Hydrobiologia (2006) 566:91–105

M.T. Furse, D. Hering, K. Brabec, A. Buffagni, L. Sandin & P.F.M. Verdonschot (eds), The Ecological Status of European Rivers: Evaluation and Intercalibration of Assessment Methods
DOI 10.1007/s10750-006-0068-5

A comparison of the European Water Framework Directive physical typology and RIVPACS-type models as alternative methods of establishing reference conditions for benthic macroinvertebrates

John Davy-Bowker[1,*], Ralph T. Clarke[1], Richard K. Johnson[2], Jiri Kokes[3], John F. Murphy[1] & Svetlana Zahrádková[4]
[1]*Centre for Ecology & Hydrology, Winfrith Technology Centre, Dorchester, Dorset DT2 8ZD, United Kingdom*
[2]*Department of Environmental Assessment, Swedish University of Agricultural Sciences, P.O. Box 7050 SE-750 07 Uppsala, Sweden*
[3]*T.G.M. Water Research Institute, Drevarska 12, 657 57 Brno, Czech Republic*
[4]*Department of Zoology and Ecology, Faculty of Science, Masaryk University Brno, Kotlářská 2, 611 37 Brno, Czech Republic*
*(*Author for correspondence: E-mail: jobo@ceh.ac.uk)*

Key words: reference condition, physical typology, RIVPACS, $\mathrm{SWEPAC_{SRI}}$, PERLA

Abstract

The EU Water Framework Directive requires European Union Member States to establish 'type-specific biological reference conditions' for streams and rivers. Types can be defined by using either a fixed typology (System-A), defined by ecoregions and categories of altitude, catchment area and geology, or by means of an alternative characterisation (System-B) that can use a variety of physical and chemical factors. Several European countries also have existing RIVPACS-type models that give site (rather than stream type) specific predictions of benthic macroinvertebrate communities. In this paper we compare the Water Framework Directive (WFD) System-A physical typology and three existing European multivariate RIVPACS-type models as alternative methods of establishing reference conditions. This work is carried out in Great Britain – using RIVPACS, Sweden – using $\mathrm{SWEPAC_{SRI}}$ and the Czech Republic – using PERLA. We found that in all three countries, all seasons and season combinations, and for all biotic indices tested, RIVPACS-type models were more effective (lower standard deviations of O/E ratios) than models based solely on the WFD System-A variables or null models (based on a single expectation for all sites). We also investigated the explanatory power of whole groups of WFD System-A variables and RIVPACS-type model variables, and the explanatory power of individual variables. We found that variables used in the RIVPACS-type models were often better correlates of macroinvertebrate community variation than the WFD System-A variables. We conclude that this is primarily because while the latter use very broad categories of map-derived variables, the former are based on continuous variables selected for their ecological significance.

Introduction

The EU Water Framework Directive (Council of the European Communities, 2000), hereafter referred to as the WFD, requires Member States of the European Union (EU) to assess, monitor and, where necessary, improve the ecological quality status of its surface waters. This landmark piece of environmental legislation seeks to achieve at least 'good ecological status' for all surface water by 2015 and, for the first time, recognises the importance of the aquatic biota in determining the quality of fresh and marine waters (Sweeting, 2001; Logan & Furse, 2002). In placing the aquatic biota at the forefront of European environmental assessment the WFD recognises the importance of biogeographical drivers of species distribution patterns in setting targets for the biota (e.g., Illies, 1978). To achieve this, the WFD has established a hierarchical water body typology. Within any given part of this typology, it is assumed that biological communities at undisturbed sites will be broadly similar and will therefore constitute a type-specific biological target. The WFD typology (WFD, Annex II, Section 1) is organised by firstly placing surface water bodies into broad categories (rivers, lakes, transitional/coastal waters, artificial water bodies or heavily modified water bodies), and secondly, within these categories, by differentiating water bodies into types. This is achieved by using either a fixed typology, 'System-A', which in the case of rivers categorises sites based on ecoregion, altitude, catchment area and geology (Table 1), or an alternative typology, 'System-B', comprising a mixture of obligatory and optional factors.

In contrast to the *a priori* stream typologies set out in the WFD, RIVPACS-type predictive models are not based on predefined physical categories. Indeed, the site classification step within the development of a RIVPACS-type model makes no reference to physical variables (Clarke et al., 2003). Reference sites (sites considered to be of high ecological and physicochemical quality) that have been selected to encompass the full range of river types within a geographical area, are first classified into groups based solely on their macroinvertebrate fauna. Secondly, discriminant analysis is used to derive predictive equations that relate a range of recorded environmental variables to the biological classification. New sites are tested by applying the discriminant equations to the environmental variables recorded at the test sites. The macroinvertebrate communities expected to occur in the absence of environmental stress are then predicted (the expected fauna, E). The observed macroinvertebrate fauna (O) from test sites can then be compared to the expected fauna (E) by calculating observed/expected (O/E) ratios for a biotic index (e.g., O/E Number of Taxa). These are equivalent to the Ecological Quality Ratios (EQRs) described in the Water Framework Directive (WFD, Annex V, Section 1.4.1.ii). RIVPACS-type models therefore differ from *a priori* typologies in the use they make of physical data. Firstly, and perhaps most fundamentally, they use only biological data for site classification. Secondly they do not make *a priori* judgements about which variables are good correlates of community composition. And thirdly, RIVPACS-type models utilise multiple environmental variables (including local site characteristics such as substrate composition) to reveal correlations with macroinvertebrate communities rather than a restricted set of large-scale bio-geographical or physical factors.

The WFD requires EU Member States to establish 'type-specific biological reference conditions' for each water body type (WFD, Annex II, Section 1.1.iv), where reference condition equates to the definitions of high status in Annex V of the Directive. The choice of using either a System-A or System-B typology is left for individual Member States to decide. However, if choosing System-B, 'Member States must achieve at least the same degree of differentiation as would be achieved using System-A' (WFD, Annex II, Section 1.1.iv). In this paper we compare the relative effectiveness of the WFD System-A and RIVPACS-type multivariate models as alternative approaches for setting biological reference states. This comparison is made in three EU Member States, namely, Great Britain – using RIVPACS (River InVertebrate Prediction And Classification System), Sweden – using $SWEPAC_{SRI}$ (Swedish Prediction And Classification system for Stream Riffle Invertebrates), and the Czech Republic – using PERLA (after the Plecoptera genus).

