Ecological Indicators 10 (2010) 896–904

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Short communication

Comparing the sensitivity of diverse macroinvertebrate metrics to a multiple stressor gradient in Mediterranean streams and its influence on the assessment of ecological status

M.M. Sánchez-Montoya *, M.R. Vidal-Abarca, M.L. Suárez

Department of Ecology and Hydrology, University of Murcia, 30100 Espinardo, Murcia, Spain

ARTICLE INFO

Article history:
Received 13 July 2009
Received in revised form 12 January 2010
Accepted 17 January 2010

Keywords:
Macroinvertebrate metrics
Mediterranean streams
Ecological status
Water Framework Directive
GUADALMED Project

ABSTRACT

Biological indicators are increasingly being used as integrative measures of the ecosystem health in streams, particularly those using macroinvertebrate assemblage composition. Monitoring biological quality of rivers has not a long tradition in some Mediterranean European countries like Spain. Several macroinvertebrate metrics have been recently proposed to assess ecological status in Mediterranean streams, so it is necessary to compare the use of proposed biological quality metrics to select the most appropriate ones.

In the present work, two classic richness metrics (total number of families and number of the Ephemeroptera, Plecoptera and Trichoptera families), three indices (IBMWP, IASPT and t-BMWQ) and two multimetric indices, recently proposed to be used in Mediterranean streams (ICM-9 and ICM-11a or IMMi-L), were compared by the analysis of the sensitivity of these metrics to a multiple stressor gradient which reflected the main pressures present in the study area. For this purpose, data from 193 sites sampled in spring (95 reference sites and 98 disturbed sites) belonging to five different Mediterranean stream types present in 35 basins were studied.

The results showed that the adjusted regression coefficients (r^2) for all seven metrics in the exponential regression models were higher than linear ones, thus indicating an exponential relationship between metrics and the environmental alteration. The two studied multimetric indices presented higher regression coefficients ($r^2 = 0.590–0.669$) than the three indices ($r^2 = 0.524–0.574$) and the two metrics ($r^2 = 0.471–0.525$), therefore showing a better response to a stressor gradient in Mediterranean streams. Within the multimetric indices group, ICM-11a showed the highest regression coefficients. Based on the results obtained, we suggest using the ICM-11a, apart from the IBMWP, to assess ecological status in Mediterranean streams.

1. Introduction

The European Water Framework Directive (WFD-European Commission, 2000) represents a huge effect for European freshwater management since it outlines the innovative concept of bioassessment: surface water classification through ecological status assessment (Annexe V, WFD). Biological as well as supporting hydromorphological and physico-chemical quality elements are used by Member States in ecological status evaluation (European Commission, 2003).

Ecological status of water bodies is defined by comparing the biological community composition with the reference condition. According to the WFD, reference conditions refer to sites that show no, or a minimum, human impact (European Commission, 2000, 2003), and should be linked to stream typology (Barbour et al., 1996). Therefore, the establishment of reference conditions, which are type-specific, and the definition of ecological quality class boundaries are closely interconnected. Both components must be established by Member States for all surface water bodies and all relevant quality elements. Ecological status must be classified into five quality classes (high, good, moderate, poor and bad). This classification is based on ecological quality ratios (EQRs: O/E scores) which are derived from biological quality values (European Commission, 2000, 2003). The boundary between good and moderate status is especially important because it sets the targets for restoration plans within the programmes of measures of water bodies which fail the environmental objectives of achieving good ecological status (Heiskanen et al., 2004).

One main WFD approach is the use of biotic indicators (phytobenthos, macrophytes, macroinvertebrates and fishes) in the stream assessment. This is a novelty in many European countries and efforts are being made to adapt national programmes to these

* Corresponding author. Tel.: +34 868 884653; fax: +34 868 883963.
E-mail address: marsanch@um.es (M.M. Sánchez-Montoya).

1470-160X/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ecolind.2010.01.008

WFD requirements (Birk and Hering, 2006). In general, there is a general consensus about the value of macroinvertebrates as water quality biological indicators (Chessman and McEvoy, 1998) and therefore the community structure of benthic invertebrates has been frequently used in the environmental monitoring and assessment of aquatic systems (Reynoldson and Metcalfe-Smith, 1992).

The European Commission suggests the use of multimetric approaches for establishing ecological status (Hering et al., 2006) and proposes the use of multimetric indices to detect diverse types of pressures (i.e., organic, hydromorphological alteration, acidification, etc.). The multimetric approach attempts to provide an integrated analysis of the biological community of a site by deriving a variety of biological measures and knowledge of a site's fauna (Karr, 1981, 1999).

Diverse macroinvertebrate indices have been proposed to assess water quality in Mediterranean streams and rivers in the WFD context. Among others, the ICM-9 was proposed by the Mediterranean Geographical Intercalibration Group (Med-GIG) as an appropriate index for Mediterranean streams (European Commission, 2007), and the ICM-11a (also called IMMi-L), developed since the Spanish experience in the Med-GIG, has been recently developed specifically for macroinvertebrate communities inhabiting Mediterranean river systems (Munné and Prat, 2009). Particularly in Spain, the IBMWP index (Alba-Tercedor and Sánchez-Ortega, 1988; Alba-Tercedor et al., 2004) is the most widespread and extensive index employed in Mediterranean rivers, and indeed the Spanish Environmental Ministry has recently proposed this index to assess ecological status in Spanish Mediterranean streams (MMARM, 2008). Also the t-BMWQ proposed by Camargo (1993) in Spain seems to be a good candidate to determine the ecological status in Mediterranean streams (Navarro, 2006).

Monitoring the biological quality of rivers has not a long tradition in some Mediterranean countries such as Spain. In this context, we aim to compare the use of diverse macroinvertebrate quality metrics to select the most appropriate ones based on their response to a stressor gradient which reflects the main pressures presented in Mediterranean streams. For this propose, a typology of Mediterranean streams (Sánchez-Montoya et al., 2007) and an approach to select reference sites in Mediterranean streams (Sánchez-Montoya et al., 2009a) were applied, both of which were developed in this study area. These methodologies were developed in the context of the GUADALMED II Project (Prat, 2004) in Spain.

Therefore, the specific aims of this study were to: (1) identify a stressor gradient and its component stressor, (2) analyse the relationship of the stressor gradient with diverse macroinvertebrate metrics in order to identify those indicators which better assess ecological status, (3) define the four quality class boundaries according to their relationship with the stressor gradient and (4) compare the classification ecological status by using those metrics which present a better adjustment to the stressor gradient.

