0.5381
TON (conc.mg/l)	All ('a'-'f')	Spring	0.7019
TON (conc.mg/l)	All ('a'-'f')	Autumn	0.4452
TON (conc.mg/l)	'a' and 'b'	Year	0.6590
TON (conc.mg/l)	'a' and 'b'	Spring	0.7754
TON (conc.mg/l)	'a' and 'b'	Autumn	0.6319
Models using 76 BMWP taxa as independent input variables			
TRP (conc.mg/l)	'a' and 'b'	Spring	0.4934
TRP (conc.mg/l)	'a' and 'b'	Autumn	0.5354
TON (conc.mg/l)	'a' and 'b'	Spring	0.7407
TON (conc.mg/l)	'a' and 'b'	Autumn	0.5835

The results clearly show that:

a) models based upon data from GQA classes 'a' and 'b' performed better than those based on all GQA classes;
b) models based on 76 taxa and 10 environmental variables performed better than those based on just 76 taxa;
c) the TON models were noticeably better than TRP models;
d) the spring TON models were better than the autumn ones, whereas the converse was the case for the TRP models; and
e) the best correlation coefficients achieved were 0.7754 for TON in the spring and 0.5467 for TRP in the autumn.

It should be noted that all of the correlation coefficients quoted in Table 4.1 were based on dependent tests, and were significant at $p < 0.00005$, owing to the large sample sizes involved.

4.2 Neural Networks

The initial tests using multiple linear regression indicated that it might be necessary to test a wide range of models in order to properly assess the predictability of TRP and TON. Consequently, it was decided to use a multi-layered perceptron (a neural network based on supervised learning), because this offered a quick means of developing and testing a large number of non-linear models. Two series of tests were carried out. In the first series, the networks were trained and tested using the same data. Although this meant that the performance tests were based on dependent data, and not therefore a proper test of overall performance, it did enable a range of key factors to be investigated quickly in a like-for-like manner. In the second series, the most promising models from the first series were further investigated using cross-validation. That is, the data were randomly divided into two equal sets (y and z) so that the models could be trained on one and independently tested on the other, and vice versa. In all cases, the network used was a multi-layered perceptron with a single output variable (TRP or TON) and 87 input variables. The input variables included: the recorded abundance levels (absence being represented by zero) of the 76 BMWP families listed in Table 1.2; the number of families found (NFAM) in the sample; and the recorded values of the 10 environmental variables listed in Table 1.3. In order to avoid over parameterisation, the networks had just one hidden layer consisting of 16 nodes in all first-series networks, and eight nodes (TRP models) or 10 nodes (TON models) in the second series networks. The number of hidden nodes used was governed by the size of the training set. The training schedule used in the first series of tests was Save Best (100,000/10,000/10). This means that during training the network was tested (i.e. on the test data, which in this case was the same as the training data) every 10,000 cycles, and if a test showed improved performance the network was saved. Training was allowed to continue until 10 successive tests showed no improvement in performance or the total number of training cycles reached 100,000. In this particular case, training was bound to reach 100,000 cycles. The final network was the last one to be saved by this process. The training schedule used in the second series of tests was Save Best (200,000/4,000/20).

Tables 4.2 and 4.3 give the results of the first series of tests for TRP and TON respectively. Separate models were developed for spring and autumn, first using all of the data (i.e. all biological GQA classes) and then by using smaller subsets of the data. The subsets were first confined to GQA classes 'a' and 'b', and then further confined to recorded values of TRP or TON lying below given threshold concentrations.

The results of the second series of tests are given in Tables 4.4 and 4.5 for TRP and TON respectively. Since the models were cross-validated, there are two models for each test, one trained on data set 'y' and tested on data set 'z' (e.g. model psab05yz) and one trained on 'z' and tested on 'y' (e.g. model psab05zy). The main benefit of this procedure is that it ensures that each model is not over-trained on the training set, because the Save Best procedure stops the training when the network's performance on the independent test set begins to deteriorate. Once trained, the networks can be tested using the test and training sets separately, thus providing performance figures for independent and dependent data respectively. However, it is the performance on independent data that is the true test of the model. Tables 4.4 and 4.5 give the correlation coefficients achieved using both dependent and independent data. Since these varied slightly between the 'yz' and 'zy' models, owing to difference in the 'y' and 'z' data sets, the average was used as the overall measure of performance.

Table 4.2 Results of first series performance tests on neural network predictors of TRP. Input vector: 76 BMWP families, NFAM and 10 environmental variables.

	Input Data Specification				Correlation
Model	Concentration Threshold	GQA quality classes	Taxonomic data type	No. of cases	Coefficient (Dependent)
			Spring Models		
psafl6	None	All ('a'- 'f')	Abundance	3255	0.3210
psabpr16	None	'a' and 'b'	Present/Absent only	2127	0.4755
psab16	None	'a' and 'b'	Abundance	2127	0.4703
psabx216	< 2.0 mg/l	'a' and 'b'	Abundance	2119	0.5620
psabx116	< 1.0 mg/l	'a' and 'b'	Abundance	2092	0.6180
psab0516	< 0.5 mg/l	'a' and 'b'	Abundance	2013	0.6630
psab0216	< 0.2 mg/l	'a' and 'b'	Abundance	1671	0.6814
psab02r16*	< 0.2 mg/l	'a' and 'b'	Abundance with Replicates*	2887*	0.7698*
			Autumn Models		
paafl6	None	All ('a'- 'f')	Abundance	3440	0.4462
paabpr16	None	'a' and 'b'	Abundance	2186	0.5751
paab16	None	'a' and 'b'	Abundance	2186	0.5601
paabx216	< 2.0 mg/l	'a' and 'b'	Abundance	2025	0.6417
paabx116	< 1.0 mg/l	'a' and 'b'	Abundance	1817	0.6100
paab0516	< 0.5 mg/l	'a' and 'b'	Abundance	1589	0.5866
paab0216	< 0.2 mg/l	'a' and 'b'	Abundance	1238	0.5733

* In this test the input data were extended by replicating samples in data deficient areas of the original set. Samples in the ranges 0.2 – 0.5, 0.5 – 1.0, 1.0 – 2.0 were made to occur twice, four times and eight times respectively. The outcome of this test is discussed in Section 6.3.

Levels of significance were not derived for these results, because the neural network software did not automatically provide them. However, as with the results of every MLR model developed in this study they were clearly highly significant (owing to the large sample sizes), and their separate derivation could not be justified given the size of the contract. This statement applies to all of the neural network results given in Tables 4.2 to 4.10

Table 4.3 Results of first series performance tests on neural network predictors of TON. Input: 76 BMWP taxa, NFAM and 10 environmental variables.

Model	Input Data Specification				Correlation Coefficient (Dependent)
	Concentration Threshold	GQA quality classes	Taxonomic data type	No. of cases	
Spring Models					
nsaf16	None	All ('a'- 'f')	Abundance	3255	0.7110
nsabpr16	None	'a' and 'b'	Present/Absent only	2127	0.8232
nsab16	None	'a' and 'b'	Abundance	2127	0.8115
nsab1516	< 15 mg/l	'a' and 'b'	Abundance	2097	0.8228
nsab1016	< 10 mg/l	'a' and 'b'	Abundance	1877	0.8052
Autumn Models					
naaf16	None	All ('a'- 'f')	Abundance	3440	0.4052
naabpr16	None	'a' and 'b'	Present/Absent only	2186	0.6613
naab16	None	'a' and 'b'	Abundance	2186	0.6402
naab1516	< 15 mg/l	'a' and 'b'	Abundance	2171	0.6716
naab1016	< 10 mg/l	'a' and 'b'	Abundance	2086	0.7039

The main points to note from Tables 4.2 and 4.3 are that:

a) on a like-for-like basis the neural networks generally performed better than MLR models;

b) the performance of the models was generally improved by the removal of data with high concentrations of TRP or TON, and this resulted in the TRP models performing better in spring than autumn once samples with TRP ≥ 1.0 mg/l had been removed.

A fuller discussion of these results is given in Section 6.3

Tables 4.4 and 4.5 give the results of the second series of tests in which the models were cross-validated to permit both independent and dependent testing of selected models. The models used were those with TRP less than 0.5 mg/l and TON less than 10 mg/l.

Table 4.4 Results of second series performance tests on neural network predictors of TRP. Input: 76 BMWP taxa, NFAM and 10 environmental variables.

Model	Input Data Specification			Dependent Tests		Independent Tests	
	Conc. Threshold	GQA Classes	Taxonomic Data Type	Test Cases	Correl. Coeff.	Test Cases	Correl. Coeff.
Spring Models							
psab05yz	< 0.5 mg/l	'a' & 'b'	Abundance	1007	0.6867	1006	0.5965
psab05zy	< 0.5 mg/l	'a' & 'b'	Abundance	1006	0.6542	1007	0.6203
				Avg	***0.6705***	*Avg*	***0.6084***
psab05pyz	< 0.5 mg/l	'a' & 'b'	Present/Absent	1007	0.6528	1006	0.5852
psab05pzy	< 0.5 mg/l	'a' & 'b'	Present/Absent	1006	0.7047	1007	0.6082
				Avg	***0.6788***	*Avg*	***0.5967***
Autumn Models							
paab05yz	< 0.5 mg/l	'a' & 'b'	Abundance	794	0.5906	793	0.4223
paab05zy	< 0.5 mg/l	'a' & 'b'	Abundance	793	0.5236	794	0.5309
				Avg	***0.5571***	*Avg*	***0.4776***
paab05pyz	< 0.5 mg/l	'a' & 'b'	Present/Absent	794	0.6259	793	0.3999
paab05pzy	< 0.5 mg/l	'a' & 'b'	Present/Absent	793	0.5649	794	0.5047
				Avg	***0.5954***	*Avg*	***0.4524***

Table 4.5 Results of second series performance tests on neural network predictors of TON. Input: 76 BMWP taxa, NFAM and 10 environmental variables.

	Input Data Specification			Dependent Tests		Independent Tests	
Model	Conc. Threshold	GQA Classes	Taxonomic Data Type	Test Cases	Correl. Coeff.	Test Cases	Correl. Coeff.
			Spring Models				
nsab10yz	< 10 mg/l	'a' & 'b'	Abundance	939	0.8331	938	0.7791
nsab10zy	< 10 mg/l	'a' & 'b'	Abundance	938	0.8258	939	0.7933
				Avg	***0.8295***	*Avg*	***0.7862***
nsab10pyz	< 10 mg/l	'a' & 'b'	Present/Absent	939	0.8129	939	0.7620
nsab10pzy	< 10 mg/l	'a' & 'b'	Present/Absent	939	0.7942	939	0.7725
				Avg	***0.8036***	*Avg*	***0.7673***
			Autumn Models				
naab10yz	< 10 mg/l	'a' & 'b'	Abundance	1043	0.7088	1043	0.6812
naab10zy	< 10 mg/l	'a' & 'b'	Abundance	1043	0.7221	1043	0.6448
				Avg	***0.7155***	*Avg*	***0.6630***
naab10pyz	< 10 mg/l	'a' & 'b'	Present/Absent	1043	0.7088	1043	0.6812
naab10pzy	< 10 mg/l	'a' & 'b'	Present/Absent	1043	0.7122	1043	0.6199
				Avg	***0.7105***	*Avg*	***0.6506***

The main points to be drawn from these results are that:

a) once again the correlation coefficients for TON were noticeably greater than those for TRP;

b) the independent correlation coefficients are generally between 0.06 to 0.08 lower than the dependent values in the case of TRP, and between 0.04 and 0.06 below them in the case of TON;

c) the models based upon present / absent data sometimes performed better than the models based upon abundance data when tested dependently, but in all cases the abundance based models came out on top when tested independently.

4.3 MIR-max Pattern Recognition

MIR-max is based on unsupervised learning, which means that it simply recognises patterns in the data without reference to any particular 'target' variable that it aims to predict. In this respect it is like a child learning to recognise the faces of its parents and other relatives without knowing who they are. Only later does it learn to add the labels 'Mummy', 'Daddy' etc. In the same way, MIR-max models identify different patterns in the data and later it is found that these reflect certain states or conditions of the underlying system (e.g. different states of health of a river as represented by the concentrations of key pollutants, like ammonia, TRP and TON). A detailed description of the MIR-max based River Pollution Diagnostic System (RPDS) was given by Walley and O'Connor (2001) and O'Connor and Walley (2001), and details can also be found on the CIES web site (http://www.cies.staffs.ac.uk). Readers seeking a more detailed explanation of MIR-max models are directed to these sources. The results of tests on RPDS that were carried out during National R&D Project E1-056 (Walley *et al.*, 2002) are given in Table 4.6 for TRP and TON.

In this present project, separate spring and autumn MIR-max models were developed using two different input vectors. The first had 86 variables, consisting of the 76 BMWP taxa listed in Table 1.2 and 10 environmental variables listed in Table 1.3. The second had 40 variables, consisting of the BMWP families (NFAM) plus the 34 BMWP families and 5 environmental variables listed in Table 4.7. These biological and environmental variables had earlier been identified as the top 40 indicators of TRP, using impact analysis on one of our initial neural network predictors of TRP. Two sets of these models were developed, one based on the abundance data (absent being represented by zero) and one based on present / absent data (i.e. 1 or 0). Since MIR-max models are based upon discrete data, the environmental input variables were transformed into five discrete bands. All models had: a) 100 output classes located in a hexagonal output space having 127 grid locations; and b) were based on data from biological GQA classes 'a' and 'b' sites only.

Despite the discrete nature of the inputs to MIR-max models, the predicted values are continuous-valued variables. The results of the performance tests carried out on the MIR-max models are summarised in Table 4.8. Since the models were cross-validated using two equal subsets of the data ('x' and 'y' in this case), there are two sets of results for each test, one each for the 'xy' and 'yx' models. The average of the two correlation coefficients is given in bold type, since it represents the best measure of overall performance.

Table 4.6 Results of performance tests on RPDS for TRP and TON.

Variable	Quality Class	Season	Site Types	Correlation Coefficient (Dependent*)
TRP (conc)	All	Spring	All	0.480
	All	Autumn	All	0.488
	All	Year	All	0.529
TON (conc)	All	Spring	All	0.742
	All	Autumn	All	0.501
	All	Year	All	0.615

* The results of these tests are not fully dependent (in the sense of being developed and tested on the same target outputs), because MIR-max is based on unsupervised learning, so does not depend on targets. However, the same inputs were used for training and testing, so the tests were not independent either.

Table 4.7 List of the 40 variables used in the reduced input MIR-max model. The BMWP taxa are shown in normal type and the other variables in italics.

Number of Families	Ephemeridae	Gerridae	Phryganeidae
Planariidae	Nemouridae	Nepidae	Limnephilidae
Dendrocoelidae	Capniidae	Aphelocheiridae	Goeridae
Viviparidae	Perlodidae	Corixidae	Lepidostomatidae
Physidae	Perlidae	Gyrinidae	Sericostomatidae
Ancylidae	Chloroperlidae	Dryopidae	*Width*
Erpobdellidae	Coenagriidae	Elmidae	*Boulders*
Asellidae	Calopterygidae	Rhyacophilidae	*Sand*
Siphlonuridae	Cordulegasteridae	Polycentropidae	*Silt*
Heptageniidae	Libellulidae	Hydropsychidae	*Discharge*

Table 4.8 Results of cross-validated performance tests on MIR-max models with respect to continuous valued predictions of TRP & TON. The average correlation coefficients derived from models 'xy' and 'yx' are given in bold type.

