PRIMARY RESEARCH PAPER

A predictive model for freshwater bioassessment (Mondego River, Portugal)

Maria João Feio · Trefor B. Reynoldson ·
Verónica Ferreira · Manuel Augusto S. Graça

Abstract We sampled macroinvertebrates at 75 locations in the Mondego river catchment, Central Portugal, and developed a predictive model for water quality assessment of this basin, based on the Reference Condition Approach. Sampling was done from June to September 2001. Fifty-five sites were identified as "Reference sites" and 20 sites were used as "Test sites" to test the model. At each site we also measured 40 habitat variables to characterize water physics and chemistry, habitat type, land use, stream hydrology and geographic location. Macroinvertebrates were generally identified to species or genus level; a total of 207 taxa were found. By Unweighted Pair Group Method with Arithmetic mean (UPGMA) clustering and analysis of species contribution to similarities percentage (SIMPER), two groups of reference sites were established. Using Discriminant Analysis (stepwise forward), four variables correctly predicted 78% of the reference sites to the appropriate group: stream order, pool quality, substrate quality and current velocity. Test sites' environmental quality was established from their relative distance to reference sites, in MDS ordination space, using a series of bands (BEAST methodology). The model performed well at upstream sites, but at downstream sites it was compromised by the lack of reference sites. As with the English RIVPACS predictive model, the Mondego model should be continually improved with the addition of new reference sites. The adaptation of the Mondego model methodology to the Water Framework Directive is possible and would consist mainly of the integration of the WFD typology and increasing the number of ellipses that define quality bands.

Keywords Reference condition · Predictive model · Water quality · Macroinvertebrates · Mondego catchment

M. J. Feio (✉) · V. Ferreira · M. A. S. Graça
Department of Zoology, and IMAR, University of Coimbra, Largo Marquês de Pombal, 3004-517 Coimbra, Portugal
e-mail: mjf@ci.uc.pt

T. B. Reynoldson
National Water Research Institute, Environment Canada, Acadia Centre for Estuarine Research, Box 115 Acadia University, B4P 2R6 Wolfville, NS, Canada

Introduction

The European Water Framework Directive (WFD) suggests a holistic approach to catchment management and requires ecological objectives to be set for surface waters (Directive 2000/60/CE, 2000). The Reference Condition Approach–RCA (Reynoldson et al., 1997) has been used in the

United Kingdom (Wright et al., 1984; Wright, 1995), Canada (e.g. Reynoldson et al., 1995; Rosenberg et al., 2000), Australia (e.g. Parsons & Norris, 1996; Pollard & Huxham, 1998) and the United States of America (Barbour & Yoder, 2000) to set numeric biological objectives for streams and rivers. In the RCA the reference condition is defined by a set of reference sites representing undisturbed or the best available biological conditions of a region.

Benthic macroinvertebrates are commonly used in water quality assessment programs and are also recommended in the WFD (Rosenberg & Resh, 1993; Zamora-Muñoz & Alba-Tercedor, 1996; Yoder & Rankin, 1998; Directive 2000/60/CE). Their sedentary nature, high abundance, ease of sampling and identification and their longevity, long enough to capture cumulative effects of stress and exposure over time, are good reasons for their use as bioindicators (Hellawell, 1977; De Pauw & Vanhooren, 1983). Yet, macroinvertebrate communities, as all biological systems, change over time and space, creating difficulties in setting numeric biological objectives (Reynoldson et al., 2000).

Biological assessment using the RCA is based on the comparison of the observed community with the community expected to occur according to reference sites. There are two main multivariate predictive approaches to do the assessment: RIVPACS and BEAST. Both define reference sites based on a set of environmental variables, classify reference assemblages into groups and predict the possibility of a new site to belong to each group based on environmental features. The major difference between the BEAST and RIVPACS models is the evaluation of test sites. The original approach, developed by RIVPACS predicts the taxa expected at a test site, compares it with the taxa expected at reference conditions and calculates the distance to reference as the deviation of the ratio Observed/Expected from 1. The second approach, the BEAST predictive model, was developed and applied in Canada (Reynoldson et al., 1995; 1997), and tests whether a new site falls within the given confidence limit in a specified MDS-ordination space, defined based on the reference-site group that the site is most likely to belong to (using the environmental

descriptors only). A banding system, with four categories of increasing distances to the ordination centroid, permits the visual evaluation of the biological quality of test sites. This method allows a direct classification of the river water quality and permits following the evolution of the quality of a site over time, by adding data obtained in new dates to the ordination space.

The Mondego is the largest river whose catchment is entirely contained within Portugal. The lower section of the catchment is densely populated while the remaining area has low to moderate human impacts. However, dam construction and industrial developments are potential threats to the environmental quality of running waters (Marques et al., 2002). It is necessary to develop a bioassessment method to monitor and evaluate the conditions of the river basin.

The Mondego catchment is very different from those systems where the BEAST approach has been applied to (Rosenberg et al., 2000; Reynoldson et al., 2000), because of a high variability of environmental conditions within its small area and a transitional climate between Temperate and Mediterranean. Our main goals were: 1) to describe the macroinvertebrate communities of reference sites in the Mondego catchment; 2) to test the efficiency of the BEAST model under the new set of environmental conditions; and 3) to discuss the potential use of the Mondego predictive model for the assessment and monitoring of the river, according to the WFD requirements (Directive 2000/60/CE).

Methods

Study area and selection of sites

The Mondego river catchment is located in the centre of Portugal, between 39°46′ and 40°48′ N, and 7°14′ and 8°52′ W. The 6670 km² catchment includes a wide range of environmental conditions from mountainous areas in the upper and middle reaches to a large alluvial plain where the river discharges to the Atlantic Ocean (Marques et al., 2002). More information on the geology, hydrology and biology of the Mondego basin can be found in Pardal et al. (2002).

To identify reference conditions describing the natural heterogeneity of the basin, sample sites were chosen *a priori* on 1:250,000 military maps, using to the following criteria: (1) two sites per UTM quadrat (100 km^2) (Universal Transverse Mercator grid); (2) if there was more than one sub-catchment in the quadrat the sites were assigned to different sub-catchments; (3) avoiding urban areas, impoundments and pollution sources; (4) covering altitudinal, geological and river order gradients of wadeable water courses. Sites were sampled between July and September 2001 (summer) as this is a period of low water levels when almost all streams are accessible and kick-sampling is possible. Seventy-five sites (Fig. 1) were sampled and characterized (corresponding to one site per 89 km^2). A site was defined as a 20 m stream reach for sampling for macroinvertebrates and environmental characterization. The data set thus acquired was used to build the predictive models described in this paper.