Materials and methods

An assessment of biotic community variance within the WFD system-A typology and existing RIVPACS-type models was performed by assessing their ability to predict the observed values of biotic indices for the reference sites (Fig. 1) in all separate seasons and all combinations of seasons. We also sought to explore the relative explanatory power of the environmental variables (as correlates of community composition) in the WFD

Table 1. Distribution of RIVPACS-type model sites within WFD System-A stream types (as a percentage within each Country)

	Country	Great Britain RIVPACS III+			Sweden SWEPAC$_{SRI}$			Czech Republic PERLA											
	Ecoregion	18			14, 20, 22			9			10			11			14		
	# Sites in model	$n=614$			$n=389$			$n=238$			$n=33$			$n=8$			$n=21$		
Site altitude (m)	Catchment size (km^2)	*Geology* Siliceous	Calcareous	Organic	Siliceous	Calcareous	Organic	Siliceous	Calcareous	Organic	Siliceous	Calcareous	Organic	Siliceous	Calcareous	Organic	Siliceous	Calcareous	Organic
< 200	10–100	17.3	22.0	0.8		1.0	41.4	0.3											
	> 100–1000	14.5	26.2	1.8		0.5	6.0	0.3											
	> 1000–10,000	2.4	6.0					0.3											
	> 10 000							0.3						0.7					
200–800	10–100	2.6	3.9	1.0	11.1	5.1	24.9												
	> 100–1000	1.0	0.3	0.2	3.3	1.3	5.1	26.9	3.0		0.7	6.0			0.3		3.7	0.7	
	> 1000–10,000							20.5	1.3		0.3	2.3		1.0	0.3		1.7		
	> 10 000							20.9	0.3		1.0			0.3			0.7		
> 800	10–100					0.3		4.3	0.3		0.7						0.3		
	> 100–1000																		
	> 1000–10,000							0.3											
	> 10 000							0.3											

Ecoregion 9, Central highlands; 10, The Carpathians; 11, Hungarian lowlands; 14, Central plains; 18, Great Britain; 20, Borealic uplands and; 22, Fenno-Scandian shield.

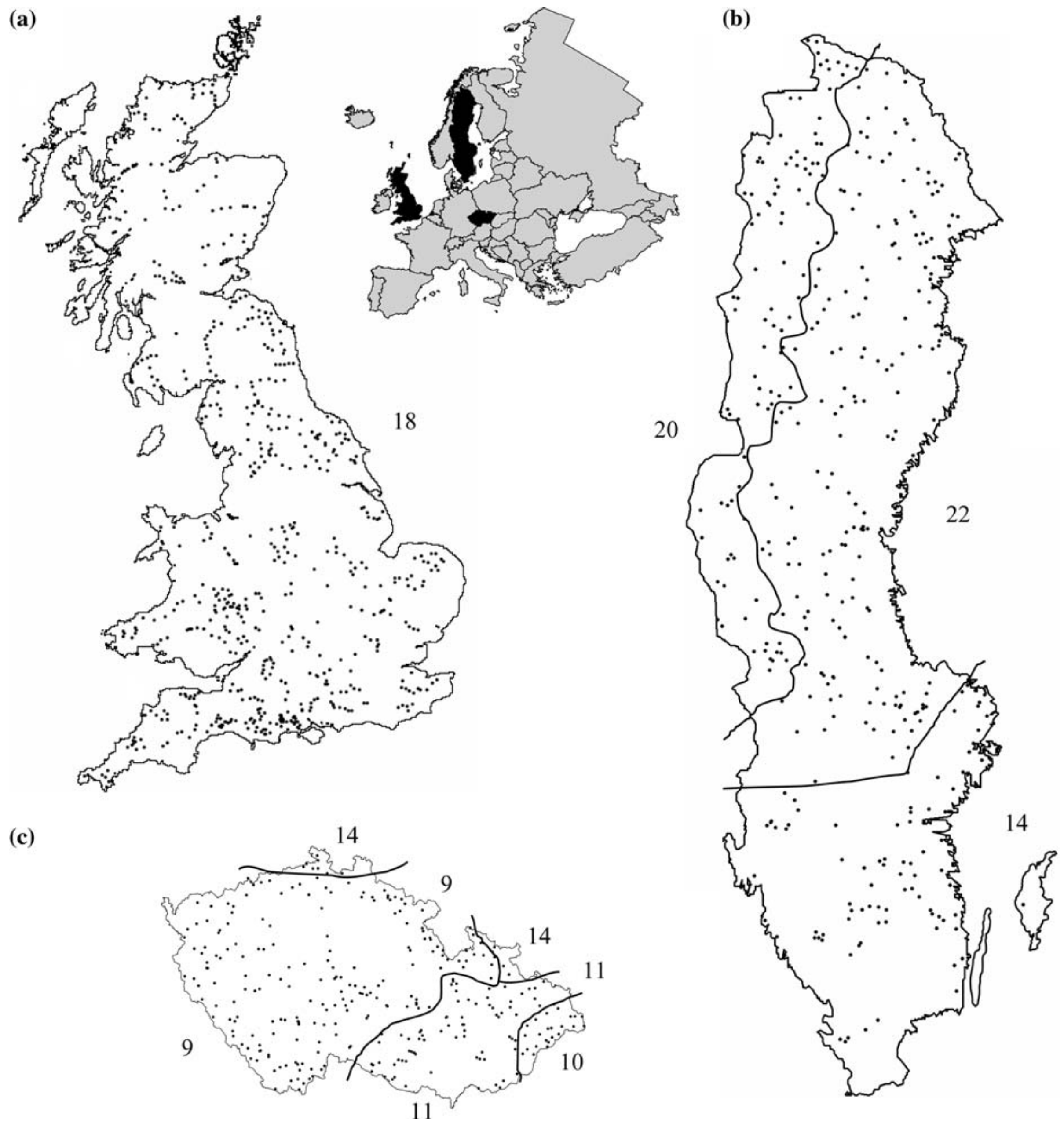

Figure 1. Geographical distribution of (a) the 614 RIVPACS III+ reference sites in Great Britain (ecoregion 18), (b) the 389 $SWEPAC_{SRI}$ reference sites in Sweden (ecoregion 14, Central plains; 20, Borealic uplands; 22, Fenno-Scandian shield), and (c) the 300 PERLA reference sites in the Czech Republic (ecoregion 9, Central highlands; 10, The Carpathians; 11, Hungarian lowlands; 14, Central plains).

System-A and RIVPACS-type models. Our analysis was done in four stages:

1. Calculation of expected biotic index values (WFD System-A typology)
2. Calculation of expected biotic index values (RIVPACS-type models)
3. Assessment of relative prediction accuracy
4. Analysis of correlates of community composition.

These stages are described below.

Expected biotic index values (WFD system-A typology)

The WFD system-A typology classifies rivers across Europe into 25 ecoregions, three altitude categories (< 200 m, 200–800 m and > 800 m), four catchment size categories (10–100 km^2, > 100–1000 km^2, > 1000–10,000 km^2 and $> 10{,}000$ km^2) and three geology categories (siliceous, calcareous and organic). There are some uncertainties in the application of the System-A typology to streams and rivers. Firstly, the ecoregions set out in the WFD are defined at a very crude scale making the interpretation of local boundaries between ecoregions difficult. Secondly, it is unclear whether the geological class should be based on the geology underlying the biological sampling site itself or the geology of the upstream catchment, and thirdly there appears to be no legislative provision for streams with catchment areas less than 10 km^2; an important source of biodiversity (Furse, 2000).