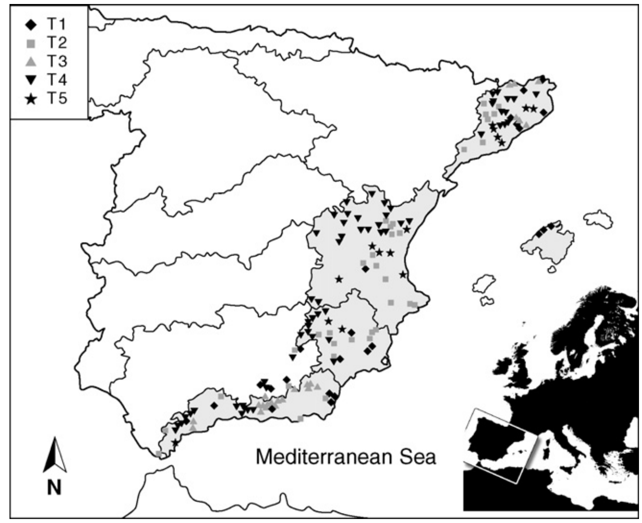

Fig. 1. Distribution of the 193 sampling sites indicating their typology (T1: temporary streams, T2: evaporite calcareous at medium altitude, T3: siliceous headwaters at high altitude, T4: calcareous headwaters at medium and high altitude, and T5: large watercourse).

2. Materials and methods

2.1. Study area and sampling sites

This study was carried out in 35 catchments of the Iberian Mediterranean coast covering approximately 84,400 km^2, from large catchments such as the Segura and Júcar rivers (>14,000 km^2) to small ones (<200 km^2) like Chillar and Pollença. An altitudinal, thermal and pluviometric gradient is present from north to south and from the mountains to the coast. The annual range of temperatures is between −2 °C and 42 °C and annual precipitation ranges from 280 mm to 1000 mm (MIMAM, 2000). Given its Mediterranean climate, this area is characterised by hot dry summers and cool wet winters and rivers show high seasonality with annual and interannual variability in discharge regimes and frequent periods of flooding and drying (Gasith and Resh, 1999). A total of 193 sampling sites (approximately 100 m) were studied (Fig. 1). These sites belong to five different stream types (Table 1) defined in this study area in a previous work (see Sánchez-Montoya et al., 2007), following the basic criteria in System B of the WFD. Although the environmental variables used to characterize these ecotypes (drainage area, the type of flow regime, the altitude, and the geology of the drainage area) are similar to those in the case of the river types proposed by the ECOSTAT intercalibration group for the Mediterranean rivers using the System A of the WFD (European Commission, 2005), there is not a clear correspondence between both typologies in relation to the types obtained.

Table 1
Number of sampling sites studied in each stream type, distinguishing the number of disturbed sites and reference sites according to Mediterranean Reference Criteria defined by Sánchez-Montoya et al. (2009a).

Stream type	Description	Catchment size range (km^2)	Hydrological state	Total sampling sites	Disturbed sites	Reference sites
Stream type 1 (T1)	Temporary streams	(12–18,136)	Intermittent/Ephemeral	31	15	16
Stream type 2 (T2)	Evaporite calcareous at medium altitude	(58–18,136)	Perennial seasonal	35	25	10
Stream type 3 (T3)	Siliceous headwaters at high altitude	(156–6532)	Perennial seasonal	33	12	21
Stream type 4 (T4)	Calcareous headwaters at medium and high altitude	(101–18,136)	Perennial seasonal	79	36	43
Stream type 5 (T5)	Large watercourse	(4994–18,136)	Perennial seasonal	15	10	5
Total				193	98	95

Table 2
Macroinvertebrate metrics, indices and multimetric indices calculated in all sampling sites ($n = 193$).

Type of metric	Metric	Definition
Taxonomic composition metrics	NFAM	Total number of families
	EPT	Number of Ephemeroptera, Plecoptera and Trichoptera families
Sensitive taxa indices	IBMWP	Iberian Biological Monitoring Working Party
	IASPT	IBMWP value/number of families
	T-BMWQ	Biological Monitoring Water Quality
Multimetric indices combination of taxonomic composition, sensitive taxa and major taxonomic group metrics	ICM-9	Multimetric index based on qualitative data used in the Med-GIG intercalibration exercise
	ICM-11a	Index proposed to be used in Mediterranean rivers

Sampling sites represent different grades of degradation, from undisturbed to highly disturbed, depending on the number of reference criteria fulfilled in accordance with the approach defined by Sánchez-Montoya et al. (2009a).

We examined two different datasets depending on the specific objective: (1) to establish stream type-specific reference conditions, data from spring 2003 of only the reference sites ($n = 95$) were considered. Only those samples collected in spring were used, since some of the studied metrics showed seasonal differences in some of the studied stream types (Sánchez-Montoya et al., 2009b); (2) to determine the stressor gradient and ecological status class boundaries, the total number of sites ($n = 193$) was used. Of the total sites, 161 were sampled in the spring of 2003, while the rest ($n = 32$) were sampled in the spring of 2000.

2.2. Macroinvertebrate data and metrics

A multi-habitat semi-quantitative kick-sample was taken from each site using the protocol of Jáimez-Cuéllar et al. (2004). Macroinvertebrate samples were collected from all the habitats with a kick-net (250–400 μm), and they were preserved in 90% ethanol. Samples were examined under a stereoscope in the laboratory. At least 200 individuals in each sample were randomly picked and identified at the family level, except for Ostracoda, Oligochaeta and Hydracarina. Large uncommon individuals were picked individually. Two benthic macroinvertebrate single metrics, three indices and two multimetric indices were calculated (Table 2).

2.3. Establishing reference conditions

In this study, the selection of reference sites was carried out according to the methodology proposed by Sánchez-Montoya et al. (2009a) in Mediterranean streams, which includes 20 *a priori* criteria that reflect the characteristics of Mediterranean streams and their most frequent disturbances. These reference criteria are related to riparian vegetation zone, introduced species, point sources of pollution, diffuse sources of pollution and land uses, river morphology and habitat conditions and finally hydrological conditions and regulation.

For stream types 1, 2, 3 and 4, only those sites which fulfilled the 20 *a priori* criteria were selected as reference sites. However, the higher presence of pressure in the large watercourse stream type (stream type 5) meant that there were no reference sites for that type. Accordingly, those sites which met more than 15 of the 20 criteria were considered the least disturbed and selected as reference sites. This approach matches the 'Best Attainable Condition' concept defined by Stoddard et al. (2006), which indicates that reference sites are selected where the impact of inevitable land use on biota has been minimised.

After this selection, the median values of all the studied metrics of reference sites were used to quantify the specific-types reference conditions for the purpose of calculating the Ecological Quality Ratio (EQR). The 25th percentile of the distribution of reference sites was selected for all the metrics in each stream type to identify the high-good boundary. This percentile is considered a slight deviation from the reference condition (Pollard and van De Bund, 2005). Boxplots of the reference EQR values for all the quality metrics in the five stream types analysed were performed to study their variability (Fig. 2).