Model	Season	Number of Taxa	Data Type	GQA Classes	Conc. Threshold	Correl. Coeff. (Independent) TRP	TON
s86Axy	Spring	86	Abund.	'a' & 'b'	None	0.5559	0.7616
s86Ayx	Spring	86	Abund.	'a' & 'b'	None	0.5468	0.6222
						0.5514	***0.6919***
a86Axy	Autumn	86	Abund.	'a' & 'b'	None	0.3871	0.5564
a86Ayx	Autumn	86	Abund.	'a' & 'b'	None	0.4049	0.5564
				Avg		***0.3960***	***0.5564***
s86Pxy	Spring	86	Pres/Abs	'a' & 'b'	None	0.5496	0.7248
s86Pyx	Spring	86	Pres/Abs	'a' & 'b'	None	0.6215	0.7303
				Avg		***0.5856***	***0.7276***
a86Pxy	Autumn	86	Pres/Abs	'a' & 'b'	None	0.3594	0.5292
a86Pyx	Autumn	86	Pres/Abs	'a' & 'b'	None	0.3580	0.4349
				Avg		***0.3587***	***0.4821***
s40Axy	Spring	40	Abund.	'a' & 'b'	None	0.5201	0.6870
s40Ayx	Spring	40	Abund.	'a' & 'b'	None	0.5912	0.6746
				Avg		***0.5557***	***0.6808***
a40Axy	Autumn	40	Abund.	'a' & 'b'	None	0.4192	0.4956
a40Ayx	Autumn	40	Abund.	'a' & 'b'	None	0.4041	0.5047
				Avg		***0.4117***	***0.5002***
s40Pxy	Spring	40	Pres/Abs	'a' & 'b'	None	0.5388	0.6729
s40Pyx	Spring	40	Pres/Abs	'a' & 'b'	None	0.6055	0.6385
				Avg		***0.5722***	***0.6557***
a40Pxy	Autumn	40	Pres/Abs	'a' & 'b'	None	0.3177	0.4635
a40Pyx	Autumn	40	Pres/Abs	'a' & 'b'	None	0.3390	0.5025
				Avg		***0.3284***	***0.4830***

The main points to note from these results are that:

a) despite being unsupervised models they confirm that the TON models perform better than the TRP models;

b) the present / absent models perform almost as well as the abundance models overall, but better in some cases;

c) the reduced input model based on abundance data performed better than the full model with respect to TRP but not with respect to TON.

The latter was almost certainly due to the fact that the reduced input vector was designed specifically for the prediction of TRP. In fact it was based on the results of impact tests on one of the initial neural network predictors of TRP in spring, thus it is likely that the reduced input model was biased towards predictions of this variable. This could also account for the relatively poor correlation coefficients produced by the reduced-input model for both TRP and TON in autumn.

It should be noted that the results of the independent tests (Table 4.8) cannot be compared on a like-for-like basis with those of the neural networks (Tables 4.4 and 4.5), because the latter were developed and tested using data with TRP < 0.5 mg/l and TON < 10 mg/l, whereas the former had no restrictions on these variables. However, if these independent test results are compared with the dependent test results derived from the equivalent neural networks (Tables 4.2 and 4.3), it is apparent that the TRP results for spring (0.5514) are extremely good, since they are better than the dependent test results (0.4703) from the neural network. This too reflects the bias of the input vector towards the prediction of TRP in spring. Unfortunately, the MIR-max models were developed at a relatively early stage in the project, and it was not possible in the time available to repeat these tests using a more balanced input vector for the reduced model. In fact, the task of deriving the best indicator taxa for TRP and TON, which is key to the optimisation of these models, proved to be more challenging than anticipated, as will be explained in Section 5.

4.4 Bayesian Belief Networks

The River Pollution Bayesian Belief Network (RPBBN) that was developed in R&D Project EI-056 (Walley *et al.*, 2002) was capable of predicting five key water quality variables, namely AMTN (total ammoniacal nitrogen), pH, DO(%), TRP and TON. The structure of the RPBBN model is shown in Figure B-1 in Appendix B. Table 4.9 gives the results of performance tests carried out on RPBBN with respect to its predictions of TRP and TON. The input data used to produce these results were the abundance levels of the 76 BMWP families, plus the season in which the sample was taken (spring or autumn) and the Site Type (as defined by Walley *et al.*, 2002)

Table 4.9 Results of dependent tests on RPBBN in terms of its ability to predict TRP and TON from biological samples (i.e. abundancies of 76 BMWP taxa), Site Type and season. The results are given for the year as a whole (i.e. spring and autumn results combined). Note: RPBBN used smoothed probabilities.

Variable	GQA Classes	Site Types	Threshold	Season	Dependent Corr. Coeff.
TRP (mg/l)	All	All	None	Whole Year	0.638
TON (mg/l)	All	All	None	Whole Year	0.803

For the purpose of this project, a dedicated Eutrophication Bayesian Belief Network (EBBN) was developed using cross validation to permit independent testing of the model. That is, the data were randomly divided into two equal subsets, S1 and S2, so that one each could be used to develop a model, leaving the other one free for use as an independent test set. The structure of the EBBN model is shown in Figure B-2 in Appendix B. The conditional probability matrices were based upon the raw conditional probabilities derived from each of the data subsets. Ideally, the raw conditional probability distributions should have been smoothed to remove the lumpiness that invariably occurs in such distributions, but time did not permit this. Past experience has shown that the smoothing of the distributions results in an increase in the correlation coefficients between the predicted and actual values of between 6 to 15 percent (Walley *et al.*, 2002). The results of the performance tests on EBBN are given in Table 4.10

Table 4.10 Results of dependent and independent tests on the dedicated Eutrophication BBN (EBBN) in terms of its ability to predict TRP and TON from biological samples (i.e. abundancies of 76 BMWP), Site Type and season. The results are given for the year as a whole (i.e. spring and autumn results combined). Note: EBBN used raw probabilities, not smoothed ones.

Variable	Model	GQA Classes	Site Types	Threshold	Season	Dependent Corr. Coeff.	Independent Corr. Ceoff.
TRP (mg/l)	S1/S2	'a' & 'b'	All	<0.5 mg/l	Spring	0.8518	0.5506
TRP (mg/l)	S2/S1	'a' & 'b'	All	<0.5 mg/l	Spring	0.8249	0.5834
					Average	***0.8384***	***0.5670***
TRP (mg/l)	S1/S2	'a' & 'b'	All	<0.5 mg/l	Autumn	0.7631	0.3287
TRP (mg/l)	S2/S1	'a' & 'b'	All	<0.5 mg/l	Autumn	0.8053	0.3526
					Average	***0.7842***	***0.3407***
TON (mg/l)	S1/S2	'a' & 'b'	All	<10 mg/l	Spring	0.9156	0.7236
TON (mg/l)	S2/S1	'a' & 'b'	All	<10 mg/l	Spring	0.9050	0.7325
					Average	***0.9103***	***0.7281***
TON (mg/l)	S1/S2	'a' & 'b'	All	<10 mg/l	Autumn	0.9243	0.5933
TON (mg/l)	S2/S1	'a' & 'b'	All	<10 mg/l	Autumn	0.9191	0.5827
					Average	***0.9217***	***0.5880***

The main points to note from these results are that:

a) the dependent correlation coefficients are easily the highest achieved by any model; and
b) the independent correlation coefficients are very noticeably lower than the corresponding dependent values, and between 6.8 and 11.3 percent lower that the like-for-like values achieved by the neural network (Tables 4.4 and 4.5)

These two effects were only to be expected, since the conditional probability matrices were not smoothed but based on raw probabilities derived from the data. Smoothing would have reduced the dependent correlation coefficients and increased the independent ones by 6 to 15 percent. It is, therefore, anticipated that had smoothing been applied, the overall performance of EBBN would have at least matched that of the corresponding neural networks.

5. IDENTIFICATION OF INDICATOR TAXA

Attempts were made to identify the principal indicator taxa for TRP and TON using the results from the MLR and neural network models. The effects of site type were removed by including the 10 environmental variables in the input vectors of the TRP and TON models, then the taxa were ranked in order of their impacts on the predictions. However, the resulting lists displayed considerable differences, both between models and between seasons.

5.1 Results from MLR Models

The results derived from the MLR models are given in Tables 5.1 and 5.2. The taxa listed are ranked in order of the absolute magnitude of the standardised regression coefficient (Beta). They have been separated into positive and negative indicators based on the sign (+ve or –ve) of their beta coefficients. All of those listed were significant at the $p < 0.005$ level. Some taxa contributed to both the spring and autumn models, but others only contributed to one of them. Where a taxon occurred in both the spring and autumn models its beta value in the table is based on the average of the two, but where it only occurred in one of the two models its beta value is based on that one alone. The table indicates which of the models (spring or autumn or both) each taxon occurred in.

Table 5.1 TRP Indicator Taxa derived from the spring (S) and autumn (A) MLR models based on 76 BMWP taxa and 10 environmental variables.

Positive Indicators of TRP			Negative Indicators of TRP		
Taxon	Season	Avg. Beta	Taxon	Season	Avg. Beta
Asellidae	S & A	0.1259	Baetidae	S	-0.1089
Calopterygidae	S & A	0.1239	Physidae	S & A	-0.0733
Coenagriidae	S & A	0.1141	Naucoridae	S	-0.0726
Aeshnidae	S	0.1030	Corixidae	A	-0.0653
Glossiphoniidae	S & A	0.0848	Polycentropidae	A	-0.0653
Ancylidae	S	0.0784	Rhyacophilidae	S	-0.0638
Hydropsychidae	S & A	0.0726	Goeridae	A	-0.0635
Platycnemidae	A	0.0644	Beraeidae	A	-0.0635
Oligochaeta	S	0.0615	Leuctridae	S	-0.0615
Sphaeriidae	A	0.0586	Valvatidae	S & A	-0.0609
Libellulidae	A	0.0564	Sericostomatidae	S & A	-0.0603
Lymnaeidae	S	0.0504	Haliplidae	S	-0.0596
Simuliidae	S	0.0411	Elmidae	A	-0.0577
Dendrocoelidae	S	0.0407	Sialidae	S	-0.0554
			Tipulidae	S	-0.0548
			Limnephilidae	S & A	-0.0516
			Nemouridae	A	-0.0515
			Ephemeridae	A	-0.0441
			Planariidae	S	-0.0437
			Aeshnidae	A	-0.0423
			Gerridae	S	-0.0389

Table 5.2 TON indicator taxa derived from the spring (S), autumn (A) MLR models based on 76 BMWP taxa and 10 environmental variables.

Positive Indicators of TON			Negative Indicators of TON		
Taxon	Season	Beta	Taxon	Season	Beta
Haliplidae	S	0.0774	Heptageniidae	S & A	-0.1009
Gammaridae	A	0.0713	Leuctridae	S & A	-0.0958
Sialidae	S	0.0713	Platycnemidae	S	-0.0911
Hydropsychidae	S & A	0.0678	Chloroperlidae	S	-0.0737
Ephemerellidae	S & A	0.0643	Nemouridae	S & A	-0.0728
Scirtidae	S	0.0592	Beraeidae	A	-0.0699
Coenagriidae	S	0.0566	Polycentropidae	A	-0.0689
Calopterygidae	A	0.0546	Chironomidae	S	-0.0668
Planorbidae	S & A	0.0523	Perlodidae	S & A	-0.0644
Asellidae	S & A	0.0517	Oligochaeta	S & A	-0.0613
Goeridae	S & A	0.0502	Leptoceridae	S & A	-0.0584
Erpobdellidae	S	0.0498	Leptophlebiidae	S	-0.0577
Ancylidae	S & A	0.0426	Gerridae	S & A	-0.0509
Tipulidae	S	0.0419	Naucoridae	S & A	-0.0483
Dendrocoelidae	A	0.0395	Hydrophilidae	A	-0.0422
Hydrometridae	A	0.0383	Sphaeriidae	S	-0.0402
Nepidae	A	0.0380	Lymnaeidae	A	-0.0400
Unionidae	S	0.0350	Libellulidae	S	-0.0383
			Planariidae	S & A	-0.0368
			Capniidae	A	-0.0361
			Aeshnidae	S & A	-0.0344
			Potamanthidae	S	-0.0305

The main feature to note from Table 5.1 is the three strong (i.e. Beta > 0.1) positive indicators of TRP that apply in both seasons, namely Asellidae, Calopterygidae and Coenagriidae. It is interesting to note that one of these, Asellidae, is a very commonly occurring taxon, whereas the other two are less common. Table 5.2 indicates that there are two fairly powerful negative indicators of TON that apply to both seasons, namely Heptageniidae and Leuctridae. Both tables list several taxa that were only found to be indicators of TRP and/or TON in just one of the two seasons.

5.2 Results from Neural Network Models

Once a neural network has been trained, the relative contribution that each input variable made to the final predictions was determined using a procedure called 'impact analysis'. In this procedure, each input variable in turn is disabled and the performance of the model determined at each step. The percentage reduction in the correlation coefficient below its baseline value (i.e. with no disabled inputs) is recorded for each disabled variable. The variables that cause the greatest percentage impact are clearly the most important input variables in the model. Some variables may have negative impacts, indicating that their removal from the model would improve its performance. Thus, impact analysis is generally used to optimise the input vector by the progressive removal of variables with very weak or negative impacts. This involves several cycles of the impact analysis procedure, which can be time-consuming if the network has a large number of input variables.

In this particular case, however, a single cycle of impact analysis was used to rank the taxa in terms of their importance as predictors of TRP and TON. The networks used for these analyses were:

a) the spring and autumn TRP models based on GQA classes 'a' and 'b' and a concentration threshold of 0.5 mg/l (i.e. models psab0516 and paab0516); and
b) the spring and autumn TON models based on GQA classes 'a' and 'b' and a concentration threshold of 10 mg/l (i.e. models nsab1016 and naab1016).

The results of the analyses are given in Table 5.3.

Table 5.3 Results of impact analyses showing the taxa listed in order of their maximum impacts on the TRP and TON models. Only taxa with a maximum impact greater than two percent are listed. Their impacts on the spring and autumn models are given, together with their average impact.