In Great Britain all sites were assigned to the WFD ecoregion 18 'Great Britain.' WFD System-A altitude categories were taken as the altitude at each RIVPACS reference site. Under licence from the Centre for Ecology and Hydrology (CEH), the Environment Agency and the Scottish Environmental Protection Agency used the CEH Intelligent River Network (Dawson et al., 2002) to determine WFD catchment size categories. The agencies derived WFD geology categories by overlaying the RIVPACS reference sites on the British Geological Survey 1:625,000-scale solid geology GIS map, defining geology categories from the geology in the immediate vicinity of each site. The WFD geology classes where categorised as 'calcareous' where the bedrock was wholly or partially composed of calcium carbonate, 'siliceous' where the bedrock was acid igneous or there was other bedrock that did not contain calcium carbonate, or 'organic' where surface deposits were composed primarily of peat. In Sweden, the Swedish Environmental Protection Agency is responsible for implementing the WFD, although much of the work has been subcontracted to academics and consultants. A Nordic-funded project is also responsible for harmonising work among the Nordic countries (Johnson et al., 2001). A System-A typology-based classification is presently being used, consisting of ecoregion delineation, altitude, catchment size and geology classification. Following the recommendations of the WFD, Illies ecoregions (Illies, 1978) are used (regions 14, 20 and 22) and site altitude and catchment size are classified into three and four WFD classes, respectively. In Sweden, geological categories are defined as siliceous (alkalinity < 0.2 meq/l), calcareous (alkalinity 0.2–1.0 meq/l) and organic (absorbance 420 nm of filtered water $> 0.06 \approx 30$ mgPt/l). In the Czech Republic, individual PERLA sites were categorised as either calcareous or siliceous depending upon the calcareous/siliceous category of the overall water body (designated by the Czech Geological Institute, where calcareous rock types were defined as those with an equivalent content of alkaline elements (Na, K, Mg, Ca) > 5.2, and water bodies with $> 40\%$ calcareous rock types in their basin were categorised as calcareous).

Expected biotic index values based on the WFD System-A typology was derived by averaging index values for all references sites in each WFD System-A group. Using this method, the biotic indices Number of Taxa (number of BMWP scoring families) and ASPT (average score per taxon of BMWP families) (National Water Council, 1981) were derived for all reference sites. Number of Taxa and ASPT were chosen because while these are the indices currently used to report the biological quality of streams and rivers in Great Britain, they can be easily applied to all three countries. Additionally, in Great Britain, the Lotic Invertebrate Index for Flow Evaluation (LIFE), an index describing macroinvertebrate sensitivity to low flow (Extence et al., 1999), was predicted for spring summer and autumn samples, and in the Czech Republic, a Saprobic index was predicted (based on an index originally developed by Pantle & Buck (1955)) with the addition of taxon weightings (Marvan, 1969) and using standard Saprobic values and weights (CSN 75 7716, 1998). These indices were included to extend the generality of our analyses beyond the indices typically used when comparing the relative effectiveness of different models. Indices were calculated separately for spring, summer and autumn samples and (where models permitted) for all seasonal combinations of samples (spring and summer;

spring and autumn; summer and autumn; and spring, summer and autumn combined).

Expected biotic index values (RIVPACS-type models)

The current version of RIVPACS (RIVPACS III+) is based 614 reference sites. For a full review of RIVPACS see (Wright et al., 1984; Moss et al., 1987; Wright, 2000; Clarke et al., 2003). RIVPACS predicted biotic index values were obtained by running RIVPACS III+ on the 614-reference site dataset (generating predicted faunal lists and biotic index values for each reference site). Number of Taxa and ASPT were predicted for spring, summer and autumn and for all season combinations (spring and summer; spring and autumn; summer and autumn; and spring, summer and autumn combined). The LIFE index was predicted for spring, summer and autumn samples.

$SWEPAC_{SRI}$ models (Johnson, unpublished) were calibrated following the procedure outlined in Johnson & Sandin (2001), but combining data from all three ecoregions (14 – Central Plains; 20 – Borealic Uplands and 22 – Fenno-Scandian Shield). Data consisted of benthic macroinvertebrate samples collected in the year 2000 national stream survey (Wilander et al., 2003). Sites deemed to be affected by liming, agriculture ($>10\%$ agricultural land use and $TP > 8$ or 10 $\mu g/l$ – compensated for humic P) or acidification ($pH \leq 6.0$ – compensated for natural acidity) were removed from the dataset resulting in 389 'unperturbed' sites distributed across the country. Predicted faunal lists and biotic indices were calculated using the combined $SWEPAC_{SRI}$ model. $SWEPAC_{SRI}$ predicted Number of Taxa and ASPT values were calculated for autumn samples ($SWEPAC_{SRI}$ is currently based on autumn samples).

The PERLA predictive system (Kokes et al., 2006) is based on 300 reference sites throughout the Czech Republic and is programmed into the HOBENT software package. Chironomidae identifications were not available for all PERLA reference sites in summer and autumn. Chironomidae were therefore excluded from the PERLA predictive model in summer and autumn, although they were included in the spring model used in our analyses. Number of Taxa, ASPT and the Saprobic index were predicted for the 238 reference sites in Ecoregion 9 (Central highlands). These were predicted separately for spring, summer and autumn samples.

Assessment of relative prediction accuracy

The effectiveness of the WFD System-A typology was then compared directly with the RIVPACS-type model predictions of expected index values. Prediction accuracy was measured as the standard deviation (SD) of the ratios of the observed (O) to expected (E) values of each biotic index for the reference sites. Van Sickle et al. (2005) introduced the idea of a null model in which the predicted reference condition index value for a site is the average observed value of the index for all reference sites. The SD of O/E, hereafter denoted SD(O/E), based on such as null model provides a useful upper limit to the SD(O/E) based on any model (an effective model should achieve a lower SD(O/E) than that of the null model). The relative sizes of the SD(O/E) enable a comparison to be made of the relative effectiveness of the WFD System-A typology and the RIVPACS-type models in terms of their ability to predict the observed values of the biotic indices and hence define reference conditions.

