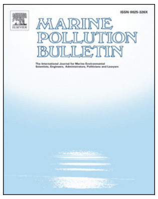
Marine Pollution Bulletin 131 (2018) 507–514

Contents lists available at ScienceDirect

Marine Pollution Bulletin

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



Assessing the effectiveness of surrogates for species over time: Evidence from decadal monitoring of a Mediterranean transitional water ecosystem

S. Bevilacqua[a,d,*], M. Mistri[e], A. Terlizzi[b,c,d], C. Munari[e]

[a] Laboratory of Zoology and Marine Biology, Department of Biological and Environmental Sciences and Technologies, University of Salento, 73100 Lecce, Italy
[b] Department of Life Sciences, University of Trieste, 34127 Trieste, Italy
[c] Stazione Zoologica Anton Dohrn, 80121 Napoli, Italy
[d] CoNiSMa, Piazzale Flaminio 9, 00196 Roma, Italy
[e] Department of Chemical and Pharmaceutical Sciences, University of Ferrara, 44121 Ferrara, Italy

ARTICLE INFO

Keywords:
BestAgg
Coastal lagoons
Long-term monitoring
Macro-invertebrates
Taxonomic sufficiency
Temporal variations

ABSTRACT

The use of higher taxa or alternative approach to species surrogacy, such as the BestAgg procedure, could represent cost-effective solutions to avoid expensive species-level identifications in monitoring activities, especially on the long term. However, whether a set of surrogates would be effective in subsequent reiteration of the same assessment remains largely unsolved. We used a long-term dataset on macro-benthic assemblages to test the hypothesis that family-level and BestAgg surrogates which are effective for a limited period of monitoring could be successfully applied to quantify community patterns also in subsequent monitoring programmes. The effectiveness of surrogates in detecting temporal variations in assemblage structure as at species level remained basically unaffected over a decade. Recognizing once and for all if species surrogacy may have a practical value for monitoring will strongly depend on future assessments of the potential of surrogates to reflect community changes and to retain this prerogative over time.

1. Introduction

Environmental and biological monitoring is at the core of applied ecological research, providing invaluable insights on patterns and processes underlying the dynamics of ecosystems, and producing sets of data that are instrumental for progresses in theoretical ecology (Lovett et al., 2007; Lindenmayer et al., 2012). Monitoring is also essential for environmental policy, since systematic collections of data are necessary to inform the adaptive management of environmental issues (Lindenmayer and Likens, 2009), whether concerning the assessment and mitigation of human impacts (Bustamante et al., 2012; Ellingsen et al., 2017), the effectiveness of conservation strategies (Fraschetti et al., 2012), the success of restoration actions (Block et al., 2001), or the surveillance of the ecological quality status of ecosystems (Borja and Dauer, 2008).

The concept of monitoring intrinsically implies performing replicated observations through time, since single assessments cannot provide a comprehensive characterization of systems being investigated. This because communities and ecosystems are not static entities, which are subject to a complex interplay of processes acting at a range of spatial, but also, temporal scales, and historical data are often a prerequisite for a deeper understanding of mechanisms driving ecological changes (Lovett et al., 2007; Lindenmayer et al., 2012; Mieszkowska et al., 2014). However, sustaining data collection on the long term is expensive, requiring the availability of adequate and continuous funding, which often represents a critical aspect for the maintenance of effective monitoring programmes (Hewitt and Thrush, 2007) and stimulates advances in optimization strategies (Mueller and Geist, 2016; Ellingsen et al., 2017). Major cost components in monitoring activities, particularly when focusing on invertebrate taxa, concern species-level identifications of organisms (Mandelik et al., 2010). Reducing expense related to the achievement of fine taxonomic resolution, therefore, may allow allocating additional budget to extend the spatial coverage and/or the time span of routine monitoring programmes (Mueller et al., 2013).

Although several approaches have been proposed to save efforts during sample processing both reducing the number of operational units and simplifying their identification, including the use of species subsets, cross-taxon congruence, or morphological groups, the analysis of communities at taxonomic levels higher than species (e.g., genus, family) has been the mainstream procedure so far, especially in aquatic environments (Dauvin et al., 2003; Jones, 2008; Sánchez-Moyano et al., 2017). This last approach relies on the concept of taxonomic sufficiency (Ellis, 1985), which assumes that, to some extent, ecological patterns

---

[*] Corresponding author at: Department of Biological and Environmental Sciences and Technologies, University of Salento, I-73100 Lecce, Italy.
E-mail address: stanislao.bevilacqua@unisalento.it (S. Bevilacqua).

https://doi.org/10.1016/j.marpolbul.2018.04.047
Received 23 February 2018; Received in revised form 6 April 2018; Accepted 18 April 2018

can be quantified at coarse levels of taxonomic resolution without substantial loss of information (Ferraro and Cole, 1995). Taxa higher than species, and particularly genera and families, have been successfully employed as surrogates for species in a wide range of environmental contexts in marine, freshwater and terrestrial ecosystems (Bevilacqua et al., 2012). However, taxonomic sufficiency implies resolving organisms to a single higher taxonomic level, irrespective of their ecological importance or actual difficulty in their identification, often leading to unnecessary and undesirable loss of taxonomic detail and related ecological information (Groc et al., 2010). More importantly, the sufficient taxonomic level for a given study is decided empirically, in the absence of any control for statistical uncertainty (Mellin et al., 2011).

An alternative approach to identify suitable sets of surrogates, the 'Best practicable aggregation of species' (BestAgg; Bevilacqua et al., 2013; Bevilacqua and Terlizzi, 2016), has been proposed based on evidence that the relationships between taxonomic relatedness of species and their ecological response to natural and human-driven changes could be largely neutral, and that the effectiveness of higher taxa to retain species-level information patterns might depend on numerical relationships across ranks of the Linnaean taxonomic hierarchy (Neeson et al., 2013; Rosser, 2016). BestAgg utilizes null models of decreasing information on species-level patterns at decreasing number of operational units used to substitute species, in order to identify the minimum number of surrogates, regardless of their type, needed to quantify community patterns as at species level. In this framework, surrogate selection is unconstrained to strict taxonomic criteria, allowing the use of any meaningful operational units (e.g., taxonomic, functional, morphological, mixed) in order to maximize ecological information (Groc et al., 2010; Bevilacqua et al., 2013). Hence, the approach determines statistical thresholds for the application of surrogates with associated estimates of uncertainty (Mueller et al., 2013), providing a more formal procedure for surrogate selection (Bevilacqua and Terlizzi, 2016) if compared to the classic approach of taxonomic sufficiency (Jones, 2008).

