



# The relative impact of co-occurring stressors on the abundance of benthic species examined with three-way correspondence analysis

Pieter M. Kroonenberg<sup>a,\*</sup>, Stephen Widdicombe<sup>b,c</sup>, Paul Somerfield<sup>b,c</sup>

<sup>a</sup> Faculty of Social and Behavioural Sciences, Leiden University, Leiden, the Netherlands

<sup>b</sup> Plymouth Marine Laboratory, Prospect Place, West Hoe, Plymouth PL1 3DH, United Kingdom

<sup>c</sup> Plymouth Environmental Research Centre, (Department of Biological Sciences), University of Plymouth, Drake Circus, Plymouth PL4 8AA, United Kingdom

## ARTICLE INFO

### Keywords:

Analysis of benthic species abundance  
Environmental sensitivity to stressors  
Organic enrichment  
Physical disturbance  
Factorial data experiment  
Three-way correspondence analysis

## ABSTRACT

This paper presents a novel application of three-way correspondence analysis as a technique to analyse three-way contingency tables with abundance scores of several species. The example data analysis presented was taken from a previous mesocosm experiment and consists of a two-factor experimental design with physical disturbance and organic enrichment as factors, applied to sediment collected from the Oslofjord, Norway. The focus of the original research was to evaluate the influence of the two factors and their interactions on the abundance of the species present in the sediment. In the current paper we demonstrate that by using a three-way correspondence approach it is possible to undertake simultaneous analysis of the species, identifying and evaluating their relative sensitivity to the environmental factors thus adding additional insight than was possible in the original analysis. In particular, this new approach allowed even relatively scarce species to be included in the analysis and evaluated together with abundant species. This paper demonstrates how three-way correspondence analysis can be a useful analytical tool in teasing out effects and interactions from multi-factorial studies.

## 1. Introduction

In the natural environment, organisms are constantly exposed to many co-occurring stresses, and the responses of individual species to these relative stressors will ultimately influence the structure and diversity of the resident community, as discussed, and extensively documented in detail, in the introduction of [Widdicombe and Austen \(2001\)](#). Whilst some stressors can be considered natural occurrences e.g. the physical disturbance caused by storms or changes in temperature driven by seasons or daily cycles, a great many biological communities are experiencing rapidly increasing pressures associated with anthropogenic activities. This, in turn, is leading to a growing societal need to predict the potential impacts of man's multiple activities on the structure and diversity of the communities affected. There are currently two major impediments to making these community level predictions. Firstly, whilst numerous studies have been conducted to assess how individual organisms will respond to specific environmental drivers, little evidence currently exists how such organisms will respond when exposed to multiple factors. This is of particular interest to those researchers studying the impacts of climate change (temperature, desiccation, ocean acidification etc) and its interaction with other existing

environmental impacts (e.g. physical disturbance, pollution, resource exploitation etc.). Secondly, where such data do exist, they are often generated from experiments in which the species being studied is in isolation, and therefore removed from the ecological interactions (e.g. competition, predation) that it would normally experience. Consequently, predicting the structure of communities for an environment that has multiple changing parameters, using species specific response data that do not include ecological factors, is extremely difficult.

To address the issues described above, and generate data needed to make the predictions required, it is essential that new studies are conducted that observe the response of communities, and the species they contain, to combinations of different stressors. However, whilst quantifying the community level responses (such as diversity and community composition) in these experiments is relatively straightforward, teasing out the individual species responses is less so. Particularly when diversity in biological communities is primarily made up of species with low numerical abundances. In this paper we use data from a previously published two-factor mesocosm experiment ([Widdicombe and Austen, 2001](#)) conducted on a marine sediment community, to propose a novel application of three-way correspondence analysis. This approach could allow the influence of the two factors and their interactions on the

\* Corresponding author.

E-mail address: [P.M.Kroonenberg@fsw.leidenuniv.nl](mailto:P.M.Kroonenberg@fsw.leidenuniv.nl) (P.M. Kroonenberg).

abundance of the individual species within the community to be better determined, and provide additional insight, especially for low abundance species.

## 2. Materials and methods

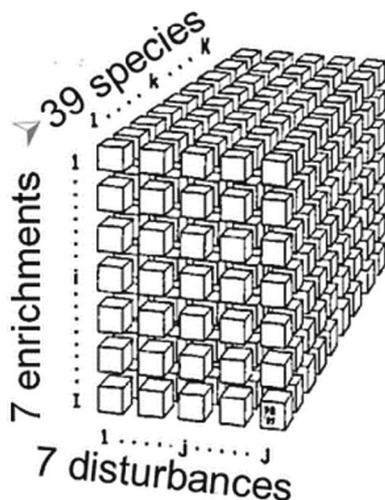
### 2.1. Data characteristics: abundance, enrichment, and disturbance

In this paper we present a secondary analysis of a data set (here referred to as the *Oslofjord benthic data*) earlier analysed by Widdicombe and Austen (2001) and Widdicombe et al. (2004). A comprehensive overview will be presented of the results of an experiment with a  $7 \times 7$  lay-out with two environmental factors: increasing organic enrichment and increasing physical disturbance. The prime focus of the original research was on ecologic diversity, but Widdicombe and Austen (2001) also reported separately on the 16 most abundant species and their reactions to the treatments. Here, the emphasis is especially on the interaction effects of the two factors on 39 species simultaneously, including the positions of 23 less frequently observed species.

The data will be analysed at species level rather than at some higher level of taxonomic aggregation even though some species have small levels of abundance in the data set. The basis for doing this can, for instance, be found in Legendre and Legendre (1998).

Before presenting this analysis, a straightforward demonstration based on abundance frequencies is supplied to illustrate the problem of aggregating over species to examine the influence of enrichment and disturbance on the entire benthic community.