TRP – Percentage Impacts					TON – Percentage Impacts				
Taxon	Sprg	Aut	Avg	Max	Taxon	Sprg	Aut	Avg	Max
Perlidae	2.06	10.72	6.39	10.72	Gerridae	2.76	10.21	6.48	10.21
Potamanthidae	0.63	9.91	5.27	9.91	Leuctridae	8.37	1.94	5.15	8.37
Calopterygidae	1.43	6.85	4.14	6.85	Chloroperlidae	8.03	0.11	4.07	8.03
Nemouridae	1.22	6.55	3.89	6.55	Perlodidae	5.58	2.76	4.17	5.58
Planariidae	0.72	6.23	3.47	6.23	Taeniopterygidae	1.01	5.17	3.09	5.17
Chloroperlidae	6.14	2.54	4.34	6.14	Heptageniidae	4.86	1.88	3.37	4.86
Viviparidae	0.57	5.74	3.15	5.74	Platycnemidae	4.59	-0.13	2.23	4.59
Taeniopterygidae	0.22	5.71	2.97	5.71	Capniidae	4.34	2.32	3.33	4.34
Dryopidae	0.00	5.53	2.77	5.53	Beraeidae	1.31	4.33	2.82	4.33
Lepidostomatidae	1.11	4.90	3.00	4.90	Siphlonuridae	3.78	0.00	1.89	3.78
Libellulidae	-0.52	4.59	2.04	4.59	Libellulidae	3.25	1.37	2.31	3.25
Haliplidae	1.29	4.18	2.74	4.18	Nemouridae	3.25	1.25	2.25	3.25
Heptageniidae	2.98	4.17	3.57	4.17	Notonectidae	2.97	0.16	1.57	2.97
Goeridae	4.13	1.10	2.62	4.13	Perlidae	1.89	2.71	2.30	2.71
Aeshnidae	1.83	4.11	2.97	4.11	Aeshnidae	2.69	-0.11	1.29	2.69
Unionidae	-0.59	4.08	1.75	4.08	Polycentropidae	0.37	2.63	1.50	2.63
Gerridae	0.21	4.06	2.13	4.06	Gyrinidae	0.31	2.48	1.40	2.48
Leuctridae	3.83	0.11	1.97	3.83	Potamanthidae	2.19	2.31	2.25	2.31
Odontoceridae	-0.54	3.44	1.45	3.44	Lepidostomatidae	2.07	0.75	1.41	2.07
Aphelocheiridae	3.40	1.80	2.60	3.40					
Beraeidae	-0.08	3.26	1.59	3.26					
Neritidae	-0.34	3.14	1.40	3.14					
Brachycentridae	1.78	3.04	2.41	3.04					
Polycentropidae	1.00	2.87	1.93	2.87					
Hydrometridae	-0.48	2.66	1.09	2.66					
Perlodidae	0.55	2.48	1.51	2.48					
Cordulegasteridae	2.35	-1.41	0.47	2.35					
Rhyacophilidae	2.10	0.81	1.46	2.10					

The main points to be noted from these results are that:

a) there is considerable variability between the spring and autumn impacts of each taxon, some even producing negative impacts in one season and strong positive ones in the other;

b) the two most important indicators in the TRP model (i.e. Perlidae and Potamanthidae) were not even identified as important indicators by the MLR model;
c) only 11 of the 28 taxa identified as good (+ve or –ve) indicators of TRP by impact analysis also appeared in the list of 35 taxa identified by the MLR model;
d) the top four indicators of TON (i.e. Gerridae, Leuctridae, Chloroperlidae and Perlodidae) that were identified by impact analysis were also identified by MLR as good indicators of TON in both seasons; and
e) 11 of the 19 taxa identified as good (+ve or –ve) indicators of TON by impact analysis also appeared in the list of 41 taxa identified by the MLR model.

Thus, there was better agreement between the TON lists than between the TRP lists. However, the total number of taxa common to the two lists was partly governed by the decision to confine the impact analysis list to those taxa with greater than two percent impact.

Table 5.4 lists those taxa that impact analysis identified as good indicators (+ve or –ve) of both TRP and TON. The indicator ranking of each taxon for TRP and TON is given, together with its overall ranking as a joint indicator. It is worth noting that the top four places in the joint ranking list are taken by stoneflies, namely Chloroperlidae, Taeniopterygidae, Perlidae and Nemouridae. However, it is well known that these creatures are sensitive to several other river quality factors, so how reliable are these results? This issue will be discussed in the Section 6.

Table 5.4 The BMWP taxa that were found to be good indicators of both TRP and TON by impact analysis on neural network models.

Taxon	TRP Ranking	TON Ranking	Joint Ranking
Heptageniidae	13	6	6
Potamanthidae	2	18	7
Taeniopterygidae	8	5	2
Nemouridae	4	12	4
Leuctridae	18	2	8
Perlodidae	26	4	11
Perlidae	1	14	3
Chloroperlidae	6	3	1
Aeshnidae	15	15	13
Libellulidae	11	11	9
Gerridae	17	1	5
Polycentropidae	24	16	14
Beraeidae	21	9	12
Lepidostomatidae	10	19	10

5.3 Combined MLR and Neural Network Results

It was thought desirable to identify and rank those taxa that were found to be important indicators by both the MLR and neural network models. Thus, the taxa that were common to both lists were extracted and separately ranked according to their importance in each model. That is, they were ranked by the absolute value of their beta values (MLR models) and by their percentage impacts (neural network models). This was done separately for spring and autumn. The rankings derived for each model were then combined to produce overall rankings (based on the average of the two) for TRP and TON in spring and autumn. The results of this analysis are given in Table 5.5

Table 5.5 Positive and negative indicator taxa for TRP and TON in spring and autumn derived from a combination of multiple linear regression and neural network models. The rank order of the indicators (best = 1) in each season are given.

Indicator Taxon	TRP			TON		
	Type	Indicator Ranking		Type	Indicator Ranking	
		Spring	Autumn		Spring	Autumn
Planariidae	- ve	10		- ve	13	17
Dendrocoelidae	+ ve	11		+ ve		20
Valvatidae	- ve		9			
Lymnaeidae				- ve		18
Planorbidae				+ ve		15
Ancylidae				+ ve		19
Oligochaeta				- ve	9	9
Glossiphoniidae	+ ve		3			
Asellidae	+ ve	6				
Gammaridae				+ ve		4
Heptageniidae				- ve	2	7
Leptophlebiidae				- ve	10	
Ephemerellidae				+ ve		8
Potamanthidae				- ve	14	
Nemouridae	- ve		5	- ve	6	10
Leuctridae	- ve	3		- ve	1	1
Capniidae				- ve		11
Perlodidae				- ve	5	5
Chloroperlidae				- ve	3	
Platycnemidae				- ve	4	
Calopterygidae	+ ve	2	1	+ ve		16
Aeshnidae	+ ve	1		- ve	12	
Libellulidae	+ ve		6	- ve	7	
Gerridae				- ve	8	3
Naucoridae	- ve	7		- ve	15	13
Haliplidae	- ve	5				
Hydrophilidae				- ve		12
Rhyacophilidae	- ve	4				
Polycentropidae	- ve		2	- ve		6
Limnephilidae	- ve	9	10			
Beraeidae	- ve		4	- ve		2
Leptoceridae				- ve	11	
Goeridae	- ve		7	+ ve		14
Sericostomatidae	- ve	8	8			
Total Indicators		11	10		15	20

The main points to note from these results are that:

a) there are almost twice as many indicators listed for TON as for TRP, indicating better agreement between the MLR and neural network models for TON than TRP;
b) two of the highly ranked positive indicators of TRP are surprisingly quite rare (e.g. Aeshnidae and Libellulidae, with only 31 and 32 occurrences respectively), but since they are negative indicators of TON they appear to thrive in low TON / high TRP conditions (which are quite rare);
c) Calopterygidae, a fairly common taxon, appears to be the best indicator (+ve) of TRP;
d) Leuctridae was clearly identified as the best indicator (-ve) of TON; and
e) seven taxa were identified as positive indicators of TON, but only in the autumn, the highest ranked of which was Gammaridae.

5.4 Additional Information from Data Analysis

In view of the comments made about Aeshnidae and Libellulidae above, it was decided to determine the average TON/TRP ratios of GQA class 'a' and 'b' sites at which each BMWP taxon occurred. The results of this analysis are given in Tables C-1 to C-4 in Appendix C. In Table C-1 the results are listed in taxonomic order, whereas in Tables C-2, C-3 and C-4 they are listed in order of average TRP, average TON and the average TON/ average TRP ratio respectively. These tables appear to contain valuable information that could help experienced limnologists to interpret the lists of indicator taxa given above, and thereby provide new insights or pointers for further research. For example, Coenagriidae, Hydrometridae, Platycnemidae and Phryganeidae (all having over 120 occurrences) tend to occur at low TON / TRP ratios. Does this have any ecological significance or explanation?

5.5 Relationships between TRP, TON and NFAM

Figure 5.1 shows the number of BMWP families (NFAM) found in samples from GQA class 'a' or 'b' sites plotted against TRP and TON. Although both graphs show considerable scatter, due to other factors affecting NFAM, some tentative conclusions can be drawn based on trends in the average values. Increases in TRP appear to be associated with a slight decrease in NFAM. Initial increases in TON, up to about 5 mg/l, are associated with an increase in NFAM from about 20 to 25. This is followed by a levelling off and then a steady decline of about 5 families per 10 mg/l of TON when TON concentrations exceed 7.5 mg/l.

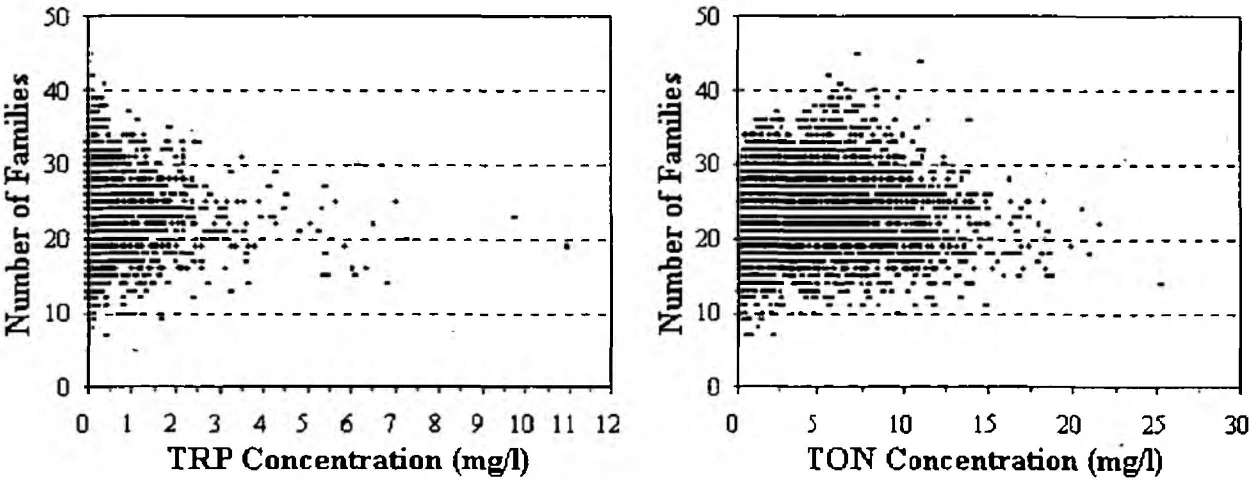

Figure 5.1 Graphs showing the effect TRP and TON on NFAM.

6. DISCUSSION AND SUMMARY OF RESULTS

6.1 Distribution Analyses

The results of the distribution analyses were discussed after presenting them in Sections 2.1 (Spring and Autumn Concentration of TRP and TON), 2.2 (Geographical Distributions of TRP and TON) and 2.3 (Distribution of Recorded Occurrences of *Cladophora*). The main findings were that:

- TRP concentrations were, on average, over three times higher in the autumn of 1995 than they were in the spring.
- TON concentrations were approximately the same in the spring and autumn of 1995.
- High concentration of TRP and TON were, as expected, confined to lowland regions, with noticeable clusters occurring in Cheshire and a band from Somerset to Cambridge.
- 'Severe' intensities of *Cladophora* occurred most frequently in GQA class 'd' rivers, whereas 'Moderate' and 'Light' intensities occurred most frequently in GQA classes 'c' and 'b' respectively.
- 'Light' and 'Moderate' intensities of *Cladophora* were quite common in GQA class 'a' rivers.

It is also worth noting that the clusters of high TRP and TON mentioned above approximately correspond to the location of two important aquifers, the Bunter Sandstone in Cheshire and the northern edge of the Chalk from Swindon to south of Cambridge.

6.2 Relationships between Nutrients, *Cladophora* and Dissolved Oxygen

The results of these analyses were briefly discussed after presenting them in Sections 3.1 (Nutrient Levels and *Cladophora*) and 3.2 (Nutrient Levels and Dissolved Oxygen). The main findings are listed below.

- Threshold concentrations of TRP and TON could not be identified in relation to the occurrence of *Cladophora*.
- No 'Severe' intensities of *Cladophora* occurred below 0.06 mg/l of TRP or 1.2 mg/l of TON.
- MLR models for the prediction of DO indicated that TRP is inversely associated with DO(%), but TON is not.
- The MLR models for GQA class 'a' and 'b' sites indicate that the inverse association between DO(%) and TRP is about one-third the strength of that between DO(%) and AMTN (total ammoniacal nitrogen). One extra mg/l of TRP is associated with a reduction in DO(%) of about 3%, whereas an extra mg/l of AMTN is associated with a reduction of about 9%.

6.3 TRP and TON Predictive Models

The development of models to predict TRP and TON from macroinvertebrate and environmental data form the main part of the project. Four different modelling techniques were used: MLR; neural networks; MIR-max pattern recognition; and BBN. In view of the short duration of the project, the majority of the work was based upon neural networks because these provided the quickest means of developing appropriately complex non-linear models. The MLR models were developed to provide an initial set of baseline results against which to compare the later models. Although independent tests provide the only true measure of the overall performance of a model, they are more demanding of data and time consuming to perform. In view of this, and the fact that most of the tests were carried out to provide like-for-like comparisons and not absolute measures of performance, the majority of the tests were based on dependent tests (i.e. the training data). Thus, the results contain a mixture of dependent and independent performance figures, so one must be aware of which is being referred to. The independent performance figures are always lower than the dependent ones, but they provide the only true measure of performance. This section aims to provide a concise summary of the results of the tests on all four model types. Brief comments on the results of the individual model types are given in Sections 4.1 (Multiple Linear Regression), 4.2 (Neural Networks), 4.3 (MIR-max Pattern Recognition) and 4.4 (Bayesian Belief Networks). The following bulleted sections highlight the conclusions that can be drawn from the tests on TRP and TON models.

- The first and most important conclusion to be drawn is that TON is far more predictable from macroinvertebrate and environmental data than is TRP. Table 6.1 gives the maximum dependent and independent correlation coefficients achieved by any of the predictive models of TRP and TON.

Table 6.1 Maximum dependent and independent correlation coefficients achieved by any of the TRP and TON models.

	Maximum Dependent Correlation Coefficient		Maximum Independent Correlation Coefficient	
Model Type	Spring	Autumn	Spring	Autumn
TRP	0.8384	0.7842	0.6084	0.4476
TON	0.9103	0.9217	0.7862	0.6630

In order to put these figures into perspective, it is worth noting that performance tests on neural network predictors of ASPT and NFAM, that were trained and tested on cross-validated subsets of the RIVPACS database, produced independent correlation coefficients of 0.8261 and 0.5860 for ASPT and NFAM respectively (Walley and Fontama, 1998). It is not possible to quote corresponding figures for the RIVPACS model itself, because it was not developed using cross-validated data. However, Walley and Fontama (1998) did demonstrate that the neural networks performed marginally better than RIVPACS III when tested dependently. Thus, it can be reasonably assumed that these correlation coefficients are a fair reflection of the performance of RIVPACS III on independent data.