The rationale underlying species surrogacy, irrespective of the approach employed, is that community patterns can be quantified equally well using operational units other than species, while reducing costs and time needed for fine taxonomic identifications (Ellis, 1985; Ferraro and Cole, 1995; Warwick, 1993). Once such operational units (i.e., surrogates for species) are properly identified from baseline data, they are assumed to be effective in comparable systems, or in reiteration of the same study such as, for instance, in repeated monitoring programmes (Olsgard and Somerfield, 2000; Mueller et al., 2013). Most of studies on surrogates, nonetheless, explored their performance in quantifying spatial changes in assemblage structure (e.g., Olsgard et al., 1998; Olsgard and Somerfield, 2000; Bevilacqua et al., 2009; Terlizzi et al., 2009; Heino, 2014), with limited efforts to extend the application of a given set of surrogates from a local to a wider geographic context (e.g., Bevilacqua et al., 2015). Very few works have examined the effectiveness of surrogates to reflect temporal variations (e.g., Musco et al., 2011; Cabral et al., 2017), and no attempts have never been done, except for rare short-term evaluations (e.g., Magierowski and Johnson, 2006; Thiault et al., 2015), to assess whether the ability of a given set of surrogates to reproduce species-level patterns is maintained over time. Here, we used long-term data on macro-benthic assemblages from a coastal lagoon system to test the hypothesis that surrogates defined on the basis of a given period of monitoring can be successfully applied to quantify patterns of temporal variation also in successive monitoring programmes.

2. Material and methods

2.1. Study area and sampling design

The study focused on a decadal dataset of soft bottom macrobenthos from the Valli di Comacchio (N Adriatic Sea), a complex of shallow (depth range 0.5–1.5 m) brackish lagoons extending over 100 km^2 in the southern part of the Po River delta (NE Italy). Data were collected yearly during several sampling campaigns carried out from 1997 to 2006 (Munari and Mistri, 2010), with four (every 3–4 months) times of sampling (hereafter referred to as dates) per year, except for 1999, 2000, 2005, and 2006, in which only three dates were available. In each date, four stations representative of the whole lagoon system were sampled. One sample, consisting of three replicate Van Veen grabs (0.06 m^2) of soft sediments, was collected in each station, for a total of 144 samples. Samples were sieved (0.5 mm-mesh size) and macro-invertebrates were sorted, identified at the species level and counted.

2.2. Simulated monitoring programmes

The identification of surrogates requires exhaustive pilot assessments at species level, able to characterize the structure of assemblages under study and to provide data as representative as possible of their spatial and temporal patterns of variation (Terlizzi et al., 2003; Heino and Soininen, 2007). In our case, preliminary multivariate analyses using species-level data showed that at least four years of sampling were needed to detect significant infra- and interannual variations in macro-invertebrate assemblages (see Table A1 in Appendix A), suggesting that, in the investigated system, a five-years programme may be sufficient for routine monitoring of temporal variations at species level and, seemingly, long enough also for the identification of successful surrogates. Therefore, we split our ten-years dataset in two halves, simulating the realistic situation in which two rounds of monitoring of five years each have been carried out to assess temporal trends in macro-invertebrate assemblages of the lagoon system. The first five years of the dataset, from 1997 until 2001, represented the 'pilot' study that was employed to define and validate surrogates for species. Data from 2002 until 2006, instead, were used as a 'test' study to assess whether surrogates would have been effective for quantifying community patterns in a successive monitoring programme.

2.3. Defining surrogates for species

Surrogates for species were defined based on pilot data (1997–2001) following the BestAgg procedure (Bevilacqua et al., 2013; Bevilacqua and Terlizzi, 2016), and the traditional approach of taxonomic sufficiency (Ellis, 1985; Olsgard et al., 1998; Dauvin et al., 2003). BestAgg focuses on determining to which extent the original S species in a data matrix can be aggregated (i.e., grouped and their abundance summed) into a smaller number of G surrogates, irrespective of their type (i.e., whether taxonomic, morphological, functional, or even mixed), before a substantial loss of information on species-level community patterns occurred. For a given sampling design, and for any term of interest in the analysis, the BestAgg procedure generates a null model of decreasing information, quantified as the Spearman's correlation ρ between the species-level matrix and the corresponding aggregated matrix, at increasing level of species aggregation, expressed as $\varphi = G/S$. The null model is built by random aggregations of the original S species in decreasing numbers of G groups, and obtaining 1000 randomly-aggregated matrices for each set of G groups. For each aggregated matrix, the correlation value (ρ) with the original species-level matrix is then calculated, and a distance-based permutational multivariate analysis of variance (PERMANOVA; Anderson, 2001) is also performed to test for the main term of interest in the analysis (in this case the variability among years). This allows assessing that the information retained in the surrogate matrices and, consequently, the ability of surrogates to reflect species-level community patterns, is correlated to φ. In addition, the null model identifies the minimum number of surrogates G_{min}, sufficient to obtain results consistent with those obtained using species, with a probability of type-I error (i.e., the probability of considering G_{min} as sufficient when actually it is not) of

$P < 0.05$ (for further details on the procedure see Bevilacqua et al. (2013) and Bevilacqua and Terlizzi (2016). Once the overall G_{min} is determined, the approach aims to maximize ecological information by identifying surrogates based on three macro-criteria: (1) general (e.g., conservation or commercial value, indicator, alien species) and/or specific (e.g., species most contributing to the observed patterns) ecological importance for the structure and functioning of the investigated system (*relevance*); (2) difficulty of taxonomic identification during sample processing (*easiness*); and (3) shared characteristics among organisms that allow meaningful species grouping (*resemblance*). General information on species ecology (e.g., sensitiveness/tolerance to environmental variations, functional role, etc.) was mined from the scientific literature (see Borja et al., 2000; Mistri and Munari, 2008; Bevilacqua et al., 2015 and reference therein), whereas relevant species for the specific community patterns under study were identified through SIMPER analysis (Clarke, 1993) on the species-level matrix (see Table A2 in Appendix A). Easiness was defined based on expert opinion (Bevilacqua et al., 2015), whereas resemblance (i.e., species grouping) was decided by combining the relevance and easiness of species with the available taxonomic and ecological knowledge of groups in order to obtain the final set of BestAgg surrogates. Although the number of surrogates in BestAgg can be even higher than G_{min}, we aggregated species into a set of surrogates equal to this value to test the performance of BestAgg in the most conservative condition.

A more classic approach to species surrogacy based on the concept of taxonomic sufficiency was also applied to define effective surrogates for species. In this approach, the identification of surrogates is restricted to the use of taxa higher than species, and relies simply on an empirical determination of the sufficient taxonomic resolution for the analysis. Species-level data are generally aggregated at coarser level of taxonomic resolution (e.g., genus, family, order, etc.) and analysed in order to find the taxonomic level at which results are still consistent with those obtained using species (e.g., Ferraro and Cole, 1995; Olsgard et al., 1998; Heino, 2014). Previous studies on macro-benthic assemblages from coastal lagoons in the same geographic region, and in other regions of the Mediterranean Sea, suggested that species-level community patterns can be successfully elucidated at the level of family (Mistri and Munari, 2008; Bevilacqua et al., 2015), which was tested here as sufficient taxonomic resolution for analysis.

2.4. Validating surrogates on the pilot programme

Data from the pilot programme (1997–2001) were analysed using species, families, and the BestAgg operational units in order to verify that the selected surrogates were actually able to describe patterns of temporal variation in the structure of macro-invertebrate assemblages as at species level.

A PERMANOVA based on Bray–Curtis dissimilarity calculated on species-level data was carried out. No transformation of data was applied and each term in the analysis was tested using 2000 permutations. The design for the analysis consisted of two factors, Years [Y, 5 levels, fixed] and Date [Da(Y), 3–4 levels, random and nested in Y)] with $n = 4$. Species-level data were aggregated respectively at (1) the level of family and (2) using the BestAgg surrogates. PERMANOVA was carried out separately for the two aggregated matrices, and results compared with those obtained using species. Pair-wise correlation (Spearman's ρ) between species and aggregated matrices was also calculated.

