Species surrogacy in environmental impact assessment and monitoring: extending the BestAgg approach to asymmetrical designs

Stanislao Bevilacqua[1,*], Antonio Terlizzi[1,2]

[1]Laboratory of Zoology and Marine Biology, Department of Biological and Environmental Sciences and Technologies, University of Salento, CoNISMa, 73100 Lecce, Italy
[2]Stazione Zoologica Anton Dohrn, 80121 Napoli, Italy

ABSTRACT: The mainstream approach to define effective surrogates for species in routine biomonitoring focuses on the use of coarse levels of taxonomic resolution. A new approach to species surrogacy, the Best Practicable Aggregation of Species (BestAgg), which uses pilot data to generate null models of decreasing information at increasing levels of aggregation of variables, has been recently proposed. The approach has been tested in different environmental contexts, providing several advantages compared to the analysis at taxonomic levels higher than species. However, BestAgg still lacks a framework for its application to asymmetrical designs, thus limiting its general use in environmental impact assessment. Here, we implemented a new procedure to fill this gap and provided a specific R code for BestAgg in asymmetrical analysis. We presented an example of the application of this procedure to a real case study assessing the potential impact of harbour activities on subtidal sessile assemblages. Results from taxonomically fine-resolved data showed a significant effect of harbour activity on assemblage structure, although variable in time. Such patterns were consistent up to order level. Results based on surrogates from BestAgg were also aligned with those obtained at fine taxonomic resolution, but led to retention of much more information on original patterns and increased timesaving in sample processing compared to a classic approach based on the use of higher taxonomic levels. BestAgg represents a more formal procedure to species surrogacy than the empirical determination of the sufficient taxonomic resolution. The approach increases cost-efficiency while maximizing ecological information, and can be used under a wide range of experimental settings, now including asymmetrical designs.

KEY WORDS: Asymmetrical designs · Harbour impact · Null models · Sessile assemblages · Species surrogacy · Taxonomic surrogates

INTRODUCTION

Environmental monitoring has long been considered to be of minor scientific value compared to manipulative experiments and theoretical work in ecology and is often seen as an expensive and unrewarding practice (Legg & Nagy 2006, Lovett et al. 2007). Mounting evidence that the systematic collection of reliable environmental and/or biological data may provide important insights into natural and human-driven changes to ecosystems (Lovett et al. 2007) has increased awareness that adequate programs for continued data collection are necessary to achieve effective environmental management (e.g. Babcock et al. 2010, Fraschetti et al. 2012, 2013, Bates et al. 2014), leading to integration of routine assessment and monitoring procedures in regulatory frameworks at international level (e.g. Borja et al. 2010).

Indeed, environmental monitoring plays a key role in ecology, providing crucial information for hypo-

thesis formulation and model testing, allowing the evaluation of management actions, and guiding adaptive environmental policies (Lindenmayer & Likens 2009). Monitoring programs, however, generally require huge investments for sample collection and processing, especially when they are carried out over large spatial and temporal scales. The most time-consuming, and hence most expensive, activities concern the identification of organisms down to species level (Lampadariou et al. 2005, Mandelik et al. 2010). This is particularly the case when monitoring programs focus on less well-known speciose systems (such as, for instance, marine benthic communities), which involve many taxonomically difficult groups of invertebrates (Cardoso et al. 2011) and require engaging specialists not always easily available (Wägele et al. 2011). In this respect, besides fuelling intense research activity to combine optimal sampling with powerful experimental designs (e.g. Fairweather 1991, Underwood 1992, Benedetti-Cecchi 2001), the need for cost-effective monitoring procedures has also stimulated attempts to reduce efforts in taxonomic identification (Beattle & Oliver 1994, Bohmann et al. 2014).

The traditional approach to overcome such taxonomic impediments in environmental assessment and monitoring relies on the use of coarse levels of taxonomic resolution (e.g. Warwick 1988, Vanderklift et al. 1996, Olsgard et al. 1997, see Dauvin et al. 2003, Terlizzi et al. 2003, Jones 2008, Bevilacqua et al. 2012 for reviews). The basic concept of this framework of 'taxonomic sufficiency' (Ellis 1985) is to identify organisms at taxonomic levels higher than species without causing substantial loss of information on the investigated ecological patterns. Evidence from the bulk of studies on the use of taxonomic surrogates in marine, freshwater, and terrestrial systems shows that, in most cases, species-level community patterns are consistent up to family level (see Bevilacqua et al. 2012 for a review), supporting the potential application of the approach at least in routine monitoring programs. However, the taxonomic resolution sufficient to detect community responses to environmental drivers to a degree comparable to species level is determined empirically, in the absence of any statistical tool to control for error (Mellin et al. 2011). A further limit of classic taxonomic sufficiency stems from the fact that the approach does not account for the ecological relevance of species, nor for difficulties in their identification. It only implies the static aggregation of species in taxa from a single higher taxonomic rank, which may cause unnecessary and undesirable

depletion of taxonomic details and related ecological information. In this respect, quantifying community patterns using taxa higher than species has raised major concerns about the subsequent potential loss of information on the underlying ecological processes (e.g. Somerfield & Clarke 1995, Lasiak 2003, Dethier & Schoch 2006, Jones 2008, Bevilacqua et al. 2009).

Over the past 3 decades, the rationale of taxonomic sufficiency has experienced no substantial advancements from the original intuition that, to some extent, higher taxa might depict species-level community patterns consistently. Case studies on the use of taxonomic surrogates have amassed, while very few attempts have been made to shed light on related issues, such as reduced statistical power (Ferraro & Cole 1990), transferability of inference across taxonomic groups and/or among habitats (Anderson et al. 2005, Bevilacqua et al. 2009, Vilmi et al. 2016) or over spatial and temporal scales (Musco et al. 2011, Heino 2014), or to model cost-efficiency at varying effort in taxonomic identification of organisms (Bennett et al. 2014). As a consequence, uncertainties on the application of this approach still remain largely unsolved (Mellin et al. 2011), thus probably limiting its adoption as a standard method to species surrogacy despite the undeniable advantages.

Most ambiguities on the use of higher taxa as surrogates for species stem from the lack of a unifying theoretical framework on potential mechanisms underlying their effectiveness, which could allow modelling information loss and determining effective taxonomic surrogates through a more formal and transparent procedure (Jones 2008). Recent attempts to shed light on mechanisms causing correlations between community patterns at species and coarser taxonomic resolution outlined neutral relationships between taxonomic relatedness of species and their ecological response to natural or human-driven environmental changes (Bevilacqua et al. 2012, Siqueira et al. 2012), increasing doubts on the role of higher taxa as ecologically meaningful units (Bertrand et al. 2006). The use of null models of species aggregation demonstrated that the loss of information on assemblage structure when using taxonomic levels higher than species may largely depend on numerical relationships across ranks of the Linnaean taxonomic hierarchy, rather than on taxonomic relatedness of species (Bevilacqua et al. 2012, Thiault et al. 2015), confirming empirical evidence from several studies (e.g. Giangrande et al. 2005, Dethier & Schoch 2006, Neeson et al. 2013).

The Best Practicable Aggregation of Species (Best-Agg) has been implemented as a novel approach to

species surrogacy (Bevilacqua et al. 2013) based on this theoretical perspective. The approach relies on modelling information loss associated with surrogates and the ensuing decrease in the probability of consistently detecting species-level community responses against the level of species aggregation, ϕ, expressed as $\phi = G/S$, i.e. the ratio of the number of surrogates, G, to the number of species, S (see Fig. S1A in Supplement 1 at www.int-res.com/articles/suppl/m547 p019_supp/). Specifically, BestAgg exploits null models based on real data from pilot assessments in order to identify thresholds in ϕ, below which the loss of information on species-level community patterns may significantly affect the effectiveness of surrogates. BestAgg, therefore, quantifies the effect of aggregating the original set of variables (e.g. species) of a multivariate data matrix into a new set of variables (i.e. surrogates) on congruencies between the information contained in the original versus the corresponding aggregated matrix (Bevilacqua et al. 2013). This allows formal estimates to be obtained of the extent to which species (or any other operational units) could be packed into surrogates until the consequent loss of information becomes limiting. Ultimately, the procedure identifies the sufficient (i.e. minimum) number of surrogates, irrespective of their nature, needed to quantify community patterns at species level. The choice of surrogates in BestAgg is not constrained to specific selection schemes, allowing mixing of different surrogate types (e.g. taxonomic, morphological, functional, etc.), and relies on a conceptual framework (see Fig. S1B in Supplement 1) aiming to select surrogates by taking into account their ecological importance (relevance), difficulty of identification (easiness), and shared characteristics among organisms (resemblance) (see Bevilacqua et al. 2012, 2013 for a full theoretical introduction to the approach).