It appears, therefore, that TON can be predicted from macroinvertebrate and environmental data to a degree of reliability almost as good as that of RIVPACS III predictions of ASPT. However, the best of the TRP models could only just match the reliability of RIVPACS III predictions of NFAM. In both cases, the spring models perform better than the autumn models when tested independently. Since Walley and Fontama (1998) concluded that the poor reliability of NFAM predictions cast doubt on its suitability for use in water quality classification, we feel that a similar conclusion must be drawn in relation to the TRP. That is, we conclude that TRP cannot be predicted to an acceptable degree of reliability from the 76 BMWP families and 10 environmental variables used in this study. It may, however, be possible to achieve an acceptable degree of reliability using macroinvertebrate data to species level, or by including appropriate macrophyte data in the input vector.

- The results of tests carried out on models based on abundance data and then present / absent data indicate that the abundance-based models performed only marginally better than present / absent models. This was contrary to expectations and previous experience, and the authors would wish to carry out a more extensive study of this issue before confirming this initial finding.

- The issue of whether the relationships between nutrient concentrations and the composition of macroinvertebrate communities is strongest in the cleaner rivers was investigated by comparing the performances of several dependent models, each based on different subsets of the input data.

  The results demonstrated that models based on data from biological GQA classes 'a' and 'b' performed noticeably better than those based upon all GQA classes (See Tables 4.1, 4.2 and 4.3). This was as expected, because the removal of GQA classes 'c' to 'f' significantly reduced the confounding effect of organic pollution. However, attempts to achieve further improvements by removing samples having high concentration of TRP or TON did not produce a clear-cut conclusion.

  TRP Models

  Table 4.2 gives the results of tests on neural network models that were based on subsets of the GQA classes 'a' to 'b' data in which samples having TRP concentration greater than 2.0, 1.0, 0.5 and 0.2 mg/l were progressively removed. The results for the spring models show that the correlation coefficient progressively increased from 0.562 to 0.681 as the threshold for exclusion was lowered, but those for the autumn models decreased from 0.642 to 0.573. Factors that may contribute to this confusing behaviour are:

  a) the fact that average TRP levels are three times as great in the autumn as they are in the spring, thus the autumn data have many more samples above the stated thresholds; and
  b) many models, including neural networks, are unable to adequately model data that are heavily skewed, because during their training they see very few cases from the sparse end of the distribution.

  The autumn model had many more samples with TRP greater than 0.2 mg/l so was able to model high TRP cases better than the spring model. Thus, the autumn models with thresholds of 'None' and 2.0 mg/l produced higher correlation coefficients (0.560 and 0.642) than did their corresponding spring models (0.470 and 0.562). However, once these high TRP samples had been removed, the spring models out-performed the autumn models, implying that the spring data had a higher information content. It must also be remembered that the autumn data set was much depleted by the removal of the high TRP samples.

  Techniques that aim to overcome the problems of highly skewed data include: a) the logarithmic (or similar) transformation of the data; and b) the replication of rare cases in the training data set. A series of tests were carried out in which neural network models were trained using log-transformed spring TRP data with concentration thresholds ranging from 'None' down to 0.2 mg/l. The correlation coefficients between their predicted and actual TRP values were all less than those of their non-transformed counterparts by between 2 and 5 percent. A neural network based on a training set that included replication of rare cases (see footnote to Table 4.2 for details) produced a significant improvement in the correlation coefficient compared to its non-replicated counterpart (i.e. 0.770 compared with 0.681). However, the validity of this approach is questionable, especially since the training data contained some 1200 repeat cases of high TRP samples. Since some samples were repeated eight times, it is possible that the neural network was over-trained on these cases.

TON Models

Table 4.3 gives the results of tests on neural network models that were based on subsets of the GQA classes 'a' to 'b' data in which samples having TON concentration greater than 15 and 10 mg/l were progressively removed. The test results showed that in spring the model with a 15 mg/l threshold gave the best overall correlation, whereas in autumn it was the 10 mg/l model that gave the best result. However, the variations from the no-threshold case were less than achieved by the corresponding TRP models. This was probably due to the fact that the TON distributions were less skewed than the TRP distributions (See Figures 2.1 and 2.2).

Summary

The tests on both the MLR and neural network models confirmed the belief that correlation between nutrient concentrations (TRP and TON) and the composition of the macroinvertebrate community is strongest for good quality rivers (GQA classes 'a' and 'b'). The spring TRP models produced results that appeared to confirm the finding of McGarrigle (1998) that the relationship breaks down at higher TRP concentration, but the autumn models contradicted this. In addition, the authors believe that the behaviour of the spring models was largely due to the highly skewed nature of the spring distribution of TRP values (see Figure 2.1). Thus, we are not able to confirm or reject the suggestion that the relationship breaks down at TRP concentration above 0.2 mg/l. Further detailed analyses may have provided a more conclusive answer to this question, but time and funding constraints did not permit this.

6.4 Identification of Indicator Taxa

The results of the tests carried out to identify indicator taxa for TRP and TON were not particularly conclusive, because there were noticeable differences between the taxa identified by the MLR models and the neural network impact analyses. These analyses are described and briefly discussed in Sections 5.1 (MLR models), 5.2 (Neural network impact analyses) and 5.3 (Combined MLR and neural network results). On reflection, it is not surprising that there were such differences since the MLR models were linear, whereas the neural networks were non-linear and hence more able to model the non-linear features of the data. However, neural networks do have some weaknesses, including a tendency not to adequately represent highly skewed data, as mentioned earlier in relation to the autumn distribution of TRP. In addition, some of the input variables were highly skewed, and may have benefited from logarithmic transformation. Despite these weaknesses, the neural networks probably still gave the more reliable results. Nevertheless, there is considerable scope for further improvement and extension of these analyses, which unfortunately we were not able to carry through in the short timeframe of the project.

Despite these comments, some tentative conclusions can be drawn and there are several aspects of the results that are worthy of further comment. When reading through these conclusions and comments, remember that they refer to models based on abundance data, so much of the indicator value of a taxon may derive from differences between abundance levels, not just between presence and absence.

- All seven BMWP families of PLECOPTERA (Stoneflies) occurred in one or more of the lists of indicator taxa. All were found to be negative indicators of TRP and/or TON. That is, they all tended to occur at low concentrations of TRP and TON. This was not simply due to the fact they tend to prefer upland sites, where TRP and TON concentrations just happen to be lower than average. Site characteristics were well represented in all models by the10 environmental variables, including altitude, slope and the nature of the substrate, so site type was effectively accounted for in the models. It is also worth noting that all of the families had high TON/TRP ratios (See Table C-4 in Appendix C), thus indicating that they are more indicative of low TRP concentrations than low TON concentrations.
- All of the eight BMWP families of EPHEMEROPTERA (Mayflies), except Caenidae, occurred in one or more of the lists of indicator taxa. All were found to be negative indicators of TRP and/or TON, except Ephemerellidae which was found by MLR to be a fairly strong positive indicator of TON in both spring and autumn. Further investigation of this finding revealed that Ephemerellidae is not so much an indicator of high TON concentrations as high TON/TRP ratios, thus its value appears to be as an indicator of relatively high TON at sites having relatively low TRP concentrations.
- 11 of the 16 BMWP families of TRICHOPTERA (Caddisflies) occurred in one or more of the lists of indicator taxa. These included three of the five caseless caddisflies (Rhyacophilidae, Polycentropidae and Hydropsychidae) and eight of the cased caddisflies (Limnephilidae, Beraeidae, Odontoceridae, Leptoceridae, Goeridae, Lepidostomatidae, Brachycentridae and Sericostomatidae). All, except Hydropsychidae and Goeridae, were found to be negative indicators of TRP and/or TON. Hydropsychidae was found by MLR to be a fairly strong positive indicator of TON in both spring and autumn. Goeridae is unusual in that it was found by MLR to be a fairly good positive indicator of TON in both spring and autumn, but also a fairly strong negative indicator of TRP in autumn. It is a little surprising therefore that its TON/TRP ratio (20.5), although higher that the median for all 76 taxa (13.8), is not noticeably different from those of the other caddisflies. It is also interesting to note the very wide range of TON/TRP ratios in which caddisflies occur, from 4.8 for Phryganeidae to 49.1 for Philopotamidae. Phryganeidae appears to be very tolerant of high TRP conditions (average TRP = 1.10 mg/l), whereas Philopotamidae appears to be very intolerant of high TRP conditions (average TRP = 0.04 mg/l), yet neither of these two families has been identified as an indicator of TRP. Since this seems inexplicable, at least from this rather simple evidence, there is clearly a need for further investigation of this behaviour.
- All six BMWP families of ODONATA (Damselflies and Dragonflies) occurred in one or more of the lists of indicator taxa, although one, Cordulegasteridae, was only identified as a relatively weak indicator of spring TRP by the neural network impact analyses. Two of the damselflies, Calopterygidae and Coenagriidae, were identified by MLR as strong positive indicators of TRP in both spring and autumn, but the impact analyses only

identified Calopterygidae. This could be due to the fact that if two strong indicators are highly correlated, a neural network may give higher preference to one than the other, thus an impact analysis only identifies the favoured one as a key indicator. The dragonfly Aeshnidae was also identified by MLR as a strong positive indicator of TRP in spring, but also as a weak negative indicator of TRP in autumn. However, Aeshnidae only occurred 31 times in the whole database so the validity of this rather strange result is questionable. The damselfly Platycnemidae and the dragonfly Libellulidae were identified as positive indicators of TRP in autumn, but, like Aeshnidae, Libellulidae occurred very rarely (32 times). However, all of the ODONATA, except Cordulegasteridae, occurred at sites having very high average TRP values and very low TON/TRP ratios (See Table C-2 and C-4 in Appendix C), thus confirming their roles as positive indicators of TRP. Cordulegasteridae, on the other hand, occurred at sites having relatively low TRP and TON concentrations and moderately high TON/TRP ratios. The indicator values of ODONATA with respect to TON were not found to be as strong or conclusive as those for TRP. Coenagriidae and Calopterygidae were again identified by MLR as positive indicators, but only for one season (spring and autumn respectively). The less frequently occurring families, Platycnemidae (120 occurrences) , Libellulidae and Aeshnidae were identified as negative indicators of TON, the latter in both spring and autumn and the other two in spring only. However, Tables C-3 and C-4 reveal that these three families occurred at moderately high TON concentrations but very low TON/TRP ratios. Thus their value as indicators seems to apply only to high TRP sites, where their presence implies lower than average TON values for such sites.

- All seven of the BMWP families of HEMIPTERA (Bugs) occurred in one or more of the lists of indicator taxa. Gerridae and Naucoridae were identified by MLR as good negative indicators of TON in both spring and autumn. The importance of Gerridae was confirmed by the fact that impact analysis identified it as the strongest indicator of TON, but Naucoridae did not even appear in the list of indicators derived by this means. Gerridae and Naucoridae were also identified by MLR as negative indicators of TRP, but only in the spring. Impact analysis confirmed that Gerridae is a good indicator of TRP, but mainly in the autumn. Corixidae was found by MLR to be a strong negative indicator of TRP, but only in the autumn. The listings of the other four families were not particular noteworthy. Hydrometridae and Nepidae were identified by MLR as weak positive indicators of TON in the autumn. Impact analysis identified Aphelocheiridae and Hydrometridae as weak indicators of TRP in the spring and autumn respectively, and Notonectidae as a weak indicator of TON in the spring. Evidence from Tables C-2 and C-3 confirms that Hydrometridae and Notonectidae are positive indicators of TRP and TON respectively.

- Five of the seven BMWP families of COLEOPTERA (Beetles) occurred in one or more of the lists of indicator taxa, the most significant of which was Haliplidae. This family was identified by MLR as the strongest positive indicator of TON and a good negative indicator of TRP, but only in the spring in both cases. Impact analysis also identified Haliplidae as a good indicator of TRP, but especially in the autumn. The other cases are less noteworthy. MLR identified Scirtidae as a good positive indicator of TON in spring, Hydrophilidae as a negative indicator of TON in autumn and Elmidae as a negative indicator of TRP in autumn. However, there is little evidence to support any of these assertions in Tables C-2 and C-3.

- Nine of the ten BMWP families of Mollusca (Snails, Limpets and Mussels) occurred in one or more of the lists of indicator taxa. However, only one family (Unionidae) was identified by both MLR and impact analysis, but for different reasons. MLP identified Unionidae as a weak positive indicator of TON in spring, whereas impact analysis identified it as a positive indicator of TRP in autumn. Thus, in this case, there is no agreement between MLP and impact analysis. The most significant indicators identified by MLR were Physidae and Valvatidae which were found to be fairly strong negative indicators of TRP in both spring and autumn, and Planorbidae and Ancylidae which were found to be positive indicators of TON in both spring and autumn. There is little evidence in Tables C-2 and C-3 to support these assertions, but a plot of TRP against abundance of Physidae showed that TRP declines with increased abundance, thus confirming that Physidae is a negative indicator of TRP. All of the other families were found to be relatively weak indicators of TRP or TON in one season only.

- The most significant findings with respect to the remaining BMWP families where that:
  a) Asellidae was identified by MLR as the strongest positive indicator of TRP and a positive indicator of TON in spring and autumn in both cases, but these findings were not confirmed by the results of impact analysis; and
  b) Glossiphoniidae was identified by MLR to be a fairly strong positive indicator of TRP in both spring and autumn, but this too was not confirmed by impact analysis.

It is clear from the above that in some cases there was good agreement between the indicators identified by MLR and impact analysis, but in other cases there was very little agreement. We believe that the truth lies somewhere between the two. However, we also believe that it is possible, with hindsight, to develop improved neural network predictors of TRP and TON and a better method of identifying the key indicators from them. Preliminary tests on a modified version of impact analysis have produced some promising results, but we need to develop software to automate the process so that it can be comprehensively tested.

6.5 General Comments on the Approach and Methods Used

This project aimed to answer several basic questions about the relationships between macroinvertebrates and nutrient concentrations in rivers using AI techniques. A general principle underlying the AI approach to the development of 'expert systems' is to start by observing the way in which the relevant experts perform their tasks and then to encode this behaviour into a software system. Thus, the developer has to first gain a sound understanding of the field in question, its essential features and relationships, and the way in which the experts use these to draw conclusions. The developer may enhance these finding, and hence the performance of the final system, by deriving features and relationships from relevant databases, if they exist. In this particular project, the subject under investigation was one in which even the 'experts' have very limited knowledge of the precise relationships between the variables, hence its focus on database analysis. However, the limited time and resources available for the project meant that it was not possible to complete the essential groundwork, like the identification of key indicators of TRP and TON, before work had to commence on the pattern recognition (MIR-max) and plausible reasoning (BBN) models. Consequently, the development of these models was not complete at the time of testing. For example, it was not possible to design the input vectors to the MIR-max and EBBN models specifically for the prediction of TRP and TON, because at that stage the key indicator taxa for each were not known. In addition, there was not time to smooth the 76 conditional probability matrices of the EBBN model, thus it was based on the raw conditional probabilities. Although this resulted in it producing the best dependent results of all the models tested, it undermined its performance on independent data, as explained in Section 4.4.