Non-metric multidimensional scaling ordinations (nMDS) of Da(Y) centroids were plotted to visualize patterns of variation through time based on species-level data and surrogates. Centroids were obtained calculating principal coordinates (PCO) on the basis of the Bray–Curtis dissimilarity matrix among all pairs of samples. A Procrustean randomization test (PROTEST; Jackson, 1995) was performed to compare of ordination plots between species and each of the two types of surrogates (i.e., families and BestAgg). The m^2 statistic in PROTEST ranges between 0 and 1, representing a measure of the goodness-of-fit between ordination plots (the lower is the m^2, the more the two configurations match), with the associated probability that the observed m^2 is significantly smaller than what expected by chance (Jackson, 1995).

2.5. Applying surrogates to successive monitoring programme

Surrogates defined and validated based on pilot data (1997–2001) were employed to analyse data from 2002 to 2006 (test programme), in order to investigate their performance when applied to subsequent periods of monitoring. Three separated PERMANOVAs (design and settings as above) were done using species, families, and BestAgg surrogates respectively, and results compared. Patterns of temporal variations were plotted with nMDS ordinations, and configuration plots compared through PROTEST, as previously described.

The analyses were repeated based on the full data series, from 1997 until 2006, to assess whether surrogates were effective in representing temporal trajectories of assemblages over the whole investigated period as at species level. In this case, the design for the analysis consisted of two factors, Years [Y, 10 levels, fixed] and Date [Da(Y), 3–4 levels, random and nested in Y] with $n = 4$. Temporal trajectories of assemblages over the decade were visualized through nMDS of Y centroids and compared using PROTEST.

Finally, a Canonical Analysis of Principal Coordinates (CAP; Anderson and Willis, 2003) based on the Bray-Curtis dissimilarity BestAgg matrix (full data series) among dates was performed for factor Year (Y) to show main groups of organisms responsible of the observed interannual variations in the structure of assemblages. Correlations of variables with the two canonical axes (r_1 and r_2) were then represented as lines in the CAP plot. Variables were included in the plot only if exceeding an arbitrarily threshold of correlation (i.e., $\sqrt{r_1^2 + r_2^2} \geq 0.3$).

Analyses were performed using R (R Development Core Team, 2016) and the software PRIMER 6 with the add-on package PERMANOVA+.

3. Results

The information retained in the randomly-aggregated matrices, expressed as the Spearman's correlation ρ with the original species matrix, decreased at decreasing ϕ following a semi-log model (Fig. 1a), indicating that the information on species-level patterns was progressively lost as the number of surrogates (G) decreased. The null model predicted that the minimum number of surrogates G_{min}, allowing the 95% of PERMANOVA tests for significant interannual variations to give results consistent with those obtained at species level, was $G_{min} = 35$ (Fig. 1b).

A set of 35 BestAgg surrogates were selected based on the pilot programme (see Table A3 in Appendix A). Most surrogates were taxonomic, although in some cases they corresponded to ecological or indicator groups (Table A3), with 12 species (*Arcuatula senhousia*, *Alitta succinea*, *Armandia cirrhosa*, *Carcinus aestuarii*, *Cerastoderma glaucum*, *Chironomus salinarius*, *Corophium orientale*, *Dyspanopeus sayi*, *Ficopomatus enigmaticus*, *Hydroides dianthus*, *Melita palmata*, *Monocorophium insidiosum*), 10 families (Capitellidae, Eunicidae, Hesionidae, Hydrobiidae, Nereididae, Other Serpulidae, Palaemonidae, Phyllodocidae, Tellinidae, Syllidae), six orders (Cumacea, Isopoda, Other Amphipoda, Spionida, Tanaidacea, Terebellida), four classes (Anthozoans, Other Bivalvia, Other Gastropoda, Oligochaeta), two morphological groups (Barnacles, Flat worms), and one phylum (Nemertea).

As far as the pilot programme, PERMANOVA detected a significant variability in species-level assemblage structure among dates and among years (Table 1). Patterns of variation were depicted in the nMDS ordination of Da(Y) centroids, which showed a clear separation between 1997–1998–1999 and 2000–2001, suggesting that most of

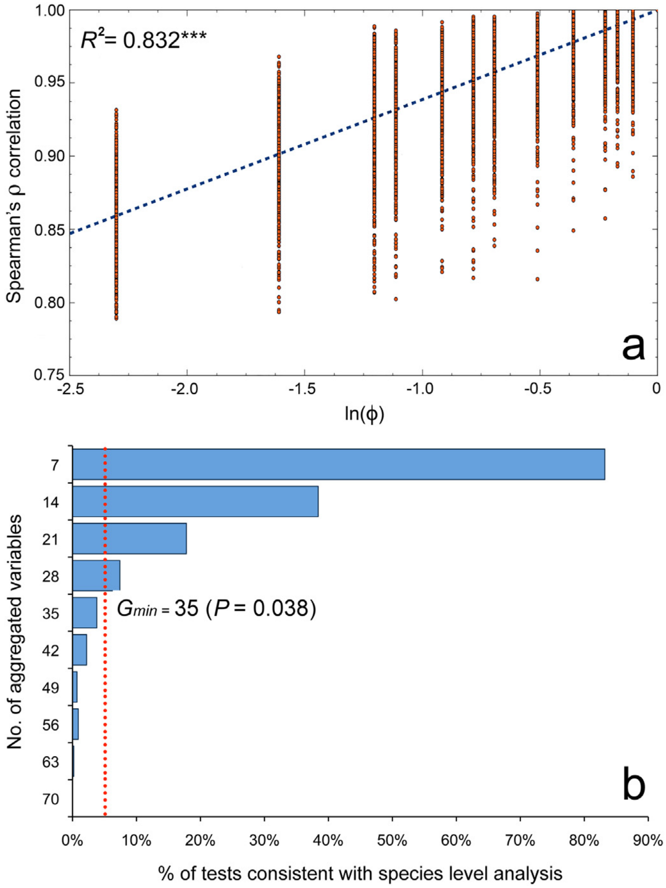


Fig. 1. (a) Semi-log plot of ρ values (Spearman's correlation) between the species-level matrix and each randomly-aggregated matrix against the corresponding ϕ values based on data from 1997 to 2001, which served as pilot programme to quantify the number of BestAgg surrogates G_{min} (see Section 2.3 for further details). The dotted line indicates the linear regression fitted to the data points. R^2 from the regression analysis is also reported. *** = $P < 0.001$. (b) Percentage of tests for significant interannual variations during the pilot programme (1997–2001) based on matrices in which the original species data were aggregated into a decreasing number of G groups (see Section 2.3 for further details). Tests were performed using PERMANOVA on Bray-Curtis dissimilarity of untransformed data, with 2000 permutations. The dotted line indicates the limit of 5% of tests consistent with species-level analysis. The minimum number of surrogates G_{min} and the associated P-value are also reported.

variations were between these two groups of years (Fig. 2). PERMANOVA on data aggregated using the BestAgg surrogates gave equivalent results, and the same occurred when using families (Table 1). PROTEST resulted in a value of m^2 of 0.018 and 0.016 ($P < 0.01$) for BestAgg surrogates and families respectively, indicating that in both cases ordination plots were almost interchangeable with that obtained using species (Fig. 2). The Spearman's correlation ρ with the species-level matrix was 0.985 for BestAgg and 0.983 for families ($P < 0.001$ in both cases), indicating that the selected set of 35 BestAgg surrogates allowed retaining the same amount of information on species-level patterns as the 42 families, although employing a lower number of operational units.