Before going into the details, a brief discussion of the terminology is necessary to characterise the data matrices and methods used in the paper. In particular, it should be clear what in the literature is meant by the words *way* and *mode*. There exists a subtle difference between them when describing matrices and methods. Generally, *way* indicates the dimensionality of a matrix. For instance, a three-way matrix has the shape of a box consisting of rows, columns and levels or slices, and each *way* may refer to the same or a different type of entities. Fig. 1 shows the three-way data matrix examined in this study, where each way refers to different entities: *Level of enrichment* (rows) by *level of disturbance* (columns) by *species* (slices). On the other hand, in a standard correlation matrix the rows and the columns consist of the same entities i.e., *variables*. Therefore, it can be characterised as a two-way one-mode matrix, the word *mode* refers to the content of the two ways.



**Fig. 1.** Three-way contingency matrix/table consisting of three modes: For the current dataset that would equate to 7 Enrichments by 7 Disturbances by 39 Species. (Note, the precise number of cubes in each ‘mode’ in this figure is used to represent the general theoretical structure and does not to represent the specific number of treatment levels of species in the current data set)

When describing techniques to analyse three-way data as in this paper, the word *way* in its name does not a priori imply the entities in the ways are equal or different, while the word *mode* suggests that the entities are different. Unfortunately, in the literature this distinction has not always been rigorously made.

### 2.2. Abundance data and correspondence analysis

The interaction between enrichment and disturbance on the community will be analysed with a three-way generalisation of correspondence analysis. This type of analysis is especially suitable to investigate the response of the species jointly, rather than approaching each of them separately; for a detailed exposition see Carlier and Kroonenberg (1996, 1998). Moreover, a larger number of species can be taken into consideration, even those with a relatively low abundance. As such, it will be a useful addition to the toolbox of the numerical ecologist.

One of the characteristics of abundance data, such as ours, is the non-occurrence of many species under the conditions generated in the experiment. This absence of species is problematic for many multivariate analysis techniques, but Legendre and Legendre (1998) argue, that in calculating the distance between two sites, the  $\chi^2$ -distance provided in correspondence analysis is not affected by double zeroes, i.e. the absence of both species. Thus the technique can be used directly with sites described by species abundances. Correspondence analysis preserves the Euclidean distances between “profiles of weighted conditional probabilities”, i.e. it preserves the  $\chi^2$ -distance. In fact, the Legendre and Legendre (1998) recommended correspondence analysis for reduced space ordination of species abundances when the data contain many zero cell frequencies.

Ter Braak (1985) explicitly discussed the justification of correspondence analysis on abundance data tables, showing that the  $\chi^2$ -distance preserved through correspondence analysis is an appropriate model for species with unimodal distributions along environmental gradients.

On the other hand, Clarke (1993) raised objections to analysing abundance data with correspondence analysis on the grounds that rare species have undue influence on the outcome of the analysis. This observation is discussed by Greenacre (2013), who shows that, in a correspondence analysis, plots rare points are not necessarily the most influential points. Rare species do often lie at the border of correspondence analysis plots because their locations are often only relatively weakly determined. To eliminate this visually distracting feature of correspondence analysis, it is generally advisable to put some lower limit on the abundance of species to be included in the analysis.

Given the earlier analyses of Widdicombe and Austen (2001) on the abundance of 16 most abundant species of this data set, we intend to discuss how an additional 23 less abundant species relate to more abundant ones. Note, however, that an additional analysis of a species two-factor interaction is recommended for detailed investigation of the biological reasons for its sensitivity to enrichment and disturbance.

### 2.3. A brief introduction in the correspondence analysis biplot

The major instrument for displaying the results of correspondence analyses is the so-called *biplot*. The part *bi-* refers that to the simultaneous portraying of the rows and the columns, as well as their relationship which can be expressed as the *scalar product* of the row profile  $i$  ( $r_i$ ) and the column profile  $j$  ( $c_j$ ). The row and the column profiles are depicted as vectors which go through the centre of the biplot coordinate axes. This centre represents both the row and the column means, i.e. their marginal frequencies. The coordinates of a biplot in correspondence analysis are either standardised, i.e. have length one and zero mean, or they are principal coordinates with length equal to the singular values. i.e. the square root of the eigenvalues. Details about the coordinate system in correspondence analysis can, for instance, be found in Beh and Lombardo (2014).

The scalar product  $d_{ij}$  of a row vector  $r_i$  and a column vector  $c_j$  is

defined as ‘length of the projection of the row vector onto the column vector’ times ‘length of the column vector’. In formula this is

$$d_{ij} = [ | \mathbf{r}_i | \cos(\psi_{ij}) ] \times | \mathbf{c}_j |,$$

where  $\psi_{ij}$  is the angle between the two vectors. This formula is illustrated in Fig. 2.

A species  $\times$  enrichment biplot of abundances shows the abundance of the species for the levels of enrichment. Fig. 2 show that *Capitella* thrives with high enrichment, while *Anobothrus gracilis* prefers just below average enrichment as can be seen from their projections onto the enrichment vector. Insightful descriptions as well as the mathematical details of the interpretation of biplots can, for instance, be found in Greenacre (1993) and Beh and Lombardo (2014).

#### 2.4. Informal description of the methodology of three-way correspondence analysis

Widdicombe & Austin (2001) analysed the disturbance-by-enrichment data per species via a two-way analyses of variance. Their results consist of 16 separate plots showing the two-way enrichment-  $\times$  -disturbance interaction for each species (see their Fig. 1). However, when there are many species, an interpretational problem arises because each species is represented by its own plot, and each of the separate plots need to be inspected for similarities and differences to discover the various two-factor interactions for the species.

Three-way correspondence analysis aims to combine the two-way effects per species in a single analysis for all species together. Informally, one could say that the technique performs simultaneous component analysis on all biplots, resulting in a set of components for the species (rows) on which all enrichment-  $\times$  -disturbance combinations (columns) will have their own coordinates.

This approach is unlike *multiple correspondence analysis* (MCA), which aims to analyse the structure of the relationships between several categorical variables in a way which is similar to principal component analysis for quantitative variables. MCA is also known under various other names, such as dual scaling, optimal scaling, and homogeneity analysis (see e.g. Abdi and Valentin (2007)).