The bulk of the work was carried out using another AI technique, a supervised neural network known as the multi-layered perceptron. This provided the quickest means of developing a range of complex non-linear models that would enable us to perform the required investigations. In fact, these models produced the best overall performances on independent data, but this might not have been the case had time permitted the full development of the MIR-max and EBBN models. Nevertheless, supervised neural networks are very powerful modelling tools, and the authors have made extensive use of them in the past. However, they are 'black-box' models and it is difficult for the user to see how or why they produce certain results or conclusions. Care is also required in their uses, especially when modelling highly skewed data, as explained earlier in Section 6.3. BBN and MIR-max models handle skewed data in a far more satisfactory way, and are far more transparent in their operation. In addition, MIR-max models a powerful data visualisation facility that often enables users to gain greater insight into the behaviour of the real system (e.g. river ecology).

7. Conclusions and Recommendations

7.1 Conclusions

- The relationship between the composition of the macroinvertebrate communities and nutrient (TRP and TON) concentrations was found to be stronger for good quality rivers (biological GQA classes 'a' and 'b') than for rivers covering all quality classes (GQA classes 'a'-'f').
- The relationship was found to be noticeably stronger for TON than TRP, and it was concluded that TON could be predicted from environmental and macroinvertebrate family-based data to an acceptable degree of reliability, but that TRP could not.
- The models based on present / absent data performed almost as well as those based on abundance data.
- There was some evidence to support the belief that the relationship between macro-invertebrates and TRP breaks down at the concentrations above 0.2 mg/l TRP, but this was not conclusive because the effect could have been due to the highly skewed nature of the distribution of TRP.
- A series of MLR models of DO based on spring, autumn and annual data consistently indicated that TRP does have a negative effect (most probably indirectly) on DO (% sat).
- No relationship could be identified between recorded occurrences of *Cladophora* and concentrations of TRP and TON.
- Increasing concentrations of TRP were found to have a slight but continuous negative impact (probably indirectly) on NFAM, whereas increasing concentrations of TON had a noticeable positive impact initially followed by a noticeable negative impact at higher concentrations.
- An analysis of the average TRP and TON concentrations of the sites at which each BMWP family occurred showed wide variations in average TRP values and TON/TRP ratios, but far less variation in average TON values.
- Lists of indicator taxa for TRP and TON were derived but these were not as conclusive as hoped owing to relatively poor agreement between the two methods used, but some tentative conclusions have been drawn.

7.2 Recommendations

Our main recommendation is that:

- A more substantial study should be carried out to explore in depth the issues addressed in this preliminary study, and to complete the development and testing of the AI models.

We recommend that the following should be included in the proposed project.

- Models should be developed and tested using an enhanced database having macro-invertebrate, macrophyte and algae data, preferably with some taxa identified to species or genera level. This is particularly important in the case of the TRP models, where there is clearly a need for further improvement in performance.

- Further analyses should be carried out to identify key indicator taxa. These should include neural network impact tests based on the modified procedure mentioned earlier and analyses based on information theory. This is an important prerequisite to the completion of the MIR-max and BBN models.

8. Acknowledgement

The authors wish to thank our Project Manager from the Environment Agency, Dr. Rachael Dils, for all the help and guidance that she has given us throughout the project.

9. References

Martin R. W. and Walley W. J. (2000) Distribution of perceived stresses in English and Welsh rivers based on the 1995 survey. R&D Technical Report E126. Environment Agency, Bristol.

McGarrigle M. L. (1998) Impact of eutrophication on Irish river water quality. In: Wilson J. G. (ed). Eutrophication in Irish waters, 82-90. Dublin, Royal Irish Academy.

O'Connor M. A. and 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.

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 Research.* **32**, 3. 613-622.

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 O'Connor M. A. (2001) Unsupervised pattern recognition for the interpretation of ecological data. *J. of Ecological Modelling* **146**, 1-3. 219-230.

Walley W. J. O'Connor M. A., Trigg D. J. and Martin R. W. (2002) 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.

Appendix A

Geographical Distributions

of

TRP, TON and *Cladophora*

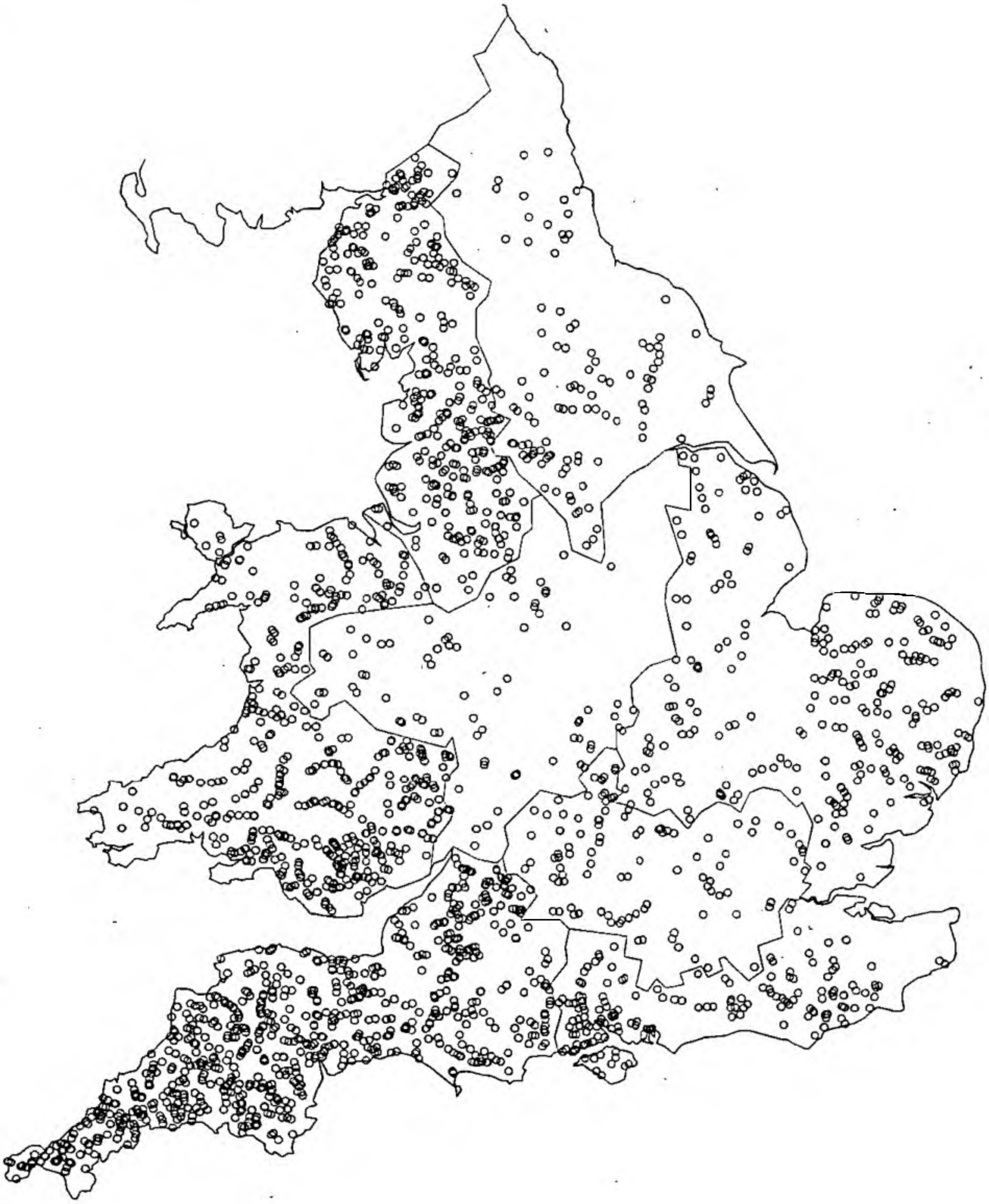

Figure A-1. Spring distribution of Biological Monitoring Sites in England and Wales at which the concentration of TRP was less than 0.2 mg/l over the three months preceding biological sampling.

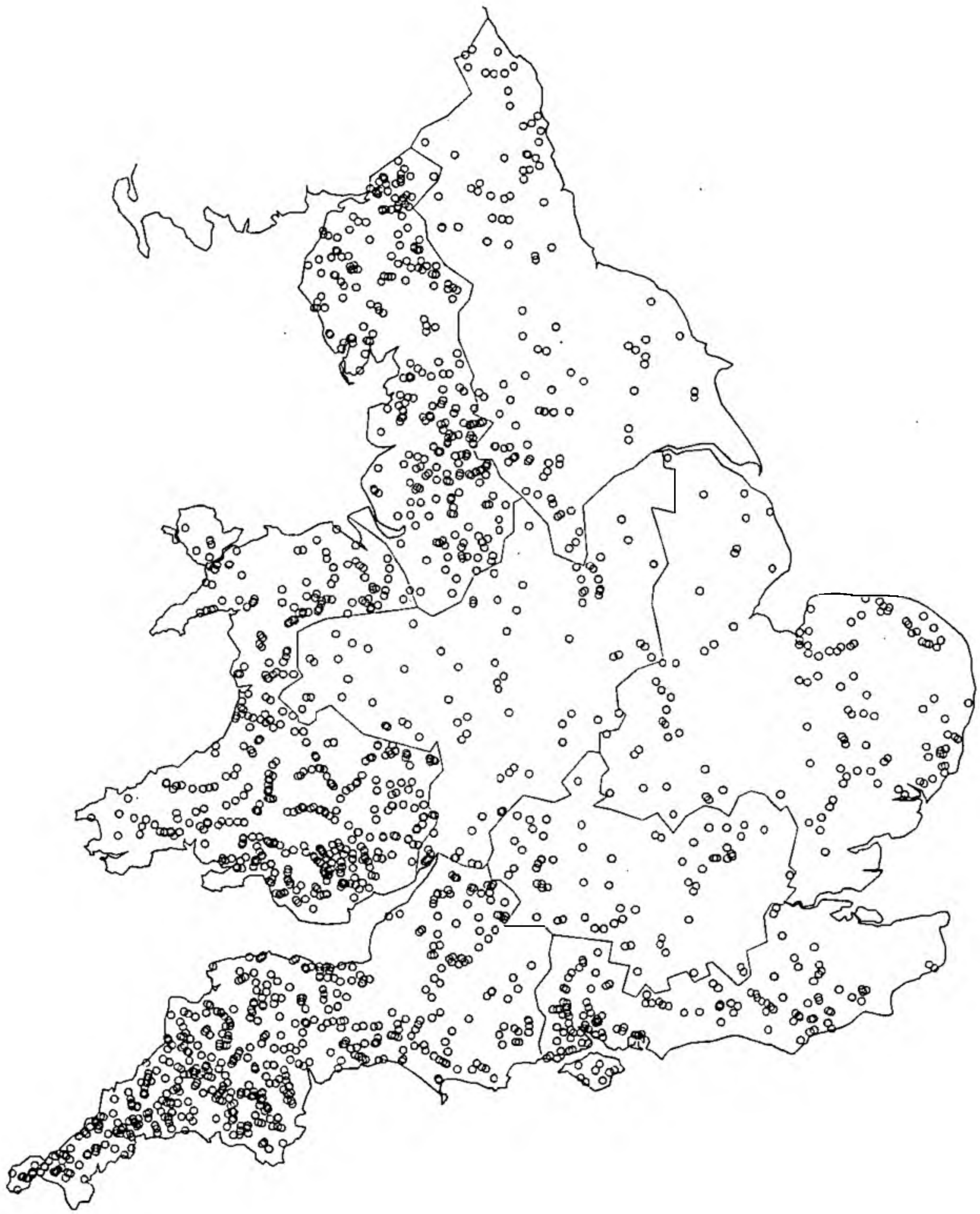

Figure A-2. Autumn distribution of Biological Monitoring Sites in England and Wales at which the concentration of TRP was less than 0.2 mg/l over the three months preceding biological sampling.

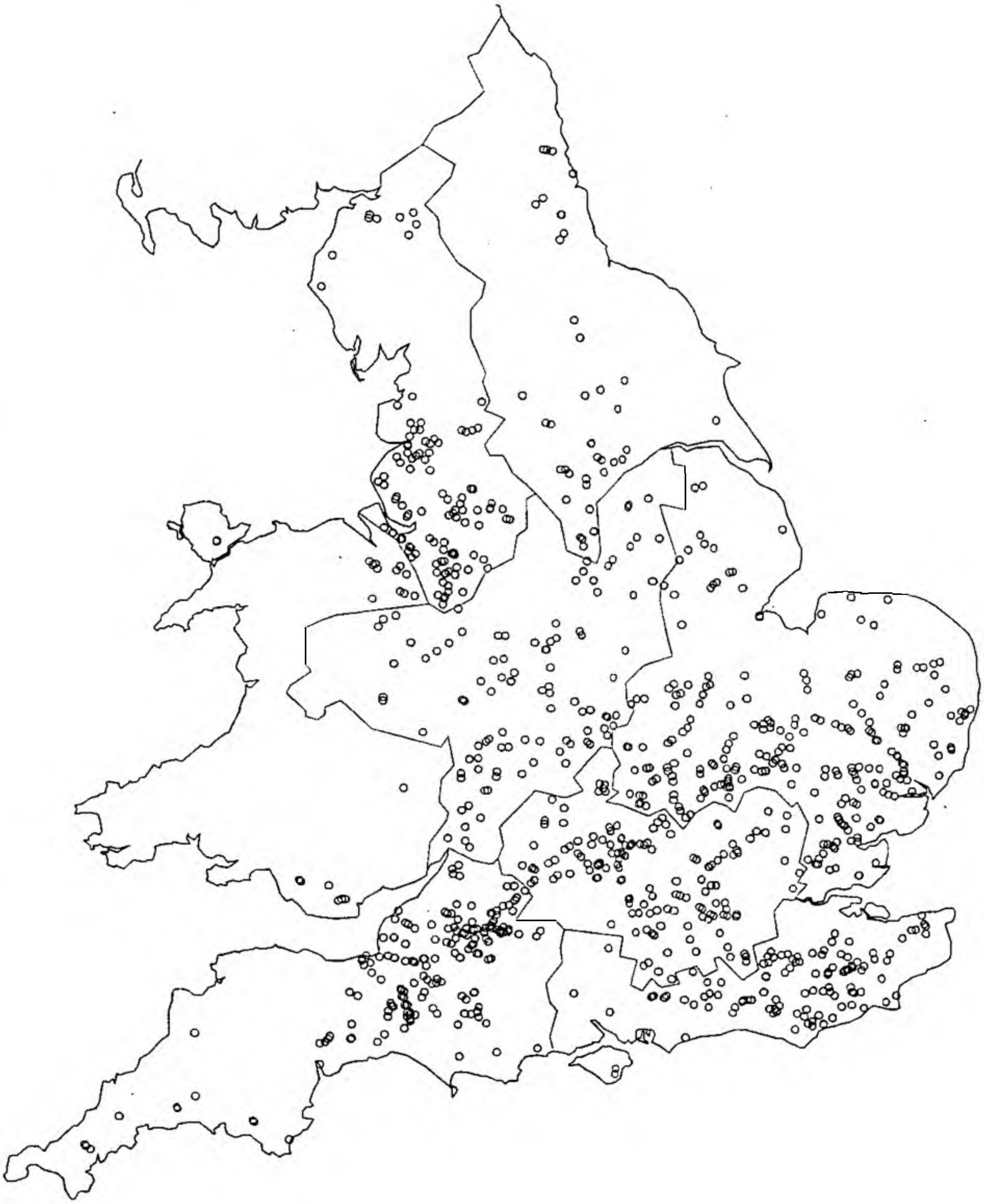

Figure A-3. Spring distribution of Biological Monitoring Sites in England and Wales at which the concentration of TRP equalled or exceeded 0.2 mg/l over the three months preceding biological sampling.