The BestAgg approach may be applied to any type of assemblage involving one or more groups of organisms from any environmental context. Also, R codes for the whole procedure are available (Bevilacqua et al. 2013), allowing employment of the approach in environmental studies involving the most common multifactorial designs (up to 3 nested and/or crossed factors), with fixed and/or random factors. However, a framework for the application of BestAgg when experimental settings require asymmetrical designs is lacking so far, thus limiting the general use of the approach in impact assessment and monitoring. The major problem when assessing environmental impacts is that there is usually a single, putatively impacted site (Underwood 1992). In such cases, the use of beyond-BACI (before/after-control/impact) designs (Underwood 1991, 1992) represents the most powerful approach for the detection of impacts associated with human activities. Such designs involve an asymmetrical ANOVA contrasting multiple controls and the impacted location sampled several times before and after the occurrence of the impact. Beyond-BACI designs can be modified in after control/impact (ACI) designs and applied to test impacts in the absence of 'before' data (Glasby 1997, Terlizzi et al. 2005), which is another recurrent condition in impact studies.

Here, we used an illustrative case study to extend the application of the BestAgg approach to asymmetrical ACI designs. Specifically, we (1) assessed the impact of a harbour on subtidal sessile macrobenthos through a typical asymmetrical ACI design, (2) developed a procedure for the application of the BestAgg approach to this type of experimental design, (3) checked the effectiveness of surrogates from BestAgg in providing results consistent with species-level analysis, and (4) employed a classical method for species surrogacy based on taxonomic sufficiency and compared the performance of the 2 approaches.

MATERIALS AND METHODS

Case study

Data came from a 2 yr monitoring of the putative impact of the Otranto harbour (SE Italy, south Adriatic Sea) on sessile macrobenthic assemblages from subtidal rocky reefs. The harbour consists of a natural embayment (about 0.7 km²) partially enclosed through breakwaters and characterized by sandy/muddy bottoms reaching the maximum depth of 7 m. This multipurpose harbour accounts for about 400 moorings which are generally used for recreational and fishing boating; the harbour also occasionally hosts small cargos and ferries.

Sessile macrobenthic assemblages were sampled in 3 locations: 1 putatively impacted (I), adjacent to the mouth of the harbour, and 2 control locations 2 to 3 km away (C1 and C2), 1 at the north and 1 at the south of the harbour, respectively. Control locations were randomly selected from a set of locations virtually unaffected by the harbour and characterized by comparable environmental conditions (e.g. substrate features, exposure, types of assemblage) with respect to the impacted location. A total of 3 sites (approximately 100 to 300 m apart) were randomly selected at each location and 10 photographic samples, each measuring 16 × 23 cm, were taken in each site at ver-

tical reefs at 5 to 8 m depth. Sampling was repeated on 3 random occasions during the monitoring program (2004 to 2005), yielding a total of 270 units of observations.

Photographic samples were analysed under magnification by superimposing a transparent grid of 24 sub-quadrats to help with the identification and visual estimates of the percentage cover of sessile organisms.

Species-level analysis

A permutational multivariate ANOVA (PERM-ANOVA; Anderson 2001) was performed to test the null hypothesis of no effect of the harbour in modifying the structure of sessile assemblages. The analysis was based on the original data matrix, which accounted for the finest taxonomic resolution (about 85 % of organisms identified at species/genus level) (see Table S1 in Supplement 1 at www.int-res.com/articles/suppl/m547p019_supp/).

The design for the analysis consisted of 3 factors; namely, time (T, 3 levels, random), location (L, 3 levels, random and crossed), and site (S[L], 3 levels, random and nested in L), with n = 10 replicates. Owing to the presence of a single purportedly impacted location (I) and 2 control locations (C1 and C2), the design was asymmetrical (Underwood 1991, 1994). Therefore, for the analysis, the location term was partitioned into 2 portions: the 1 df contrast of the impacted versus control locations (I-vs-Cs) and the variability between control locations (Cs). The same partitioning was applied to the overall mean square of the S(L) term, which was divided into the variability among sites within I (S[I]) and within Cs (S[Cs]), and to the residual variability, which was divided into the small-scale (i.e. among replicates) variability within I (Res I) and within Cs (Res Cs). All terms in the analysis involving the L and S(L) terms were partitioned following the same logic. Appropriate denominators for *F*-ratios were identified from expected mean squares and *F*-tests were constructed following the logic of asymmetrical designs (see Terlizzi et al. 2005 for further details).

In light of the precautionary principle (Gray 1990, Fairweather 1991) and the relatively small number of degrees of freedom for some of the tests, a significance level of $\alpha = 0.10$ was set *a priori* for interpreting statistically significant results. The analysis was based on Bray-Curtis dissimilarities calculated on untransformed data and each term was tested using 4999 random permutations.

Multivariate patterns were visualized through non-metric multidimensional scaling (NMDS) ordination of T × L centroids calculated using principal coordinates from the full Bray-Curtis dissimilarity matrix among all pairs of observations.

Classic taxonomic sufficiency

Taxonomic sufficiency, like any other surrogacy approach, needs representative pilot data at species level to determine the sufficient, i.e. the coarsest, level of taxonomic resolution able to reflect species-level patterns consistently. The approach, therefore, is based on analysing data at different taxonomic resolution in order to find the taxonomic level at which results are still consistent with those obtained using species. This sufficient taxonomic level is then assumed to be effective for quantifying community patterns in very similar experimental contexts or in the reiteration of the same study (e.g. monitoring programs).

In this framework, multivariate analyses were repeated at decreasing taxonomic resolution, following the logic of taxonomic sufficiency. Species-level abundance data from the original data matrix were aggregated (i.e. grouped and summed) into higher taxa of single ranks of the Linnaean taxonomic hierarchy, from genus to phylum (Table S1). Then, asymmetrical PERMANOVA (design as above) based on each taxonomically aggregated matrix was carried out, and results compared with those obtained analysing species-level data. We anticipated that results from species-level analysis would be consistent up to the level of order (see 'Results'), which was considered the sufficient taxonomic level to compare against the set of surrogates from the BestAgg approach. Multivariate patterns at order level were plotted through the NMDS ordination of T × L centroids.

BestAgg

In contrast to taxonomic sufficiency, which relies on the empirical determination of the sufficient taxonomic level, the BestAgg approach uses pilot data to generate a null model of decreasing information at increasing levels of aggregation of variables. This allows identification of the maximum level of aggregation of the original variables, *S* (i.e. species, taxa, etc.), which is expressed by the lowest ϕ value (ϕ_{low}), below which the loss of information on species-level community patterns becomes prejudicial for the

effectiveness of surrogates (see Fig. S1A). Consequently, the procedure identifies the minimum number of surrogates, G, i.e. G_{min} (where $G_{min} = \phi_{low} \times S$), 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 the construction of the null model, the BestAgg procedure randomly aggregates the S species from the original species-level matrix in decreasing numbers of G groups, simulating decreasing values of ϕ. The number of groups, G, is progressively decreased through a step-wise reduction of fixed decrement, d (equal to about 5 to 10% of the original number of variables) starting from $G = S$ (see Bevilacqua et al. 2013). For each set of G groups, the correlation value (ρ) between each of 1000 randomly aggregated matrices and the original species-level matrix is then calculated, and 1000 PERMANOVAs are also performed, one for each aggregated matrix, to test for the term of interest in the analysis. Finally, ϕ_{low} is determined as the lowest ϕ value allowing 95% of tests to give results consistent with those obtained at species level.

