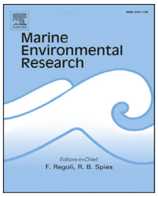
Marine Environmental Research 87-88 (2013) 112–121

Contents lists available at SciVerse ScienceDirect

Marine Environmental Research

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



Multivariate analysis applied to agglomerated macrobenthic data from an unpolluted estuary

Anxo Conde[a,b,*], Júlio M. Novais[a,1], Jorge Domínguez[b]

[a] IBB-Institute for Biotechnology and Bioengineering, Center for Biological and Chemical Engineering, Instituto Superior Técnico (IST), Av. Rovisco Pais, 1049-001 Lisbon, Portugal
[b] Departamento de Ecoloxía e Bioloxía Animal, Universidade de Vigo, Vigo E-36310, Spain

ARTICLE INFO

Article history:
Received 20 February 2013
Received in revised form
15 April 2013
Accepted 19 April 2013

Keywords:
Taxonomic sufficiency
Functional groups
Environmental gradients
Ordination methods
Model selection
Minho estuary
Iberian Peninsula

ABSTRACT

We agglomerated species into higher taxonomic aggregations and functional groups to analyse environmental gradients in an unpolluted estuary. We then applied non-metric Multidimensional Scaling and Redundancy Analysis (RDA) for ordination of the agglomerated data matrices. The correlation between the ordinations produced by both methods was generally high. However, the performance of the RDA models depended on the data matrix used to fit the model. As a result, salinity and total nitrogen were only found significant when aggregated data matrices were used rather than species data matrix. We used the results to select a RDA model that explained a higher percentage of variance in the species data set than the parsimonious model. We conclude that the use of aggregated matrices may be considered complementary to the use of species data to obtain a broader insight into the distribution of macrobenthic assemblages in relation to environmental gradients.

1. Introduction

The distribution of macrobenthic assemblages in relation to environmental drivers has been often described at the species level (Rodrigues et al., 2006; Ysebaert et al., 1998; Zajac and Whitlatch, 1982). Agglomeration or aggregation of species at low resolution levels has also been used to describe environmental gradients. According to the concept of taxonomic sufficiency, species data are aggregated into higher taxonomic levels deemed to be sufficient for the purposes of a study. The term 'taxonomic sufficiency' (coined by Ellis, 1985) has been used to identify the effects of pollution on marine communities (e.g. Ferraro and Cole, 1990). The family level has been proposed as an appropriate level of description for pollution studies (Gómez Gesteira et al., 2003), although higher levels may also be used (Warwick, 1988). However, it has been argued that the likelihood of detecting a stressor at higher taxonomic levels will depend on the severity of the stressor (Ferraro and Cole, 1990), and therefore the taxonomic resolution used in a study must take into consideration the environmental problem in question. Similar conclusions were reached in long term monitoring studies that propose periodical analysis at species level rather than at lower taxonomic resolution, such as the family level (Musco et al., 2011).

Agglomeration of species into guilds (usually feeding guilds) has also been considered (Gaudêncio and Cabral, 2007; García-Arberas and Rallo, 2002). The use of functional groups (or functional traits) subdivides the aggregated data at a finer functional level. For instance, the carnivore guild may be further subdivided in accordance with the motile or sessile characteristic of the species, reflecting different ecological niches that are not indicated by the 'carnivore' grouping alone. Similarly, although species in the same feeding guild commonly compete for the same food resource, the interaction is not necessarily reflected in a taxonomic approach (Rosenberg, 2001). Improvements in computational methods enable more complex analyses that allow species, biological traits and environmental matrices to be considered simultaneously (Dolédec et al., 1996; Legendre et al., 1997). Analysis and visualization of patterns based on species agglomeration matrices in relation to environmental factors may be accomplished by common ordination techniques such as Non-metric Dimensional Scaling (NMDS, Terlizzi et al., 2008) and Redundancy Analysis (RDA, Boström et al., 2006).

A few studies have shown that the family level provides an adequate description of macrobenthic assemblages along natural

---

* Corresponding author. IBB-Institute for Biotechnology and Bioengineering, Center for Biological and Chemical Engineering, Instituto Superior Técnico (IST), Av. Rovisco Pais, 1049-001 Lisbon, Portugal. Tel.: +351 218419124; fax: +351 218419062.

E-mail addresses: aconde@uvigo.es, aconde@ist.utl.pt (A. Conde).

[1] Deceased.

0141-1136/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.marenvres.2013.04.005

gradients (De Biasi et al., 2003; Dethier and Schoch, 2006; Wlodarska-Kowalczuk and Kedra, 2007, and references therein). The family level is a better descriptor than higher taxonomic levels or trophic groups for explaining the distribution of the assemblages (Dethier and Schoch, 2006), although a lower level of data resolution may also be useful (Wlodarska-Kowalczuk and Kedra, 2007).

Aggregation of species has been underexplored as a means of describing natural gradients (Wlodarska-Kowalczuk and Kedra, 2007). The aim of this study was to assess the usefulness of agglomerated data sets in explaining the distribution of macrobenthic assemblages along a natural gradient in an unpolluted estuary. As estuaries may display horizontal, vertical or cross-sectional spatial gradients, among others (McLusky, 1993), the study was conducted in an area of an estuary between the zones most affected by marine and freshwater influences. We measured a number of abiotic variables (redox potential, total nitrogen content of the sediments, salinity and others) with the aim of explaining the distribution of the macrobenthic assemblages. We compared the results of applying multivariate analyses to the species dataset and the results of applying the same analyses to agglomerated data grouped at low level of resolution (order and class taxonomic level and gross feeding guilds), which are rather overlooked in comparison with finer aggregation types such as the family taxonomic level (De Biasi et al., 2003) or subdivisions of feeding guilds (Fauchald and Jumars, 1979). The underlying null hypothesis of the study was that there was no relation between multivariate patterns in agglomerated data sets and the species data set.