Significant variations among years characterized the species assemblages during the test monitoring programme (2002–2006) but, unlike the previous five years of monitoring, no within-year (among dates) variability was detected (Table 1). Major variations seemed to separate assemblages in 2002 and 2005 from those characterizing the remaining years (Fig. 2). PERMANOVA results based on family-level data and BestAgg surrogates were aligned with those at species level (Table 1). In both cases, configurations of Da(Y) centroids in the nMDS were still comparable with species-level ordination (Fig. 2). The value of m^2 from PROTEST was 0.187 ($P < 0.001$) for BestAgg surrogates and 0.248 for families, indicating a better performance of the former in depicting patterns of temporal variation in assemblage structure.

The analyses on the whole time-series gave equivalent results, whether using species, BestAgg surrogates or families (Table 1). At species level, the ordination of centroids of years showed three distinct clusters, suggesting that assemblage structure mostly varied from 1997–1998–1999 to 2000–2001–2002, and again in the last years of the decade, with the exception of 2005 that seemed to cluster with the second group of years (Fig. 3). The configuration of centroids in the nMDS based on BestAgg surrogates closely matched ($m^2 = 0.079$, $P < 0.01$) the species-level ordination plot (Fig. 3). In contrast, family-level ordination showed a worse representation ($m^2 = 0.227$, $P < 0.01$) of temporal trajectories in species-level assemblages (Fig. 3).

CAP plots for factor Year largely overlapped between species and BestAgg (Fig. 4), reflecting the observed temporal pattern of variation among the three main groups of years and, namely, the first three years (1997–1998–1999) of the decade (right side of the plots), a second group including 2000–2001–2002 and 2005 (upper side), and the last group of years 2003–2004–2006 (left side). The analysis based on BestAgg surrogates showed that macro-invertebrate assemblages in the first group of years were mostly characterized by opportunistic/tolerant [*Alitta succinea* (ALI), Spionida (SPI), *Cerastoderma glaucum* (CER)], alien [*Dyspanopeus sayi* (DYS)] and pollution indicator [*Cossura* spp. (COS)] organisms, whereas the two polychaetes families of Nereididae (NER) and Hesionidae (HES), sensitive organisms [Anthozoans (ANT), Nemertea (RW), Hydrobiidae (HYD)], and few opportunistic/tolerant taxa [Oligochaeta (OLI), *Arcuatula senhousia* (ARC)] characterized the

Table 1
Summary of PERMANOVAs testing for interannual variations in macrobenthic assemblages using species, BestAgg surrogates, and families during the pilot programme (years 1997–2001), the test programme (years 2002 to 2006) and the whole decade (all years from 1997 to 2006). Analyses were based on Bray-Curtis dissimilarities (data untransformed), with 2000 permutations. *** = $P < 0.001$; ** = $P < 0.01$; * = $P < 0.05$; ns = not significant.

Program	Source of variation	df	Species		BestAgg		Family	
			MS	Pseudo-*F*	MS	Pseudo-*F*	MS	Pseudo-*F*
Pilot	Year = Y	4	9792.9	2.27***	8389.7	1.99**	8108.5	1.91**
(1997–2001)	Date(Y) = Da(Y)	13	4321.4	1.36**	4204.7	1.34*	4238.4	1.36*
	Residual	52	3181.1		3137.4		3126.6	
Test	Year = Y	4	10,324.0	2.99***	8702.6	2.80***	8224.6	2.64***
(2002–2006)	Date(Y) = Da(Y)	13	3454.2	1.22^{ns}	3109.3	1.12^{ns}	3120.7	1.13^{ns}
	Residual	52	2838.9		2766.3		2755.4	
All years	Year = Y	9	13,399.0	3.45***	12,431.0	3.40***	11,977.0	3.26***
(1997–2006)	Date(Y) = Da(Y)	26	3887.8	1.29**	3657.0	1.24**	3679.5	1.25**
	Residual	106	3013.2		2955.4		2944.5	