2.4. Stressor gradient and environmental data

To create the stressor gradient, 27 different variables belonging to different categories (physico-chemical, land uses, hydrological variables and quality indices) were studied for each site to establish a real axis of degradation from all the 193 sampling sites (Table 3). Land use variables were calculated by firstly using a digital terrain model (DTM 30 m × 30 m; Centro Geográfico del Ejército, Ministerio de Defensa, Spain, 2005) and the Arc/Info software (version 9.0, ESRI, Redlands, California, USA, 2005) to delimit and calculate the water drainage area upstream to each site. The percentages of different land uses within the calculated drainage area were calculated on the basis of the Corine Land Cover 2000 Programme (Spain; Scale 1:100,000; Instituto Geográfico Nacional. Centro Nacional de Información Geográfica). The Potential Hydrological Alteration Index was used to quantify a possible impact of the presence of reservoirs upstream to the sites (CEDEX, 2004).

A Principal Component Analysis (PCA) was performed to generate the stressor gradient using the Statistica program (Stat Soft, 1999). Previously, all the 27 variables were transformed to ascertain the normal distribution of each variable (Table 3). Pearson's correlation was performed to identify strongly correlated variables (correlation coefficient >0.8), and such variables were removed so as not to be included redundant information in the PCA. Because the variables used in the PCA were quantified in very different measures, correlation matrices based on the standardised variables (median = 0 and variance = 1) were used to perform this analysis. The first ordination axis was selected as the stressor gradient, and the projected site values were standardised along this axis (from 0 to 1).

2.5. Establishing quality class boundaries

Linear and exponential regressions were performed, using the Statistica program (Stat Soft, 1999), to analyse the response of the seven metrics to the stressor gradient. For this end, the EQR values of the 193 sites were calculated and represented against the gradient to identify the better adjustment of these values by means of the adjusted regression coefficient (r^2). After, the four quality class boundaries were defined. The 25th percentile of the distribution of reference values was selected as the upper anchor, defining the class boundary between high and good. When the best adjustment between the metric and the stressor gradient is achieved by a linear regression, the width of the four remaining

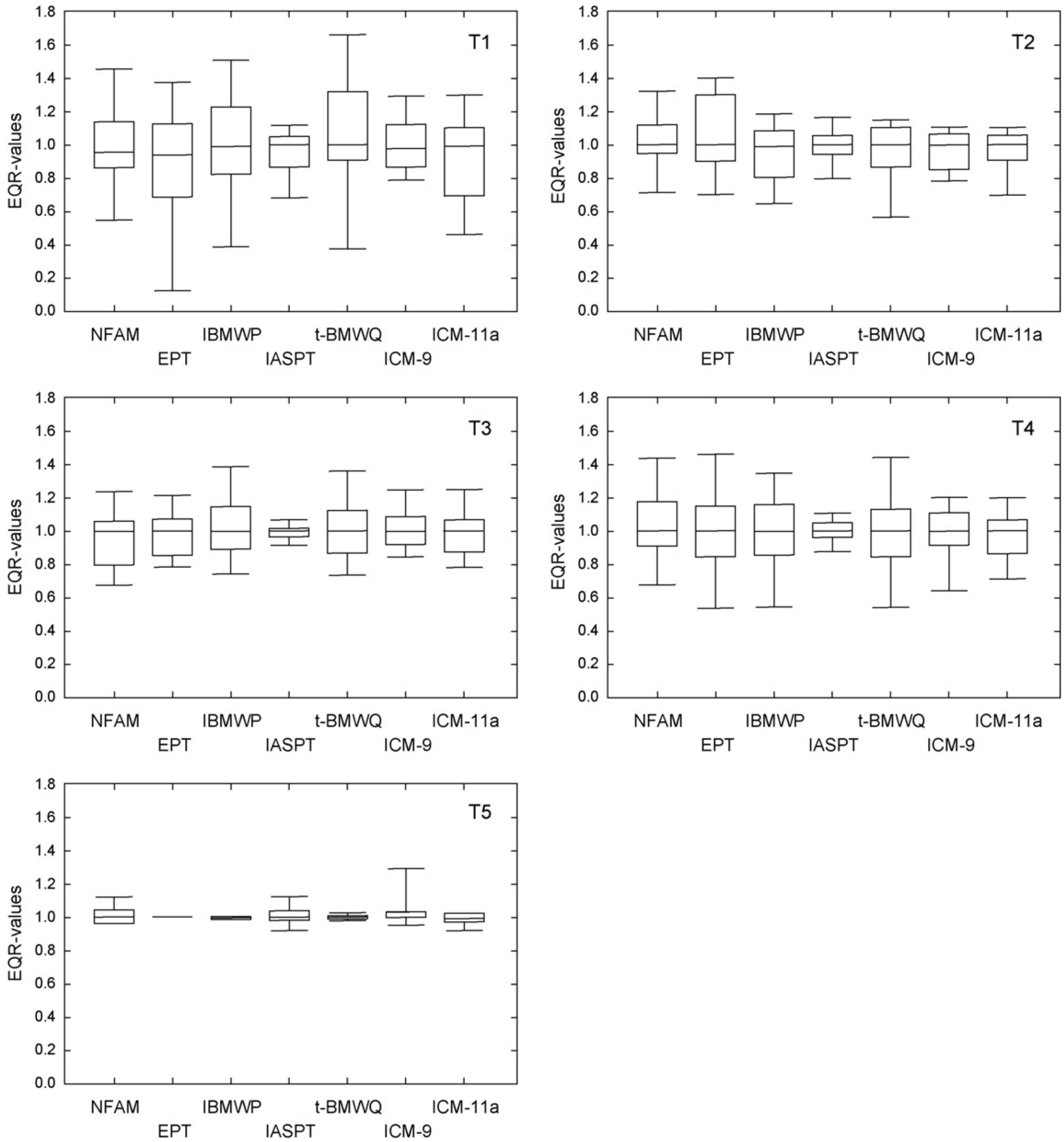

Fig. 2. Distribution of the EQR values of reference sites in the five streams type for the seven studied metrics. The median values (central line), 25th and 75th percentile values (box) and the maximum and minimum values are shown.

classes is evenly spaced over the remaining interval (European Commission, 2003). If the relationship is not linear, however, an equitable division of the EQR is not appropriate. In such cases, the stressor gradient value corresponding to the class boundary between high and good (established by the 25th percentile) is calculated. Starting from this value the rest of stressor gradient is divided into four equitable fragments, and the class boundary is the EQR value which coincides with these four established intervals. When the four class boundaries were calculated for the metrics, the quality class for all the sampling sites was estimated.

In order to examine how well the metrics which presented a better adjustment performed in the detection of anthropogenic pressures, the Type I error or α and the Type error II or β were estimated. To estimate the Type I error, the percentage of reference sites which were classified below the good level (moderate, poor and bad) was calculated. The Type II error was estimated by calculating the percentage of highly disturbed sites (sites that fulfilled only five criteria or less) that were classified above the moderate level (high and good).