Analysis of correlates of community composition

After assessing the relative performance of the WFD System-A and RIVPACS-type models, we also sought to identify which of the environmental variables (both collectively and individually) used in the WFD System-A and RIVPACS-type models were the best correlates of macroinvertebrate community composition across the reference sites. This was investigated with canonical correspondence analysis (CCA) using the CANOCO 4.5 software package (ter Braak & Smilauer, 2002). The biological data sets used in our analyses were prepared in the same way as those used in the biological classifications underpinning each of the RIVPACS-type models. For RIVPACS III+, spring, summer and autumn data were combined into a 3 seasons combined dataset where family ($\log_{10}$ categories) and species (presence/absence) records included any of the species and families occurring in any separate season. Family $\log_{10}$ abundances were taken as the maximum $\log_{10}$

abundance (except where all three $\log_{10}$ abundances were the same, in which case the $\log_{10}$ abundance was increased by 1 category). The $SWEPAC_{SRI}$ biological classification was based on autumn sample, species level, presence/absence data while that for PERLA was based spring sample, species level, abundance data. Preliminary detrended correspondence analyses (DCA) revealed that the rate of turnover of macroinvertebrate taxa across the sites on the first axis of variation was >3 (DCA axis 1 length 3.01, 3.16 and 4.47 for RIVPACS, $SWEPAC_{SRI}$ and PERLA, respectively). The unimodal model within CCA was therefore considered to be appropriate for use with these datasets (ter Braak & Prentice, 1988). Within each country we determined the proportion of variation in the biotic data that could be accounted for by:

(1) The WFD System-A environmental variables as a whole
(2) The RIVPACS-type model environmental variables as a whole
(3) Both sets of variables combined.

This was calculated as the sum of all conditional effect eigenvalues (the collective contribution of the first variable and each successive variable to a forward selection model) as a proportion of total inertia (the total extent of variation in the macroinvertebrate communities across the reference sites). Three further CCAs (using both sets of variables combined) were used to determine the individual explanatory power of each environmental variable. These were the marginal effect eigenvalues, i.e., the variance explained when a particular variable is the only explanatory variable (ter Braak & Smilauer, 2002).

Results

In all three countries the reference sites were unevenly distributed throughout the WFD System-A stream types (Table 1). In Great Britain 91% of the reference sites were below 200 m altitude and 92% had catchment areas of 1000 km^2 or less. In terms of geology, 58% were calcareous and 38% were siliceous while only 4% were organic. In Sweden the overall percentage of geologically organic sites was much higher (77%) and calcareous sites were much less common (8%). The Swedish reference sites were distributed evenly between WFD altitude categories <200 m (49%) and 200–800 m (51%) and mainly drained 10–100 km^2 catchments (83%), with no sites from catchments above 1000 km^2. In the Czech Republic the reference sites were distributed across four ecoregions with 79% in ecoregion 9 (the central highlands) and only 21% in the other three ecoregions combined. While the Czech reference sites were distributed over a fairly broad range of catchment area categories (only catchment area 10–100 km^2 having a low number of sites), the Czech sites had altitudes predominantly between 200 and 800 m (92%), and were mainly siliceous (85%). In general, the matrix of possible WFD stream types (Table 1) contained many blank cells where combinations of ecoregion, altitude, catchment size and geology were not represented in each country. Where WFD System-A stream types were represented, the proportion of sites in each type was often imbalanced towards a few predominant types.

For all countries (Great Britain, Sweden and the Czech Republic), all indices (Number of Taxa, ASPT, LIFE and Saprobic) and for all seasons and season combinations, the SD(O/E) ratios were consistently highest for the null models, indicating relatively high uncertainty, and consistently lowest for the RIVPACS-type models, indicating relatively lower uncertainty (Table 2). In Great Britain, SD(O/E) ratios were lowest in the combined season RIVPACS models compared to the separate season models. For all three-model types (null, WFD System-A and RIVPACS-type models) the SD(O/E) ratios were lower for ASPT than for Number of Taxa (Fig. 2). In Great Britain, the percent reductions in SD(O/E) for the LIFE index were also greater than (or in one case equal to) those for ASPT. In Sweden there was a greater reduction in SD(O/E) for ASPT than for Number of Taxa. In the Czech Republic, while the percentage reductions in SD(O/E) were always greatest for PERLA, the reduction in ASPT was greatest in the spring (22%), but not so great in summer (4%) or autumn (7%). This could be because Chironomidae, which make up 108 of the 564 taxa in the spring dataset, were included in the spring model but excluded from the summer and autumn models. Although all the species in the

Table 2. Standard deviations of the observed/expected ratios of biotic indices for RIVPACS-type model reference sites based on null models, WFD System-A models and the RIVPACS-type models used in Great Britain, Sweden and the Czech Republic (% reduction in SD(O/E) compared to null models in parentheses)

Country ecoregion index	Season	Prediction method		
		Null model	WFD System-A	RIVPACS type model
Great Britain				
Ecoregion 18 (Great Britain)				RIVPACS III +
No. of Taxa	Spring	0.244	0.231 (5)	0.198 (19)
	Summer	0.255	0.243 (5)	0.196 (23)
	Autumn	0.263	0.251 (5)	0.218 (17)
	Spr + Sum	0.205	0.193 (6)	0.156 (24)
	Spr + Aut	0.206	0.193 (6)	0.156 (24)
	Sum + Aut	0.211	0.200 (5)	0.161 (24)
	Spr + Sum + Aut	0.188	0.176 (6)	0.138 (27)
ASPT	Spring	0.125	0.107 (14)	0.075 (40)
	Summer	0.122	0.109 (11)	0.081 (34)
	Autumn	0.132	0.113 (14)	0.086 (35)
	Spr + Sum	0.109	0.094 (14)	0.062 (43)
	Spr + Aut	0.109	0.096 (12)	0.064 (41)
	Sum + Aut	0.112	0.096 (14)	0.067 (40)
	Spr + Sum + Aut	0.103	0.088 (15)	0.057 (45)
LIFE	Spring	0.081	0.072 (11)	0.048 (41)
	Summer	0.091	0.085 (7)	0.059 (35)
	Autumn	0.085	0.077 (9)	0.055 (35)
Sweden				
Whole-country model				$SWEPAC_{SRI}$
No. of Taxa	Autumn	0.355	0.345 (3)	0.312 (12)
ASPT	Autumn	0.139	0.122 (12)	0.102 (27)
Czech Republic				
Ecoregion 9 (Central highlands)				PERLA
No. of Taxa	Spring	0.337	0.295 (12)	0.221 (34)
	Summer	0.335	0.289 (14)	0.265 (21)
	Autumn	0.383	0.317 (17)	0.276 (28)
ASPT	Spring	0.096	0.093 (3)	0.075 (22)
	Summer	0.081	0.079 (2)	0.078 (4)
	Autumn	0.091	0.089 (2)	0.085 (7)
Czech Saprobic	Spring	0.407	0.336 (17)	0.225 (45)
	Summer	0.461	0.345 (25)	0.266 (42)
	Autumn	0.452	0.382 (15)	0.279 (38)

family Chironomidae have the same BMWP score (2), the Chironomidae species differ in their sensitivity to organic pollution (Armitage & Blackburn, 1985), and hence models including the Chironomidae species in their biological classification may be better able to distinguish sites that differ in their natural nutrient levels.