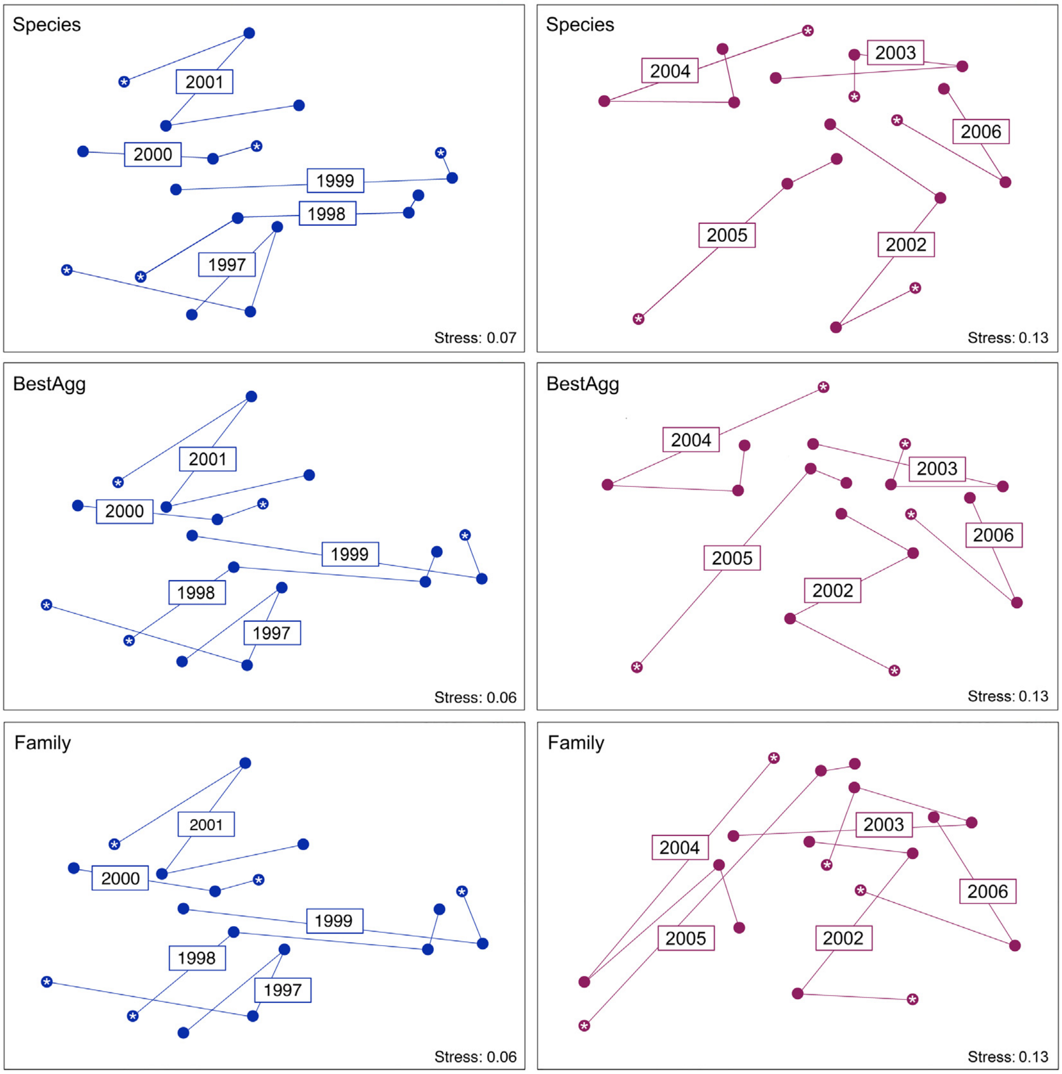


Fig. 2. Non-metric multidimensional scaling ordinations (nMDS) of Da(Y) centroids for the pilot programme (1997–2001, panels on the left side) and for the test programme (2002–2006, panels on the right side) based on Bray–Curtis dissimilarity for species-level data, and data analysed using BestAgg surrogates and families. Centroids of dates from the same year were connected by solid lines, starting from the first date of sampling (indicated with asterisks).

second group of years (Fig. 4). Finally, isopods [*Melita palmata* (MEL), Other Isopoda (ISO)], opportunistic/tolerant [*Monocorophium insidiosum* (MON), *Ficopomatus enigmaticus* (FIC), Terebellida (TER)] and alien [*Hydroides dianthus* (HYD)] organisms were of particular importance in structuring assemblages during the last years of the decade (Fig. 4).

4. Discussion

The scientific literature on species surrogacy typically focused on single-time assessments and the congruence of responses between species and surrogates in subsequent reiterations of the same study has been generally advocated rather than explicitly documented. Indeed, several processes such as, for instance, natural fluctuations in environmental factors or variations in the magnitude of human-driven alterations to the investigated system, could cause changes in composition and/or shifts in relative abundance of species, so that surrogates proved to be effective for specific time intervals might not perform equally well in subsequent periods (Musco et al., 2011; Thiault et al., 2015). The few studies accounting for multiple times of sampling showed that temporal variations in species assemblages may be still noticed analysing genera or families (e.g., Musco et al., 2011; Cabral et al., 2017), suggesting that the use of coarser taxonomy could be a promising approach to enhance cost-effectiveness of monitoring. However, most evidence concerned the effects of taxonomic resolution on ordination of data or short-term experiments (e.g., Cabral et al., 2017) and, above all, involved a range of environmental contexts too limited to consolidate the integration of species surrogacy in routine monitoring programmes. The substantial lack of studies that have tested the performance of a given set of surrogates in successive monitoring campaigns further increases the uncertainty. Rare attempts in this direction showed that taxonomic (i.e., genera, families) and BestAgg surrogates from a single time of sampling were effective also in subsequent samplings when assessing human impacts on sessile macrobenthos (Bevilacqua et al., 2013) or the effects of protection on fish assemblages (Thiault et al., 2015), but referred to monitoring programmes of 2 years or less. In this study, surrogates for species obtained and validated on the basis of a definite period of monitoring were formally tested in a successive independent period showing that, once surrogates were appropriately defined, their ability to detect infra- and interannual variations in assemblage structure as at species level remained basically unchanged over a decade. Results of species-level analyses on the pilot monitoring programme were reproduced consistently using families, which were able to depict temporal patterns

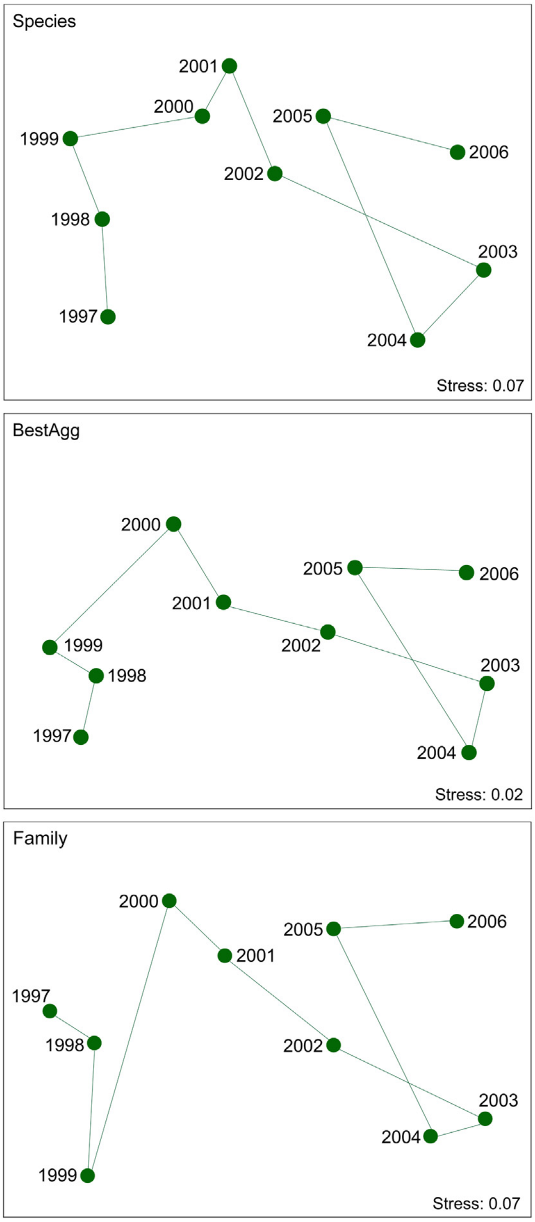


Fig. 3. Non-metric multidimensional scaling ordinations (nMDS) of centroids of years for the whole decade (1997–2006) based on Bray–Curtis dissimilarity for species-level data, and data analysed using BestAgg surrogates and families.

also in the following five years of monitoring, and the same occurred when species were substituted with BestAgg surrogates. Such findings unambiguously demonstrated that species surrogacy could play a crucial role in optimizing trade-offs between the need for saving costs and effective monitoring, especially in long-term assessments. Therefore, surrogates based on a pilot monitoring at species level long enough to capture variations at all the relevant spatial and temporal scales, could be reasonably considered as effective for a decade, before requiring recalibration. However, the 'expiring period' for surrogates could vary depending on the studied system and further investigation are needed in this direction.

The approach of taxonomic sufficiency and the BestAgg procedure were both effective in providing suitable surrogates for species. However, relying on null models and flexible criteria aiming at maximizing ecological information for the identification of surrogates, as in BestAgg, can be a more profitable strategy for species surrogacy than the static use of single taxonomic ranks of the Linnaean hierarchy (Bevilacqua et al., 2013). In our case, a total of 53 families were found during the whole decade of monitoring, which represented the set of operational units to use based on sufficient taxonomy. In contrast, a total of 40 BestAgg surrogates were sufficient to substitute the pool of 98 species found in the study area during the whole decade (see Table A3 in Appendix A). Since temporal trajectories of species assemblages were approximated more closely when using BestAgg surrogates rather than families, and cost savings derived from the application of surrogates are proportional to the reduction in the number of operational units (Ferraro and Cole, 1995), BestAgg represented a more efficient and accurate procedure for species surrogacy if compared to the traditional higher-taxon approach. Such findings confirmed the results of other studies on macro-invertebrate assemblages from coastal lagoons and subtidal rocky reefs where BestAgg surrogates were 20–50% more efficient in saving time for sample processing and identifications with respect to family-level analysis (Bevilacqua et al., 2015; Bevilacqua and Terlizzi, 2016).