In a typical asymmetrical ACI design involving several times of sampling, the test for the higher-order interaction of T × I-vs-Cs should be considered, before testing for the main effect of the contrast I-vs-Cs term. In presence of the significant T × I-vs-Cs interaction, the impact, although variable in time, is nevertheless detected and interest may lie also in quantifying patterns of temporal variations in the difference between I-vs-Cs. In such cases, the sufficient number of surrogates should be defined separately for each time in order to ensure their effectiveness in detecting temporal changes in I-vs-Cs, besides detecting the main effect of T × I-vs-Cs (see Bevilacqua et al. 2013).

In light of this, we developed a specific procedure for the application of BestAgg in asymmetrical ACI designs based on (1) implementing separated null models for the identification of ϕ_{low} for each time of sampling and (2) selecting the most conservative (i.e. the highest) ϕ_{low} value among those obtained. The procedure estimates the ϕ_{low} sufficient for the detection of significant differences between I-vs-Cs to be at species level, and can be applied irrespective of whether the term of interest is I-vs-Cs or T × I-vs-Cs (i.e. impact consistent in time or not, respectively), or when a single time of sampling is available. In addition, because in asymmetrical analysis the mean squares to be used as denominator of the F-ratio when testing for I-vs-Cs may change depending on

results for other terms in the analysis, the procedure allows selection of the appropriate denominator following the logic of asymmetrical designs (see Terlizzi et al. 2005). The whole procedure can be performed through the R code provided in Supplement 2 at www.int-res.com/articles/suppl/m547p019_supp/. Example data along with a brief user manual are also provided (see Supplement 2, Sample Files 1 & 2).

This extension of BestAgg to asymmetrical designs was applied to the case study presented here. A total of 3 separated null models, one for each time of sampling, were built following the procedure described above. To construct each model, the 53 taxa in the original matrix were randomly aggregated in decreasing numbers of G groups, through a step-wise reduction of a fixed decrement $d = 3$. For each model, the relationship between the level of aggregation and information loss was checked by fitting a linear regression of ρ values from random aggregations against the corresponding $\ln(\phi)$. Then, ϕ_{low}, and therefore G_{min} (i.e. the sufficient number of surrogates), was determined for each time of sampling, except for T3, as in this last case no impact was detected at the level of species (see 'Results'). Finally, the highest value of G_{min} was considered sufficient for analyses. Analyses were performed using R (R Core Team 2015).

Once the overall G_{min} is determined, the next step in BestAgg concerns the selection of surrogates based on their relevance, easiness, and resemblance (Fig. S1B in Supplement 1), in order to maximize ecological information while reducing as much as possible the difficulty in the identification of organisms. Note that G_{min} represents a threshold value (i.e. the minimum number of groups for aggregation), and therefore the number of surrogates obtained after the selection procedure, hereafter referred to as $G_{BestAgg}$, can be even higher than G_{min}. Relevant species were identified based on their general ecological importance for the structure and functioning of sessile assemblages (e.g. habitat formers, bioconstructors, protected species), and/or their particular relation with harbour impacts (e.g. tolerant, indicator, sensitive species). Information on species or taxa was mined from the literature (see Table S2 in Supplement 1). Also, species or taxa relevant in the specific case study were selected from those most contributing to the difference between I-vs-Cs at each time of sampling, thus taking into account potential interannual variations in species that were important to discriminate I-vs-Cs. Such taxa were identified through SIMPER on the original data matrix, by considering only species whose contribution to dissimilarities was

≥3% (see Table S3 in Supplement 1). Easiness was defined based on expert opinion in relation to the difficulty of identification of the organism and the sampling method (photographic samples), whereas resemblance (species groupings) was decided by combining the relevance and easiness of species (or taxa) with the available taxonomic and ecological knowledge of groups (Table S2).

Finally, species-level data were aggregated in the BestAgg surrogates and used to perform an asymmetrical PERMANOVA (design as above). Results were then compared with those obtained using species-level data. Multivariate patterns from BestAgg surrogates were plotted through the NMDS ordination of T × L centroids.

Comparing the performance of the 2 approaches

The correlation (Spearman's ρ) between the species-level matrix and the matrix in which species were aggregated in the BestAgg surrogates should fall at least within random expectations from the null model, as BestAgg assumes that surrogates are random subsets of the original pool of species. If the selected surrogates lead to a significantly higher correlation, surrogates are then able to reflect species-level community patterns better than would be expected by chance, whereas ρ values below expectations likely indicate inappropriate sets of surrogates. The BestAgg approach allows this to be checked through a randomization test based on the general null model. Variables in the original matrix were randomly aggregated in $G_{BestAgg}$ groups (i.e. the number of BestAgg surrogates obtained after the selection procedure; see Table S2), with 1000 random aggregations. Correlation values (ρ) between the original species-level matrix and each randomly aggregated matrix were then calculated obtaining a frequency distribution against which testing $\rho_{BestAgg}$, i.e. the correlation between the original species-level matrix and the matrix aggregated using the BestAgg surrogates. The same procedure was also employed to test the correlation between the original matrix and the matrix aggregated using orders (Bevilacqua et al. 2013), which was the sufficient taxonomic level obtained following the classic approach based on taxonomic sufficiency.

A second-stage NMDS ordination was plotted to visualize differences among similarity matrices obtained using different levels of taxonomic resolution (from species to phylum) and BestAgg surrogates (Somerfield & Clarke 1995).

The time required to process samples using the sufficient taxonomic level (i.e. order, see above, see also 'Results') and the BestAgg surrogates was analysed in order to compare the performance of the 2 approaches in terms of associated timesaving. Two independent operators analysed independent sets of samples, randomly selected from all those available (270 samples), by identifying and counting organisms using different operational units, namely species, orders, and BestAgg surrogates. For each operator and each operational unit type, n = 5 randomly selected samples were processed and the time to complete processing was recorded. ANOVA was employed to analyse sample processing times. The design for the analysis consisted of factors operational unit (OU, 3 levels, fixed) and operator (O, 2 levels, random and crossed to OU). Post hoc Student-Newman-Keuls (SNK) tests were done for multiple comparisons of the mean sample processing time among different operational units.

RESULTS

PERMANOVA on species-level data detected a significant T × I-vs-Cs interaction (Table 1). Analyses carried out separately for each time of sampling showed that assemblages at I differed from those characterizing Cs in T1 and T2, but not in T3 (Table 1). Such patterns were clearly depicted in the NMDS plot of T × L centroids (Fig. 1).

Results of PERMANOVA at decreasing level of taxonomic resolution and those obtained analysing species were consistent up to order level, indicating that assemblages might be analysed using orders, while still detecting the same patterns of difference in I-vs-Cs through time as at species level (Table 2). Multivariate patterns at order level were visualized in the NMDS plot of T × L centroids, which, however, showed a less clear separation of I from Cs in T1 and T2 with respect to the NMDS ordination at species level (Fig. 1).