3. Results

3.1. Establishing reference conditions

Of the 193 sampling sites, 95 were selected as reference sites (Table 1). More than 50% of sites belonging to T1, T3 and T4 were selected as reference sites. This percentage was lower for both T2 and T5 stream types. In order to quantify reference conditions for the purpose of calculating EQRs, the median value of reference sites was used as reference value. Table 4 shows reference values of each metric for the five stream types. The reference values of all the metrics were very similar for T2, T3 and T4. However, T5 (large watercourses) showed lower reference values than those previously mentioned, and temporary streams (T1) presented the lowest values in all metrics. The EQR-25th percentile of the distribution of all the values of each metrics is also shown in Table 4. These values varied from 0.75 for the EPT in the case of T1, to 1 for the EPT in the case of T5. Fig. 2 shows the graphic representation of the variability of the EQR values of all the reference sites in the five stream types for the seven metrics. The

Table 3
Variables studied in each sites in order to establish a stressor gradient. The methodology used to calculate the variables is shown. Variables marked with an asterisk were removed because they were highly correlated with others.

Categories	Variable	Code	Transf	Description	Methodology
Physico-chemical variables	Oxygen	Oxy	Log10 $(x+1)$	Concentration of oxygen (mg/l)	In situ
	Conductivity	Conduc	Log10 $(x+1)$	Water conductivity (μS/cm)	
	pH	pH	Log10 $(x+1)$	Water pH	
	Nitrate	NO3	Log10 $(x+1)$	Concentration of N-NO_3^- (mg/l)	Photometer
	Nitrite	NO2	Log10 $(x+1)$	Concentration of N-NO_2^- (mg/l)	(LASA 100) or
	Ammonium	NH4	Log10 $(x+1)$	Concentration of NH_4^+ (mg/l)	Standard Protocol
	Phosphate	PO4	Log10 $(x+1)$	Concentration of PO_4 (mg/l)	APHA (1998)
Land uses	% Dry land agriculture*	%Dry	arc$\left[\sqrt{x/100}\right]$	Percentage of dry land farming in drainage area (cereal, vineyard and tree crops as olive)	GIS data
	% Irrigated farming	%Irrig	arc$\left[\sqrt{x/100}\right]$	Percentage of intensive irrigated farming in drainage area (rice field, irrigated vineyards and others irrigated fruit trees)	
	% Burnt vegetation	%Burn-veg	arc$\left[\sqrt{x/100}\right]$	Percentage of burned vegetation in drainage area	
	% Impervious area	%Imperv	arc$\left[\sqrt{x/100}\right]$	Percentage of impervious area in drainage area (urban, industrial, communication infrastructures and recreation areas are included)	
	% Natural land use*	%Natur	arc$\left[\sqrt{x/100}\right]$	Percentage of natural land use in drainage area	
	% Non-natural land use	%Nonatur	arc$\left[\sqrt{x/100}\right]$	Percentage of non-natural land use in drainage area	
Hydrological variables	Potential Hidrological Alteration	PHA	Log10 $(x+1)$	Indicator of regulation in base on the dimensions of the reservoirs allocated up-streams	(CEDEX, 2004)
Quality indices	QBR index	QBR	Log10 $(x+1)$	Total score of the Riparian Vegetation Quality index	QBR index
	Riparian Cover	QBR-Cove	Log10 $(x+1)$	Proportion of riparian area covered by trees and shrubs	(Munné et al., 2003)
	Riparian Structure	QBR-Stru	Log10 $(x+1)$	Proportion of the riparian vegetation composed of trees and shrubs separately	
	Riparian Quality	QBR-Qual	Log10 $(x+1)$	Absence of introduced species, garbage and other human impact on riparian vegetation	
	Riparian Naturality	QBR-Nat	Log10 $(x+1)$	Human impact altering channel form	
	IHF Index	IHF	Log10 $(x+1)$	Total score of the Habitat Fluvial Index	IHF index (Pardo et al., 2004)
	Embeddedness	IHF-Embed	Log10 $(x+1)$	Percentage of embeddedness in riffles or sedimentation in pools	
	Riffles vs. pools	IHF-R/L	Log10 $(x+1)$	Frequency of riffles in sampling reach: distance between riffles/stream width	
	Substrate composition	IHF-Subst	Log10 $(x+1)$	Percentage of boulders, stones, pebbles, gravel, sand and clay	
	Flow and depth regimes	IHF-Flow	Log10 $(x+1)$	Number of classes present in sampling reach: slow-depth, slow-shallow, fast-depth and fast-shallow	
	Shade	IHF-Shade	Log10 $(x+1)$	A score running from not shaded to completely shaded	
	Heterogeneity elements	IHF-Heterog	Log10 $(x+1)$	Percentage of leaf litter, presence of wood and branches, tree roots and natural dams	
	Instream vegetation	IHF-Inst-veg	Log10 $(x+1)$	Types and abundance of different instream vegetation formations: % of plocon, pecton and macrophytes	

mean of the EQR-25th percentile values for all metrics within each stream type, indicates that the most variable group of reference sites is that belonging to temporary streams (T1), presenting a mean value of 0.85, while T2, T3 and T4 presented a higher value (0.88). Finally, T5 presented the least variable group of reference sites with a mean value close to 1 (0.97).

3.2. The stressor gradient and its relationship with the studied metrics

Pearson's correlation showed that only 3 of the 27 variables selected to generate the stressor gradient were highly correlated ($r > 0.8$). Natural land use and dry land agriculture were removed and the remaining 25 variables, previously transformed, were used to perform the PCA analysis (Table 3). Fig. 3 shows the projection of all the variables in the two new first axes which absorbed a total variance of 37.7%. The least disturbed sites were located to the right of this axis and presented high oxygen values, undisturbed riparian vegetation and good fluvial habitat. Sites with a high nutrient concentration, altered hydrological conditions and higher percentages of agriculture, along with other non-natural land uses, were ordered in the left-hand margin.

Table 4
The median as reference value and the EQR-25th percentile as class boundary between high and good (which is shown within a parenthesis) are shown for the seven metrics in each stream type. Also the number of reference sites considered (n) is shown.