While the environmental variables used by the three RIVPACS-type predictive models differ (Table 3), several variables (or types of variables), either in their $\log_{10}$ transformed or untransformed form, are used in either two or all three of the models (latitude, longitude, altitude, air temperature, distance from source, slope, depth, width,

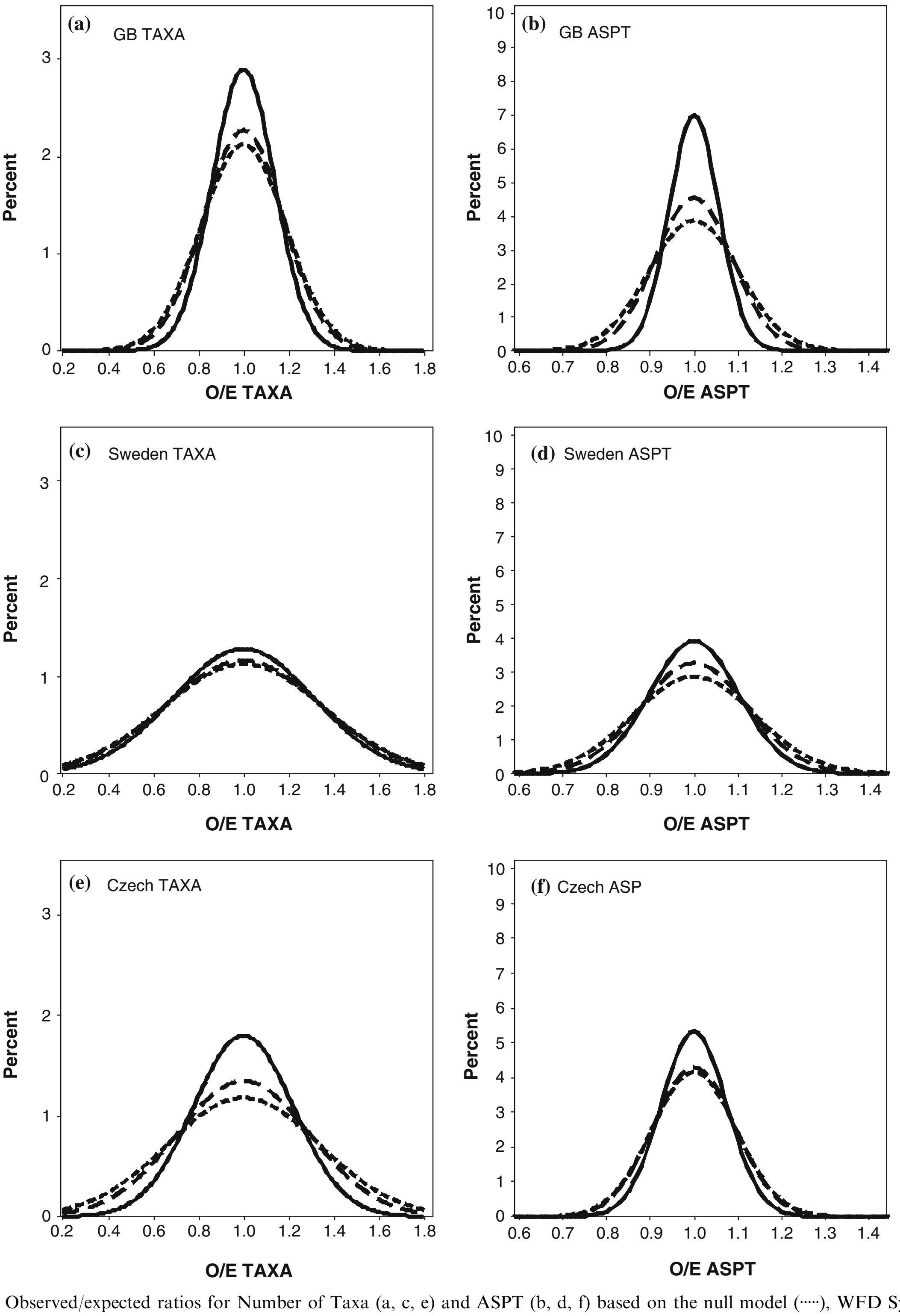

Figure 2. Observed/expected ratios for Number of Taxa (a, c, e) and ASPT (b, d, f) based on the null model (·····), WFD System-A typology (– – – –) and RIVPACS-type models (——). O/E Ratios based on spring summer and autumn data combined data for Great Britain, autumn data for Sweden and spring data for the Czech Republic.

substratum, discharge/velocity and alkalinity). In RIVPACS (and in the case of $\bar{x}$ catchment air temperature in $SWEPAC_{SRI}$) several variables were used in their $\log_{10}$ form as this has been found to improve their strength as correlates of macroinvertebrate community composition.

In all three countries the variables used by the RIVPACS-type predictive models could explain a larger proportion of the variation in macroinvertebrate communities than the WFD System-A variables (Table 4). In Great Britain and the Czech Republic, in the combined analyses of WFD System-A and RIVPACS-type model variables, the proportion of variance explained was slightly higher than that achieved by the RIVPACS-type model variables alone. This suggests that while the RIVPACS-type model variables are more effective as environmental predictors, the WFD System-A variables may be contributing a small amount of unique explanatory power not already encapsulated within the variables used by these RIVPACS-type models. In Sweden, the combined analyses of WFD System-A and RIVPACS-type model variables explained 14.2% of the total variation in community data compared to 8.0 and 9.9% for the WFD System-A and $SWEPAC_{SRI}$ model variables, respectively. This suggests that the System-A variables and the $SWEPAC_{SRI}$ variables are somewhat more distinct in the aspects of community variation they describe.

The individual explanatory power of each environmental variable (from a pool of RIVPACS-type model variables and WFD System-A variables within a given country) is shown in Table 5. In Great Britain, the RIVPACS variables were without exception better correlates of macroinvertebrate community composition than the WFD System-A variables. In several cases the greater explanatory power of a RIVPACS variable versus its equivalent WFD System-A variable was probably because that variable (or a closely related variable) was used in continuous rather than categorical form (e.g., the RIVPACS continuous variable $\log_{10}$ altitude versus WFD System-A altitude category, and RIVPACS $\log_{10}$ distance from source (which is highly correlated with catchment size, r_s 0.806, $p < 0.001$) versus WFD System-A catchment size category).