2. Material and methods

2.1. Study site

The River Minho is the longest river in the Northwest Iberian Peninsula. The annual average discharge of the River Minho is 13,560 h m^3, with a monthly flow average ranging from 100 m^3 s^{-1} in August to 1000 m^3 s^{-1} in February (deCastro et al., 2006). The estuarine part of the river (Fig. 1) lies between Portugal and Galicia (Spain). The Minho estuary has mesotidal features and is partially mixed, although it tends to be a salt wedge estuary when high floods occur (Sousa et al., 2005). Numerous small tributaries drain into the estuary (Fig. 1).

A number of studies have highlighted the low level of anthropogenic pressure on the Minho estuary (Monteiro et al., 2007; Moreira et al., 2006). Other authors have recommended using this estuary as a pristine reference site for comparison with other metal polluted estuaries (Reis et al., 2009). However, Sousa et al. (2008a) have identified up to 11 well-established alien species of invertebrates and fish in the Minho estuary. Among these, the invasive bivalve *Corbicula fluminea* (Müller, 1774) has achieved a key position in the benthic assemblages because of its high abundance, biomass and production (Sousa et al., 2008b).

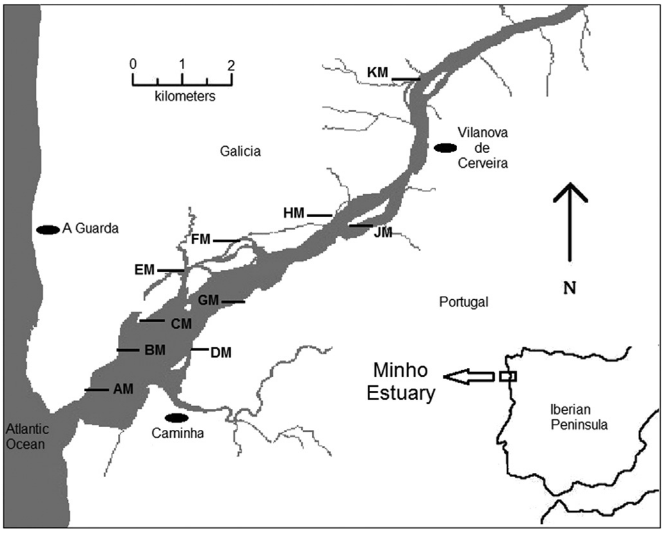


Fig. 1. Map of the Minho estuary showing the location of the sampling sites.

2.2. Sampling and laboratory procedures

Sampling was conducted on both banks of the estuary during a summer spring tide, between 23 and 26 August 2010. We sampled ten sites and assigned them codes in alphabetical order from AM, BM… to KM, excluding IM (Fig. 1). We conducted intertidal sampling during low spring tides immediately above the water edge. These sampling surveys included sites along the main axis of the estuary, in salt marshes and on an estuarine island (Boega Island, site JM). The length of the sampling area was approximately 13 km, extending from the mouth of the estuary to a few kilometres upstream of the village of Vilanova da Cerveira (Fig. 1).

We used a standard field probe (WTW 340i) to measure temperature, salinity, oxygen, redox potential in the interstitial water, at a depth of approximate 10 cm in the sediment. We sampled the top 3 cm of the sediment (making composite sample from three replicate samples per site) to determine the sediment grain size by the dry sieving method. We defined the grain size fractions as follows: gravel (>2 mm), coarse and medium sand (2–0.250 mm), fine sand (0.250–0.063 mm) and finest grains (<0.063 mm), following the method of Rodrigues et al. (2006) and Silva et al. (2006), although these authors considered coarse and medium sand content separately. At each sampling site, we placed samples of the top 3 cm of the sediment in plastic containers and kept them in a cooler at 4 °C for subsequent determination of total organic carbon and total nitrogen (on dried samples) in an elemental analyser (LECO CN2000).

We inserted a corer of inner diameter 95 mm to a depth of 25 cm in the sediment (7 replicates = 0.05 m^2) to sample infaunal organisms. We sieved all samples through a 1 mm mesh and preserved the material retained on the mesh in 70% ethanol. We used a dissecting microscope to sort, count and identify samples of benthic fauna to the lowest possible taxonomic level.

2.3. Taxa aggregation and categories

We agglomerated the taxa into four aggregation types, each constituted by different categories, as explained below. We first aggregated the taxa into two different taxonomic levels: order and class; the number of categories for both aggregation types depended on the fauna found in each sampling site. We also aggregated the taxa into four categories of trophic guilds: suspension feeders, deposit-feeders (including surface and sub-surface-deposit feeders), carnivores (or predators) and omnivores. This trophic aggregation, with the exception of the omnivore category, follows that used by Chardy and Clavier (1988). We categorized feeding guilds in accordance with Ysebaert et al. (1998), Cummins et al. (1989) and Mancinelli et al. (2005). The feeding guilds are hereinafter referred to as guilds. For the final aggregation, we combined class and guilds in the same data matrix, to produce a mixed data matrix. This approach is based on that used by Pagola-Carte and Saiz-Salinas (2001), although in the present study each taxon occurs twice, as a category of both class and guilds. For example, the species *Hediste diversicolor* occurs in the polychaete category (class) and the omnivore category (guild). We labelled taxa that did not fit any of the categories within a specific aggregation type as *other* (e.g. we