Aquatic macroinvertebrates samples

Sampling procedure

At each site invertebrates were sampled with a kick net (0.30 × 0.30 m opening and 500 μm mesh size). The samples were collected by 3-min kicking of the riverbed substrate across the river, perpendicular to the banks, at regular distances from bank to bank in the 20 m stream reach that comprises a sampling site. Each sample was a composite of either three or six sub-samples, depending on whether a stream was <3 or >3 m wetted width, respectively. The sub-samples were obtained by kicking the substrate upstream from the net (in an area of approximately 1 m long × net width) for 0.5 min. Therefore the total sampling time was 3 × 0.5 min (= 1.5 min) in small streams and 6 × 0.5 min in large streams (>3 m width). All counts were converted to individuals captured/minute.

Fig. 1 Mondego River basin in Portugal and sampling sites distribution in the basin
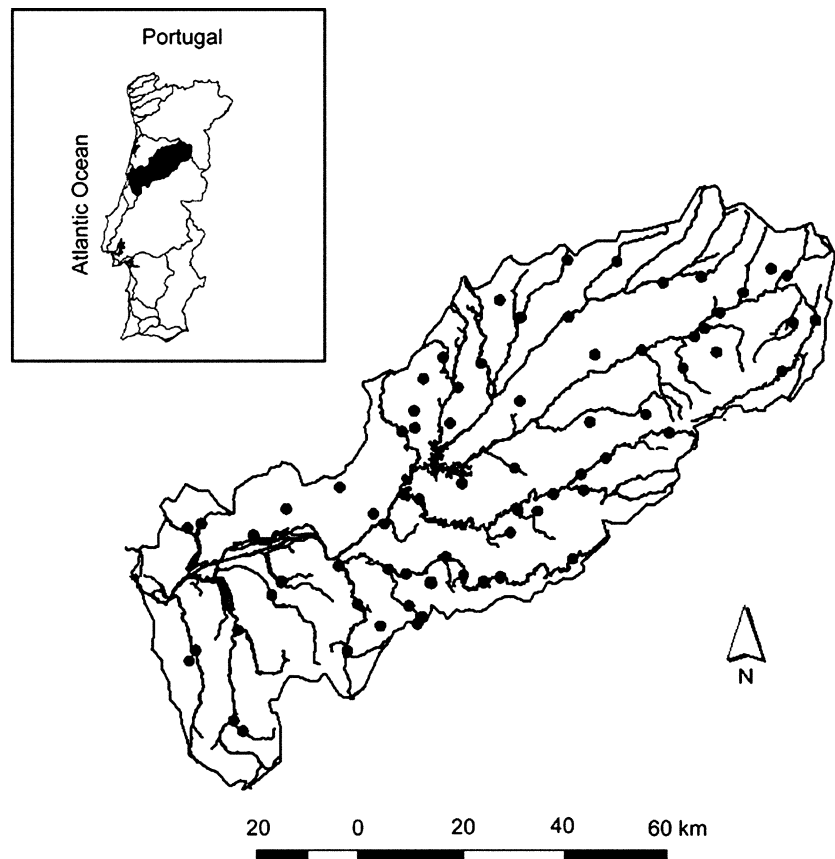
Sample processing

Benthic samples were preserved in 10% formalin. Before sorting, samples were washed and subdivided into three fractions (retained by >2 mm, 1 mm and 0.5 mm mesh sieves) to increase sorting efficiency. If the number of invertebrates was too high, the fraction retained in the 0.5 mm mesh sieved was sub-sampled by area (1/4 or 1/3) depending on the total number of animals in the sample (if >50 and <100 animals were found in the first square of the grid, that corresponds to 1/12 of the fraction, than 1/3 of the fraction was counted; if >100 animals were found in the first square of the grid that 1/4 of the fraction was counted). Sorted invertebrates were then preserved in 70% ethanol for further identification. Animals were counted and identified to the lowest practical level, mostly genus and species. Hydracarina, Colembolla, Copepoda, Ostracoda and Hydridae were counted at the order level and included in the analysis.

Site characterization

For each site 40 environmental variables were measured either in the field (e.g. related to the stream morphology and hydrology or riparian vegetation), in the laboratory (e.g. nutrients in the water or alkalinity) or obtained from cartographic and digital information (e.g. using Geographic Information Systems techniques), and later used in the data analysis. These variables were used to identify reference sites and to predict the reference conditions (Table 1).

The stream name, altitude (m) and distance to source (km), and stream order (Strahler system) were obtained from 1:25,000 and 1:250,000 maps, respectively (Instituto Geográfico do Exército 1998a; b). Latitude and longitude were obtained by GPS. Percentages of forest, industrial, urban and degraded areas, agriculture and eucalyptus, were measured for an area of 1 km radius around each sampling site from digital sources (1:25,000 data from ''Ministério do Ambiente e Ordenamento do Território'' 2002). The mean annual temperature (°C) and the precipitation (mm and days with precipitation/year) were obtained from Instituto do Ambiente (2003).

Field measurements included: stream width (m), mean depth (m), current velocity along the transect where macroinvertebrates were collected (m s^{-1}; n = 3 or 6 for <3 m or >3 m stream width, respectively), dissolved oxygen (% and mg l^{-1}), pH, conductivity (μS cm^{-1}) and total dissolved solids (mg l^{-1}), water temperature (°C) and mean substrate particle size (mm; n = 9 or 18 dominant substrate particles from the stream bed for <3 m or >3 m stream width, respectively). The mean discharge (D; m^3 s^{-1}) was calculated later as: D = stream width × mean depth × mean current velocity.

The habitat quality was assessed according to Barbour et al. (1999) by classifying sites to one of four categories (poor, marginal, sub-optimal and optimal) by visual evaluation. The same approach (Barbour et al., 1999) was used to measure variables describing habitat complexity (diversity of habitat types, e.g. boulders, branches, aquatic vegetation), pool quality (in mixture of deep/ shallow, large/small pools) and substrate quality (in mixture of substrate types, e.g. gravel cobble, sand).

Other variables evaluated in the field by visual observation were the valley form (categories: 1 for V shapes and troughs; 2 for U shape, meander and plain floodplain), total width of the riparian vegetation (m), woody vegetation (%), shading at zenith in the stream (%) and the lithology (categories: 1 for sedimentary rocks; 2 for sedimentary + metamorphic rocks; 3 for plutonic rocks).