Different species will obviously have different values on the components, as will the enrichment-  $\times$  -disturbance combinations, so that they need to be displayed in the same plot. The location of a species will be determined by its typical values on its enrichment-  $\times$  -disturbance interaction. Whereas in classical correspondence analysis the biplot

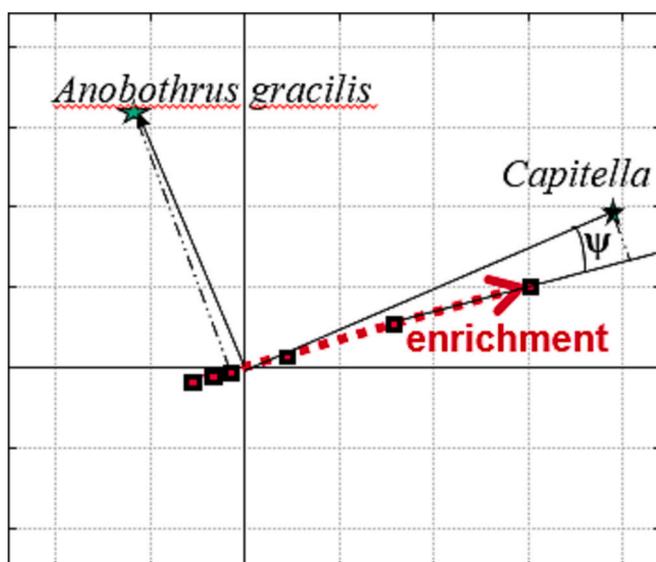


Fig. 2. Elementary biplot with basic interpretational elements.

displays the two-way interaction between the enrichments and the disturbances, three-way correspondence analysis displays the interaction between the species and the two-way enrichment-  $\times$  -disturbance combination in an *interactive biplot*. Thus the technique makes it possible to portray in this biplot the structure of the enrichment-  $\times$  -disturbance combinations in the space of the species.

#### 2.5. Experimental design

The experiment to generate the data was carried out during 1996 at the NIVA marine station Solbergstrand, Oslofjord, Norway. The methods are described in detail in Widdicombe and Austen (2001) and summarised here. Two sets of plastic buckets with homogenised sediment from the Oslofjord were each arranged in  $7 \times 7$  blocks, thus in two times 49 buckets. During a 12-week period, each bucket was subjected to one of seven levels of organic enrichment combined with one of seven frequencies of physical disturbance. The seven *treatment levels* (P0 to P6) of the enrichment were equivalent to 0, 12.5, 25, 50, 100, 200, and 400 g carbon  $m^{-2}$ . The seven levels of disturbance were no disturbance (D0), once every four weeks (D1), once every 2 weeks (D2), once a week (D3), twice a week (D4), three times a week (D5), and every day (D6). Note that these two treatments have a strong ordinal character (see Section 6 for comments on this). The data of the two replicates were taken together and analysed as if there was only one bucket.

#### 2.6. Data and their treatment

Correspondence analysis typically operates on non-negative numbers, mostly frequencies or some transformation of them. Widdicombe and Austen (2001) transformed the abundances with a fourth-square root (i.e. double square root) and subsequently multiplied by a 1000 and rounded; see e.g. Field et al. (1982) and Clarke (1993), for a justification for this procedure with abundances. Their transformed data remained non-negative values and were treated as if they were frequencies. For a further discussion of the double square root transformation can be found in Gorja (1992).

The 16 most abundant species were individually analysed by Widdicombe and Austen (2001). The most abundant species analysed were *Heteromastus filiformis* (22,608 individuals) and *Chaetozone setosa* (3121 individuals), while the least abundant species analysed was *Eteone flava* (88 individuals). In this study we also included the 23 next most abundant species which had medium frequencies from 87 to 10. It was decided not to include the remaining 41 species due to their rarity. They occurred only 1–9 times, i.e. on average only in one bucket or less. In addition, the frequencies of the most abundant species (*Heteromastus filiformis*) were divided by 10 before taking the double square root thereby making their number more comparable with the other species and preventing it from dominating the analysis. This also allowed for easier graphing. For the correspondence analysis, the data were arranged as a three-way contingency table of 39 species, 7 levels of enrichment and 7 levels of disturbance (Fig. 1). Note that the two factors are ordinal, but this aspect was only taken into consideration during the interpretation and not during the analysis.

To summarise, the main analysis was carried out on a three-way contingency table in which each cell contained the double square root of the abundance of species found in a bucket with enrichment regime  $j$  ( $P_{(j-1)}$ ) and disturbance level  $k$  ( $D_{(k-1)}$ ).

### 3. Three-way correspondence analysis<sup>1</sup>

Important properties of three-way correspondence analysis (three-way CA) are that the dependence (or interactions) between the variables

<sup>1</sup> This section is based on Carlier and Kroonenberg (1998) with kind permission from the Academic Press

in a three-way table can be measured, modelled, and displayed. Moreover, it shares and extends many properties of ordinary (two-way) correspondence analysis by aggregating the  $K$  levels of the third variable. Some other papers in which three-way tables were analysed without reducing them to two-way tables are Choulakian (1988), Kroonenberg (1989), Dequier (1974), Carlier and Kroonenberg (1996, 1998), and Lombardo et al. (2016), Lombardo et al., 2021). The latter also discuss extensions of the standard technique (see Section 6).

In this section we will successively discuss the measuring, modelling, and displaying of the dependence in a three-way table basing ourselves on the ‘standard’ version of three-way correspondence analysis.

### 3.1. Measuring dependence

In three-way tables one can distinguish (1) global dependence based on all interactions, i.e. it consists of the deviation from the three-way independence model, (2) marginal dependence due to the two-way interactions, and (3) three-way dependence due to the three-way interaction alone.