The ability of BestAgg surrogates to outperform the use of higher taxa as surrogates for species has been highlighted in several studies that analysed changes in macro-invertebrate assemblages related to human impacts (Bevilacqua et al., 2013; Bevilacqua and Terlizzi, 2016; Jiang et al., 2017) and natural environmental gradients (Bevilacqua et al., 2013; Milošević et al., 2014) in marine and freshwater ecosystems. Once BestAgg surrogates were defined for a given area, they were also effective in detecting spatial and spatio-temporal patterns of assemblage variation in neighbouring systems (Bevilacqua et al., 2013, 2015). Our findings showed that this aptitude of BestAgg surrogates applied when they are defined based on a given period and extended to future assessments. This better performance originates from the procedure for surrogate selection on which BestAgg is structured. As first, and unlike the higher-taxon approach in which the sufficient taxonomic resolution dictates the number of operational units to use (i.e., the number of taxa belonging to the sufficient taxonomic rank), null models in BestAgg only fixes the minimum number of surrogates needed. Therefore, a number of surrogates even higher than this threshold value can be selected, if this allows retaining much more ecological information on species assemblages without increasing identification efforts. Although this last aspect seems to introduce some level of subjectivity, surrogate selection is far from being arbitrary: (i) the identity of surrogates is determined based on objective macro-criteria, (ii) the relevance of case-specific surrogates relies on statistical analyses (SIMPER), (iii) whereas general ecological relevance is to be substantiated by scientific evidence, and (iv) the selected set of surrogates in BestAgg can be tested against null expectations (Bevilacqua et al., 2013). Also, as null models involve thousands of simulated scenarios of species aggregation, the predictive power of BestAgg is more robust to the confounding effects of natural variations in space and time with respect to the one-off empirical determination of the sufficient taxonomic level that characterize the higher-taxon approach (Bevilacqua et al., 2015; Thiault et al., 2015). Finally, the BestAgg procedure prioritizes the selection of surrogates that are ecologically relevant and easy to identify, and restricts species grouping to cases in which this is justified by particular difficulties in identifications. In this view, any taxon or group of organisms (e.g., morphological groups), even single species or genera if of easy identification, can be included in the set of BestAgg surrogates, which in turn may result in a close match between surrogate and species-level community patterns (Magierowski and Johnson, 2006). In addition, considering relevant and easy operational units (by retaining important species of simple identification, and/or lumping difficult species into easier groups that are still ecologically meaningful) allows reducing the expertise needed for classification of organisms. Therefore, even non-expert taxonomist could process samples, after being adequately trained to identify the selected set of operational units. The fact that BestAgg places emphasis on the ecological information that surrogates could convey further enhances their sensitivity to relevant environmental variations (Milošević et al., 2014;

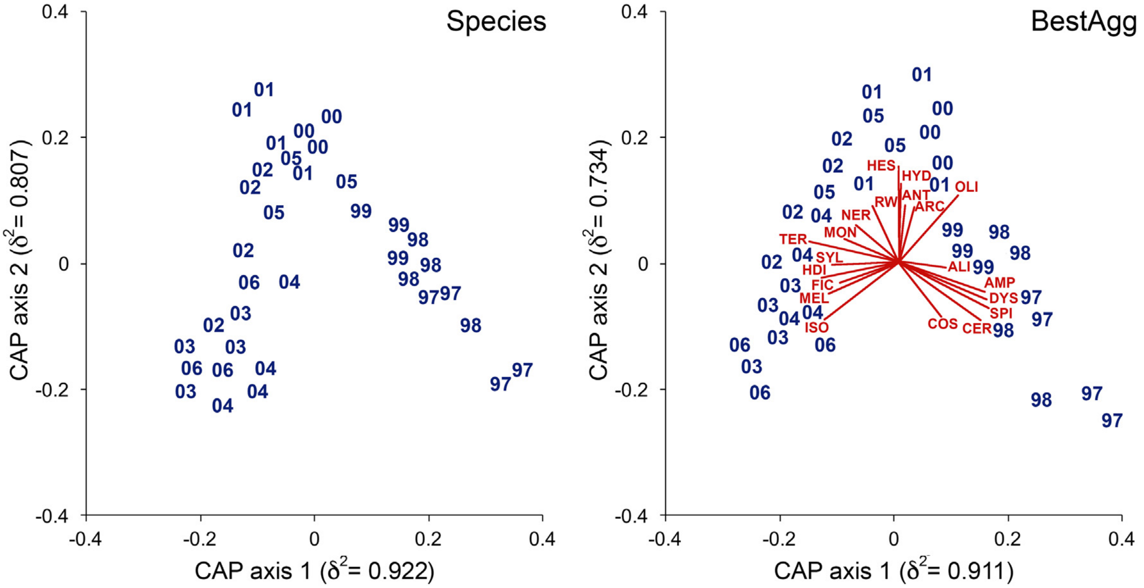


Fig. 4. CAP plots for the factor Year based on the distance matrix among dates of the whole time-series, at species level and using the BestAgg surrogates. Numbers indicates years from 1997 until 2006. Individual surrogates highly correlated with canonical axes are reported (vectors). *A. succinea* (ALI), *A. sarsi* (AMP), Anthozoans (ANT), *A. senhousia* (ARC), *C. glaucum* (CER), *Cossura* spp. (COS), *D. sayi* (DYS), *F. enigmaticus* (FIC), Hesionidae (HES), Hydrobiidae (HYD), *H. dianthus* (HYD), *M. palmata* (MEL), *M. insidiosum* (MON), Nereididae (NER), Oligochaeta (OLI), Other Isopoda (ISO), Nemertea (RW), Spionida (SPI), Syllidae (SYL), Terebellida (TER).

Jiang et al., 2017), and help the interpretation of the observed community patterns. In our case, for example, many species were retained in the final set of BestAgg surrogates because of easy identification and ecologically relevant, being opportunistic, indicator, alien, or commercial organisms (see Table A3); this ecological information would have been lost in the context of taxonomic sufficiency, where species are mechanically aggregated into higher taxa.

The presence of opportunistic/tolerant groups as a dominant component of macroinvertebrate assemblages characterized the whole investigated period, and particularly the first years of the decade, from 1997 to 1999, suggesting a general condition of environmental deterioration in the lagoon system. This was likely the legacy of multiple impacts, including eutrophication and contamination events related to sewage discharge from aquaculture and runoff from neighbouring farming areas, which strongly affected the system in the 90s and culminated with cyanobacterial blooms and the ensuing anoxic crisis (Sorokin et al., 1996), leading to a shift from a macrophytes-dominated to phytoplankton-dominated system (Mistri et al., 2000; Munari and Mistri, 2012). At the beginning of 2001, the improvement in water circulation, due to dredging and enlargement of one of the three channels that connect the lagoon system to the open sea (Munari et al., 2003), probably mitigated harsh environmental conditions, and resulted in the decreased dominance of opportunistic groups in favour of more sensitive ones as observed in the central part of the decade. Recovery, if any, appeared very slow and indicators of disturbed conditions still characterized macro-invertebrate assemblages at the end of the study period.