Periphyton samples were obtained by scraping one 45 mm diameter circle from each of 3 randomly collected stones from the river bed. The sample was then washed into a 300 ml water bottle and filled with clean water. The water was filtered (100 ml) through GFC fibre-glass filters (Whatman) and used to measure chlorophyll-a, according to APHA (1995). The remaining 200 ml were discarded. Stream water was collected in 250 ml plastic bottles at the time of invertebrate sampling, for later analyses.

Sample processing

After processing the invertebrate samples, the remaining coarse particulate organic matter (>1 mm) was dried at 60°C for 24 h, weighed

Table 1 Environmental variables used to characterize each sampling site in the summer of 2001 and the respective transformation applied

Site related variables:	Stream morphology and riparian vegetation:
Stream order	**Valley form**
Distance to source (km; log $_{10}$)	Riparian vegetation (total width; m; $\sqrt{\ }$)
Decimal latitude and longitude	Woody vegetation (%)
Altitude (m; $^4\sqrt{\ }$)	Shading at zenith (%)
Lithology	**Land use:**
Site morphology and hydrology:	Forest (%)
Mean annual temperature (°C)	Industrial, urban and degraded areas (%; $\sqrt{\ }$)
Mean annual total precipitation (mm)	Agriculture (%)
Mean annual precipitation (days/year)	Eucalyptus (%; arcsine ($^6\sqrt{\ }$/100))
Mean stream width (m; 1 / $\sqrt{\ }$)	**Water characteristics and dissolved substances:**
Mean stream depth (m)	Water temperature (°C; 1/$\sqrt{\ }$)
Current velocity (m s^{-1}; $\sqrt{\ }$)	pH
Mean discharge (m^3 s^{-1}; 1/$^4\sqrt{\ }$)	Conductivity (uS cm^{-1}; 1/$\sqrt{\ }$)
Periphyton:	O$_2$ (% and mg l^{-1})
Chlorophyll in periphyton (mg m^{-2}; $\sqrt{\ }$)	Total dissolved solids (mg l^{-1}) (TDS; 1/$\sqrt{\ }$)
Habitat:	Chloride (Cl$^-$) (mg l^{-1}; log$_{10}$)
Substrate quality	Nitrate (NO$_3^{2-}$) (mg l^{-1}; (log x + 1))
Mean substrate size (mm)	Nitrite (NO$_2^-$) (mg l^{-1}; $^4\sqrt{\ }$ (log x + 1))
Habitat complexity	Sulphate (SO$_4^{2-}$) (mg l^{-1}; log$_{10}$)
Pool quality	Phosphate-P (P-PO$_4^{3-}$)(mg l^{-1}; $\sqrt{\ }$ (log x + 1))
Organic matter in the substrate:	Ammonia-N (mg l^{-1}; 1/ $^6\sqrt{\ }$ (log x + 1))
CPOM > 1 mm (g; $\sqrt{\ }$)	Alkalinity (mg l^{-1}; 1/$\sqrt{\ }$)

In bold are the variables used in Discriminant Analysis (SYSTAT 8.0)

and finally ashed at 550°C and reweighed. Ash free dry mass was calculated by the difference between dry weight and ashed weight.

The water samples were analyzed for chloride (mg l^{-1}), nitrate (mg l^{-1}), nitrite (mg l^{-1}), sulphate (mg l^{-1}), phosphate-phosphorous (mg l^{-1}) and ammonia-nitrogen (mg l^{-1}) using an Ion Chromatograph and alkalinity (mg l^{-1}) by analytic titration to pH of 4.5 (APHA, 1995).

Data analysis

All environmental variables were tested for normality using the Kolmogorov - Smirnov normality test and the non-normal variables were transformed to approximate normality using a square root, logarithm or arcsine transformation (Zar, 1996). When normality was not obtained after such transformations, power transformations were tried until the best approximation to normality was achieved, as suggested by Zamora-Muñoz & Alba-Tercedor (1996) (see Table 1). These transformed data were used in all further analysis.

Taxa accounting for ≤0.01% of the total abundance in all samples were not considered in the data analysis because they tend to add unwanted noise to the classification analysis (Gauch, 1982; Rosenberg et al., 2000). The value of 0.01 was used after ensuring that the eliminated taxa were not abundant at any single site, and thus characteristic of a rare community. A fourth root transformation was applied to macroinvertebrate abundances in all analyses, except where otherwise indicated.

Selection and definition of reference sites

Reference sites were identified from their environmental characteristics. To ensure that when elevated levels of nutrients, or other variables, were observed they were due to natural conditions, the sites were grouped into their ten sub-catchments. For each sub-catchment, outlier sites (sites outside the grouping of the majority of sites of a sub-catchment in the plot) were identified using Principal Components Analysis (PCA) with PRIMER software (2001, version 5.2.6, PRIMER-E

Lda., Plymouth, UK; Clarke and Warwick, 2001) and then excluded from the reference site data set. In the PCA we only used variables susceptible to modification by human activity. These included: nitrates, nitrites, ammonia-N, phosphates-P, sulphates, chlorides and chlorophyll concentrations, pH, TDS, conductivity, alkalinity and % agriculture. As an additional criterion, sites with pH, conductivity, nitrates, nitrites and phosphate-P values outside the national accepted limits for waters for multiple uses (Instituto do Ambiente, 2003) were also rejected as reference sites. Sites that were not considered "reference" were used as "test sites".

Reference site grouping and model building

Identification of the biotic communities and building the predictive models that link habitat attributes to those groups is a two-step process. However, the process is iterative and continues until an optimal model is produced. In the first step, pattern analysis was used to investigate the biological structure of the data at the reference sites. The Bray-Curtis coefficient was the association measure used. An unweighted pair group method with arithmetic mean (CLUSTER analysis, group average linking, PRIMER software, 2001, version 5.2.6, PRIMER-E Lda, Plymouth, UK; Clarke & Warwick, 2001) was performed on the biotic data and, based on the dendogram, candidate groups of similar sites were identified using different cut levels of similarity. The consistency of the alternative groups were examined through additional analysis: the SIMPER procedure (PRIMER software) was used to check the similarity within and between groups in terms of invertebrate communities, as well as to examine the contribution of each taxa to the average Bray-Curtis similarity within a group (Clarke & Warwick, 2001). SIMPER assumes that fundamental information on the multivariate structure of an abundance matrix is summarized in the Bray-Curtis similarities between samples and it is by disaggregating these that one most precisely identifies the species responsible for particular aspects of the multivariate description (Clarke, 1993; Clarke & Warwick, 2001).