In Sweden the System-A variable ecoregion (which divides northern Sweden into the Borealic

Table 3. Environmental variables used in RIVPACS-type models in Great Britain, Sweden and the Czech Republic (•, untransformed; □, $\log_{10}$ transformed)

Variable type	Environmental variable	PERLA Czech Republic	RIVPACS III+ Great Britain	$SWEPAC_{SRI}$ Sweden
Geographical	Latitude	•	•	
	Longitude	•	•	
Altitude	Altitude	•	□	
Meteorological	$\bar{x}$ Air temperature		•	
	$\bar{x}$ Catchment air temperature			□
Catchment dimensions	Distance from source	•	□	
	Catchment area size			•
Gradient	Slope at site	•	□	
Channel dimensions	$\bar{x}$ Water depth	•	□	
	$\bar{x}$ Water width	•	□	•
Substratum	$\bar{x}$ Substratum composition	•	•	
	% Fine sediment			•
	% Floating-leaved vegetation			•
Hydrological	River discharge category		•	
	Stream velocity			•
Alkalinity	Alkalinity		•	•
	Alkalinity		□	
Catchment vegetation	% Forest in catchment			•

Table 4. Percentage of the total variation across macroinvertebrate communities explained collectively (conditional effects) by the WFD System-A variables, the RIVPACS-type model variables, and both groups of variables combined

	WFD System-A variables	RIVPACS-type model variables	WFD System-A and RIVPACS-type model variables combined
Great Britain	7.2	18.9	20.0
Sweden	8.0	9.9	14.2
Czech Republic	4.3	11.3	12.6

Uplands (20) in the west and the Fenno-Scandian Shield (22) in the east, and separates southern Sweden as the Central Plains (14) – Fig. 1) was a particularly strong correlate of macroinvertebrate community variation. While our analyses of reference sites in Great Britain and the Czech Republic did not traverse ecoregion boundaries, in Sweden these boundaries appear to have useful ecological meaning. The WFD System-A variable altitude was also a good descriptor. However, unlike the RIVPACS-type models in Great Britain and the Czech Republic, the continuous variables latitude, longitude and altitude are not part of the combined $SWEPAC_{SRI}$ model. The high descriptive power of the categorised ecoregion and altitude variables suggest that ecoregion and altitude

Table 5. Percentage of the total variation across macroinvertebrate communities explained (% VE) when a particular variable is the only explanatory variable (marginal effect). The % VE for all WFD System-A and RIVPACS-type predictive model variables are presented

Great Britain – Ecoregion 18	% VE	Sweden – Ecoregions 14, 20 & 22	% VE	Czech Republic – Ecoregion 9	% VE
Alkalinity	7.0	WFD Ecoregion	8.4	Distance from source	3.4
$\bar{x}$ Substratum composition	6.4	WFD Altitude category	3.9	WFD Catchment size category	3.2
Log_{10} alkalinity	5.9	Percent fine sediment	3.2	$\bar{x}$ Water width	2.8
Log_{10} Slope	5.9	WFD Geology – organic category	2.3	$\bar{x}$ Water depth	2.7
Longitude	5.4	$\bar{x}$ Water width	2.0	Slope	2.6
Log_{10} distance from source	4.3	Stream velocity	1.8	Altitude	2.1
Log_{10} Altitude	3.7	Percent floating-leaved vegetation	1.6	$\bar{x}$ Substratum composition	1.4
Log_{10} $\bar{x}$ water depth	3.7	WFD Geology – siliceous category	1.6	Longitude	1.4
Latitude	3.7	Alkalinity	1.4	Latitude	1.3
Log_{10} $\bar{x}$ water width	3.2	Percent forest in catchment	1.1	WFD Altitude category	0.6
River discharge (flow) category	3.2	Log_{10} $\bar{x}$ catchment air temperature	1.0	WFD Geology – calcareous category	0.4
$\bar{x}$ Air temperature	3.2	WFD Geology – calcareous category	0.7	WFD Geology – siliceous category	0.4
WFD Catchment size category	2.7	Catchment size	0.4	WFD Geology – organic category	0.0
WFD Geology – calcareous category	2.7	WFD Catchment size category	0.4		
WFD Geology - siliceous category	2.1				
WFD Altitude category	1.6				
WFD Geology – organic category	0.5				

as continuous variables (latitude, longitude and altitude) could be even stronger correlates. The way in which the boundary between ecoregions 20 and 22 also represents an altitude delineation (ecoregion 20 generally being at higher altitude than 22, and both being higher than 14) perhaps accounts for why both these variables are good predictors in Sweden. The percentage of fine sand was the third strongest variable. This is (to some extent) equivalent to $\bar{x}$ substratum size, the second strongest variable in Great Britain. However, in contrast to Great Britain and the Czech Republic, in Sweden the WFD System-A variable geology (organic category) was the fourth strongest variable. Sweden is also the country with the highest proportion of organic sites (77% versus 4% in Great Britain and 0% in the Czech Republic).

In the Czech Republic the WFD System-A variables were generally relatively weak descriptors of macroinvertebrate community variation, with the notable exception of the WFD System-A variable catchment size category. The high explanatory power of catchment size category is probably due to the relatively even spread of PERLA reference sites across the catchment size categories 100–1000 km^2, 1000–10,000 km^2 and $>10{,}000$ km^2 in comparison to the spread of sites across other WFD System-A variable categories. It is also interesting to note that distance from source (highly correlated with System-A catchment size category, r_s 0.875, $p<0.001$) was the strongest explanatory variable. The continuous variable altitude was also a stronger descriptor than the WFD System-A variable altitude category.

While the RIVPACS-type model variables were collectively always better descriptors of community variation, the usefulness of individual variables differed between countries. In general the RIVPACS-type model variables were individually better descriptors, although some WFD System-A variables in particular countries were also relatively strong correlates of macroinvertebrate community variation. In all cases where continuous and categorical variables (or closely correlated variables such as catchment size and distance from source) were present in the CCA analyses within in country, the continuous variables were better descriptors of macroinvertebrate community variation.

Discussion

Attempts to use fixed *a priori* stream typologies (especially ecoregions) to define biotic communities have had mixed results. For example, Verdonschot & Nijboer (2004) in their analysis of 889 streams across eight European countries concluded that large-scale typological factors explained most of the variation in macroinvertebrate assemblages. Similarly, Rabeni & Doisy (2000) and Feminella (2000) found that benthic macroinvertebrate assemblages in Missouri and parts of the southeastern USA coincided well with existing ecoregions. However, Sandin & Johnson (2000) in their study of Swedish streams concluded that ecoregion classifications need to be augmented with other factors such as altitude, stream size and catchment characteristics to discriminate macroinvertebrate communities and Waite et al. (2000) found that there was a large variation in macroinvertebrate community composition across the Mid-Atlantic Highlands of the USA both within and between ecoregions and that ordination did not reveal a distinct clustering of sites by ecoregion. Further, Van Sickle and Hughes (2000), in their study in Western Oregon, USA argue that geographic partitions can be expected to account for only a minor proportion of the total variation seen in macroinvertebrate communities across a large region and based on their study of five *a priori* landscape classifications in several regions of the USA, Hawkins & Vinson (2000) conclude that benthic macroinvertebrates vary continuously along environmental gradients so that methods of bioassessment that seek to place sites into discrete categories are fundamentally limited compared to approaches that recognise biological continua. These studies indicate that while fixed typologies provide a useful large-scale framework for setting ecological targets, they do not necessarily account for all sources of observed biological variation. Ecological targets set solely in terms of fixed typologies may not be precise enough to accurately define target communities.