Above all, distinct variations seemed to separate assemblage structure among the initial, the middle, and the last years of the decade. Yet, surrogates from the pilot monitoring programme, which included only the first five years of the decade, were able to detect variations occurring in the remaining five-year period as well, despite the instability of the ecosystem due to past human impacts and the high temporal variability that naturally characterizes transitional water environments such as coastal lagoons (Pérez-Ruzafa et al., 2007). This highlighted the potential robustness of surrogates to idiosyncrasies in temporal patterns, a desirable advantage for any tool aimed at increasing cost-efficiency of long-term assessments that, during their accomplishment, could face unpredictable changes of the monitored ecosystems.

The use of higher taxa as surrogates for species in ecological research has been explored since long time (e.g., Ellis, 1985), recognizing the associated benefits at least for routine assessments, once that comprehensive baseline information on the investigated system were available (Terlizzi et al., 2003; Dauvin et al., 2003). Beyond the consistent reduction in costs and time spent for difficult species-level identification, the approach offered undeniable advantages also when comparing biodiversity data from different areas (Fontaine et al., 2015), optimizing indicators of ecological status (Tweedley et al., 2014; Sánchez-Moyano et al., 2017), and resolving incoherent taxonomy across large and heterogeneous data series (Ellingsen et al., 2017). Fresh theoretical perspectives disentangled species surrogacy from a strict taxonomic framework (Bevilacqua et al., 2012; Neeson et al., 2013; Rosser, 2016) and allowed the implementation of new methods, such as the BestAgg approach, which attempted to compensate the absence of statistically sound procedures for the identification and validation of surrogates for species (Jones, 2008) and could provide further developments in the optimization of ecological indices. However, the use of surrogates for species has been rarely integrated in environmental regulatory frameworks or formally applied in monitoring programmes (e.g., Mistri and Munari, 2008), probably due to the lack of solid indications on their reliability on the long term (Magierowski and Johnson, 2006; Mellin et al., 2011). We demonstrated that surrogates for species based on data from a limited period can be applied also in future monitoring, even when the investigated system is subjected to large and inconsistent temporal fluctuations. Future research should persevere on consolidating evidence on the ability of surrogates to detect temporal changes as at species level, and to retain this prerogative over time, to give to species surrogacy a chance to have practical value for monitoring programmes.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.marpolbul.2018.04.047.

References

Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. Aust. Ecol. 26, 32–46.
Anderson, M.J., Willis, T.J., 2003. Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. Ecology 84, 511–525.
Bevilacqua, S., Terlizzi, A., 2016. Species surrogacy in environmental impact assessment

and monitoring: extending the BestAgg approach to asymmetrical designs. Mar. Ecol. Prog. Ser. 547, 19–32.

Bevilacqua, S., Fraschetti, S., Musco, L., Terlizzi, A., 2009. Taxonomic sufficiency in the detection of natural and human-induced changes in marine assemblages: a comparison of habitats and taxonomic groups. Mar. Pollut. Bull. 58, 1850–1859.

Bevilacqua, S., Terlizzi, A., Fraschetti, S., Claudet, J., Boero, F., 2012. Taxonomic relatedness does not matter for species surrogacy in the assessment of community responses to environmental drivers. J. Appl. Ecol. 49, 357–366.

Bevilacqua, S., Claudet, J., Terlizzi, A., 2013. Best practicable aggregation of species: a step forward for species surrogacy in environmental assessment and monitoring. Ecol. Evol. 3, 3780–3793.

Bevilacqua, S., Terlizzi, A., Mistri, M., Munari, C., 2015. New frameworks for species surrogacy in monitoring highly variable coastal ecosystems: applying the BestAgg approach to Mediterranean coastal lagoons. Ecol. Indic. 52, 207–218.

Block, W.M., Franklin, A.B., Ward, J.P., Ganey, J.L., White, G.C., 2001. Design and implementation of monitoring studies to evaluate the success of ecological restoration on wildlife. Restor. Ecol. 9, 293–303.

Borja, A., Dauer, D.M., 2008. Assessing the environmental quality status in estuarine and coastal systems: comparing methodologies and indices. Ecol. Indic. 8, 331–424.

Borja, A., Franco, J., Pérez, V., 2000. A marine biotic index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments. Mar. Pollut. Bull. 40, 1100–1114.

Bustamante, M., Bevilacqua, S., Tajadura, J., Terlizzi, A., Saiz-Salinas, J.I., 2012. Detecting human mitigation intervention: effects of sewage treatment upgrade on rocky macrofaunal assemblages. Mar. Environ. Res. 80, 27–37.

Cabral, A.F., Buosi, P.R.B., Segovia, B.T., Velho, L.F.M., Bini, L.M., 2017. Taxonomic sufficiency in detecting hydrological changes and reproducing ordination patterns: a test using planktonic ciliates. Ecol. Indic. 82, 227–232.

Clarke, K.R., 1993. Non-parametric multivariate analysis of changes in community structure. Aust. J. Ecol. 18, 117–143.

Dauvin, J.-C., Gomez Gesteira, J.L., Salvande Fraga, M., 2003. Taxonomic sufficiency: an overview of its use in the monitoring of sublittoral benthic communities after oil spills. Mar. Pollut. Bull. 46, 552–555.

R Development Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna Available at. http://www.R-project.org.

Ellingsen, K.E., Yoccoz, N.G., Tveraa, T., Hewitt, J.E., Thrush, S.F., 2017. Long-term environmental monitoring for assessment of change: measurement inconsistencies over time and potential solutions. Environ. Monit. Assess. 189, 595.

Ellis, D., 1985. Taxonomic sufficiency in pollution assessment. Mar. Pollut. Bull. 16, 459.

Ferraro, S.P., Cole, F.A., 1995. Taxonomic level sufficient for assessing pollution impacts on the southern Californian Bight macrobenthos—revisited. Environ. Toxicol. Chem. 14, 1031–1040.

Fontaine, A., Devillers, R., Peres-Neto, P.R., Johnson, L.E., 2015. Delineating marine ecological units: a novel approach for deciding which taxonomic group to use and which taxonomic resolution to choose. Divers. Distrib. 21, 1167–1180.

Fraschetti, S., Bevilacqua, S., Guarnieri, G., Terlizzi, A., 2012. Idiosyncratic effects of protection in a remote marine reserve. Mar. Ecol. Prog. Ser. 486, 21–34.

Groc, S., Delabie, J.H.C., Longino, J.T., Orivel, J., Majer, J.D., Vasconcelos, H.L., Dejean, A., 2010. A new method based on taxonomic sufficiency to simplify studies on Neotropical ant assemblages. Biol. Conserv. 143, 2832–2839.

Heino, J., 2014. Taxonomic surrogacy, numerical resolution and responses of stream macroinvertebrate communities to ecological gradients: are the inferences transferable among regions? Ecol. Indic. 36, 186–194.

Heino, J., Soininen, J., 2007. Are higher taxa adequate surrogates for species-level assemblage patterns and species richness in stream organisms? Biol. Conserv. 137, 78–89.

Hewitt, J.E., Thrush, S.F., 2007. Effective long-term ecological monitoring using spatially and temporally nested sampling. Environ. Monit. Assess. 133, 295–307.