In the second step of the data analysis, the observed biological structure was related to the environmental characteristics using Discriminant Analysis. Variables most likely to be influenced by human activity or which are instantaneous measures were eliminated from the analysis: this included all water characteristics and dissolved substances, land use and riparian vegetation related variables and periphyton (see Table 1). Stepwise Forward Discriminant Analysis (Hair et al., 1998) with Jackknife cross-validation classification was performed in SYSTAT (1998, version 8.0, SPSS Inc. Chicago, IL.) with various combinations of groups (Alpha-to-Enter = 0.150 and Alpha-to-Remove = 0.150) to determine which variables best discriminated the selected reference biotic groups. The performance of a number of models using different site groupings and sets of predictor variables was evaluated based on how many sites were predicted to their appropriate reference group after Discriminant Analysis.

Assessment

The purpose of the model is to assess the quality of sites suspected of being disturbed. The assessment method used was the BEAST approach (BEnthic Assessment of SedimenT) developed in Canada by Reynoldson et al. (1995, 1997), which assumes that the quality of a test site is determined by the degree of similarity between a test site and the reference sites. So, first, test sites are assigned a probability of belonging to a reference group. This is done using the discriminant model developed in Step 2 (by running a complete Discriminant Analysis in SYSTAT) that predicts the reference site group to which each test site should belong, based on its environmental variables. The biotic data from each test site are then compared with the invertebrate assemblages of the reference site group to which it most probably belonged. This comparison is done using non-metric Multi Dimensional Scaling ordination (MDS in PRIMER software) of the reference sites and respective test sites. To obtain lower stress levels, three-dimensional MDS was used as recommended by Clarke & Warwick (2001); see

also Reece et al. (2001). The test sites were attributed to the worst position on the three plots (the greatest difference from reference, see as an example Fig. 3 in Results section).

To minimize the effect of the test site on the ordination space of the reference sites alone, test sites were plotted individually against reference sites. If the test site fell within the cluster of sites representing reference sites then it was considered to be equivalent to reference. If it fell outside the cluster it was considered to be different from reference (Reynoldson et al., 2001). To define bounds around the reference sites' cluster, 3 Gaussian bivariate probability ellipses (90, 99 and 99.9%, Altman, 1978; Owen & Chmielewski, 1985) are drawn in SYSTAT using the Scatterplot procedure (Reynoldson et al., 2001). The resulting ellipses are centered on the sample means of the x and y variables. The unbiased sample standard deviations of x and y determine its major axes and the sample covariance between x and y its orientation. The 90% ellipse contains (on average) 90% of the reference sites. The three ellipses describe four quality bands that define a gradient of impact based on reference sites. The test site quality is described in terms of location in these bands. Band 1, located inside the inner ellipse (90%), represents sites equivalent to reference, Band 2 represents sites possibly different from reference, Band 3 sites different from reference and Band 4, outside the third ellipse (99.9%), sites very different from reference (Reynoldson et al., 2000). Three representations of the ordination were produced for each test site in two dimensions (axis 1 vs. axis 2, axis 2 vs. axis 3 and axis 1 vs. axis 3). The test site was attributed to its worst position on the three plots, thus any conclusions as to the state of a test site were conservative.

Results

From the 75 sites sampled in the Mondego river basin during the summer 2001, we identified 207 taxa with a mean sample size of 1,579 individuals (ranging from 8 to 16,774 individuals). From the reference sites ($n = 55$; see below), 190 taxa were identified, representing 92% of the total number of taxa and 47% of the individuals sampled in the all basin. Ten taxa, from reference sites, alone account for 72% of the total number of animals sampled and 23 taxa accounted for 90% of all invertebrates. The remaining 167 taxa represented together only 10% of the individuals.

Eight sites were excluded from the reference data set based on Principal Component Analysis on the habitat data, by sub-catchment, as they were outside the cluster formed by the majority of the sites included in the same sub-catchment. The analysis of the Principal Component (PC) scores and the coefficients in the linear combinations of variables making up PC's revealed that those eight outliers had high levels of 1 or more of the variables likely of being modified by human activity (e.g. chloride, sulphate, ammonium). Twelve other sites were also excluded because of high levels of nitrates, other nutrients, or conductivity outside the national accepted limits for waters for multiple uses (Instituto do Ambiente, 2003).

Reference sites groups and model building

Based on CLUSTER analysis six combinations of 2–6 groups of reference sites were considered. Discriminant analysis (stepwise, forward) was used with these combinations: six groups (1a, 1b, 2a, 2b, 2c and 3), four groups option A (1a + 1b + 3, 2a, 2b and 2c), four groups option B (1a + 1b, 2a, 2b and 2c, group 3 was eliminated), three groups (1, 2 and 3), 2 groups option A (2a + 2b + 2c and 1a + 1b + 3) and two groups option B (2a + 2b + 2c and 1a + 1b, Group 3 was eliminated) (Fig. 2). Table 2 shows the predictive model performances from discriminant analysis and selected variables for each model.

From these six options, four of them were excluded for having small groups with only four (Group 3), five (Group 1a) and six sites (Group 2c), that should not be used in a model according to Reynoldson & Wright (2000), which recommended a minimum group size of ten sites. For the remaining two options, two groups A and two groups B, the SIMPER analysis showed similar results and higher similarities within the groups rather than between them (Table 3). However,
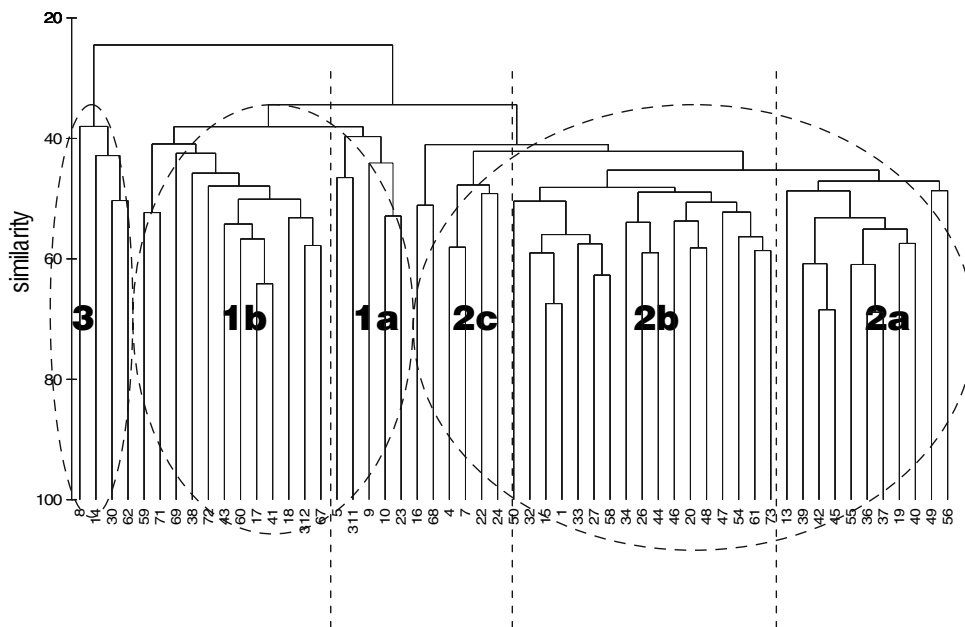
Fig. 2 Classification of reference sites according to their macroinvertebrate communities