The hierarchical water body typology set out in the EU Water Framework Directive defines streams and rivers across Europe in terms of ecoregions (Illies, 1978) and broad categories of altitude, catchment area and geology. It is implicit within this typology that macroinvertebrate communities at undisturbed sites should be broadly similar and therefore predictable. Our results from three European countries suggest that while the WFD System-A typology is more effective than a null model as a means of predicting reference values for macroinvertebrate biotic indices, it is considerably less effective than the site-specific multivariate RIVPACS-type models already in place in Great Britain (RIVPACS), Sweden ($SWEPAC_{SRI}$) and the Czech Republic (PERLA). This is probably due to the inclusion of a wider range of continuous (rather than categorical) variables, which are both map derived and site/sampling date specific. While there are some exceptions in the case of individual variables (particularly in Sweden), as a group the RIVPACS-type model variables have greater ecological significance than the WFD System-A variables and therefore enable more effective predictive models to be built.

Of the environmental variables (or types of variables) used in either two or all three of the RIVPACS-type models developed in Europe, many of these are also used in RIVPACS-type models developed for other geographical areas. For example, many of the Australian AUSRIVAS models utilise latitude, longitude, alkalinity, altitude, distance from source, slope, width, substratum, and discharge (Simpson & Norris, 2000) and eight of the 10 variables used in the Canadian BEAST model are altitude, longitude, substratum composition (three separate measures), depth, velocity and alkalinity (Rosenberg et al., 2000). Similarly in predictive models developed in California (Hawkins et al., 2000), seven of the 11 variables included in a species level model (longitude, altitude, depth, latitude, distance from source, width and slope) and six of the nine variables in a family level model (depth, longitude, altitude, slope, distance from source and width) are the same or similar to the variables used in two or all three of the European RIVPACS-type models. The variables used in each of the models above have themselves been selected from longer lists of candidate variables, and it is interesting to consider how the same or similar variables tend to emerge as good predictors of macroinvertebrate communities within models developed to serve such geographically widespread areas.

Within Europe, while the variables used by the three RIVPACS-type models may be broadly similar, they are not exactly the same. It is the use of multiple variables and the selection of these variables for their ecological significance in a given region that is key to the success of these models. In contrast, the WFD System-A typology is constrained in its use of a limited number of categorical variables. While in some cases these variables are good predictors, in many cases they are relatively weak correlates of variation in macroinvertebrate community composition. The variables WFD Ecoregion and WFD Altitude (in Sweden) and WFD Catchment size category (in the Czech Republic, Ecoregion 9 – Central Highlands) are examples of WFD System-A variables identified in this study as strong correlates of macroinvertebrate community variation. However, the usefulness of the WFD System-A typology is limited to certain variable type and geographical area combinations. This highlights the problem of the Europe-wide application of the same category boundaries in the WFD System-A typology. While the 238 PERLA reference sites in Ecoregion 9 of the Czech Republic are quite evenly spread among the three largest WFD System-A catchment size categories, the RIVPACS and $SWEPAC_{SRI}$ reference sites are distributed almost exclusively across the two smallest catchment size categories. The Europe-wide application of the same intervals of catchment size category appears to be useful in the Czech Republic, but has little ecological significance in Great Britain or Sweden.

Another problem with the WFD System-A typology is the use of categorical rather than continuous variables. This is exemplified by the case of altitude. In Great Britain and the Czech Republic the continuous variables $\log_{10}$ altitude and altitude, respectively, are both better predictor variables than the WFD System-A variable altitude category. The loss of predictive power by summarising the continuous variable altitude into broad categories is considerable. The same problem almost certainly contributes to the low percentage variance explained by many of the other WFD System-A variables.

A further problem with *a priori* typological approaches such as the WFD System-A is that they usually utilise variables gathered solely at large geographical scales. In Europe this can be addressed by opting for a System-B approach (which can incorporate a range of additional variables gathered at a variety of scales). Our analysis shows that substratum composition, width and depth, all of which are local scale variables measured at the time of sampling, can also be strong correlates of macroinvertebrate community composition. The importance of both large-scale and local factors as determinants of macroinvertebrate communities should therefore not be overlooked. Similar conclusions were reached by Heino et al. (2003) based on a study of macroinvertebrate diversity in headwater streams.

The percentage reduction in SD(O/E) achieved by the three European RIVPACS-type models compared to null models (based on Number of Taxa at BMWP family level) varied between 12 and 34% depending on the season model used. These generally exceeded the percent reductions in SD(O/E) achieved by a predictive model built from 86 reference sites in the Mid-Atlantic Highlands region of the USA (Van Sickle et al., 2005) where SD(O/E), based on Number of Taxa at species and genus level, was reduced by 13.7%. Another predictive model built from 209 sites in North Carolina, USA (Van Sickle et al., 2005) again based on Number of Taxa at species and genus level, achieved an impressive reduction in SD(O/E) compared to a null model of 52.5%.

The SD(O/E) of a null model is equivalent to the coefficient of variation (cv) in the observed metric values of the reference sites and reflects the natural variability in the values of the metric within a region. For example, null model Number of Taxa SD(O/E) values in Great Britain are lower than those in Sweden and the Czech Republic indicating that Number of Taxa is inherently more variable in the latter two countries. Also, different metrics have different null model SD(O/E) values within countries. For example, null model ASPT SD(O/E) values in the Great Britain are lower than null model Number of Taxa SD(O/E) values. In contrast, in relation to a null model the percent reduction in SD(O/E) obtained by using a predictive model (of any type) indicates the predictive model's effectiveness. By making reference to null models, the statistic, percent reduction in SD(O/E), allows us to compare the performance of models both between indices and between regions. The null model approach proposed by Van Sickle et al. (2005) therefore appears to provide an objective test of model performance and as such is likely to be extremely useful.

Conclusion

This study has shown that the site-specific multivariate RIVPACS-type predictive models already in place in Great Britain (RIVPACS), Sweden ($SWEPAC_{SRI}$) and the Czech Republic (PERLA) are more effective than both null models and the WFD System-A physical typology as methods of predicting macroinvertebrate reference conditions. The multivariate models are more effective primarily because they make use of continuous rather than categorical predictor variables (that have been selected for their value as good correlates of macroinvertebrate community composition) and because the multivariate RIVPACS-type models are not constrained by the use of a limited number of variables.

Acknowledgements

We would like to acknowledge Robin Guthrie (Scottish Environment Protection Agency) and John Murray-Bligh (Environment Agency) for their help defining WFD System-A types for the RIVPACS reference sites in Great Britain.