Quality metrics	T1 ($n=17$)	T2 ($n=10$)	T3 ($n=21$)	T4 ($n=43$)	T5 ($n=5$)
NFAM	22.0 (0.91)	29.5 (0.95)	34.0 (0.79)	34.0 (0.91)	26.0 (0.96)
EPT	8.0 (0.75)	10 (0.9)	14.0 (0.86)	13.0 (0.85)	9.0 (1.00)
IBMWP	106.0 (0.83)	170.5 (0.80)	188.0 (0.89)	196.0 (0.86)	133.0 (0.99)
IASPT	5.0 (0.88)	5.4 (0.94)	6.0 (0.96)	5.6 (0.96)	5.0 (0.98)
T-BMWQ	154.5 (0.91)	260.5 (0.87)	294 (0.87)	295 (0.84)	205.0 (0.99)
ICM9	11.6 (0.90)	15.2 (0.85)	16.5 (0.92)	17.1 (0.92)	13.1 (0.99)
ICM11a	11.6 (0.76)	14.6 (0.91)	16.1 (0.87)	16.0 (0.86)	11.7 (0.97)

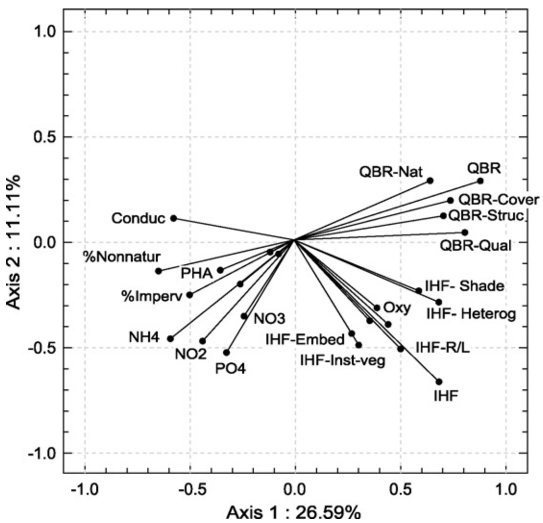

Fig. 3. Projection in the two axes of the PCA of the 25 variables used to generate the stressor gradient. Only some of the variables are shown in the graphic.

Table 5
Adjusted regression coefficients (r^2) of linear and exponential regressions used to analyse the respond of metrics to the multiple stressor gradient. All the *p*-levels were <0.001.

Quality metrics	Linear	Exponential
NFAM	0.326	0.471
EPT	0.493	0.525
IBMWP	0.437	0.574
IASPT	0.496	0.524
T-BMWQ	0.388	0.535
ICM-9	0.499	0.590
ICM-11a	0.562	0.669

To analyse the relationship of each studied metric with the stressor gradient, both linear and exponential regressions were calculated using the EQR of each metric of all 193 sites studied (Table 5). The relationship between all the metrics and the stressor gradient displayed higher regression coefficients (r^2) for exponential regressions rather for linear regressions, thus indicating an exponential response of metrics to this multiple stressor gradient. Exponential regression coefficients varied from 0.47 to 0.67, presenting the NFAM the lowest value and the ICM-11a the highest one. Fig. 4 shows the projection of the 193 sampling sites in the space formed by the stressor gradient and the EQR of each metric, indicating the exponential regression coefficient.

3.3. Setting quality class boundaries

The regression formula of the IBMWP and two multimetric indices, which showed a better adjustment, were employed to calculate the EQR values of the class boundaries, using the previously calculated equal four values of the stressor gradient from the 25th percentile value (Table 6), previously explained in the methodology. According to the established quality class boundaries, the quality classes of each site were estimated using the three metrics.

In order to examine how well the different metrics performed in detecting anthropogenic pressures in accordance with the quality class obtained, the Type I and the Type II errors were estimated for all the metrics. The number of references sites considered was 95 and the number of very disturbed site was 7. The three indices presented a very low Type I error (1.05% for IBMAP and ICM-9 and 2.1% for ICM-11a) and none of the indices showed the Type II error.

Since the IBMWP and the ICM-11a were the index and the multimetric index respectively that presented the best adjustments to the stressor gradient, the number of coincidence cases in the classification of quality classes according to these two metrics was analysed. The 26.3% (51 of 193) of sampled sites were classified into different quality classes according to the IBMWP and the ICM-11a. However, this percentage was lower (6.6%, 13 sites of 195) when we analysed the number of sites that were classified into the two different quality level groups (high and good or moderate, poor and bad). The high linear regression coefficient value (r^2 = 0.706) between these two indices accounts for the high coincidence of the classification into quality classes using the two indices (Fig. 5).

4. Discussion

4.1. Establishing the stressor gradient

The five Mediterranean stream types established in this area are subject to many pressures related to the elimination of riparian vegetation, the presence of exotic species, point and diffuse sources of pollution and land uses, river morphology and habitat conditions, and hydrological conditions and regulation (Sánchez-Montoya et al., 2009a). In this study, non-reference sites showed a combination of several pressures of which the most frequent were those related to non-natural land uses, diffuse sources of pollution and flow regulation. Hence, it was considered that water quality measured exclusively using physical and chemical parameters, provided skewed information regarding the level of impact to which the biological communities in this study area are subject. Therefore, we considered it more suitable to define the stressor gradient with multiple stressors than defining a specific stressor gradient by including all the pressures measured at each site.

4.2. Suitability of metrics to assess degradation in Mediterranean streams

In the assessment of ecological status, in which macroinvertebrates are used, one important point is the identification of metrics that can detect one or several types of human perturbation.

The exponential regression coefficients for all seven metrics were higher than linear coefficients, indicating an exponential relationship between metrics and the environmental alteration. This relationship was expected for indices such as IBMWP, IASPT and t-BMWP due to their method of calculation where taxa do not contribute equally to the final score depending on their tolerance to organic pollution. The score of these indices increases as the degree of degradation decreases and new taxa begin to appear with high scores (Munné and Prat, 2009).

The two studied multimetric indices presented higher regression coefficients (r^2 = 0.590–0.669) than the three indices (r^2 = 0.524–0.574) and the two metrics (r^2 = 0.471–0.525), showing a better response to a stressor gradient in the Mediterranean streams. This finding supports the knowledge that a single metric responds well to specific pressures, but multimetric indices provide an integrated analysis of the biological community and are capable of detecting multiple stressors (Karr, 1981, 1999). The same result was previously obtained by Ofenböck et al. (2004) in Austria, where multimetric indices showed higher discrimination efficiency than any of the individual metrics studied (e.g., EPT, NFAM, number of the Oligochaeta and Diptera taxa and the Margalef diversity index).