#### 3.1.1. Measuring global dependence

Three-way contingency tables have orders  $I, J,$  and  $K$ . The first way (rows) is indicated by  $i$  with  $i = 1 \dots I$ , the second way (columns) by  $j$  with  $j = 1 \dots J$ , and the third way (tubes) by  $k$  with  $k = 1 \dots K$ , and relative frequencies are indicated by  $p_{ijk}$ . Dependence in the table is measured with the Pearson mean square contingency coefficient  $\Phi^2$  which is

$$\Phi^2 = \sum_{ij} p_{i..} p_{.j.} \left( \frac{p_{ijk} - p_{i..} p_{.j.}}{p_{i..} p_{.j.}} \right)^2 + \sum_{ik} p_{i..} p_{.k.} \left( \frac{p_{ijk} - p_{i..} p_{.k.}}{p_{i..} p_{.k.}} \right)^2 + \sum_{jk} p_{.j.} p_{.k.} \left( \frac{p_{ijk} - p_{.j.} p_{.k.}}{p_{.j.} p_{.k.}} \right)^2 + \sum_{ijk} p_{i..} p_{.j.} p_{.k.} \left( \frac{p_{ijk} - a_{ijk}}{p_{i..} p_{.j.} p_{.k.}} \right)^2 =$$

defined as

$$\Phi^2 = \frac{X^2}{N} = \sum_{i,j,k} \frac{(p_{ijk} - p_{i..} p_{.j.} p_{.k.})^2}{p_{i..} p_{.j.} p_{.k.}} = \sum_{i,j,k} p_{i..} p_{.j.} p_{.k.} \left[ \frac{p_{ijk} - p_{i..} p_{.j.} p_{.k.}}{p_{i..} p_{.j.} p_{.k.}} \right]^2 = \sum_{i,j,k} p_{i..} p_{.j.} p_{.k.} (\Pi_{ijk})^2. \tag{1}$$

Thus  $\Phi^2$  is based on the deviations from the three-way independence model, and it contains all two-way interactions and the three-way interaction. It is the sum of the weighted cell contributions to the dependence  $\Pi_{ijk}$ . The contribution of cell  $(i,j,k)$  to the  $\Phi^2$  may be rewritten as

$$\Pi_{ijk} = \frac{Pr[ij|k] \cdot Pr[ij]}{Pr[ij] \cdot Pr[i] Pr[j]} - 1. \tag{2}$$

The quantity  $1 + \Pi_{ijk}$  is the product of (1) the ratio  $Pr[ij|k]/Pr[ij]$  which measures the relative increase or decrease in the joint probability of the categories  $i$  and  $j$  given occasion  $k$  and (2) the ratio  $Pr[ij]/Pr[i]Pr[j]$  which measures the relative increase or decrease in the deviation from the marginal independence. If the conditional probabilities for all  $k$  are equal, then  $Pr[ij|k] = Pr[ij]$  and the first ratio is 1. Then  $\Pi_{ijk} = \Pi_{ij.}$ , and the three-way table could be analysed with ordinary two-way correspondence analysis by aggregating the  $K$  levels of the third variable. The symmetric statement after permutation of the indices holds as well. Therefore, the  $\Pi_{ijk}$  measure the global dependence of the cell  $(i,j,k)$ .

$$\sum_{i,j,k} \frac{(p_{ijk} - p_{i..} p_{.j.} p_{.k.})^2}{p_{i..} p_{.j.} p_{.k.}} = .$$

The elements of the two-way marginal totals are defined as weighted sums over the third index. Thus, for the margins these elements are:

$$\Pi_{.jk} = \sum_i p_{i..} \Pi_{ijk} = \sum_i p_{i..} \left( \frac{p_{ijk} - p_{i..} p_{.j.} p_{.k.}}{p_{i..} p_{.j.} p_{.k.}} \right) = \frac{p_{.jk} - p_{.j.} p_{.k.}}{p_{.j.} p_{.k.}}. \tag{3}$$

The elements of the other two-way margins,  $\Pi_{i.k}$  and  $\Pi_{ij.}$ , are similarly defined. One-way marginal totals are summed over two indices, and they are zero due to the definition of  $\Pi_{ijk}$ ; the overall total is zero as well. For instance, the one-way row margin  $i$

$$\Pi_{i..} = \sum_j \sum_k p_{i..} p_{.j.} p_{.k.} \Pi_{ijk} = 0. \tag{4}$$

#### 3.1.2. Measuring marginal and three-way dependence

The global dependence of a cell,  $\Pi_{ijk}$ , can be split into separate contributions of the two-way interactions and the three-way interaction,

$$\Pi_{ijk} = \frac{p_{ij.} - p_{i..} p_{.j.}}{p_{i..} p_{.j.}} + \frac{p_{i.k} - p_{i..} p_{.k.}}{p_{i..} p_{.k.}} + \frac{p_{.jk} - p_{.j.} p_{.k.}}{p_{.j.} p_{.k.}} + \frac{p_{ijk} - a_{ijk}}{p_{i..} p_{.j.} p_{.k.}}, \tag{5}$$

where  $a_{ijk} = p_{ij.} p_{.k.} + p_{i.k} p_{.j.} + p_{.jk} p_{i..} - 2p_{i..} p_{.j.} p_{.k.}$ . The terms referring to the two-way margins are equivalent to those in ordinary correspondence analysis. The last term,  $a_{ijk}$ , measures the size of the three-way interaction for cell  $(i,j,k)$ . Darroch (1974) provides a comparative discussion of this additive definition of interaction and the multiplicative definition as used in loglinear analysis, see also Kroonenberg and Anderson (2006) for a discussion of additive and multiplicative modelling of contingency tables.

Due to the additive splitting of the dependence of individual cells,  $\Phi^2$ , the measure for global dependence of the table, can be partitioned as

$$\Phi^2_{IJ} + \Phi^2_{IK} + \Phi^2_{JK} + \Phi^2_{IJK} \tag{6}$$

This partitioning of the Pearson mean square contingency coefficient was introduced by Lancaster (1951) and has also been used by Dequier (1974) and Choulakian (1988). The importance of decomposition (Eq. 6) is that it provides measures of fit for each of the interactions and thus their contribution to the global dependence.