Results of linear regressions of ρ values from random aggregations against the corresponding ln(φ) carried out separately for each time of sampling were significant (p < 0.001), indicating that the loss of information strongly depended (adj. R^2 > 0.80 in all cases) on the level of aggregation following a semi-log model (Fig. 2). Predictions from the null models, based on random aggregations of the original species-level data, indicated that the lowest φ value allowing the 95% of tests on aggregated data to give results consistent with those obtained at species level

Table 1. Asymmetrical PERMANOVA testing for difference between impacted versus control locations (I-vs-Cs) based on species-level data. Results of analyses separated for each time of sampling (T1, T2, T3) are also reported. Terms involved in significant higher-order interactions were not analysed. Mean squares for denominators were chosen following the logic of asymmetrical designs (see 'Materials and methods'). All analyses were based on Bray-Curtis dissimilarities (data untransformed) and each test performed with 4999 permutations. T: time; L: location; S: site; *italic*: p-values obtained using 4999 Monte Carlo samples from the asymptotic permutation distribution; **bold**: significant p-values relevant to the hypothesis; ns: not significant; **p < 0.01, ***p < 0.001

Source of variation	df	MS	F	p
T	2	15221.0		
L	2	26089.0		
I-vs-Cs	1	36123.0		
Cs	1	16055.0		
S(L)	6	7211.1		
S(I)	2	8145.3		
S(Cs)	4	6744.0		
T × L	4	7347.0		
T × I-vs-Cs	2	11641.8	1.90	**0.0358**
T × Cs	2	3052.3	0.41	0.8912
T × S(L)	12	6122.7	4.30	0.0002
T × S(I)	4	3594.2	2.15	0.0002
T × S(Cs)	8	7386.9	5.69	0.0002
Res	243	1422.7		
Res(I)	81	1671.1		
Res(Cs)	162	1298.5		

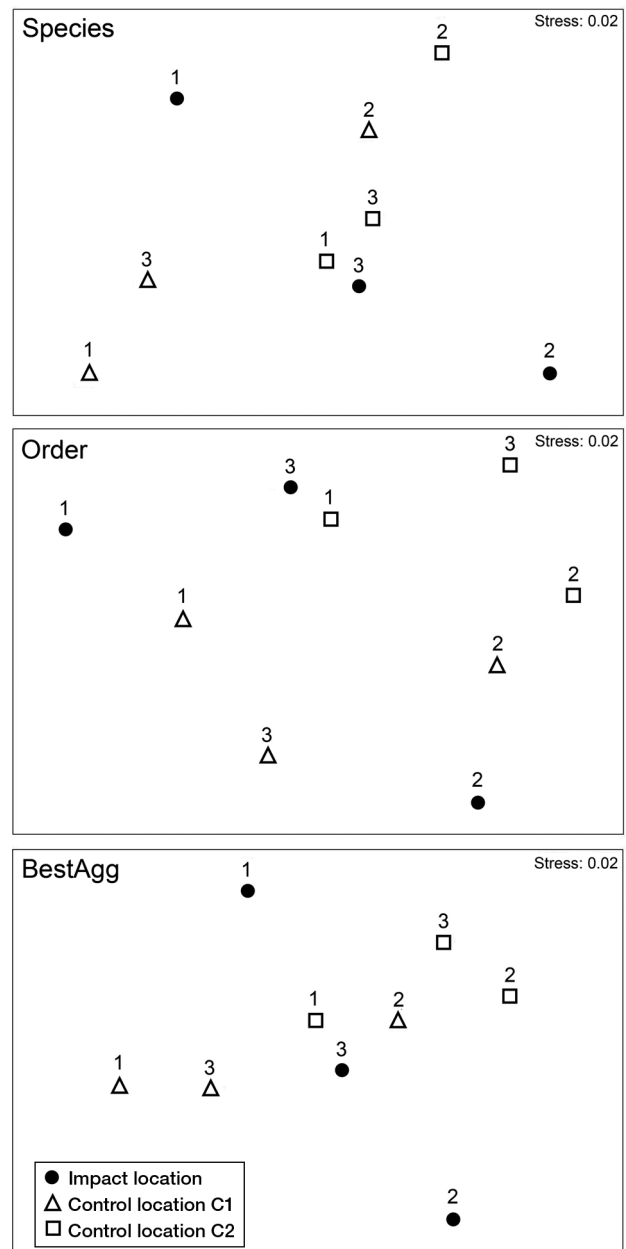|	T1	T2	T3
L	*1.98***	*1.63*ns	*2.73***
I-vs-Cs	***2.81****	***2.82****	*1.88*ns
Cs	*0.87*ns	*0.46*ns	*1.81*ns
S(L)	*4.69****	*4.79****	*4.21****

Fig. 1. Non-metric multidimensional scaling (NMDS) ordinations of T × L centroids based on Bray-Curtis dissimilarities for species-level data, and data aggregated using BestAgg surrogates and orders. Numbers indicate the 3 times of sampling (T1, T2, T3)

was ϕ_{low} = 0.60 (corresponding to G_{min} = 32) for T1, and ϕ_{low} = 0.43 (corresponding to G_{min} = 23) for T2 (Fig. 2; see also Table S4 in Supplement 1 at www.int-res.com/articles/suppl/m547p019_supp/). The most conservative ϕ_{low} therefore indicated that 32 surrogates should be sufficient to detect the impact consistently to species level.

The selection procedure of BestAgg surrogates (see Fig. S1B) identified a set of $G_{BestAgg}$ = G_{min} = 32 surrogates (see Table S2). About 50% of surrogates (15) were taxonomic, 6 of them at the species level (*Colpomenia sinuosa*, *Codium bursa*, *Halimeda tuna*, *Cladocora caespitosa*, *Chondrilla nucula*, *Halocynthia papillosa*), 5 at the genus level (*Valonia* spp., *Dictyota* spp., *Peyssonnelia* spp., *Didemnum* spp., *Microcosmus* spp.), and 4 at higher taxonomic levels (*Laurencia* complex, serpulids, barnacles, hydrozoa). The remaining surrogates (17) were morphological

(articulated corallines, unbranched green algae, filamentous algae, encrusting coralline algae, filamentous green algae, coarsely branched algae, thin tubular sheet-like algae, encrusting bryozoans, solitary anthozoans, massive sponges, solitary ascidians, colonial ascidians) or morpho-functional groups (turf-forming algae, canopy-forming algae, boring bivalves, boring sponges).

Table 2. Summary of asymmetrical PERMANOVAs testing for difference between impacted versus control locations (I-vs-Cs) based on data aggregated at decreasing taxonomic resolution, from genus to phylum. Results of analyses separated for each time of sampling (T1, T2, T3) are also reported when appropriate. Analyses were carried out under the same settings as at species level. **Bold**: significant p-values relevant to the hypothesis; ns: not significant; $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. For each taxonomic rank, the number of taxa and the corresponding aggregation ratio (ϕ) are shown

	Genus MS	Genus F	Genus p	Family MS	Family F	Family p	Order MS	Order F	Order p	Class MS	Class F	Class p	Phylum MS	Phylum F	Phylum p
T × I-vs-Cs	11645.0	1.91	**0.0350**	9793.5	1.83	**0.0468**	8702.9	1.74	**0.0738**	5784.5	2.31	0.1344	6059.9	2.17	0.1546
I-vs-Cs	T1 **	T2 **	T3 ns	T1 **	T2 *	T3 ns	T1 *	T2 *	T3 ns	T1 –	T2 –	T3 –	T1 –	T2 –	T3 –
Number of taxa	48			41			34			16			12		
ϕ	0.91			0.77			0.64			0.30			0.23		