the two group option B hypothesis was selected for further use as it was marginally the best combination, with slightly higher similarities among groups than the two groups option A. Therefore, 51 reference sites actually remained to be used in the model.

Table 2 Model performances (stepwise, forward, Discriminant Analysis, Systat 8.0) of 6 different hypotheses for grouping reference sites, based on biotic information and the best predicting variables in each case

Groups	Model performance (Jackknifed classification) (%)	Best predicting variables
2 groups option A (1 + 3, 2)	80	Stream order Substrate quality Pool quality Current velocity
2 groups option B (1, 2, without 3)	78	Stream order Substrate quality Pool quality Current velocity
3 groups (1, 2, 3)	73	Substrate quality Substrate size
4 groups A (1 + 3; 2a, 2b, 2c)	71	Substrate quality Stream width Precipitation (days)
4 groups B (without 3; 1, 2a, 2b, 2c)	71	Substrate quality Stream width Precipitation (days)
6 groups (1a, 1b, 2a, 2b, 2c, 3)	55	Distance to source Substrate quality Stream width Precipitation (days) Lithology

Table 3 Similarities within and between groups for the two best group combinations, using Simper analysis (The symbol – indicate comparisons between the same group, 100% similar, or repetitions)

Similarity within groups (%)	Groups	Similarity between groups (%)
2 groups A	2 groups A	G2
42	G1 + 3	34
50	G2	–
2 groups B	2 groups B	G2
45	G1	37
50	G2	–

Thus, and following two groups option B, the RCA model for the Mondego river basin, at the lowest taxonomic level, consisted of two reference biotic groups of 17 sites (group 1) and 34 sites (group 2). These groups were discriminated by the variables stream order, pool quality, current velocity and substrate quality, with an accuracy of 78%. The mean values of each of these variables for each group are presented in Table 4. Group 1 includes larger streams (greater stream order) with a good pool quality but less diversity of deep/shallow and large/small pools, lower current velocity and poorer substrate quality than group 2. In streams of group 1 there was a smaller number of taxa and, although most of the insect orders are represented, they are dominated by only one or two taxa (e.g. *Calamoceras marsupus* and *Mystacides azurea* for Trichoptera, *Oulimnius* for Coleoptera). Group 2 includes smaller streams with a high pool quality, greater substrate diversity and higher current velocities. In this group the classes and orders of the most representative taxa are similar to those of Group

Table 4 Mean values of the Mondego model discriminant values for each reference group

Habitat variables	Mean value	
	Group 1	Group 2
Stream order (1–4; 1:250,000 maps)	2.4	1.8
Pool quality (1–4 categories)	2.6	3.2
Current velocity (ms^{-1})	0.2	0.4
Substrate quality	2.8	3.6

1, but contain more taxa (e.g. there are five taxa of Trichoptera and Coleoptera).

Test site evaluation

Using the two-group model, the results of the complete Discriminant Analysis showed the probabilities of test site membership. Of 20 test sites, 11 sites had a higher probability of belonging to Group 1 and 9 sites to Group 2.

The probabilities of the sites belonging to a group were very high (79 and 82%, respectively) which makes us confident that test sites are being compared with the appropriate group of reference sites. Figure 3 shows one example of the BEAST plots. Of the 20 sites, of unknown ecological condition, six were evaluated as being unstressed, seven as being potentially stressed, three as stressed and four as severely stressed.

Discussion

Several factors can significantly affect the performance of a BEAST model, including the criteria and number of reference sites, sampling method and sampling season, taxonomic resolution, and the number of reference site groups established.

It is recommended that the population of reference sites represent the full range of conditions expected to occur naturally at all other sites to be assessed (Reynoldson & Wright, 2000). This was attempted in the Mondego River basin with a prior selection of sites covering altitudinal, geological and river order gradients. Nevertheless, the recognition of high quality sites was not always easy, therefore a final data examination using an iterative process was conducted and sites of dubious quality were eliminated from the reference set, as recommended by Reynoldson & Wright (2000).

The selected Mondego catchment predictive model was 78% accurate in predicting reference sites to a group. The accuracy of prediction is obviously partially dependent on the number of reference groups established in the classification. Determination of an appropriate number of groups is a balance between model accuracy

Fig. 3 Example of one test site assessment (site 28) with BEAST ellipses by comparison with reference
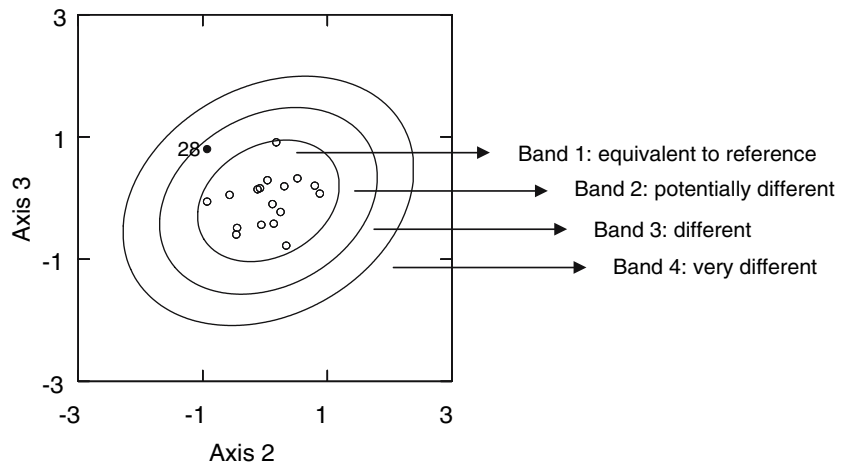
(i.e., the number of sites correctly predicted), capturing sufficient variation (i.e., a large enough number of sites in the group, to properly characterize the community), and sensitivity in detecting change. A model with a larger number of groups, besides resulting in groups with fewer sites, also has an increased difficulty in the correct prediction of a site membership. A model with a smaller number of groups means that each group has a higher intrinsic variability, limiting the sensitivity of the model (Reece et al., 2001).