### 3.2. Modelling dependence

Given measures for global dependence, marginal dependence, and three-way dependence, a model for these measures must be found with which it will be possible to construct graphs depicting the dependence. In two-way correspondence analysis, the generalised singular value decomposition (GSVD) is used for this purpose. Thus, for the three-way case, a three-way analogue of the GSVD is desired. There are, however, several candidates of which we will consider only the so-called *Tucker3 model*; (see Carlier and Kroonenberg, 1996, for other possibilities). This model is also referred to as the *three-mode component analysis model* (see Kroonenberg, 1983, 2008; Tucker, 1966), but there is no specific requirement that the three-ways of the data have to be different in content. Sets of correlations matrices could be analysed by the technique as well.

#### 3.2.1. Modelling global dependence

In the Tucker3 generalisation of the singular value decomposition each of the way has its own components, but the generalised singular values are unlike the two-way case in that they are indexed by the components of all three ways. The component matrices of three ways are respectively **A** (size:  $I \times P$ ), **B** (size:  $J \times Q$ ), and **C** (size:  $K \times R$ ), where  $P, Q,$  and  $R$  are the numbers of components (columns) of **A, B,** and **C,**

respectively. The statistical formula for the Tucker3 GSVD model then becomes:

$$\Pi_{ijk} = \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R g_{pqr} a_{ip} b_{jq} c_{kr} + e_{ijk}. \tag{7}$$

In this decomposition, the  $a_{ip}$  are the elements of the components  $\{a_p\}$  which are pairwise orthonormal with respect to the weight  $(p_i)$ . Similarly the  $b_{jq}$  are the elements of the components  $\{b_q\}$  which are pairwise orthonormal with respect to  $(p_j)$  and the  $c_{kr}$  are the elements of the components  $\{c_r\}$  which are orthonormal with respect to  $(p_k)$ . The  $g_{pqr}$  are the three-way analogues of the singular values, and they are often referred to as elements of the *core array*  $\mathbf{G}$ . The  $e_{ijk}$  represent the errors of approximation. In three-way correspondence analysis, a weighted least-squares criterion is used: the parameters  $g_{pqr}$ ,  $a_{ip}$ ,  $b_{jq}$  and  $c_{kr}$  are those which minimise

$$\sum_{i,j,k} p_{i..} p_{.j.} p_{..k} e_{ijk}^2. \tag{8}$$

As in two-way correspondence analysis,  $\Phi^2$ , the global measure of dependence can be split into a part fitted with the three-way singular value decomposition or three-way model and a not-fitted residual part.

### 3.2.2. Modelling marginal dependence

One of the attractive features of the additive partitioning of the dependence in Eq. (6) is that the single decomposition of the global dependence can be used to model the marginal dependence as well.

The marginal dependence of the columns  $j$  and tubes  $k$  is contained in a matrix  $\Pi_{JK}$  with elements  $\Pi_{jk}^{JK} = (p_{.jk} - p_j p_{..k}) / p_j p_{..k}$  with similar expressions for the other two matrices with marginal dependencies  $\Pi_{IJ}$  and  $\Pi_{IK}$ . The elements  $\Pi_{jk}^{JK}$  are derived via a weighted summation over  $i$  from the global dependence,

$$\Pi_{jk}^{JK} = \sum_i p_i \cdot \Pi_{ijk} \tag{9}$$

We have modelled the global dependence  $\Pi_{ijk}$  with the Tucker3 model in Eq. (7). To find the model equation for the marginal dependence  $\Pi_{jk}^{JK}$  of ways  $J$  and  $K$ , we have to use the same weighted summation as in Eq. (9), so that we have to use the weighted summation over the terms in Eq. (7) depending on the index  $i$  leading to summation of the summated component elements  $a_p = \sum_i p_i \cdot a_{ip}$  and the summated error terms  $e_{jk} = \sum_i p_i \cdot e_{ijk}$ .

$$\Pi_{jk}^{JK} = \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R g_{pqr} a_p b_{jq} c_{kr} + e_{jk} \tag{10}$$

Thus this formula shows that the marginal model is derived from the overall model by averaging the appropriate components (here:  $a_p$ ). This will turn out to be extremely effective in the displays we intend to make.

As Eq. (10) shows, the modelling of the partial dependence is achieved by centring the components of one of the modes (here the species way:  $a_p$ ). In Carlier and Kroonenberg (1996) it was discussed what happens when one wants to remove more than one marginal dependence at a time.

### 3.3. Plotting dependence: biplots

With respect to dependence and its modelling, the three ways of the contingency table behave in an entirely symmetric fashion. This symmetry can, however, not be maintained when graphing the dependence, as no sensible spatial representations exist to portray all three ways simultaneously in one graph. Given the effectiveness of biplots in standard correspondence analysis, biplots should be used to display the dependence or its approximation in three-way correspondence analysis. In particular, the *joint biplot*, discussed by Carlier and Kroonenberg (1996), and the *interactive biplot* discussed and used by Carlier and Kroonenberg (1998) fulfils this purpose.

The *interactive biplot* aims to portray all three ways in a single biplot.

As a biplot has only two types of points, two of the three ways must be combined into a single one. In the correspondence analysis literature, this combination process is referred to as *interactive coding*. The remaining way supplies the other set of points, and it will be referred as the *reference mode*. The choice of reference mode is very data dependent. Given that the ordered factors (ways) fulfil a comparable role in the present data they will be coded interactively, and the species will form the reference mode.