T1
$\phi_{low} = 0.60$
Adj. $R^2 = 0.83***$

T2
$\phi_{low} = 0.43$
Adj. $R^2 = 0.88***$

T3
Adj. $R^2 = 0.87***$

Ln (ϕ)
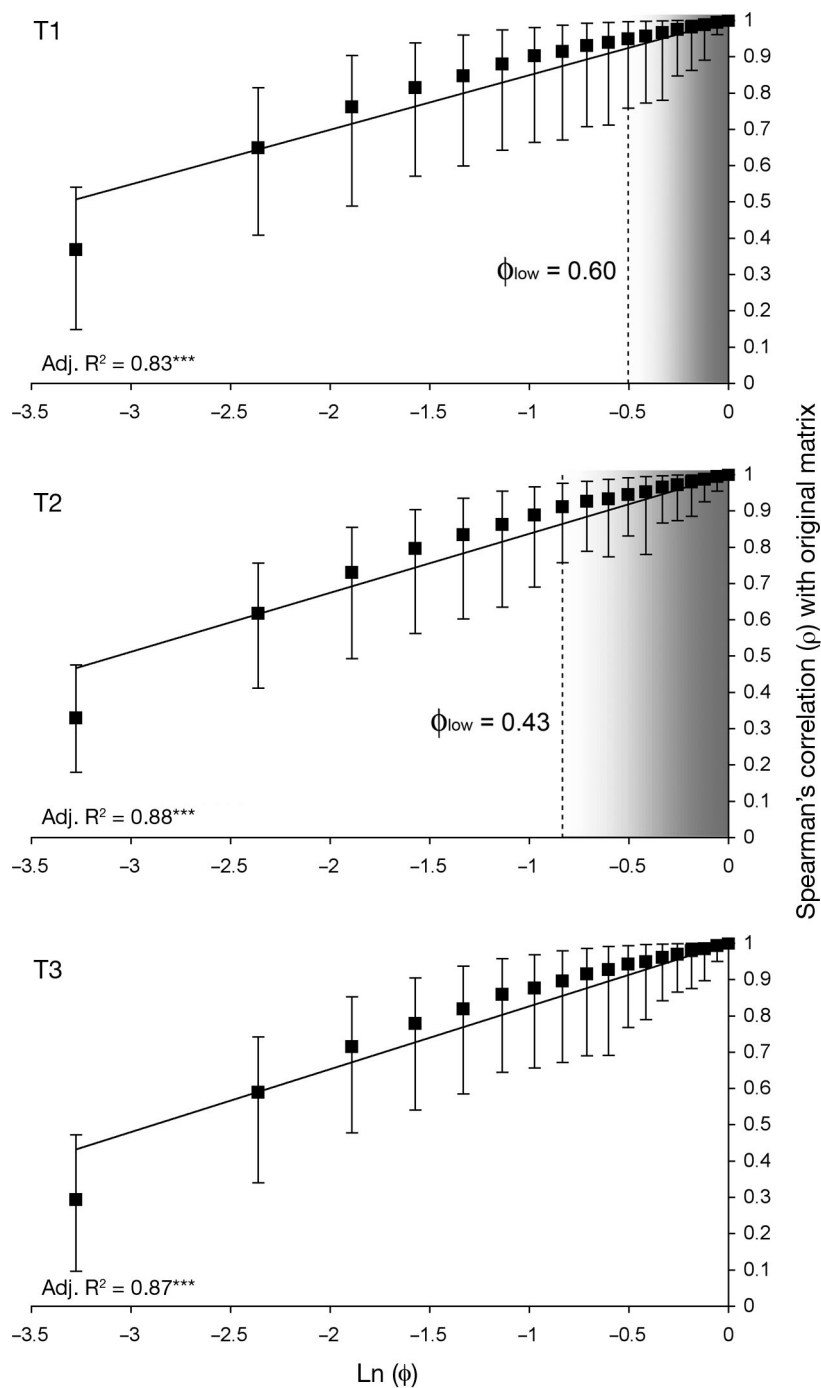
Spearman's correlation (ρ) with original matrix

Fig. 2. Semilog plots of Spearman's correlation values (ρ) between the species-level matrix and each randomly aggregated matrix against the corresponding aggregation level (ϕ) values, for each time of sampling (T1, T2, T3). Fading grey zones indicate the range of ϕ values for which results were consistent with those obtained analysing species-level data (see also Table 3). Dotted lines indicate ϕ_{low} (i.e. the lowest practicable aggregation level), sufficient to reflect species-level patterns consistently. The value of ϕ_{low} for T3 was not determined due to the lack of significant effects of the impact in this sampling time. Adjusted R^2 from regression analyses are also reported. $***p < 0.001$

PERMANOVA based on data aggregated using the selected BestAgg surrogates detected a significant T × I-vs-Cs interaction, and separated analyses for each time of sampling showed differences in I-vs-Cs through time consistent with species-level analyses (Table 3). The NMDS ordinations based on species data and BestAgg surrogates were almost interchangeable (Fig. 1).

The randomization test based on the whole data set showed that the correlation $\rho_{BestAgg}$ between BestAgg and species-level matrices fell outside the frequency distribution specific for $G_{BestAgg} = 32$, indicating that the selected BestAgg surrogates allowed retention of significantly more information (p = 0.02) than would be expected to occur by chance (Fig. 3A). In contrast, the correlation ρ_{Order} between species and order-level matrices fell within the random distribution (p = 0.07) specific for $G_{Order} = 34$ (i.e. the number of taxa at the level of order), indicating that the information on species-level pattern retained using the 34 orders is comparable to the information retained by any other random set of 34 surrogates (Fig. 3B).

Table 3. Asymmetrical PERMANOVA testing for difference between impacted versus control locations (I-vs-Cs) based on data aggregated using the selected BestAgg surrogates. Results of analyses separated for each time of sampling (T1, T2, T3) are also reported. Analyses were carried out under the same settings as at species level. T: time; L: location; S: site; Res: residual; *italic*: p-values obtained using 4999 Monte Carlo samples from the asymptotic permutation distribution; **bold**: significant p-values relevant to the hypothesis; ns: not significant; **p < 0.01, ***p < 0.001

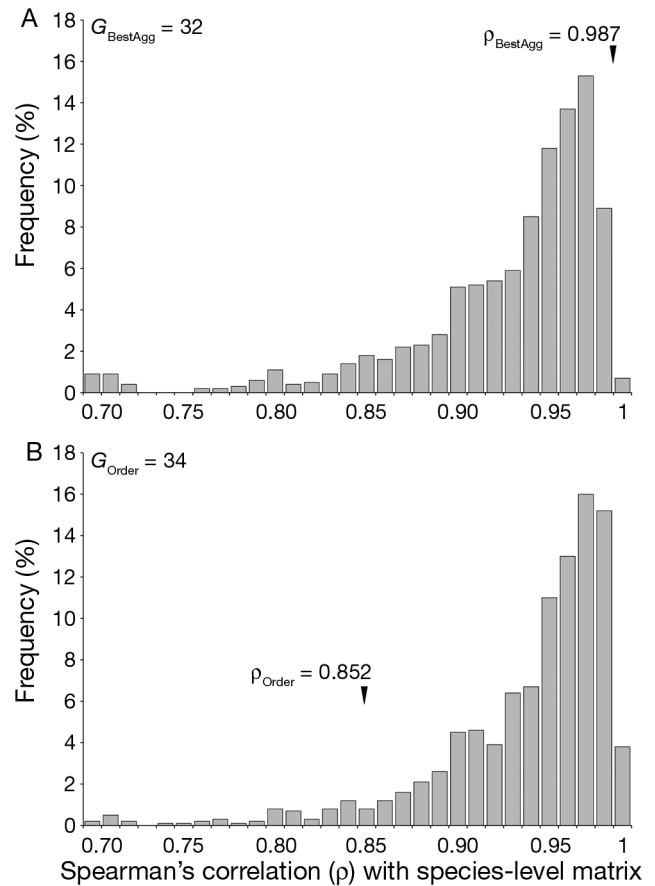| Source of variation	df	MS	F	p
T	2	13797.0		
L	2	23587.0		
I-vs-Cs	1			
Cs	1	16132.0		
S(L)	6	6786.2		
S(I)	2	7524.3		
S(Cs)	4	6417.1		
T × L	4	7293.3		
T × I-vs-Cs	2	11572.7	1.91	**0.0328**
T × Cs	2	3013.9	0.41	0.8926
T × S(L)	12	6054.5	4.41	0.0002
T × S(I)	4	3410.4	2.12	0.0004
T × S(*C*s)	8	7376.5	5.89	0.0002
Res	243	1371.2		
Res(I)	81	1609.4		
Res(Cs)	162	1252.1		
	T1	T2	T3	
L	*2.09***	*1.61*ns	*2.40***	
I-vs-Cs	***2.95****	***2.75****	*1.53*ns	
Cs	*0.91*ns	*0.47*ns	*1.80*ns	
S(L)	4.70***	4.85***	4.25***	

Fig. 3. Frequency distribution (n = 1000) of Spearman's correlation (ρ) values between the whole original species-level matrix and corresponding matrices in which species were randomly aggregated in (A) $G_{BestAgg}$ groups (i.e. the number of BestAgg surrogates) and (B) G_{Order} groups (i.e. the number of orders), to test the actual $\rho_{BestAgg}$ and ρ_{Order} (i.e. the true correlation values between the species-level matrix and the matrix aggregated using the BestAgg surrogates, or using orders, respectively)

The second-stage NMDS ordination showed a decreasing congruence of species-level matrix with higher-taxon matrices at decreasing taxonomic resolution. Genus-level data showed the strongest relationship with species, followed by families and orders, with a clear separation of classes and phyla (Fig. 4). BestAgg surrogates clustered closely with species and genus, despite the reduced number of surrogates if compared with these taxonomic levels (Fig. 4), indicating that in BestAgg, the reduction of the number of operational units might not result in a substantial loss of information, as occurred using decreasing levels of taxonomic resolution.