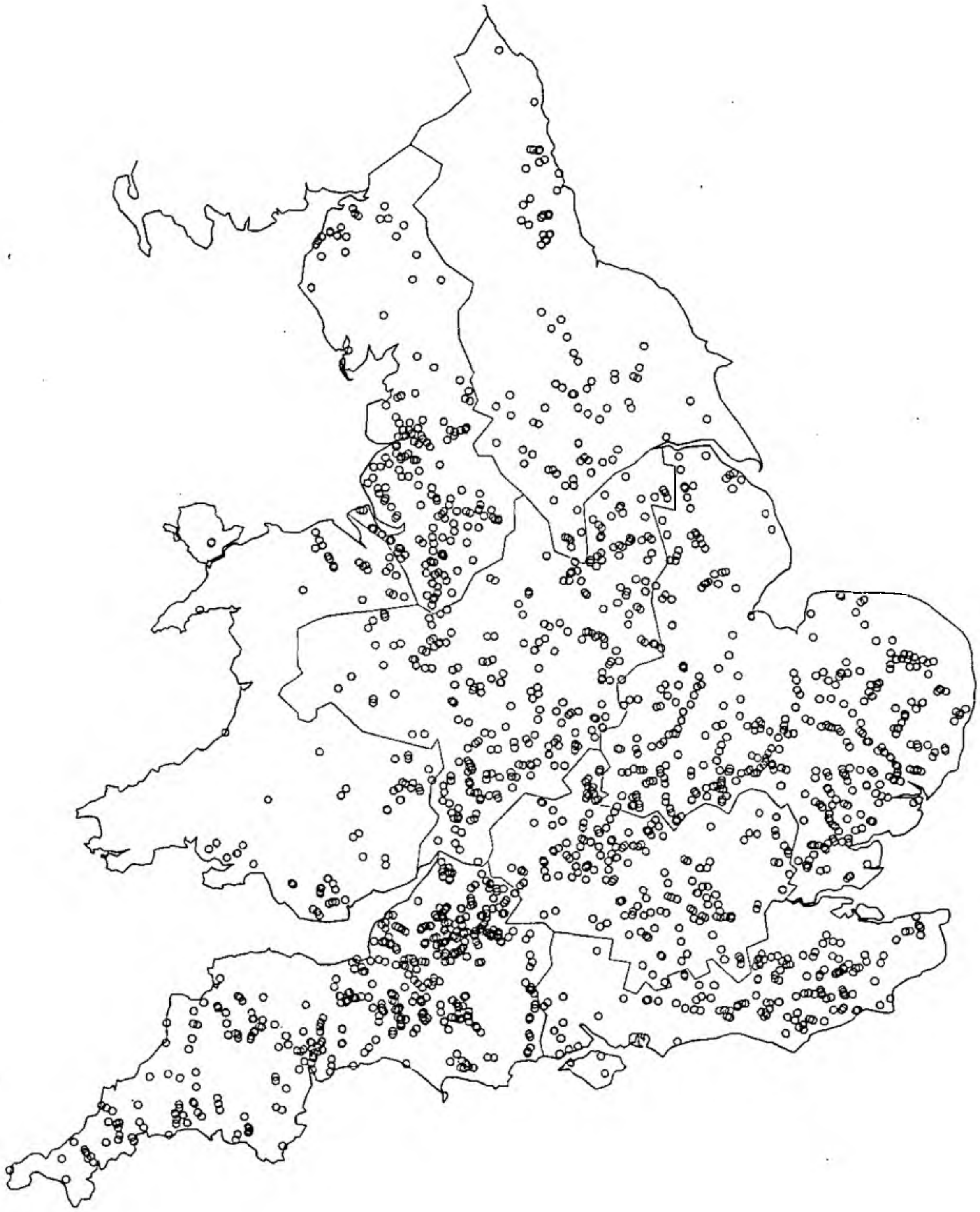

Figure A-4. Autumn distribution of Biological Monitoring Sites in England and Wales at which the concentration of TRP equalled or exceeded 0.2 mg/l over the three months preceding biological sampling.

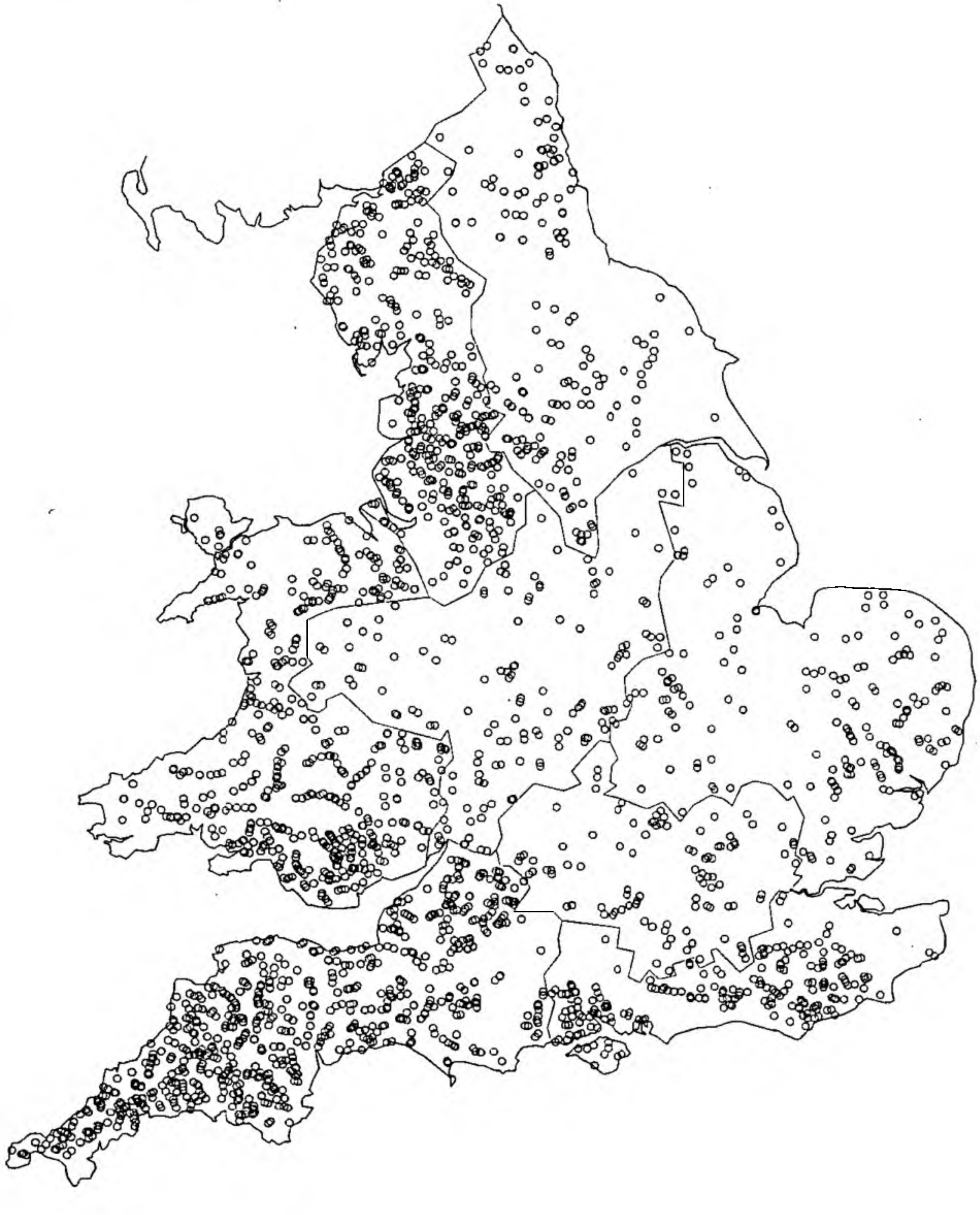

Figure A-5. Combined Spring and Autumn distribution of Biological Monitoring Sites in England and Wales at which the concentration of TON was less than 5 mg/l over the three months preceding biological sampling.

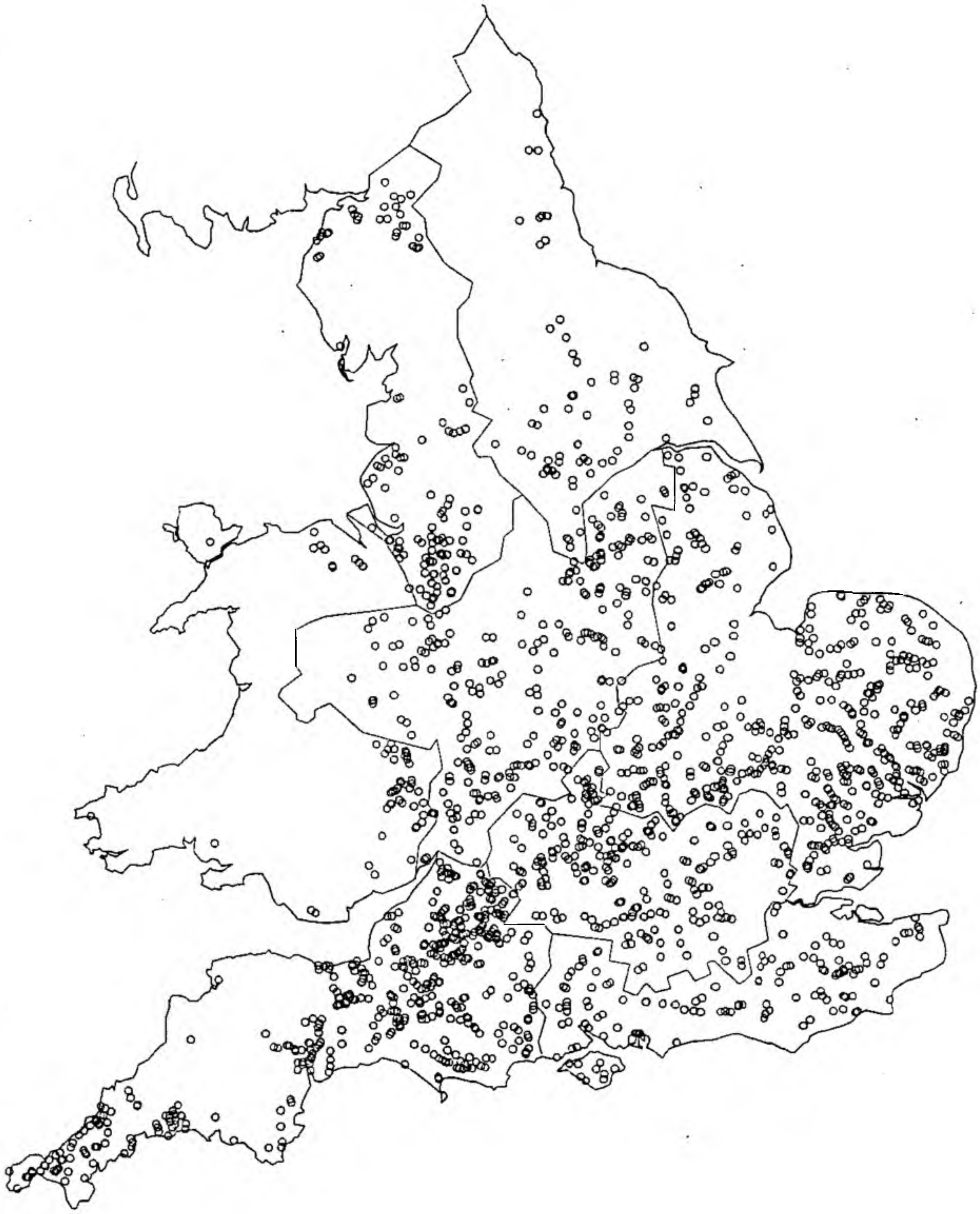

Figure A-6. Combined Spring and Autumn distribution of Biological Monitoring Sites in England and Wales at which the concentration of TON equalled or exceeded 5 mg/l over the three months preceding biological sampling.

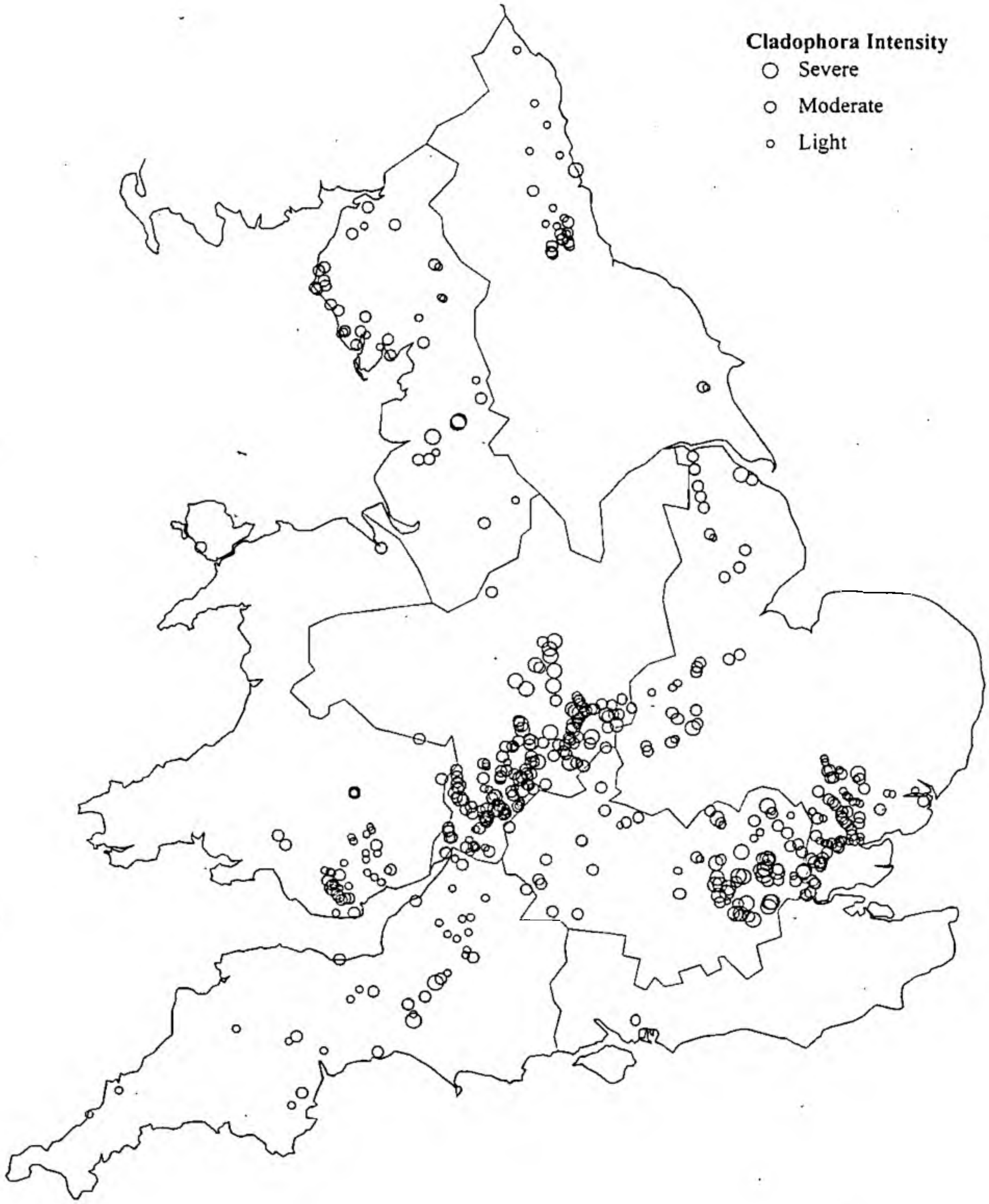

Figure A-7. Distribution of recorded occurrences of Cladophora at Biological Monitoring Sites in England and Wales, based on data from the 1995 environmental stress survey. (NB. Different recording practices within Regions/Areas produced an incomplete coverage.)