The construction of the biplot for the global dependence follows directly from three-way singular value decomposition of the global dependence:

$$\begin{aligned} \hat{\Pi}_{ijk} &= \sum_{p=1}^P \left[ \sum_{q=1}^Q \sum_{r=1}^R g_{pqr} b_{jq} c_{kr} \right] a_{ip} = \sum_{p=1}^P d_{(jk)p} a_{ip} = \sum_{p=1}^P d_{\ell p} a_{ip} \\ &= \hat{\Pi}_{\ell i}. \end{aligned} \tag{11}$$

We have replaced the  $(jk)$  with a new index  $\ell$  so that the coordinates of the  $l$ -th row points in a biplot are the  $d_{\ell p}$ , and those of the  $i$ -th column points are the  $a_{ip}$ . Note that the core elements  $g_{pqr}$  are absorbed in the coordinates of the row points. Therefore, it can be shown that the interactive biplot preserves the *row-metric* with respect to the weights  $p_{i..}$  if we assume that the points are well represented by the biplot. The number of two-dimensional biplots does not depend on  $Q$  or  $R$  but only on  $P$ , the number of components of the reference mode, and it is equal to  $P/2$  when  $P$  is even.

The interactive biplot is especially useful when the number of elements  $J \times K$  is not too large, and, when at least one of the two sets  $J$  or  $K$  is ordered, the interpretation is facilitated. By using a correspondence analyse especially designed to model ordered variables, the interpretation can be further enhanced.

## 4. Results

Following the structure of the previous sections we will discuss the results for the Oslofjord benthic data in the same order. However, first we will inspect the contingency table of abundance frequencies aggregated over species.

### 4.1. Aggregated enrichment- $\times$ -disturbance abundance

A basic preliminary question is whether a three-way analysis is necessary. To investigate this we have aggregated the frequencies of the 59 most abundance species in a two-way frequency table of 7 enrichments-  $\times$  -7 disturbances (Table 1).

In the Table only a few regularities are to be seen. (1) The column marginal disturbance totals steadily decline up to *Twice a week*, showing fewer individuals with increased disturbance. (2) The decrease stops with even more disturbance. (3) There is no distinct pattern for the number of individuals for increasing enrichment up until 100 g/m<sup>2</sup>. Further enrichment diminishes the number of individuals considerably. (4) There are dramatic differences within columns and within rows. The irregularities point to an underlying problem, i.e., the different responses of different species to the enrichment and disturbance and their interaction. In other words, there is a strong three-way species  $\times$  enrichment  $\times$  disturbance interaction. Therefore, the options for meaningful analyses are either by species or via an analysis of the three-way data. The first option was carried out in detail by Widdicombe and Austen (2001), the second option is pursued in this paper via three-way correspondence analysis.

For the present larger data set it was decided to examine a three-way analysis with  $P = 2$ ,  $Q = 2$ , and  $R = 2$ , thus two components for each of the ways. To evaluate the size of the fitted interaction of the  $2 \times 2 \times 2$  model, the size of fit the interaction of the  $4 \times 3 \times 3$  solution is also presented in Table 2., as well as the fit of the interactions. The fit of the extended solution increases with 12%, but it requires 118 extra

**Table 1**  
Overall abundance of the 59 species.

Enrichment (grams per m <sup>2</sup> )	Disturbance							Total
	None	Once in 4 weeks	Once in 2 weeks	Once a week	Twice a week	Thrice a week	Every day	
0	1355	768	861	819	372	413	918	5506
12.5	1444	1191	933	908	348	1343	256	6423
25	1425	998	643	979	402	455	431	5333
50	718	935	631	898	416	441	307	4346
100	1180	684	1098	826	368	639	438	5233
200	701	616	496	407	313	293	372	3198
400	98	579	976	150	188	327	240	2558
Total	6921	5771	5638	4987	2407	3911	2962	32597

**Table 2**  
Oslofjord benthic data: Chi-Square partitioning of the dependence and model fit based on Eqs. (6) and (7).

Effects	Φ <sup>2</sup>		2×2×2 Model <sup>a</sup>	2×2×2 Model <sup>a</sup>	4×3×3 Model <sup>a</sup>
	df	% Total Dependence	% Fitted Dependence (vertically)	% Fit/Total Dependence <i>per effect</i> (horizontally)	
	1	2	3	4	5
Constant	1	0	0		
<i>Main effects</i>					
Species (=39)	38	0%	0% <sup>b</sup>	0	0
Disturbance (= 7)	6	0%	-2% <sup>b</sup>	0	1
Enrichments (=7)	6	0%	-0% <sup>b</sup>	0	0
<i>Interactions</i>					
Enrichment-x-Disturbance	36	6%	23%	71%	78%
Species-x-Enrichment	228	20%	54%	50%	56%
Species-x-Disturbance	228	16%	13%	15%	49%
<i>Three-way Interaction</i>					
	1368	58%	10%	4%	12%
<i>Total<sup>c</sup></i>	1860	100%	100%	18%	30%

<sup>a</sup> The number of components in a  $P \times Q \times R$  model refer to the Species, Enrichments and Disturbance ways, respectively.

<sup>b</sup> For the perfectly fitting model the main effects are 0; however, for other models, the estimates for the main effects may show small deviation from zero.

<sup>c</sup>  $\Phi^2(\text{Total}; 2 \times 2 \times 2 \text{ model}) = 908,095 (100\%)$ ;  $\Phi^2(\text{Fit } 2 \times 2 \times 2) = 167,437 (18\%)$ ,  $\Phi^2(\text{Fit } 4 \times 3 \times 3) = 273,337 (30\%)$ .

<sup>d</sup> Columns 1, 2, 3 should be compared vertically to compare different interactions within the  $2 \times 2 \times 2$  model.

The blue values in columns 4 and 5 should be compared horizontally per interaction for the two different models.

parameters to be estimated and interpreted compared to the  $2 \times 2 \times 2$  solution. Therefore, it has to be decided which of the two solutions should be chosen, basing ourselves on both statistical and substantive arguments. This will be done in the next sections.

#### 4.2. Measuring dependence

The first task when conducting a three-way correspondence analysis is to assess the partitioning of the chi-squared statistic to evaluate the

contributions of the four interactions (see Table 2). As the dependence is expressed as deviation from the independence model, the main effects are theoretically zero but may be slightly different from zero in an approximate solution. The main effects do not contain relevant information about the basic problem.