Jackson, D.A., 1995. PROTEST: a PROcrustean randomization TEST of community environment concordance. Ecoscience 2, 297–303.

Jiang, X., Song, Z., Xiong, J., Proctor, H., Xie, Z., 2017. Different surrogacy approaches for stream macroinvertebrates in discriminating human disturbances in Central China. Ecol. Indic. 75, 182–191.

Jones, F.C., 2008. Taxonomic sufficiency: the influence of taxonomic resolution on freshwater bioassessments using benthic macroinvertebrates. Environ. Rev. 16, 45–69.

Lindenmayer, D.B., Likens, G.E., 2009. Adaptive monitoring: a new paradigm for long-term research and monitoring. Trends Ecol. Evol. 24, 482–486.

Lindenmayer, D.B., Likens, G.E., Andersen, A., Bowman, D., Bull, C.M., Burns, E., Dickman, C.R., Hoffmann, A.A., Keith, D.A., Liddell, M.J., Lowe, A.J., Metcalfe, D.J., Phinn, S.R., Russell-Smith, J., Thurgate, N., Wardle, G.M., 2012. Value of long-term ecological studies. Austral Ecol. 37, 745–757.

Lovett, G.M., Burns, D.A., Driscoll, C.T., Jenkins, J.C., Mitchell, M.J., Rustad, L., Shanley, J.B., Likens, G.E., Haeuber, R., 2007. Who needs environmental monitoring? Front. Ecol. Environ. 5, 253–260.

Magierowski, R.H., Johnson, C.R., 2006. Robustness of surrogates of biodiversity in marine benthic communities. Ecol. Appl. 16, 2264–2275.

Mandelik, Y., Roll, U., Fleischer, A., 2010. Cost-efficiency of biodiversity indicators for Mediterranean ecosystems and the effects of socio-economic factors. J. Appl. Ecol. 47, 1179–1188.

Mellin, C., Delean, S., Caley, J., Edgar, G., Meekan, M., Pitcher, R., Przeslawski, R., Williams, A., Bradshaw, C., 2011. Effectiveness of biological surrogates for predicting patterns of marine biodiversity: a global meta-analysis. PLoS One 6, e20141.

Mieszkowska, N., Sugden, H., Firth, L.B., Hawkins, S.J., 2014. The role of sustained observations in tracking impacts of environmental change on marine biodiversity and ecosystems. Philos. Transact. A Math. Phys. Eng. Sci. 372, 20130339.

Milošević, D., Stojković, M., Čerba, D., Petrović, A., Paunović, M., Simić, V., 2014. Different aggregation approaches in the chironomid community and the threshold of acceptable information loss. Hydrobiologia 727, 35–50.

Mistri, M., Munari, C., 2008. BITS: a SMART indicator for soft-bottom, non-tidal lagoons. Mar. Pollut. Bull. 56, 580–606.

Mistri, M., Fano, E.A., Rossi, G., Caselli, K., Rossi, R., 2000. Variability in macrobenthos communities in the Valli di Comacchio, Northern Italy, a hypereutrophized lagoonal ecosystem. Estuar. Coast. Shelf Sci. 51, 599–611.

Mueller, M., Geist, J., 2016. Conceptual guidelines for the implementation of the ecosystem approach in biodiversity monitoring. Ecosphere 7, e01305.

Mueller, M., Pander, J., Geist, J., 2013. Taxonomic sufficiency in freshwater ecosystems: effects of taxonomic resolution, functional traits, and data transformation. Freshw. Sci. 32, 762–778.

Munari, C., Mistri, M., 2010. Towards the application of the Water Framework Directive in Italy: assessing the potential of benthic tools in Adriatic coastal transitional ecosystems. Mar. Pollut. Bull. 60, 1040–1050.

Munari, C., Mistri, M., 2012. Ecological status assessment and response of benthic communities to environmental variability: the Valli di Comacchio (Italy) as a study case. Mar. Environ. Res. 81, 53–61.

Munari, C., Modugno, S., Ghion, F., Castaldelli, G., Fano, E.A., Rossi, R., Mistri, M., 2003. Recovery of the macrobenthic community in theValli di Comacchio, northern Adriatic Sea, Italy. Oceanol. Acta 26, 67–75.

Musco, L., Mikac, B., Tataranni, M., Giangrande, A., Terlizzi, A., 2011. The use of coarser taxonomy in the detection of long-term changes in polychaete assemblages. Mar. Environ. Res. 71, 131–138.

Neeson, T., van Rijn, I., Mandelik, Y., 2013. How taxonomic diversity, community structure and sample size determine the reliability of higher taxon surrogates. Ecol. Appl. 23, 1216–1225.

Olsgard, F., Somerfield, P.J., 2000. Surrogates in marine benthic investigations – which taxonomic unit to target? J. Aquat. Ecosyst. Stress. Recover. 7, 25–42.

Olsgard, F., Somerfield, P.J., Carr, M.R., 1998. Relationship between taxonomic resolution, macrobenthic community patterns and disturbance. Mar. Ecol. Prog. Ser. 127, 25–36.

Pérez-Ruzafa, A., Marcos, C., Pérez-Ruzafa, I.M., Barcala, E., Hegazid, M.I., Quispe, J., 2007. Detecting changes resulting from human pressure in a naturally quick-changing and heterogeneous environment: spatial and temporal scales of variability in coastal lagoons. Estuar. Coast. Shelf Sci. 75, 175–188.

Rosser, N., 2016. Shortcuts in biodiversity research: what determines the performance of higher taxa as surrogates for species? Ecol. Evol. 7, 2595–2603.

Sánchez-Moyano, J.E., García-Asencio, I., Donázar-Aramendía, I., Miró, J.M., Megina, C., García-Gómez, J.C., 2017. BENFES, a new biotic index for assessing ecological status of soft-bottom communities. Towards a lower taxonomic complexity, greater reliability and less effort. Mar. Environ. Res. 132, 41–50.

Sorokin, Y.I., Sorokin, P.Y., Gnes, A., 1996. Structure and functioning of the anthropogenically transformed Comacchio lagoonal ecosystem (Ferrara, Italy). Mar. Ecol. Prog. Ser. 133, 57–71.

Terlizzi, A., Bevilacqua, S., Fraschetti, S., Boero, F., 2003. Taxonomic sufficiency and the increasing insufficiency of taxonomic expertise. Mar. Pollut. Bull. 46, 544–560.

Terlizzi, A., Anderson, M.J., Bevilacqua, S., Fraschetti, S., Włodarska-Kowalczuk, M., Ellingsen, K.E., 2009. Beta diversity and taxonomic sufficiency: do higher-level taxa reflect heterogeneity in species composition? Divers. Distrib. 15, 450–458.

Thiault, L., Bevilacqua, S., Terlizzi, A., Claudet, J., 2015. Taxonomic relatedness does not reflect coherent ecological response of fish to protection. Biol. Conserv. 190, 98–106.

Tweedley, J.R., Warwick, R.M., Clarke, K.R., Potter, I.C., 2014. Family-level AMBI is valid for use in the north-eastern Atlantic but not for assessing the health of microtidal Australian estuaries. Estuar. Coast. Shelf Sci. 141, 85–96.

Warwick, R.M., 1993. Environmental impact studies on marine communities - pragmatical considerations. Aust. J. Ecol. 18, 63–80.