Appendix B

Structure of the RPBBN and EBBN

Bayesian Belief Networks

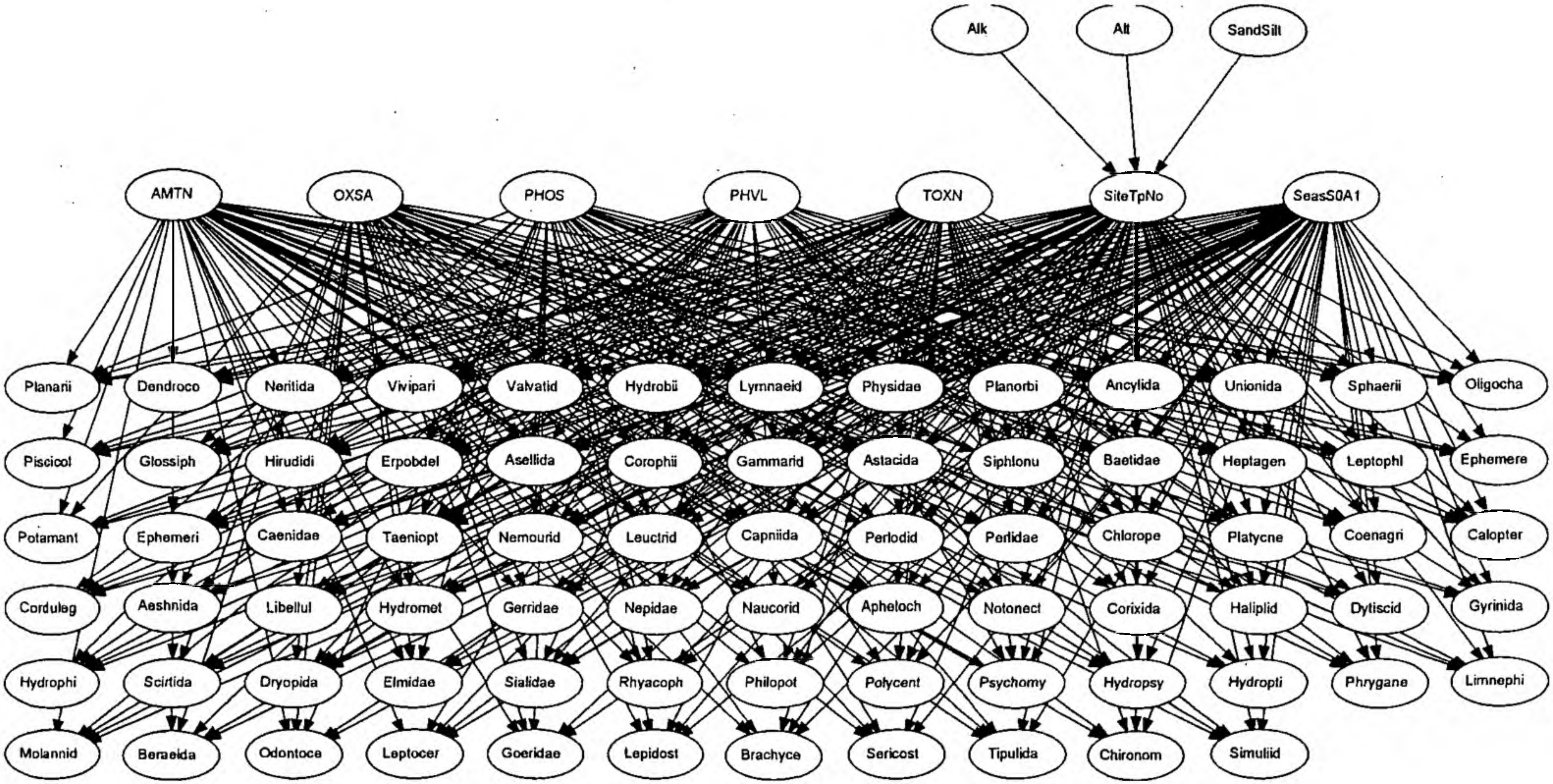

Figure B-1 Structure of the River Pollution Bayesian Belief Network (RPBBN)

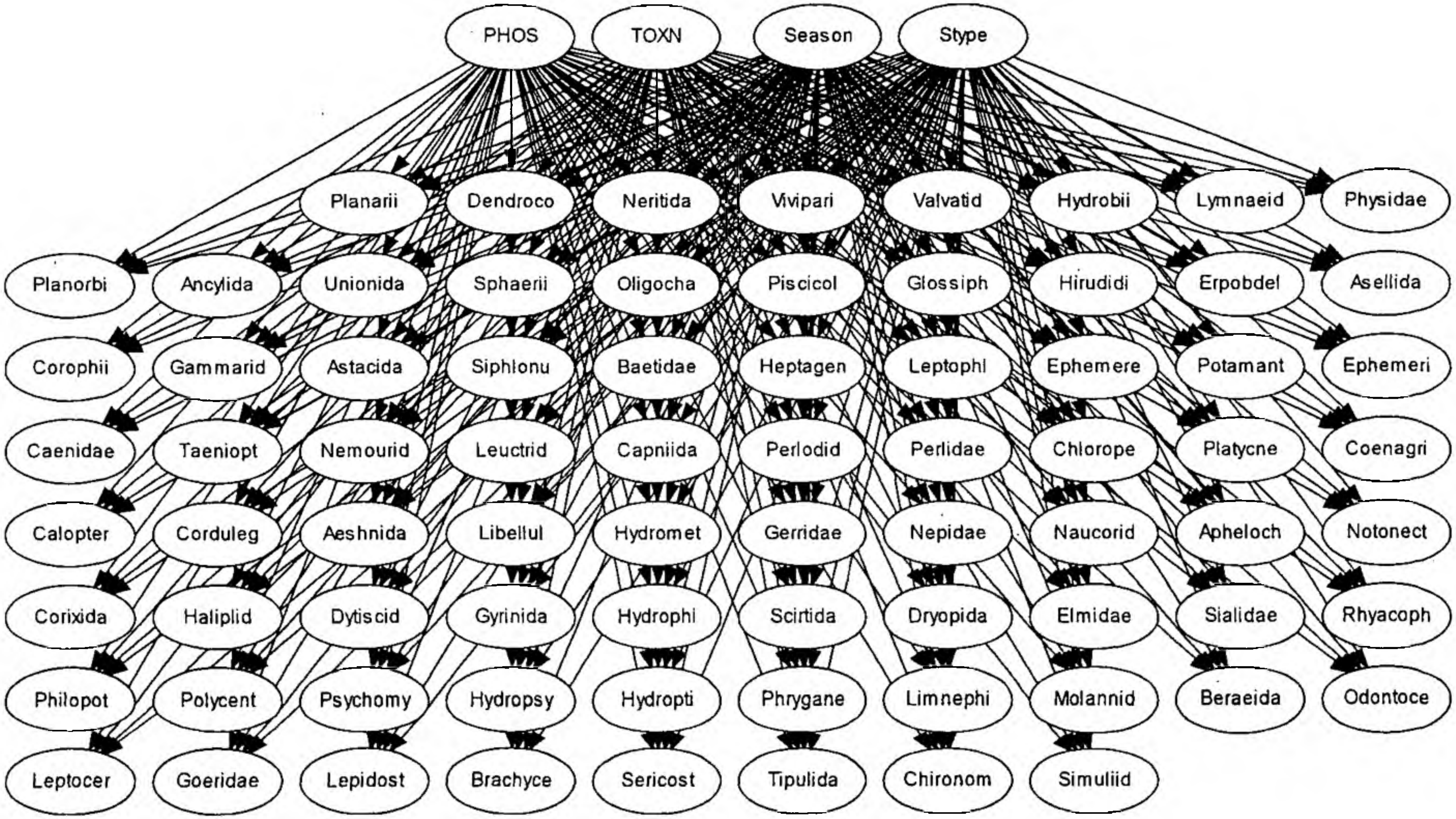

Figure B-2 Structure of the dedicated Eutrophication Bayesian Belief Network (EBBN)

Appendix C

Tables of Average TRP, TON

and

TON/TRP ratios

for

76 BMWP Families.

Table C-1 Average TRP and TON (mg/l) at GQA class A&B sites where each taxon occurred, listed in taxonomic order.

Taxon	Cases	% Occurr.	Avg. TRP	Avg. TON	TON/TRP	Taxon	Cases	% Occurr.	Avg. TRP	Avg. TON	TON/TRP
Planariidae	2273	52.70	0.298	4.14	13.90	Calopterygidae	985	22.84	0.650	5.36	8.25
Dendrocoelidae	265	6.14	0.499	5.74	11.50	Cordulegasteridae	132	3.06	0.125	3.25	25.98
Neritidae	369	8.56	0.476	5.49	11.53	Aeshnidae	31	0.72	0.819	4.56	5.57
Viviparidae	86	1.99	0.625	5.65	9.04	Libellulidae	32	0.74	1.101	5.12	4.65
Valvatidae	877	20.33	0.534	5.88	11.00	Hydrometridae	136	3.15	0.857	4.87	5.68
Hydrobiidae	3595	83.35	0.355	4.51	12.70	Gerridae	217	5.03	0.475	4.04	8.52
Lymnaeidae	2265	52.52	0.439	4.68	10.65	Nepidae	32	0.74	0.817	6.21	7.60
Physidae	835	19.36	0.570	5.35	9.39	Naucoridae	44	1.02	0.783	4.31	5.51
Planorbidae	1560	36.17	0.617	5.58	9.03	Aphelocheiridae	212	4.92	0.297	4.55	15.34
Ancylidae	2700	62.60	0.312	4.28	13.73	Notonectidae	306	7.09	0.728	5.46	7.50
Unionidae	245	5.68	0.669	5.39	8.06	Corixidae	988	22.91	0.594	5.53	9.31
Sphaeriidae	3185	73.85	0.408	4.91	12.02	Haliplidae	1160	26.90	0.537	5.26	9.78
Oligochaeta	4215	97.73	0.334	4.36	13.03	Dytiscidae	2288	53.05	0.387	4.52	11.68
Piscicolidae	824	19.11	0.448	4.94	11.02	Gyrinidae	1873	43.43	0.232	3.74	16.12
Glossiphoniidae	2574	59.68	0.453	5.10	11.25	Hydrophilidae	1887	43.75	0.218	3.32	15.21
Hirudididae	1	0.02	0.227	6.14	27.04	Scirtidae	354	8.21	0.222	4.72	21.28
Erpobdellidae	2457	56.97	0.433	4.96	11.44	Dryopidae	23	0.53	0.086	4.22	49.38
Asellidae	2146	49.76	0.527	5.45	10.34	Elmidae	3920	90.89	0.298	4.23	14.18
Corophiidae	67	1.55	0.794	5.68	7.15	Sialidae	1118	25.92	0.579	5.48	9.46
Gammaridae	3719	86.23	0.359	4.60	12.82	Rhyacophilidae	2406	55.78	0.152	3.55	23.42
Astacidae	58	1.34	0.271	5.03	18.54	Philopotamidae	154	3.57	0.044	2.18	49.14
Siphlonuridae	5	0.12	0.048	2.68	56.27	Polycentropidae	1390	32.23	0.243	3.59	14.75
Baetidae	3850	89.27	0.302	4.18	13.83	Psychomyiidae	948	21.98	0.412	5.29	12.83
Heptageniidae	2201	51.03	0.121	3.17	26.16	Hydropsychidae	3350	77.67	0.273	4.08	14.92
Leptophlebiidae	1328	30.79	0.204	4.25	20.82	Hydroptilidae	1329	30.81	0.320	4.89	15.27
Ephemerellidae	1603	37.17	0.194	4.15	21.40	Phryganeidae	168	3.90	1.048	5.03	4.80
Potamanthidae	9	0.21	0.034	3.74	111.77	Limnephilidae	2590	60.05	0.242	4.64	19.18
Ephemeridae	1348	31.25	0.304	4.80	15.79	Molannidae	239	5.54	0.734	5.62	7.66
Caenidae	2197	50.94	0.409	4.67	11.44	Beraeidae	132	3.06	0.306	4.94	16.12
Taeniopterygidae	695	16.11	0.078	3.06	39.08	Odontoceridae	387	8.97	0.101	3.11	30.84
Nemouridae	1589	36.84	0.108	3.20	29.70	Leptoceridae	2682	62.18	0.390	4.53	11.63
Leuctridae	1985	46.02	0.126	2.97	23.54	Goeridae	1529	35.45	0.194	3.98	20.46
Capniidae	23	0.53	0.150	4.90	32.60	Lepidostomatidae	1580	36.63	0.151	3.15	20.90
Perlodidae	1571	36.42	0.075	2.86	37.91	Brachycentridae	476	11.04	0.214	3.51	16.42
Perlidae	293	6.79	0.044	1.91	42.99	Sericostomatidae	2328	53.98	0.170	3.61	21.18
Chloroperlidae	812	18.83	0.051	2.47	48.78	Tipulidae	3022	70.07	0.252	4.16	16.55
Platycnemidae	120	2.78	0.927	5.17	5.57	Simuliidae	2935	68.05	0.244	4.04	16.57
Coenagriidae	703	16.30	0.813	5.69	7.00	Chironomidae	4232	98.12	0.330	4.34	13.15

Table C-2 Average TRP and TON (mg/l) at GQA class A&B sites where each taxon occurred, listed in order of average TRP.