ANOVA showed significant changes in sample processing time when using different operational units that were consistent between operators (Table 4).
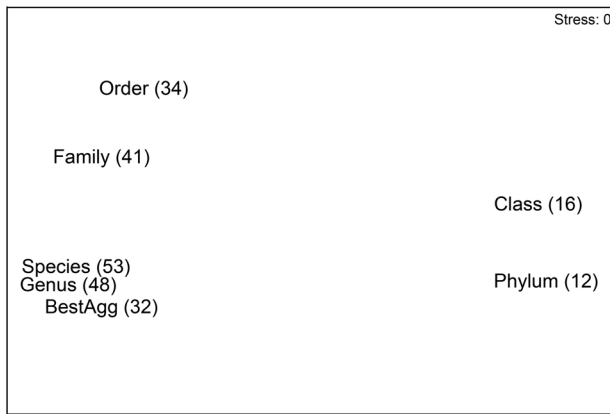
Fig. 4. Second-stage NMDS ordination of inter-matrix Spearman's rank correlations based on Bray-Curtis dissimilarities matrices at the species, genus, family, order, class and phylum level, and using the BestAgg surrogates. Numbers in brackets indicate the number of taxa or surrogates for each aggregation type

Fig. 5. Mean (±SE, n = 5) sample-processing time (min) required to analyse photographic samples of sessile assemblages using different operational units (i.e. species, orders, and BestAgg surrogates). For each operational unit, the 2 columns indicate 2 independent operators

Table 4. Results of ANOVA testing for differences among sample processing times using different operational units (i.e. species, orders, and BestAgg surrogates). Cochran's C-test for variance homogeneity not significant (C = 0.4309). Results of SNK tests are also reported. OU: operational unit; O: operator; Res: residual; **bold**: significant p-values

Source of variation	df	MS	F	p
OU	2	174 376.3	29.8	**0.0324**
O	1	2430.0	0.8	0.3718
OU × O	2	5843.7	2.0	0.1584
Res	24	2934.1		
SNK tests	Species > Order > BestAgg			

Specifically, the selected BestAgg surrogates and orders allowed time savings of about 50% and 30%, respectively, relative to species-level analysis of samples (Fig. 5).

DISCUSSION

The response of the BestAgg procedure in providing suitable sets of surrogates able to detect natural and human-driven changes to communities consistently to species-level analyses has previously been demonstrated in a wide range of aquatic environments (Bevilacqua et al. 2013, 2015, Milošević et al. 2014). Moreover, the effectiveness of the null models of species aggregation underlying the BestAgg approach has been verified on mollusc assemblages in relation to the effect of sewage discharge on sub-
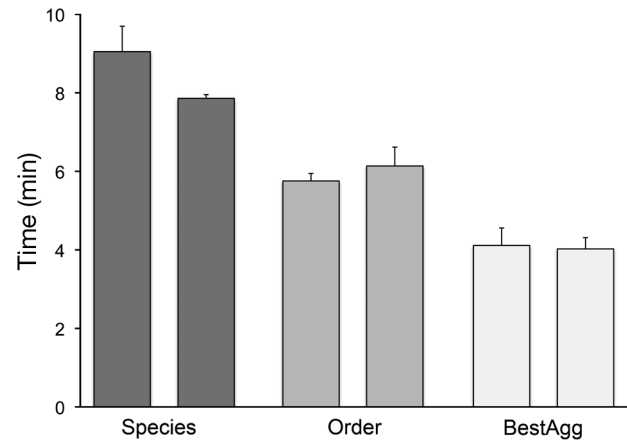
tidal rocky reefs, fish farming and coastal development on seagrass meadows (Bevilacqua et al. 2012). An emblematic example in the context of environmental impact assessment concerned the application of the approach to quantifying the effect of offshore gas platforms on macrobenthic assemblages from continental shelf mud flats in the Mediterranean. In that case study (Bevilacqua et al. 2013), BestAgg led to detection of the main effect of platforms and patterns of community changes at decreasing distance from the offshore structures at species level, despite the aggregation of the original 259 species variables into only 26 surrogates. In the present study, for the first time, BestAgg has been employed successfully in a case study involving a single impacted location and multiple controls and a specific BestAgg framework has been developed for asymmetrical ACI designs, which represents a substantial advancement to increase the range of applications of the approach in impact assessment and monitoring.

Multivariate analyses detected a significant effect of the harbour in modifying the structure of sessile assemblages, although not consistent in time. Such findings align with evidence from other studies investigating the impact of harbours of comparable size in the Mediterranean which have highlighted major effects of harbour activities on intertidal assemblages as well as more subtle changes associated with high temporal variability in subtidal habitats (e.g. Di Franco et al. 2011). The intensity of human disturbances (e.g. boating, maintenance operations, tourist frequentation) in small harbours and marinas is likely subject to seasonal fluctuations

(De Donno et al. 2008), which could underlie the observed temporal variability in I-vs-Cs. Human activities associated with harbours may affect the surrounding marine assemblages in several ways, causing alterations of sedimentary regimes due to physical modifications of the coastline (Airoldi & Beck 2007), by introducing artificial substrates (Vaselli et al. 2008), or increasing pollution and organic enrichment (Ingole et al. 2009). Our results showed that differences between I-vs-Cs were mostly due to increased abundance of turf-forming and erect algae at I with respect to Cs, where coralline encrusting forms were more abundant, suggesting a potential effect related to an increased nutrient load. Multiple human stressors, however, are likely to interact in determining the observed response of the sessile assemblage near the harbour, as heavy metal and hydrocarbon contamination was also recorded within the area (De Donno et al. 2008).

The same spatio-temporal patterns of difference between I-vs-Cs were detected while using surrogates from BestAgg, which led to almost interchangeable results with species-level analyses, from statistical tests to ordination plots. The application of the BestAgg procedure allowed us to decrease the number of operational units by selecting 32 surrogates that effectively substituted the original set of 53 species variables. On average, the use of BestAgg surrogates halved the time for sample processing, whereas the sufficient taxonomic level of order, despite a comparable decrease in the number of operational units, saved only one-third of the time required to analyse samples. Unlike orders, BestAgg surrogates showed a significantly higher correlation with species-level patterns than would be expected by chance for an equal set of surrogates. Further, BestAgg reduced the number of operational units by ~40% relative to the original set of species variables, while retaining as much information on species-level patterns as would a genus-level analysis. In contrast, genus-level analysis would reduce the number of operational units by less than 10%.

The performance of BestAgg as a timesaving and highly efficient approach to species surrogacy might be largely explained due to the surrogate selection procedure. On the one hand, it allows for ease of identification and does not disregard the use of intuitive surrogates, such as morphological groups, which greatly facilitates organism identification and counting. On the other hand, it prioritizes the choice of surrogates whose response is, or might be, more aligned than others with species-level community response to the specific source of disturbance under study, thus conferring high sensitivity of BestAgg surrogates to anthropogenic impacts. Indeed, BestAgg outperformed classical taxonomic sufficiency in detecting human impacts and exploring natural environmental gradients in highly variable aquatic systems, such as continental river basins (Milošević et al. 2014) and Mediterranean coastal lagoons (Bevilacqua et al. 2015), confirming the potential limits of taxonomic sufficiency when impacts are subtle (Ferraro & Cole 1990, 1995, Olsgard et al. 1998) and/or environmental heterogeneity is high (Olsgard et al. 1997).