The grouping process (required to use Discriminant anaysis) is one of the most subjective components of the modeling. While in some cases the groups are clear and distinct, with larger data sets groups are frequently more difficult to define as the sites tend to follow a biological continuum that represents the distributional ranges of the taxa forming the assemblage. RIVPACS/AUS-RIVAS models deal with this by estimating the expected probability of occurrence of a particular taxon as a weighted average of its probability of occurring in each of the possible classification groups (Clarke, 2000).

Recent investigations have looked at alternative approaches to the grouping process that have promise. Linke et al. (2005) and Bates Prins & Smith (2007) have used nearest neighbours in environmental ordination space to match reference and test sites. Bailey et al. (2006) have linked environmental data to biological ordination gradients to predict response to stress. These approaches do provide an alternative to the

somewhat artificial grouping procedure, and should be an area of future research in this field.

However, at this time, more traditional grouping methods were used in this study, and we consider the use of two groups appropriate as increasing the number of groups resulted in groups with few sites and the model performed well in the test sites assessments. Although, it should be noted that there is probably a valid third group, at this time, it is underrepresented and in the future more sites of this type should be sampled. In the case of the Mondego River, it is also likely that the low number of groups is the result of a study area much smaller than others used in former studies. For example, the Mondego catchment is 6670 km² compared to 250,000 km² for the Fraser River in BC which used six groups, and the RIVPACS study covering the UK (approx 240,000 km²) which used 35 groups. It is also possible that the small number of groups is a result of insufficient reference sites in the lower Mondego basin. Two sites of Group 3 belong to this area but there were few high quality sites available in this area to form a group. This absence of sites in the lower basin is also likely to compromise the effectiveness of the water quality evaluation for that area, since its characteristics, such as substrate type and lithology, are very different from the rest of the basin. The low sections of basins are frequently the areas where environmental quality may be compromised and require assessment, as they are generally more densely populated, with intensive

agriculture and frequently industrialized. Difficulties in finding reference sites for such areas have been discussed by other authors (Alba-Tercedor & Pujante, 2000; Reynoldson & Wright, 2000). The various alternatives that have been proposed include establishing "target conditions" instead of "reference conditions", where a "target" is defined as a condition that indicates the direction of improvement with respect to water management objectives (Verdonschot & Nijboer, 2000). This same approach is considered in the Water Framework Directive (Directive 2000/60/CE) for "heavily modified water bodies". Recently, Stoddard et al. (2006) have proposed the concept of Best Attainable Condition (BAC) for this type of expected condition, which can be defined as a state that is better than any in existence in a heavily modified region, but differs from either Minimally Disturbed Condition or the Reference Condition for Biological Integrity because those states might not be achievable. These methods all require the establishment of some type of hypothetical community that is perceived as an improvement of the present condition or, in the context of this approach, some location in the trajectory from current state to reference condition. Another alternative may be finding reference sites for the lower area in nearby river basins, although the other catchments often share similar problems in terms of land use.

From the test site evaluations, six sites were considered in reference condition, based on the macroinvertebrate assemblages, although they did not meet the environmental criteria we required to be considered a reference. This may mean that the environmental measures responsible for the exclusion of the test sites from the reference set reflected temporary conditions, or that the fauna was resistant to that specific type of impact. Consequently, although the test sites evaluated as being in good biological condition could theoretically be included in the model as reference sites, this decision should be taken very carefully. The model evaluations of the stressed sites reflected a decrease in species richness and evenness with the increasing band number, showing that the model is reflecting the stress incorporated in the invertebrate communities.

It should be born in mind that these models are not static and can be improved continuously. RIVPACS model development began in 1977 and went through many changes: the field protocol, the level of identification, the criteria used to define reference sites, the classification and prediction methods are some examples. The model described here for the Mondego River basin can and should be improved and it is only a first step. Increasing the number of reference sites in the Mondego basin, enlarging the sampling area to adjacent basins, and multiple year sampling of a small set of reference sites are the immediate next tasks. Moreover, among the four discriminant variables chosen for the Mondego model, three of them (pool quality, substrate quality and current velocity) should be altered in the future to include variables more resistant to human disturbances (as geological or geographic characteristics), although we considered that they did not correspond to the most common alterations in our streams, in opposition to nutrients enrichment and alterations in the riparian corridors. Further evaluation of the Mondego model should be undertaken using a diverse set of test sites in known and good ecological condition that were not used to build the model. At the same time, test sites of known poor quality as a result of pollution and/or the use of simulated stressed faunal assemblages will provide quantification of the magnitude and direction of the faunal response as well as of model sensitivity.

One theoretical limitation of the Reference Condition Approach in bioassessment is the unknown temporal stability of the reference sites. Therefore, Reynoldson & Wright (2000) recommend sampling a subset of reference sites to examine temporal change. Following this recommendation, the temporal variability of the fauna and the validity of the Mondego model predictions was analyzed in another publication and the conclusion indicate that, for models mainly built on species/genus level, data samples of new test sites should be collected in the same season as for the reference sites (Feio et al., 2006). Several other studies have examined this question and suggested the construction of seasonal models or the addition of samples from several seasons as an alternative for the limitation of the use of a

predictive model to a particular time of the year (Humphrey et al., 2000; Wright, 2000; Reece et al., 2001).

At present, RIVPACS is integrated in the Biological General Quality Assessment scheme in United Kingdom and is viewed as an important tool for the application of the WFD in the United Kingdom (Hemsley-Flint, 2000). Similarly, the Australian River Assessment Scheme (AUSRI-VAS) is applied to the entire Australian continent for routine river bioassessment (Simpson & Norris, 2000). The Mondego predictive model can also be readily modified to meet the requirements of the Water Framework Directive and to be used in the assessment of the ecological quality of Portuguese rivers. This would require rebuilding the model with a nation wide base collected for the purpose of the WFD integration in Portugal along with some methodological adaptations in the data analysis process.