In our study, the single metric EPT and the IASPT index presented very similar exponential and linear regression coefficient values (Table 5). These regression coefficients were higher

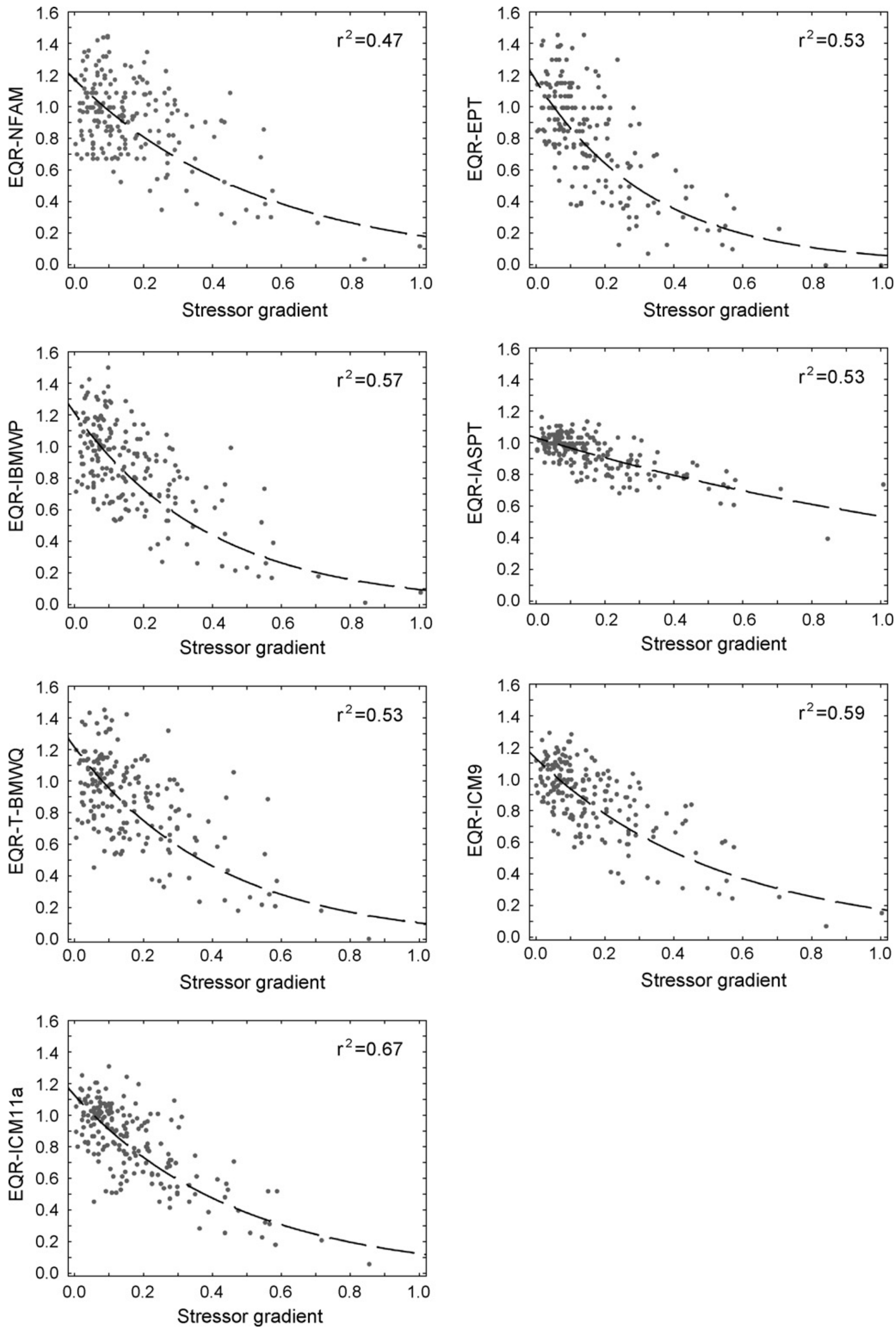

Fig. 4. Exponential regression between the stressor gradient and the two metrics (NFAM and EPT), the three indices (IBMWP, IASPT and t-BMWP) and two multimetric indices (ICM-9 and ICM-11a). The adjusted regression coefficient (r^2) is shown in each graphic.

than those obtained by Sandin and Hering in 2004 (EPT = 0.349 and ASPT = 0.428) when studying the response of these metrics to an organic pollution gradient in seven stream types of four countries from North to South Europe.

Within the group of single metrics, EPT was more sensitive to the stressor gradient than NFAM. This finding was not unexpected since NFAM is known to be a simplified measure of biological composition (Walsh, 2006). However, the species belonging to the EPT insect orders are regarded as being very sensitive to impairment, and loss of taxa richness within this group indicates perturbation (Wallace et al., 1996). A similar result was obtained by Walsh (2006) in Australian streams who compared the response of diverse metrics to an urban gradient. In that study, SIGNAL and EPT were found to be more sensitive than NFAM to impacts generated by the urbanisation of catchments, such as hydrological disturbance, water quality impairment and habitat loss.

Regarding the three indices analysed, the IBMWP showed a better adjustment with the stressor gradient than the IASPT and the t-BMWQ. The higher regression coefficient of the IBMWP in relation to the t-BMWQ is likely explained since the t-BMWQ does

Table 6
Quality class boundaries (EQR values) for IBMWP, ICM-9 and ICM-11a for the five Mediterranean stream types.

Stream type	Quality class boundary	IBMWP	ICM-9	ICM-11a
T1	High–Good	0.83	0.90	0.76
	Good–Moderate	0.48	0.60	0.48
	Moderate–Poor	0.28	0.39	0.31
	Poor–Bad	0.16	0.25	0.20
T2	High–Good	0.80	0.85	0.91
	Good–Moderate	0.46	0.57	0.55
	Moderate–Poor	0.27	0.38	0.34
	Poor–Bad	0.16	0.26	0.21
T3	High–Good	0.89	0.92	0.87
	Good–Moderate	0.51	0.61	0.53
	Moderate–Poor	0.29	0.41	0.33
	Poor–Bad	0.17	0.27	0.20
T4	High–Good	0.86	0.92	0.86
	Good–Moderate	0.49	0.61	0.54
	Moderate–Poor	0.29	0.41	0.33
	Poor–Bad	0.16	0.27	0.20
T5	High–Good	0.99	0.99	0.97
	Good–Moderate	0.55	0.57	0.53
	Moderate–Poor	0.31	0.42	0.35
	Poor–Bad	0.17	0.27	0.21

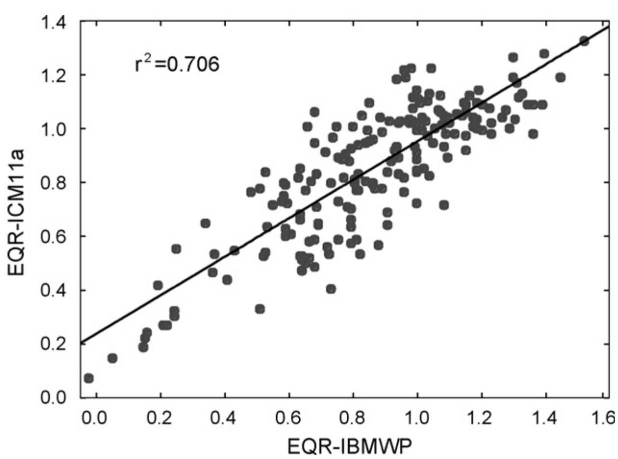

Fig. 5. Linear regression between the IBMWP and the ICM-11a. The adjusted regression coefficient (r^2) is shown.

not include taxa belonging to different orders such as Diptera (e.g., Dixidae and Stratiomyidae), Heteroptera (Gerridae, Hydrometridae, Veliidae, Mesovellidae and Pleidae), Mollusca (Thiaridae) and Coleoptera (Helophoridae), among others, which are present in the study area and some of which appear very frequently. Hence the IBMWP represents a better composition and structure of the macroinvertebrate community in our study area. However, this finding contrasts with the result obtained in the Mediterranean rivers of the Spanish Castilla la Mancha Region where the t-BMWQ was more sensitive to a gradient of organic pollution than the IBMWP (Navarro, 2006).