Note that in Table 2 we are concentrating on the descriptive significance rather than on statistical significance of the results. Thus we consider the relative size of interactions with respect to each other, within and between models. The prime reason is that we intend to show what can be learned about the data via three-way correspondence analysis with visual representations.

The information which is of prime interest in the present case is the nature of the interactions of the Disturbance  $\times$  Enrichment effect on the abundances of the individual species. These are contained in two two-way interactions and the three-way interaction.

All interactions: Species  $\times$  Enrichment, Species  $\times$  Disturbance, Three-way interaction

The size of the two-way interactions Species  $\times$  Enrichment and Species  $\times$  Disturbance shows that there are large differences between the Species in their reactions to the factors. In absolute terms the largest part of the dependence (58%) is taken up by the three-way interaction, which does not come as a surprise because next to the systematic part, it also contains the larger part of the random unstructured error in the data.

Two-way interaction: Enrichment  $\times$  Disturbance

This design interaction is the smallest of the three two-way interactions (6%) which indicates that, when summed over all species, the marginal two-way dependence does not contain very many systematic deviations from independence, as was shown in Table 1. This is due to the aggregation over species who react differently to the factors, so that differences between the species in their two-way interactions do not

come to the fore. In other words, they have been averaged out.

### 4.3. Modelling dependence

One of the problems in large contingency tables is that not only does the three-way interaction contain random errors, but the two-way interactions do as well. Therefore, it is necessary to separate the systematic and random information contained in each interaction. One of the ways to do this is to perform separate correspondence analyses on each of the interactions. However, as indicated above, an integrated analysis of the complete dependence in the table may be preferred using three-way correspondence analysis, because with this type of analysis the systematic information contained in each interaction can be modelled simultaneously using a relatively small number of parameters; see Eqs. (10) and (11). This leaves the remaining degrees of freedom for the error.

Deciding on the number of dimensions for each of the ways in the analysis requires finding a balance between an acceptable fit and good interpretability of the solution. Table 2 listed such information for two three-way correspondence analysis solutions: one with two components for each way, the  $2 \times 2 \times 2$ -model, and the  $4 \times 3 \times 3$ -model with three components for each of the factors and four for the species. The fit of the first model is 18%, using 106 parameters, and the fit of the second model is 30% using 224 parameters. As interpreting the second model is considerably more complex than the first one as it has about twice as many parameters, we will concentrate on what can be learned from the simpler  $2 \times 2 \times 2$ -model. The rather low percentages fit and the relative limited fit increase by enlarging the model suggests that the data are rather noisy. Notwithstanding, as will become clear in the following, the smaller model supplies considerable insight in the effect of the factors on the abundances. This supports relevant systematic information is

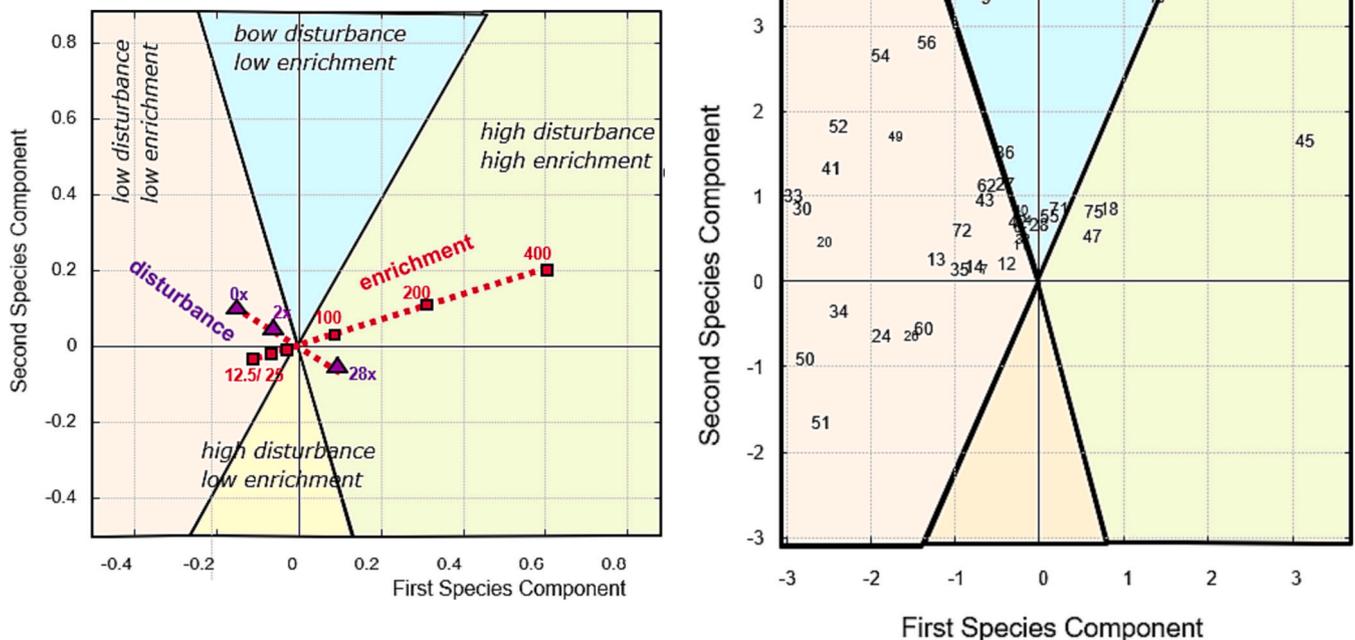


Fig. 3. a. Interactive biplot showing the central part of the two-component Species space with the two-way Species  $\times$  Enrichment and the Species  $\times$  Disturbance interactions in red.

The two black lines through the origin (i.e. average enrichment, average disturbance) are perpendicular to a two-way interaction with species. They roughly divide the space in above and below average disturbance and in above and below average enrichment.

b. Species component space.

Species positions in the Species space. Numbers correspond to list of species in the Appendix.

Species close to each other thrive generally under the same regimes of enrichment and disturbance.