Between the WFD and the BEAST approaches there are differences and similarities. The main similarity is the concept of Reference Condition, which is the basis of both approaches: the ecological status of a stream is obtained by comparing the biological community of that stream to the reference conditions (Sandin & Hering, 2004). Yet, differences arise when defining reference conditions. To describe reference conditions the WFD requires a typological framework (Directive 2000/60/CE). Each river has to be differentiated into a type, based on different environmental variables but always including altitude, catchment area and geology (Moog et al., 2004). For each type of river the reference condition is to be based on the aquatic community (fishes, macrobenthic fauna and aquatic flora). This will require validation of the initial typology with biological data collected from reference sites. A major question for ecologists is precisely whether these environmental descriptors have ecological significance (Moog et al., 2004; Verdonschot & Nijboer, 2004). For example, for the Mondego River catchment, even though altitude and geology (obligatory variables for the WFD typology) were candidate variables, none of them was selected as predictive variables of the final model. In the RCA approach the stream typology is also done during the classification step used to describe the structures of the biological data. However, this typology is based on the biological community and the environmental variables that best separate sites into the predefined groups are only obtained afterwards by Stepwise Discriminant Analysis (Reynoldson et al., 2001). This approach avoids the problem of fitting the communities into pre-established non-biological types. Yet, assuming that there is a good match between the classification of biological communities and WFD typological framework, those differences should not prevent the use of a predictive model, such as BEAST, since the WFD types can be used as groups.

Finally, the ecological status of a water body, according to the WFD, is classified into five quality classes: high, good, moderate, poor and bad (Directive 2000/60/CE; Sandin & Hering, 2004), while the BEAST ellipses, as currently used, define four categories (equivalent to reference, possibly different, different and very different). This methodological difference can be easily overcome by the addition of another probability ellipse, which will create a new band. Nevertheless, new limits would have to be tested with sites of known condition.

In conclusion, the Mondego model can be used to assess and monitor the catchment water quality, which do not exclude more future tests and improvements; the Mondego model can be modified to fit all the Water Framework requirements and its performance in the Mondego River catchment indicates that it could be adopted as a monitoring and assessment tool for Portuguese streams, based on aquatic macroinvertebrates; there is also the possibility of applying the methodology to the other quality elements used in WFD, such as fishes, or diatoms.

References

APHA, 1995. Standard methods for the examination of water and wastewater, 19th edn. American Public Health Association, Washington, D.C.

Alba-Tercedor, J. & A. M. Pujante, 2000. Running-water biomonitoring in Spain: opportunities for a predictive approach. In Wright, J. F., D. W. Sutcliffe & M. T. Furse (eds), RIVPACS and similar techniques for assessing the biological quality of freshwaters. Freshwater Biological Association and Environment Agency, Ambleside, Cumbria, UK, 207–216.

Altman, D. G., 1978. Plotting probability ellipses. Applied Statistics 27: 347–349.

Barbour, M. T. & C. Yoder, 2000. The multimetric approach to bioassessment as used in the United States of America. In Wright, J. F., D. W. Sutcliffe & M. T. Furse (eds), RIVPACS and similar techniques for assessing the biological quality of freshwaters. Freshwater Biological Association and Environment Agency, Ambleside, Cumbria, UK, 281–292.

Barbour, M. T., J. Gerritsen, B. D. Snyder & J. B. Stribling, 1999. EPA - Rapid bioassessment protocols for use in wadeable streams and rivers. Periphyton, Benthic Macroinvertebrates and Fish, 2nd Edition. EPA. 841-B-99–002. U.S. Environmental Protection Agency; Office of Water, Washington, D.C.

Bailey, R. C., T. B. Reynoldson, A. G.Yates, J. Bailey & S. Linke. 2006. Integrating stream bioassessment and landscape ecology as a tool for landuse planning. Freshwater Biology (OnlineEarly Articles). doi:10.1111/j.1365–2427.2006.01685.x.

Bates Prins, S. C. & E. P. Smith, 2007. Using biological metrics to score and evaluate sites. Freshwater Biology. Freshwater Biology 52: 98–111.

Clarke, K. R., 1993. Non-parametric multivariate analysis of changes in community structure. Australian Journal of Ecology 18: 117–143.

Clarke, R., 2000. Uncertainty in estimates of biological quality based on RIVPACS. In Wright, J. F., D. W. Sutcliffe & M. T. Furse (eds), RIVPACS and similar techniques for assessing the biological quality of freshwaters. Freshwater Biological Association and Environment Agency, Ambleside, Cumbria, UK, 39–51.

Clarke, K. R. & R. M. Warwick, 2001. Change in Marine Communities: An Approach to Statistical Analysis and Interpretation. 2nd Edition. PRIMER-e Ltd, Plymouth Marine Laboratory.

Directive 2000/60/EC, 2000. Water Framework Directive of the European Parliament and the Council, of 23 October 2000, establishing a framework for Community action in the field of water policy. Official Journal of the European Communities, L327: 1–72.

De Pauw, N. & G. Vanhooren, 1983. Method for biological quality assessment of water courses in Belgium. Hydrobiologia 100: 153–168.

Feio, M. J., T. B. Reynoldson & M. A. S. Graça, 2006. Effect of Seasonal Changes on Predictive Model Assessments of Streams Water Quality with Macro-invertebrates. International Review of Hydrobiology 91: 509–520.

Gauch, H. G., 1982. Multivariate analysis in community ecology. Cambridge, University Press, New York.

Hair, J. F., R. E. Anderson, R. L. Tatham & W. C. Black, 1998. Multivariate data analysis. 5th edn. Prentice-Hall International, Inc., U.S.A.

Hellawell, J. M., 1977. Change in natural and managed ecosystems - detection, measurement and assessment. Proceedings of the Royal Society: Biological Sciences 197: 31–57.

Hemsley-Flint, B., 2000. Classification of the biological quality of rivers in England, Wales. In Wright, J. F., D. W. Sutcliffe & M. T. Furse (eds), RIVPACS and similar techniques for assessing the biological quality of freshwaters. Freshwater Biological Association and Environment Agency, Ambleside, Cumbria, UK, 55–69.

Humphrey, C. L., A. W. Storey & L. Thurtell, 2000. AUSRIVAS: operator sample processing errors and temporal variability - implications for model sensitivity. In Wright, J. F., D. W. Sutcliffe & M. T. Furse (eds), RIVPACS and similar techniques for assessing the biological quality of freshwaters. Freshwater Biological Association and Environment Agency, Ambleside, Cumbria, UK, 144–163.

Instituto do Ambiente, 2003. Atlas Digital do Ambiente. In: http://www.iambiente.pt/atlas/est/index.jsp.