References

Armitage, P. & J. Blackburn, 1985. Chironomidae in the Pennine stream system receiving mine drainage and organic enrichment. Hydrobiologia 121: 165–172.

Clarke, R. T., J. F. Wright & M. T. Furse, 2003. RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers. Ecological Modelling 160: 219–233.

Council of the European Communities., 2000. Directive 2000/60/EC, Establishing a Framework for Community Action in the Field of Water Policy. European Commission PE-CONS 3639/1/100 Rev 1, Luxembourg.

CSN 75 7716, 1998. Water quality, biological analysis, determination of saprobic index. Czech Technical State Standard. Czech Standards Institute, Prague, 174 pp.

Dawson, F. H., D. D. Hornby & J. H. Hilton, 2002. A method for the automated extraction of environmental variables to help the classification of rivers in Britain. Aquatic Conservation: Marine and Freshwater Ecosystems 12: 391–403.

Extence, C. A., D. M. Balbi & R. P. Chadd, 1999. River flow indexing using British benthic macroinvertebrates: a framework for setting hydroecological objectives. Regulated Rivers Research and Management 15: 543–574.

Feminella, J. W., 2000. Correspondence between stream macroinvertebrate assemblages and 4 ecoregions of the southeastern USA. Journal of the North American Benthological Society 19: 442–461.

Furse, M. T., 2000. The application of RIVPACS procedures in headwater streams – an extensive and important national resource. In Wright, J. F., D. W. Sutcliffe & M. T. Furse (eds), Assessing the Biological Quality of Fresh Waters. Freshwater Biological Association, Ambleside, 79–91.

Hawkins, C. P., R. H. Norris, J. N. Hogue & J. W. Feminella, 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. Ecological Applications 10: 1456–1477.

Hawkins, C. P. & M. R. Vinson, 2000. Weak correspondence between landscape classifications and stream macroinvertebrate assemblages: implications for bioassessment. Journal of the North American Benthological Society 19: 501–517.

Heino, J., T. Muotka & R. Paavola, 2003. Determinants of macroinvertebrate diversity in headwater streams: regional and local influences. Journal of Animal Ecology 72: 425–434.

Illies, J., 1978. Limnofauna Europaea. Gustav Fisher Verlag, Stuttgart.

Johnson, R. K. & L. Sandin, 2001. Development of a Prediction and Classification System for Lake (Littoral) and Stream (Riffle) Macroinvertebrate Communities. Stencil. Department of Environmental Assessment, SLU, Uppsala.

Johnson, R. K., K. Aagaard, K. J. Aanes, N. Friberg, G. M. Gislason, H. Lax & L. Sandin, 2001. Macroinvertebrates. In Skriver, J. (ed), Biological Monitoring in Nordic Rivers and Lakes. TemaNord Environment 513: 43–52.

Kokeš, J., S. Zahrádková, D. Němejcová, J. Hodovský, J. Jarkovský & T. Soldán, 2006. The PERLA system in the Czech Republic: a multivariate approach for assessing the ecological status of running waters. Hydrobiologia 566: 343–354.

Logan, P. & M. Furse, 2002. Preparing for the European Water Framework Directive – making the links between habitat and aquatic biota. Aquatic Conservation: Marine and Freshwater Ecosystems 12: 425–437.

Marvan, P., 1969. Primechania k primeneniu statisticheskich metodov po opredeleniu saprobnosti. Symposium SEV Voprosy saprobnosti, Zivohost, 19–43.

Moss, D., M. T. Furse, J. F. Wright & P. D. Armitage, 1987. The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. Freshwater Biology 17: 41–52.

National Water Council., 1981. River Quality: The 1980 Survey and Future Outlook. National Water Council, London.

Pantle, E. & H. Buck, 1955. Die biologische Uberwachung der Gewasser und die Darstellung der Ergebnisse. Gas und Wasserfach. 96: 604.

Rabeni, C. F. & K. E. Doisy, 2000. Correspondence of stream benthic invertebrate assemblages to regional classification schemes in Missouri. Journal of the North American Benthological Society 19: 419–428.

Rosenberg, D. M., T. B. Reynoldson & V. H. Resh, 2000. Establishing reference conditions in the Fraser River catchment, British Columbia, Canada, using the BEAST (BEnthic Assessment of SedimenT) predictive model. In Wright, J. F., D. W. Sutcliffe & M. T. Furse (eds), Assessing the Biological Quality of Fresh Waters. Freshwater Biological Association, Ambleside, 181–194.

Sandin, L. & R. K. Johnson, 2000. Ecoregions and benthic macroinvertebrate assemblages of Swedish streams. Journal of the North American Benthological Society 19: 462–474.

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), Assessing the Biological Quality of Fresh Waters. Freshwater Biological Association, Ambleside, 125–142.

Sweeting, R., 2001. Classification of ecological status of lakes and rivers – biological elements in the classification. In Back, S. & K. Karttunnen (eds), Classification of Ecological Status of Lakes and Rivers. TemaNord Environment 2001:584, Nordic Council of Ministers, Copenhagen, 9.

ter Braak, C. J. F. & I. C. Prentice, 1988. A theory of gradient analysis. Advances in Ecological Research 18: 271–317.

ter Braak, C. J. F. & P. Smilauer, 2002. CANOCO Reference Manual and CanoDraw for User's Guide: Software for Canonical Community Ordination (version 4.5). Microcomputer Power (Ithaca NY, USA), 500 pp.

Van Sickle, J. & R. M. Hughes, 2000. Classification strengths of ecoregions, catchments, and geographical clusters for aquatic vertebrates in Oregon. Journal of the North American Benthological Society 19: 370–384.

Van Sickle, J., C. P. Hawkins, D. P. Larsen & A. H. Herlihy, 2005. A null model for the expected macroinvertebrate assemblage in streams. Journal of the North American Benthological Society 24: 178–191.

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

Waite, I. R., A. T. Herlihy, D. P. Larsen & D. J. Klemm, 2000. Comparing strengths of geographic and non-geographic classifications of stream benthic macroinvertebrates in the Mid-Atlantic Highlands, USA. Journal of the North American Benthological Society 19: 429–441.

Wilander, A., R. K. Johnson & W. Goedkoop, 2003. Riksinventering 2000: En synoptisk studie av vattenkemi och bottenfauna i svenksa sjöar och vattendrag. Department of Environmental Assessment, Swedish University of Agricultural Sciences, Uppsala, Report 2003, 1 pp.

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 prediction of community type using environmental data. Freshwater Biology 14: 221–256.

Wright, J. F., 2000. An introduction to RIVPACS. In Wright, J. F., D. W. Sutcliffe & M. T. Furse (eds), Assessing the Biological Quality of Fresh Waters. Freshwater Biological Association, Ambleside, 1–24.