A major strength of BestAgg is that it uses null models of species aggregation, which provides high robustness to the confounding effects of natural variability of assemblages in space and time (Bevilacqua et al. 2015, Thiault et al. 2015), and provides a statistical framework to control for uncertainty (Bevilacqua et al. 2013). As null models involve a number of simulated possibilities of species aggregation, they are probably more reliable if compared to, for example, empirical determinations of the sufficient taxonomic level, which in practice may represent only a single random scenario of aggregation (Thiault et al. 2015). With empirical determination, the ability of coarser taxonomic levels to detect differences in assemblage structure associated to anthropogenic or natural disturbance could be overestimated because of the high probability of rejecting the null hypothesis that a given taxonomic resolution is not sufficient, whereas it is actually true (Somerfield & Clarke 1995, Narayanaswamy et al. 2003). This issue is particularly relevant when the sufficient taxonomic level from empirical determinations based on pilot assessments is assumed to be effective also for subsequent monitoring. For instance, Thiault et al. (2015), analysing the response of fish fauna to protection within a marine protected area, showed that, in contrast to what occurred using null models of species aggregation, the effectiveness of taxonomic surrogates obtained empirically was not consistent across different times of sampling. However, whatever the surrogacy approach employed, the effectiveness of surrogates through time should be carefully checked (Musco et al. 2011). As general guidelines, pilot assessments representative of the temporal variability of the system being investigated and recalibrations of surrogates, especially when changes in the features of disturbance (e.g. magnitude, frequency) occur, are strongly recommended.

It is worth noting that the theoretical framework underlying the BestAgg approach makes no inference about patterns, and associated mechanisms,

leading evolutionarily related organisms to resemble, or differ from, one another. Actually, evidence of direct relationships between taxonomic or phylogenetic relatedness and ecological similarity among species are largely idiosyncratic (Wiens & Graham 2005, Losos 2008) and, therefore, difficult to include within a general framework for species surrogacy. Moreover, outcomes of the complex interplay among different processes driving species coexistence (Cadotte et al. 2013), from environmental filtering to competition, may not necessarily correlate to phylogenetic relationships (Gerhold et al. 2015), and could simulate neutral patterns of species assembly (Losos 2008, Mayfield & Levine 2010, Purves & Turnbull 2010). As a result, from an ecological perspective, higher taxa in a given species assemblage can be considered, or behave as, random groups of species, and the effectiveness of taxonomic surrogates may be simply the consequence of numerical overlapping between species and higher-taxon richness (Defeo & Lercari 2004, Giangrande et al. 2005, Dethier & Schoch 2006, Ramsay et al. 2006, Musco et al. 2009, Neeson et al. 2013), rather than assumed to be a prerogative of specific taxonomic ranks, as a consequence of the putative ecological similarities among species that taxa of such ranks might convey (Ferraro & Cole 1990, Warwick 1993).

In this view, BestAgg focuses on the pure effect of the aggregation of variables, lying outside any other putative or factual relationship among them. In analogy to Occam's razor, the use of null models of species aggregation in BestAgg simply represents a more parsimonious theoretical framework able to elucidate patterns of surrogate effectiveness without invoking complex explanations based on evolutionary processes and the links among phylogenetic relatedness, functional trait similarity and ecological coherence of species. As with other approaches aiming at selecting the best complementary suite of operational units (Tulloch et al. 2013), or integrating different sets of surrogates (Groc et al. 2010), BestAgg allows information to be maximized and the interpretability of the observed patterns to be increased through a flexible use of ecologically meaningful operational units, while minimizing their number (Bevilacqua et al. 2013).

As less is being invested in scientific research and management, the tendency is to privilege mechanistic science and predictive modelling to the detriment of monitoring and assessment programs (Schindler & Hilborn 2015). However, with the rise of the so-called 'blue economy', it is expected that oceans and seas will experience an expansion of sea-based human activities in the next decades (Fernández-Macho et al. 2015, Pinto et al. 2015), which will require local to large-scale monitoring to ensure a sustainable development (Borja et al. 2010, Visbeck et al. 2014). Under such circumstances, the implementation of standard and formal approaches to optimize environmental assessments and monitoring, especially in long-term programs, is strongly advocated. The classic approach of taxonomic sufficiency appears insufficient to satisfy the need for cost-effective monitoring procedures able to combine cost reduction, timesaving, and the collection of sound ecological data. As an alternative, BestAgg may represent a fresh perspective for species surrogacy in routine monitoring and assessment, now also extendible to experimental contexts requiring asymmetrical designs.

LITERATURE CITED

Airoldi L, Beck MW (2007) Loss status and trends for coastal marine habitats of Europe. Oceanogr Mar Biol Annu Rev 45:345–405

Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. Austral Ecol 26:32–46

Anderson MJ, Connell SD, Gillanders BM, Diebel CE, Blom WM, Saunders JE, Landers TJ (2005) Relationships between taxonomic resolution and spatial scales of multivariate variation. J Anim Ecol 74:636–646

Babcock RC, Alcala AC, Lafferty KD, McClanahan TR and others (2010) Decadal trends in marine reserves reveal differential rates of change in direct and indirect effects. Proc Natl Acad Sci USA 107:18256–18261

Bates AE, Barrett NS, Stuart-Smith RD, Holbrook NJ, Thompson PA, Edgar GJ (2014) Resilience and signatures of tropicalization in protected reef fish communities. Nat Clim Change 4:62–67

Beattle AJ, Oliver I (1994) Taxonomic minimalism. Trends Ecol Evol 9:488–490

Benedetti-Cecchi L (2001) Beyond BACI: optimization of environmental sampling designs through monitoring and simulation. Ecol Appl 11:783–799

Bennett JR, Sisson DR, Smol JP, Cumming BF, Possingham HP, Buckley YM (2014) Optimizing taxonomic resolution and sample size to design cost-effective ecological models for environmental assessment. J Appl Ecol 51:1722–1732

Bertrand Y, Pteijel F, Rouse GW (2006) Taxonomic surrogacy in biodiversity assessments, and the meaning of Linnaean ranks. Syst Biodivers 4:149–159

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

Bohmann K, Evans A, Gilbert MTP, Carvalho GR and others (2014) Environmental DNA for wildlife biology and biodiversity monitoring. Trends Ecol Evol 29:358−367

Borja A, Elliott M, Carstensen J, Heiskanen AS, van de Bund W (2010) Marine management — towards an integrated implementation of the European Marine Strategy Framework and the Water Framework Directives. Mar Pollut Bull 60:2175−2186

Cadotte M, Albert CH, Walker SC (2013) The ecology of differences: assessing community assembly with trait and evolutionary distances. Ecol Lett 16:1234−1244

Cardoso P, Erwin TL, Borges PAV, New TR (2011) The seven impediments in invertebrate conservation and how to overcome them. Biol Conserv 144:2647−2655

Dauvin JC, Gomez Gesteira JL, 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

De Donno A, Bagordo F, Lugoli F, Leopizzi MI, Fanizzi FP (2008) Dynamics of microbial autochthon populations and human impact on the harbour of Otranto (Lecce, Italy). J Coast Res 24:103−109

Defeo O, Lercari D (2004) Testing taxonomic resolution levels for ecological monitoring in sandy beach macrobenthic communities. Aquat Conserv 14:65−74

Dethier MN, Schoch GC (2006) Taxonomic sufficiency in distinguishing natural spatial patterns on an estuarine shoreline. Mar Ecol Prog Ser 306:41−49