Taxon	Cases	% Occurr.	Avg. TRP	Avg. TON	TON/TRP	Taxon	Cases	% Occurr.	Avg. TRP	Avg. TON	TON/TRP
Potamanthidae	9	0.21	0.033	3.74	111.77	Hydroptilidae	1329	30.81	0.320	4.89	15.27
Perlidae	293	6.79	0.044	1.91	42.99	Chironomidae	4232	98.12	0.330	4.34	13.15
Philopotamidae	154	3.57	0.044	2.18	49.14	Oligochaeta	4215	97.73	0.334	4.36	13.03
Siphlonuridae	5	0.12	0.048	2.68	56.27	Hydrobiidae	3595	83.35	0.355	4.51	12.70
Chloroperlidae	812	18.83	0.051	2.47	48.78	Gammaridae	3719	86.23	0.359	4.60	12.82
Perlodidae	1571	36.42	0.075	2.86	37.91	Dytiscidae	2288	53.05	0.387	4.52	11.68
Taeniopterygidae	695	16.11	0.078	3.06	39.08	Leptoceridae	2682	62.18	0.390	4.53	11.63
Dryopidae	23	0.53	0.086	4.22	49.38	Sphaeriidae	3185	73.85	0.408	4.91	12.02
Odontoceridae	387	8.97	0.101	3.11	30.84	Caenidae	2197	50.94	0.409	4.67	11.44
Nemouridae	1589	36.84	0.108	3.20	29.70	Psychomyiidae	948	21.98	0.412	5.29	12.83
Heptageniidae	2201	51.03	0.121	3.17	26.16	Erpobdellidae	2457	56.97	0.433	4.96	11.44
Cordulegasteridae	132	3.06	0.125	3.25	25.98	Lymnaeidae	2265	52.52	0.439	4.68	10.65
Leuctridae	1985	46.02	0.126	2.97	23.54	Piscicolidae	824	19.11	0.448	4.94	11.02
Capniidae	23	0.53	0.150	4.90	32.60	Glossiphoniidae	2574	59.68	0.453	5.10	11.25
Lepidostomatidae	1580	36.63	0.151	3.15	20.90	Gerridae	217	5.03	0.475	4.04	8.52
Rhyacophilidae	2406	55.78	0.152	3.55	23.42	Neritidae	369	8.56	0.476	5.49	11.53
Sericostomatidae	2328	53.98	0.170	3.61	21.18	Dendrocoelidae	265	6.14	0.499	5.74	11.50
Ephemerellidae	1603	37.17	0.194	4.15	21.40	Asellidae	2146	49.76	0.527	5.45	10.34
Goeridae	1529	35.45	0.194	3.98	20.46	Valvatidae	877	20.33	0.534	5.87	11.00
Leptophlebiidae	1328	30.79	0.204	4.25	20.82	Haliplidae	1160	26.90	0.537	5.26	9.78
Brachycentridae	476	11.04	0.214	3.51	16.42	Physidae	835	19.36	0.570	5.35	9.39
Hydrophilidae	1887	43.75	0.218	3.32	15.21	Sialidae	1118	25.92	0.579	5.48	9.46
Scirtidae	354	8.21	0.222	4.72	21.28	Corixidae	988	22.91	0.594	5.53	9.31
Hirudididae	1	0.02	0.227	6.14	27.04	Planorbidae	1560	36.17	0.617	5.58	9.03
Gyrinidae	1873	43.43	0.232	3.74	16.12	Viviparidae	86	1.99	0.625	5.65	9.04
Limnephilidae	2590	60.05	0.242	4.64	19.18	Calopterygidae	985	22.84	0.650	5.36	8.25
Polycentropidae	1390	32.23	0.243	3.59	14.75	Unionidae	245	5.68	0.669	5.39	8.06
Simuliidae	2935	68.05	0.244	4.04	16.57	Notonectidae	306	7.09	0.728	5.46	7.50
Tipulidae	3022	70.07	0.252	4.16	16.55	Molannidae	239	5.54	0.734	5.62	7.66
Astacidae	58	1.34	0.271	5.03	18.54	Naucoridae	44	1.02	0.783	4.31	5.51
Hydropsychidae	3350	77.67	0.273	4.08	14.92	Corophiidae	67	1.55	0.794	5.68	7.15
Aphelocheiridae	212	4.92	0.297	4.55	15.34	Coenagriidae	703	16.30	0.813	5.69	7.00
Planariidae	2273	52.70	0.298	4.14	13.90	Nepidae	32	0.74	0.817	6.21	7.60
Elmidae	3920	90.89	0.298	4.23	14.18	Aeshnidae	31	0.72	0.819	4.56	5.57
Baetidae	3850	89.27	0.302	4.18	13.83	Hydrometridae	136	3.15	0.857	4.87	5.68
Ephemeridae	1348	31.25	0.304	4.80	15.79	Platycnemidae	120	2.78	0.927	5.17	5.57
Beraeidae	132	3.06	0.306	4.94	16.12	Phryganeidae	168	3.90	1.048	5.03	4.80
Ancylidae	2700	62.60	0.312	4.28	13.73	Libellulidae	32	0.74	1.101	5.12	4.65

Table C-3 Average TRP and TON (mg/l) at GQA class A&B sites where each taxon occurred, listed in order of average TON.

Taxon	Cases	% Occurr.	Avg. TRP	Avg. TON	TON/TRP	Taxon	Cases	% Occurr.	Avg. TRP	Avg. TON	TON/TRP
Perlidae	293	6.79	0.044	1.91	42.99	Aeshnidae	31	0.72	0.819	4.56	5.57
Philopotamidae	154	3.57	0.044	2.18	49.14	Gammaridae	3719	86.23	0.359	4.60	12.82
Chloroperlidae	812	18.83	0.051	2.47	48.78	Limnephilidae	2590	60.05	0.242	4.64	19.18
Siphlonuridae	5	0.12	0.048	2.68	56.27	Caenidae	2197	50.94	0.409	4.67	11.44
Perlodidae	1571	36.42	0.075	2.86	37.91	Lymnaeidae	2265	52.52	0.439	4.68	10.65
Leuctridae	1985	46.02	0.126	2.97	23.54	Scirtidae	354	8.21	0.222	4.72	21.28
Taeniopterygidae	695	16.11	0.078	3.06	39.08	Ephemeridae	1348	31.25	0.304	4.80	15.79
Odontoceridae	387	8.97	0.101	3.11	30.84	Hydrometridae	136	3.15	0.857	4.87	5.68
Lepidostomatidae	1580	36.63	0.151	3.15	20.90	Hydroptilidae	1329	30.81	0.320	4.89	15.27
Heptageniidae	2201	51.03	0.121	3.17	26.16	Capniidae	23	0.53	0.150	4.90	32.60
Nemouridae	1589	36.84	0.108	3.20	29.70	Sphaeriidae	3185	73.85	0.408	4.91	12.02
Cordulegasteridae	132	3.06	0.125	3.25	25.98	Beraeidae	132	3.06	0.306	4.94	16.12
Hydrophilidae	1887	43.75	0.218	3.32	15.21	Piscicolidae	824	19.11	0.448	4.94	11.02
Brachycentridae	476	11.04	0.214	3.51	16.42	Erpobdellidae	2457	56.97	0.433	4.96	11.44
Rhyacophilidae	2406	55.78	0.152	3.55	23.42	Astacidae	58	1.34	0.271	5.03	18.54
Polycentropidae	1390	32.23	0.243	3.59	14.75	Phryganeidae	168	3.90	1.048	5.03	4.80
Sericostomatidae	2328	53.98	0.170	3.61	21.18	Glossiphoniidae	2574	59.68	0.453	5.10	11.25
Potamanthidae	9	0.21	0.033	3.74	111.77	Libellulidae	32	0.74	1.101	5.12	4.65
Gyrinidae	1873	43.43	0.232	3.74	16.12	Platycnemidae	120	2.78	0.927	5.17	5.57
Goeridae	1529	35.45	0.194	3.98	20.46	Haliplidae	1160	26.90	0.537	5.26	9.78
Simuliidae	2935	68.05	0.244	4.04	16.57	Psychomyiidae	948	21.98	0.412	5.29	12.83
Gerridae	217	5.03	0.475	4.04	8.52	Physidae	835	19.36	0.570	5.35	9.39
Hydropsychidae	3350	77.67	0.273	4.08	14.92	Calopterygidae	985	22.84	0.650	5.36	8.25
Planariidae	2273	52.70	0.298	4.14	13.90	Unionidae	245	5.68	0.669	5.39	8.06
Ephemerellidae	1603	37.17	0.194	4.15	21.40	Asellidae	2146	49.76	0.527	5.45	10.34
Tipulidae	3022	70.07	0.252	4.16	16.55	Notonectidae	306	7.09	0.728	5.46	7.50
Baetidae	3850	89.27	0.302	4.18	13.83	Sialidae	1118	25.92	0.579	5.48	9.46
Dryopidae	23	0.53	0.086	4.22	49.38	Neritidae	369	8.56	0.476	5.49	11.53
Elmidae	3920	90.89	0.298	4.23	14.18	Corixidae	988	22.91	0.594	5.53	9.31
Leptophlebiidae	1328	30.79	0.204	4.25	20.82	Planorbidae	1560	36.17	0.617	5.58	9.03
Ancylidae	2700	62.60	0.312	4.28	13.73	Molannidae	239	5.54	0.734	5.62	7.66
Naucoridae	44	1.02	0.783	4.31	5.51	Viviparidae	86	1.99	0.625	5.65	9.04
Chironomidae	4232	98.12	0.330	4.34	13.15	Corophiidae	67	1.55	0.794	5.68	7.15
Oligochaeta	4215	97.73	0.334	4.36	13.03	Coenagriidae	703	16.30	0.813	5.69	7.00
Hydrobiidae	3595	83.35	0.355	4.51	12.70	Dendrocoelidae	265	6.14	0.499	5.74	11.50
Dytiscidae	2288	53.05	0.387	4.52	11.68	Valvatidae	877	20.33	0.534	5.87	11.00
Leptoceridae	2682	62.18	0.390	4.53	11.63	Hirudididae	1	0.02	0.227	6.14	27.04
Aphelocheiridae	212	4.92	0.297	4.55	15.34	Nepidae	32	0.74	0.817	6.21	7.60

Table C-4 Average TRP and TON (mg/l) at GQA class A&B sites where each taxon occurred, listed in order of the TON/TRP ratio.

Taxon	Cases	% Occurr.	Avg. TRP	Avg. TON	TON/TRP	Taxon	Cases	% Occurr.	Avg. TRP	Avg. TON	TON/TRP
Libellulidae	32	0.74	1.101	5.12	4.65	Baetidae	3850	89.27	0.302	4.18	13.83
Phryganeidae	168	3.90	1.048	5.03	4.80	Planariidae	2273	52.70	0.298	4.14	13.90
Naucoridae	44	1.02	0.783	4.31	5.51	Elmidae	3920	90.89	0.298	4.23	14.18
Aeshnidae	31	0.72	0.819	4.56	5.57	Polycentropidae	1390	32.23	0.243	3.59	14.75
Platycnemidae	120	2.78	0.927	5.17	5.57	Hydropsychidae	3350	77.67	0.273	4.08	14.92
Hydrometridae	136	3.15	0.857	4.87	5.68	Hydrophilidae	1887	43.75	0.218	3.32	15.21
Coenagriidae	703	16.30	0.813	5.69	7.00	Hydroptilidae	1329	30.81	0.320	4.89	15.27
Corophiidae	67	1.55	0.794	5.68	7.15	Aphelocheiridae	212	4.92	0.297	4.55	15.34
Notonectidae	306	7.09	0.728	5.46	7.50	Ephemeridae	1348	31.25	0.304	4.80	15.79
Nepidae	32	0.74	0.817	6.21	7.60	Beraeidae	132	3.06	0.306	4.94	16.12
Molannidae	239	5.54	0.734	5.62	7.66	Gyrinidae	1873	43.43	0.232	3.74	16.12
Unionidae	245	5.68	0.669	5.39	8.06	Brachycentridae	476	11.04	0.214	3.51	16.42
Calopterygidae	985	22.84	0.650	5.36	8.25	Tipulidae	3022	70.07	0.252	4.16	16.55
Gerridae	217	5.03	0.475	4.04	8.52	Simuliidae	2935	68.05	0.244	4.04	16.57
Planorbidae	1560	36.17	0.617	5.58	9.03	Astacidae	58	1.34	0.271	5.03	18.54
Viviparidae	86	1.99	0.625	5.65	9.04	Limnephilidae	2590	60.05	0.242	4.64	19.18
Corixidae	988	22.91	0.594	5.53	9.31	Goeridae	1529	35.45	0.194	3.98	20.46
Physidae	835	19.36	0.570	5.35	9.39	Leptophlebiidae	1328	30.79	0.204	4.25	20.82
Sialidae	1118	25.92	0.579	5.48	9.46	Lepidostomatidae	1580	36.63	0.151	3.15	20.90
Haliplidae	1160	26.90	0.537	5.26	9.78	Sericostomatidae	2328	53.98	0.170	3.61	21.18
Asellidae	2146	49.76	0.527	5.45	10.34	Scirtidae	354	8.21	0.222	4.72	21.28
Lymnaeidae	2265	52.52	0.439	4.68	10.65	Ephemerellidae	1603	37.17	0.194	4.15	21.40
Valvatidae	877	20.33	0.534	5.87	11.00	Rhyacophilidae	2406	55.78	0.152	3.55	23.42
Piscicolidae	824	19.11	0.448	4.94	11.02	Leuctridae	1985	46.02	0.126	2.97	23.54
Glossiphoniidae	2574	59.68	0.453	5.10	11.25	Cordulegasteridae	132	3.06	0.125	3.25	25.98
Caenidae	2197	50.94	0.409	4.67	11.44	Heptageniidae	2201	51.03	0.121	3.17	26.16
Erpobdellidae	2457	56.97	0.433	4.96	11.44	Hirudididae	1	0.02	0.227	6.14	27.04
Dendrocoelidae	265	6.14	0.499	5.74	11.50	Nemouridae	1589	36.84	0.108	3.20	29.70
Neritidae	369	8.56	0.476	5.49	11.53	Odontoceridae	387	8.97	0.101	3.11	30.84
Leptoceridae	2682	62.18	0.390	4.53	11.63	Capniidae	23	0.53	0.150	4.90	32.60
Dytiscidae	2288	53.05	0.387	4.52	11.68	Perlodidae	1571	36.42	0.075	2.86	37.91
Sphaeriidae	3185	73.85	0.408	4.91	12.02	Taeniopterygidae	695	16.11	0.078	3.06	39.08
Hydrobiidae	3595	83.35	0.355	4.51	12.70	Perlidae	293	6.79	0.044	1.91	42.99
Gammaridae	3719	86.23	0.359	4.60	12.82	Chloroperlidae	812	18.83	0.051	2.47	48.78
Psychomyiidae	948	21.98	0.412	5.29	12.83	Philopotamidae	154	3.57	0.044	2.18	49.14
Oligochaeta	4215	97.73	0.334	4.36	13.03	Dryopidae	23	0.53	0.086	4.22	49.38
Chironomidae	4232	98.12	0.330	4.34	13.15	Siphlonuridae	5	0.12	0.048	2.68	56.27
Ancylidae	2700	62.60	0.312	4.28	13.73	Potamanthidae	9	0.21	0.033	3.74	111.77