It must be also emphasised, that the ICM-11a showed a better adjustment to the stressor gradient than the ICM-9. This finding demonstrates that the ICM-11a is more appropriate than the ICM-9, proposed by the Med-GIG, for the stream types analysed, as observed previously by Munné and Prat (2009) in Mediterranean streams.

Finally, in the present study metrics based on only qualitative data were used. Although abundance measures are explicitly demanded in the WFD, these are known to have high variances and are rarely used in multimetric approaches (Barbour et al., 1999). Particularly in the Mediterranean streams and rivers, it has been proved that abundance metrics discriminate worse different levels of degradation than the presence-abundance metrics due to the strong fluctuation of hydrologic regimens in these streams that produce a high variability in taxa abundance (Pinto et al., 2004).

4.3. Setting class boundaries

The definition of class boundaries is a crucial step for implementing the WFD (e.g., Buffagni et al., 2006). Given the differences in the taxonomic composition between the studied stream types (see Sánchez-Montoya et al., 2007) class boundaries must be calibrated separately for each stream type. In this study, we used the specific reference condition of each stream type and the stressor gradient for this harmonisation.

As previously mentioned, the reference values of all studied metrics were lower for the temporary stream than for the rest of permanent ones, due to the influence of natural intermittency of flow on the macroinvertebrate communities (see Sánchez-Montoya et al., 2009a,b). This finding has been also recognised by others authors in temporary streams (e.g., Boulton et al., 2000; Sheldon, 2005; Argyroudi et al., 2009). Moreover, it must be emphasized the higher variability of the EQR reference values in the temporary stream type, due to the large variability produced by different reference communities that are included in this stream type, and which may justify a further division of this stream type into subgroups to ensure the proper application of the WFD (Sánchez-Montoya et al., 2007). In contrast, the lower variability of the EQR reference values in the case of the large watercourse stream type (T5) and therefore the highest values of class boundary between high and good in all the studied metrics, might be attributed to the low number of reference sites found within this stream type. Hence, further efforts must be made to find reference sites in order to improve the ecological classification in this stream type.

5. Conclusions

The results obtained in this work support the knowledge that the IBMWP is suitable for the biological quality assessment in Spanish Mediterranean streams (Munné and Prat, 2009). Since the European Commission suggests the use of multimetric indices and based on the results obtained in this work, we also recommend the use of the multimetric index ICM-11a for Mediterranean streams, as Munné and Prat (2009) previously suggested.

Finally, apart from the general approach to analyse metrics with a multiple stressor gradient which has been used in this study for Mediterranean streams, further research into stressor-specific metrics ought to be considered with a view to possibly improving ecological status classification in streams (Ofenböck et al., 2004; Clews and Ormerod, 2009).

Acknowledgements

This study was carried out thanks to our use of the GUADALMED Project database (Phase I and II) and all researchers who worked in this project. This research was supported by the GUADALMED-2 project (REN2001-3438-C07-01) and the pre-doctoral grant to M.M. Sánchez-Montoya from the Spanish Ministry of Science and Technology. Special thanks to Victoria García and Andrés Mellado for their help in the fieldwork in the Segura Basin. Thanks to J. Pajarón for his help with the figures of this article.

References

Alba-Tercedor, J., Sánchez-Ortega, A., 1988. Un método rápido y simple para evaluar la calidad biológica de las aguas corrientes basado en el de Hellawell (1978). Limnetica 4, 51–56.