Di Franco A, Graziano M, Franzitta G, Felline S, Chemello R, Milazzo M (2011) Do small marinas drive habitat specific impacts? A case study from Mediterranean Sea. Mar Pollut Bull 62:926−933

Ellis D (1985) Taxonomic sufficiency in pollution assessment. Mar Pollut Bull 16:459

Fairweather PG (1991) Statistical power and design requirements for environmental monitoring. Aust J Mar Freshw Res 42:555−568

Fernández-Macho J, Murillas A, Ansuategi A, Escapa M and others (2015) Measuring the maritime economy: Spain in the European Atlantic Arc. Mar Policy 60:49−61

Ferraro SP, Cole FA (1990) Taxonomic level and sample size sufficient for assessing pollution impacts on the Southern California bight macrobenthos. Mar Ecol Prog Ser 67: 251−262

Ferraro SP, Cole FA (1995) Taxonomic level sufficient for assessing pollution impacts on the southern Californian Bight macrobenthos — revisited. Environ Toxicol Chem 14:1031−1040

Fraschetti S, Bevilacqua S, Guarnieri G, Terlizzi A (2012) Idiosyncratic effects of protection in a remote marine reserve. Mar Ecol Prog Ser 466:21−34

Fraschetti S, Guarnieri G, Bevilacqua S, Terlizzi A, Boero F (2013) Protection enhances community and habitat stability: evidence from a Mediterranean marine protected area. PLoS ONE 8:e81838

Gerhold P, Cahill JF, Winter M, Bartish IV, Prinzing A (2015) Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). Funct Ecol 29:600−614

Giangrande A, Licciano M, Musco L (2005) Polychaetes as environmental indicators revisited. Mar Pollut Bull 50: 1153−1162

Glasby TM (1997) Analysing data from post-impact studies using asymmetrical analyses of variance: a case study of epibiota on marinas. Aust J Ecol 22:448−459

Gray JS (1990) Statistics and the precautionary principle. Mar Pollut Bull 21:174−176

Groc S, Delabie JHC, Longino JT, Orivel J, Majer JD, Vasconcelos HL, 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

Ingole B, Sivadas S, Nanajkar M, Sautya S, Nag A (2009) A comparative study of macrobenthic community from harbours along the central west coast of India. Environ Monit Assess 154:135−146

Jones FC (2008) Taxonomic sufficiency: the influence of taxonomic resolution on freshwater bioassessments using benthic macroinvertebrates. Environ Rev 16:45−69

Lampadariou N, Karakassis I, Pearson TH (2005) Cost/ benefit analysis of a benthic monitoring programme of organic benthic enrichment using different sampling and analysis methods. Mar Pollut Bull 50:1606−1618

Lasiak T (2003) Influence of taxonomic resolution, biological attributes and data transformations on multivariate comparisons of rocky macrofaunal assemblages. Mar Ecol Prog Ser 250:29−34

Legg CJ, Nagy L (2006) Why most conservation monitoring is, but need not be, a waste of time. J Environ Manage 78:194−199

Lindenmayer DB, Likens GE (2009) Adaptive monitoring: a new paradigm for long-term research and monitoring. Trends Ecol Evol 24:482−486

Losos JB (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. Ecol Lett 11:995−1003

Lovett GM, Burns DA, Driscoll CT, Jenkins JC and others (2007) Who needs environmental monitoring? Front Ecol Environ 5:253−260

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

Mayfield MM, Levine JM (2010) Opposing effects of competitive exclusion on the phylogenetic structure of communities. Ecol Lett 13:1085−1093

Mellin C, Delean S, Caley J, Edgar G and others (2011) Effectiveness of biological surrogates for predicting patterns of marine biodiversity: a global meta-analysis. PLoS ONE 6:e20141

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

Musco L, Terlizzi A, Licciano M, Giangrande A (2009) Taxonomic structure and the effectiveness of surrogates in environmental monitoring: a lesson from polychaetes. Mar Ecol Prog Ser 383:199–210

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

Narayanaswamy BE, Nickell TD, Gage JD (2003) Appropriate levels of taxonomic discrimination in deep-sea studies: species vs family. Mar Ecol Prog Ser 257:59–68

Neeson TM, 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 PJ, Carr MR (1997) Relationships between taxonomic resolution and data transformations in analyses of a macrobenthic community along an established pollution gradient. Mar Ecol Prog Ser 149:173–181

Olsgard F, Somerfield PJ, Carr MR (1998) Relationships between taxonomic resolution, macrobenthic community patterns and disturbance. Mar Ecol Prog Ser 172:25–36

Pinto H, Cruz AR, Combe C (2015) Cooperation and the emergence of maritime clusters in the Atlantic: analysis and implications of innovation and human capital for blue growth. Mar Policy 57:167–177

Purves DW, Turnbull LA (2010) Different but equal: the implausible assumption at the heart of neutral theory. J Anim Ecol 79:1215–1225

R Core Team (2015) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. www.R-project.org/

Ramsay PM, Kent M, Reid CL, Duckworth JC (2006) Taxonomic, morphological and structural surrogates for the rapid assessment of vegetation. J Veg Sci 17:747–754

Schindler DE, Hilborn R (2015) Prediction, precaution, and policy under global change. Science 347:953–954

Siqueira T, Bini LM, Roque FO, Cottenie K (2012) A metacommunity framework for enhancing the effectiveness of biological monitoring strategies. PLoS ONE 7:e43626

Somerfield PJ, Clarke KR (1995) Taxonomic levels, in marine community studies, revisited. Mar Ecol Prog Ser 127:113–119

Terlizzi A, Bevilacqua S, Fraschetti S, Boero F (2003) Taxonomic sufficiency and the increasing insufficiency of taxonomic expertise. Mar Pollut Bull 46:556–561

Terlizzi A, Benedetti-Cecchi L, Bevilacqua S, Fraschetti S, Guidetti P, Anderson MJ (2005) Multivariate and univariate asymmetrical analyses in environmental impact assessment: a case study of Mediterranean subtidal sessile assemblages. Mar Ecol Prog Ser 289:27–42

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

Tulloch AIT, Chadès I, Possingham HP (2013) Accounting for complementarity to maximize monitoring power for species management. Conserv Biol 27:988–999

Underwood AJ (1991) Beyond BACI: experimental designs for detecting human environmental impacts on temporal variations in natural populations. Aust J Mar Freshw Res 42:569–587

Underwood AJ (1992) Beyond BACI: the detection of environmental impacts on populations in the real, but variable world. J Exp Mar Biol Ecol 161:145–178

Underwood AJ (1994) On beyond BACI: sampling designs that might reliably detect environmental disturbances. Ecol Appl 4:3–15

Vanderklift MA, Ward TJ, Jacoby CA (1996) Effect of reducing taxonomic resolution on ordinations to detect pollution-induced gradients in macrobenthic infaunal assemblages. Mar Ecol Prog Ser 136:137–145

Vaselli S, Bulleri F, Benedetti-Cecchi L (2008) Hard coastal-defence structures as habitats for native and exotic rocky-bottom species. Mar Environ Res 66:395–403

Vilmi A, Karjalainen SM, Nokela T, Tolonen K, Heino J (2016) Unravelling the drivers of aquatic communities using disparate organismal groups and different taxonomic levels. Ecol Indic 60:108–118

Visbeck M, Kronfeld-Goharani U, Neumann B, Rickels W and others (2014) Securing blue wealth: the need for a special sustainable development goal for the ocean and coasts. Mar Policy 48:184–191

Wägele H, Klussmann-Kolb A, Kuhlmann M, Haszprunar G, Lindberg D, Koch A, Wägele JW (2011) The taxonomist — an endangered race. A practical proposal for its survival. Front Zool 8:25

Warwick RM (1988) The level of taxonomic discrimination required to detect pollution effects on marine benthic communities. Mar Pollut Bull 19:259–268

Warwick RM (1993) Environmental impact studies on marine communities: pragmatical considerations. Aust J Ecol 18:63–80

Wiens JJ, Graham CH (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. Annu Rev Ecol Syst 36:519–539