Instituto Geográfico do Exército, 1998a. Carta Militar de Portugal, 1:250 000, Coimbra.

Instituto Geográfico do Exército, 1998b. Carta Militar de Portugal, 1:250 000, Viseu.

Linke, S., R. H. Norris, D. P. Faith & D. Stockwell, 2005. ANNA: A new prediction method for bioassessment programs. Freshwater Biology 50: 147–158.

Marques, J. C., M. A. S. Graça & M. A. Pardal, 2002. Introducing the Mondego river basin. In Pardal, M. A., J. C. Marques & M. A. S. Graça (eds), Aquatic ecology of the Mondego river basin. Global importance of local experience. Universidade de Coimbra, Coimbra, Portugal, 7–12. .

Moog, O., A. Schmidt-Kloiber, T. Ofenböck & J. Gerritsen, 2004. Does the ecoregion approach support the typological demands of the EU "Water Framework Directive". Hydrobiologia 516: 21–33.

Owen, J. G. & M. A. Chmielewski, 1985. On canonical variates analysis and the construction of confidence ellipses in systematic studies. Systematic Zoology 34: 366–374.

Pardal, M. A., J. C. Marques & M. A. S. Graça, 2002. Aquatic ecology of the Mondego river basin. Global importance of local experience. Universidade de Coimbra, Coimbra, Portugal.

Parsons, M. & R. H. Norris, 1996. The effect of habitat-specific sampling on biological assessment of water quality using predictive model. Freshwater Biology 36: 419–434.

Pollard, P. & M. Huxham, 1998. The European Water Framework Directive: a new era in the management of aquatic systems health? Aquatic Conservation 8: 773–792.

Reece, P. F., T. B. Reynoldson, J. S. Richardson & D. M. Rosenberg, 2001. Implications of seasonal variation for biomonitoring with predictive models in the Fraser River catchment, British Columbia. Canadian Journal of Fisheries and Aquatic Sciences 58: 1411–1418.

Reynoldson, T. B., R. C. Bailey, K. E. Day & R. H. Norris, 1995. Biological guidelines for freshwater sediment based on Benthic Assessment of SedimenT (the BEAST) using a multivariate approach for predicting biological state. Australian Journal of Ecology 20: 198–219.

Reynoldson, T. B., R. H. Norris, V. H. Resh, K. E. Day & D. M. Rosenberg, 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. Journal of the North American Benthological Society 16: 833–852.

Reynoldson, T. B. & J. F. Wright, 2000. The reference condition: problems and solutions. In Wright, J. F., D. W. Sutcliffe, & M. T. Furse (eds), RIVPACS and similar techniques for assessing the biological quality of freshwaters. Freshwater Biological Association and Environment Agency, Ambleside, Cumbria, UK, 293–303.

Reynoldson, T. B., K. E. Day & T. Pascoe, 2000. The development of the BEAST: a predictive approach for assessing sediment quality in the North American Great Lakes. In Wright, J. F., D. W. Stucliffe & M. T. Furse (eds), RIVPACS and similar techniques for assessing the biological quality of freshwaters. Freshwater Biological Association and Environment Agency, Ambleside, Cumbria, UK, 165–180.

Reynoldson, T. B., D. M. Rosenberg & V. H. Resh, 2001. Comparisons of models predicting invertebrate assemblages for biomonitoring in the Fraser River catchment, British Columbia. Canadian Journal of Fisheries and Aquatic Sciences 58: 1395–1410.

Rosenberg, D. M. & V. H. Resh, 1993. Introduction to Freshwater Biomonitoring and Benthic Macroinvertebrates. 1–9. In Rosenberg, D. M. & V. H. Resh (eds), Freshwater Biomonitoring and Benthic Macroinvertebrates. Routledge, Chapman and Hall, Inc.

Rosenberg, D. M., T. B. Reynoldson & V. H. Resh, 2000. Establishing reference conditions in the Fraser River catchment, Canada, using BEAST (Benthic Assessment of SedimenT) predictive model. In Wright, J. F., D. W. Sutcliffe & M. T. Furse (eds), RIVPACS and similar techniques for assessing the biological quality of freshwaters. Freshwater Biological Association and Environment Agency, Ambleside, Cumbria, UK, 181–194.

Sandin, L. & D. Hering, 2004. Comparing macroinvertebrate indices to detect organic pollution across Europe: a contribution to the EC Water Framework Directive intercalibration. Hydrobiologica 516: 55–68.

Simpson, J. C. & R. H. Norris, 2000. Biological assessment of river quality: development of AUSRIVAS models and outputs. In Wright, J. F., D. W. Sutcliffe & M. T. Furse (eds), RIVPACS and similar techniques for assessing the biological quality of freshwaters. Freshwater Biological Association and Environment Agency, Ambleside, Cumbria, UK, 125–142.

Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson & R. H. Norris, 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications 16: 1267–1276.

Verdonschot, P. F. M. & R. C. Nijboer, 2000. Typology of macrofaunal assemblages applied to water and nature management: a Dutch approach. In Wright, J. F., D. W. Sutcliffe & M. T. Furse (eds), RIVPACS and similar techniques for assessing the biological quality of freshwaters. Freshwater Biological Association and Environment Agency, Ambleside, Cumbria, UK, 241–262.

Verdonschot, P. F. M. & R. C. Nijboer, 2004. Testing the European stream typology of the Water Framework Directive for macroinvertebrates. Hydrobiologia 516: 35–54.

Wright, J. F., D. Moss, P. D. Armitage & M. T. Furse, 1984. A preliminary classification of running-water sites in Great Britain based on macro-invertebrate species and the prediction of community type using environmental data. Freshwater Biology 14: 221–256.

Wright, J. F., 1995. Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. Australian Journal of Ecology 20: 181–197.

Wright, J. F., 2000. An introduction to RIVPACS. In Wright, J. F., D. W. Sutcliffe & M. T. Furse (eds), RIVPACS and similar techniques for assessing the biological quality of freshwaters. Freshwater Biological Association and Environment Agency, Ambleside, Cumbria, UK, 1–24.

Yoder, C. O. & E. T. Rankin, 1998. The role of biological indicators in a sate water quality management process. Environmental Monitoring and Assessment 51: 61–88.

Zamora-Muñoz, C. & J. Alba-Tercedor, 1996. Bioassessment of organically polluted Spanish rivers, using a biotic index and multivariate methods. Journal of the North American Benthological Society 15: 332–352.

Zar, J. H., 1996. Biostatistical Analysis. 3rd edn,. Prentice-Hall International, Inc., U.S.A.