Alba-Tercedor, J., Jáimez-Cuellar, P., Álvarez, M., Avilés, J., Bonada, N., Casas, J., Mellado, A., Ortega, M., Pardo, I., Prat, N., Rieradevall, M., Robles, S., Sáinz-Cantero, C.E., Sánchez-Ortega, A., Suárez, M.L., Toro, M., Vidal-Abarca, M.R., Vivas, S., Zamora-Muñoz, C., 2004. Caracterización del estado ecológico de ríos mediterráneos ibéricos mediante el índice IBMWP (antes BMWṔ). Limnetica 21, 175–186.
APHA, 1998. Standard Methods for the Examination of Water and Waste Water, 20th ed. American Public Health Association, Washington.
Argyroudi, A., Chatzinikolaou, Y., Poirazidis, K., Lazaridou, M., 2009. Do intermittent and ephemeral Mediterranean rivers belong to the same river type? Aquat. Ecol. 43, 465–476.
Barbour, M.T., Stribling, J.B., Gerritsen, B.D., 1996. Biological Criteria: Technical Guidance for Streams and Small Rivers. EPA/822/B-96/001, US, Environmental Protection Agency, Washington, DC.
Barbour, M.T., Gerritsen, B.D., Snyder, B.D., Stribling, J.B., 1999. Rapid bioassessment protocols for use in streams and wadeable rivers. EPA/841/B-99/002, US Enviromental Protection Agency, Washington, DC.
Birk, S., Hering, D., 2006. Direct comparison of assessment methods using benthic macroinvertebrates: a contribution to the EU Water Framework Directive intercalibration exercise. Hydrobiologia 566, 401–415.
Boulton, A.J., Sheldon, F., Thoms, M.C., Stanley, E., 2000. Problems and constraints in managing rivers with variable flow regimes. In: Boon, P.J., Davies, B.R., Petts, G.E. (Eds.), Global Perspectives on River Conservation: Science, Policy and Practice. Wiley, Chichester, pp. 415–430.
Buffagni, A., Erba, S., Cazzola, M., Murray-Bligh, J., Soszka, H., Genoni, P., 2006. The STAR common metrics approach to the WFD intercalibration process: full application for small, lowland rivers in three European countries. Hydrobiologia 566, 379–399.
Camargo, J.A., 1993. Macrobenthic surveys as a valuable tool for assessing freshwater quality in the Iberian Peninsula. Environ. Monit. Asses. 24, 71–90.
CEDEX (Centro de Estudios y Experimentación de Obras Públicas del Ministerio de Fomento), 2004. Selección preliminar de posibles tramos fluviales de referencia, Versión 1, Madrid.
Chessman, B.C., McEvoy, P.K., 1998. Towards diagnostic biotic indices for river macroinvertebrate. Hydrobiologia 364, 169–182.
Clews, E., Ormerod, J., 2009. Improving bio-diagnostic using simple combinations of standard biotic indices. River Res. Appl. 25, 348–361.
European Commission, 2000. Directive 2000/60/EC of the European Parliament of the Council of 23rd October 2000 establishing a framework for community action in the field of water policy. Official J. Eur. Commun. 327, 1–72.
European Commission, 2003. Common implementation strategy for the Water Framework Directive (2000/60/EC), Working Group REFCON, Guidance document no. 10, Riversand lakes – Tipology, reference conditions and classification systems.
European Commission, 2005. Common Implementation Strategy for the Water Framework Directive (2000/60/EC). Guidance Document No. 14. Guidance on the Intercalibration Process 2004–2006.
European Commission, 2007. MedGIG Intercalibration technical report – Part 1 Rivers. Section1 Benthic Invertebrates, 15 June.
Gasith, A., Resh, V.H., 1999. Streams in Mediterranean climate region: abiotic influences and biotic responses to predictable seasonal events. Annu. Rev. Ecol. Evol. S. 30, 51–81.
Heiskanen, A.S., van De Bund, W., Cardoso, A.C., Noges, P., 2004. Towards good of ecological status of surface waters in Europe – interpretation and harmonisation of the concept. Water Sci. Technol. 49, 169–177.
Hering, D., Feld, C.K., Moog, O., Ofenböck, T., 2006. Cook book for the development of a Multimetric Index for biological condition of aquatic ecosystems: experiences from the European AQEM and STAR projects and related initiatives. Hydrobiologia 566, 311–324.
Jáimez-Cuéllar, P., Vivas, S., Bonada, N., Robles, S., Mellado, A., Álvarez, M., Avilés, J., Casas, J., Ortega, M., Pardo, I., Prat, N., Rieradevall, M., Sáinz-Cantero, C.E., Sánchez-Ortega, A., Suárez, M.L., Toro, M., Vidal-Abarca, M.R., Zamora-Muñoz, C., Alba-Tercedor, J., 2004. Protocolo Guadalmed (PRECE). Limnética 21 (2002), 187–204.
Karr, J.R., 1981. Assessment of biotic integrity using fish communities. Fisheries 6, 21–27.
Karr, J.R., 1999. Defining and measuring river health. Freshwater Biol. 41, 221–234.
MIMAM (Ministerio de Medio Ambiente), 2000. Libro Blanco del Agua en España. Ministerio de Medio Ambiente, Madrid, España.
MMARM (Ministerio de Medio Ambiente, y Medio Rural y Marino), 2008. Instrucción de Planificación. Ministerio de Medio Ambiente, Madrid, España.
Munné, A., Prat, N., Solà, C., Bonada, N., Rieradevall, M., 2003. A simple field method for assessing the ecological quality of riparian habitat in rivers and streams: QBR index. Aquat. Conserv.: Mar. Freshwat. Ecosyst. 13, 147–163.
Munné, A., Prat, N., 2009. Use of macroinvertebrate-based multimetric indices for water quality evaluation in Spanish Mediterranean rivers. an intercalibartion approach with the IBMWP index. Hydrobiologia 628, 203–225.
Navarro, C., 2006. El Estado Ecológico de los ríos de Castilla-La Mancha. PhD Thesis. University of Castilla-La Mancha, Spain.
Ofenböck, T., Moog, O., Gerritsen, J., Barbour, M., 2004. A stressor specific multimetric approach for monitoring running waters in Austria using benthic macroinvertebrates. Hydrobiologia 516, 251–268.
Pardo, I., Álvarez, M., Casas, J., Moreno, J.L., Vivas, S., Bonada, N., Alba-Tercedor, J., Jaímez-Cuéllar, P., Moyá, G., Prat, N., Robles, S., Suárez, M.L., Toro, M., Vidal-Abarca, M.R., 2004. El hábitat de los ríos mediterráneos. Diseño de un índice de diversidad de hábitat. Limnetica 21, 115–133.
Pinto, P., Rosado, J., Morais, M., Antunes, I., 2004. Assessment methodology for southern siliceous basins in Portugal. Hydrobiologia 516, 191–214.
Pollard, P., van De Bund, W., 2005. Template for the development of a boundary setting protocol for the purposes of the intercalibration exercise. Agreed version of WG 2.A Ecological Status, Version 1.2. 6 June 2005, Ispra.
Prat, N., 2004. El Proyecto GUADALMED. Limnetica 21, 1–3.
Reynoldson, T.B., Metcalfe-Smith, J.L., 1992. An overview of the assessment of aquatic ecosystem health using benthic invertebrates. J. Aquat. Anim. Health 1, 295–308.
Sánchez-Montoya, M.M., Puntí, T., Suárez, M.L., Vidal-Abarca, M.R., Rieradevall, M., Poquet, J.M., Zamora-Muñoz, C., Robles, S., Álvarez, M., Alba-Tercedor, J., Toro, M., Pujante, A., Munné, A., Prat, N., 2007. Concordance between ecotypes and macroinvertebrate assemblages in Mediterranean streams. Freshwater Biol. 52, 2240–2255.
Sánchez-Montoya, M.M., Vidal-Abarca, M.R., Puntí, T., Poquet, J.M., Prat, N., Rieradevall, M., Alba-Tercedor, J., Zamora-Muñoz, C., Toro, M., Robles, S., Álvarez, M., Suárez, M.L., 2009a. Defining criteria to select reference sites in Mediterranean streams. Hydrobiologia 619, 39–54.
Sánchez-Montoya, M.M., Suárez, M.L., Vidal-Abarca, M.R., 2009b. Seasonal and interannual variability of macroinvertebrate reference communities and its influence on bioassessment in different Mediterranean stream types. Fundam. Appl. Limnol. (Arch. Hydrobiol.) 174, 353–367.
Sandin, L., Hering, D., 2004. Comparing macroinvertebrate indices to detect organic pollution across Europe: a contribution to the EC Water Framework Directive intercalibration. Hydrobiologia 516, 55–68.
Sheldon, F., 2005. Incorporating natural variability into the assessment of ecological health in Australian dryland rivers. Hydrobiologia 552, 45–56.
Stat Soft, Inc., 1999. STATISTICA for Windows (Computer Program Manual), Stat Soft, Inc., Tulsa, OK.
Stoddard, J.L., Larse, D.P., Hawkins, C.P., Johnson, R.K., Norris, R.H., 2006. Setting expectations for the ecological conditions of streams: the concept of reference condition. Ecol. Appl. 16, 1267–1276.
Wallace, J.B., Grubaugh, J.W., Whiles, M.R., 1996. Biotic indices and stream ecosystem processes: results from an experimental study. Ecol. Appl. 6, 140–151.
Walsh, J.C., 2006. Biological indicators of stream health using macroinvertebrate assemblage composition: a comparison of sensitivity to an urban gradient. Mar. Freshwater Res. 54, 37–47.