NB. Note the scales of Fig. 3a and 3b are different to allow inspection of the two-way interactions. They can, therefore, not be superimposed. This is done in Fig. 6. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

contained in the data.

The overall fit of this-model can be split into the fit of the four interactions (red in Table 2, 4th column): 54% - species- × -enrichment interaction, 13% - species- × disturbance interaction, 23% - enrichment- × disturbance interaction and - 10% three-way interaction. This 2 × 2 × 2-model succeeds in a very good representation of the design interaction (71%; while it only represents 4% of the interaction contained in three-way interaction (blue in Table 2, 5th column). Even though the three-way interaction contains 58% of the total dependence, only 4% of the three-way interaction is contained in the present model, but this variability makes up 10% of the fit of the model to the data (see Table 2, 4th column).

#### 4.4. Plotting dependence

The basic information obtained from a three-way correspondence analysis are sets of components for each of the three ways (A, B, C) and information on the way these components are linked to each other is contained in the elements  $g_{pqr}$  of the core array (see Eq. (7)). The principal coordinates are based on the deviations from independence and in the previous section we have indicated to what extent these coordinates are based on the information of the different interactions.

As Table 2 shows the 2 × 2 × 2-model with 2 components for each way has an overall fit of 18%. It is difficult to evaluate whether this is sufficient or acceptable for these kinds of data given the lack of similar studies. In our case we will assume that there is a large random component in these data. For example from Fig. 5 it can be seen that at least for the *Ophelina modesta*, the abundance pattern is rather irregular, and thus contains considerable error. Similar patterns can be observed for the other species. We will therefore aim to eke out what variability is systematic. The percentages for the two components of each way are respectively: Species: 65%, 35%; Enrichment: 90%, 10%; Disturbance: 55%, 45%.

##### 4.4.1. Species space and two-way interactions

As explained in Section 3.3, one of the best ways to gain insight into the patterns in the interactions is to produce graphs of these interactions

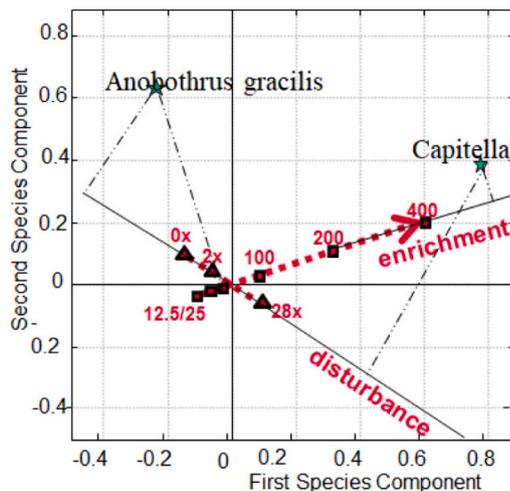


Fig. 4. How to read detail in the biplot.

The angle between Disturbance and Enrichment is approximately 45° which indicates that there is an interaction between the two factors for the species. *Anobothrus gracilis* (nr. 56) projects on the low side (around 50) for enrichment about the average of all species, but extremely low on disturbance. Thus, it cannot tolerate any disturbance and lives at rather low levels of enrichment (see also Fig. 7).

On the other hand, unlike other species, *Capitella* spp. (nr. 45) thrives with high levels of both enrichment and disturbance.

via interactive biplots. Such plots need a reference mode which provides the actual plotting space. In our case this space will be that of the species. The species will have normalised coordinates and the interactively coded environments and disturbances will be in principal coordinates.

Eq. (10) and its variants have been used calculating the marginal dependences of the species and enrichments, and the species and disturbances. The associated coordinates of the ordered levels of the two factors in the species space have been plotted and connected with red dotted lines in Fig. 1a. Projections onto these lines allow one to determine the relative position with respect to a species' enrichment- × -disturbance. They show for instance that species *Capitella* sp. (Appendix number 45) thrives with high enrichment and high disturbance, while species *Anobothrus gracilis* (Appendix number. 56) can only thrive without disturbance but it is only mildly sensitive to enrichment. This procedure is illustrated in Fig. 3. The interaction lines are also used to divide the species space roughly into four areas as indicated in the plot. These areas are created by drawing lines through the origin perpendicular to the red interaction lines (Fig. 3a).

The origin of the space represents the abundance of the average species, so that the average species (if such an individual exists) thrives best when being roughly four times disturbed and being supplied an organic enrichment of around 50 g per square meter. However, species on the rims of the plot thrive better under other regimes.

Even though the two plots in Fig. 3 portray the same species component space, they cannot be directly superimposed as presented here, because of they are scaled differently. This was done to enable detailed examination of the interactions. Their joint representation in a single graph is given in Fig. 7.

In Fig. 3b one can see that many species can be found around the origin, but primarily in the low disturbance and medium enrichment section of the plot. There are 17 species further out from the origin with two of them [*Capitella* sp. (Appendix nr. 45)  $n = 33$ ; *Nudibranch* indet A. (Appendix nr. 78):  $n = 10$ ] at the rim of the configuration; they thrive with very high enrichment levels. A more detailed explanation of the interpretation of the interactive biplot is given in Fig. 4

We can deduce from the locations of the species Fig. 3b that most species prefer lower disturbances and lower enrichment, as they are located in the north-west corner of the plot. Five species (nrs. 24, 26, 50, 51, 60) allow for more disturbances, but not more enrichment. An example is species 50 (*Ophelina modesta*) whose abundance for each value of the factors is depicted in Fig. 5.

##### 4.4.2. Three-way interactive biplot

The interactive biplot in Fig. 6 is constructed with the coordinates of the interactively coded factors from the three-way interaction. Each point represents one of the 49 enrichment-disturbance combinations. As the species space is the same as before, the species are located at the same position as in Fig. 3b. Each level of disturbance is a line marked by the 7 enrichment levels.

The biplot shows that the increase in enrichment is primarily represented by the first component, and the increase in disturbance by the second one. Note, however, the interchange between 1× and 2× disturbance. A similar inversion can be seen for enrichment between 0 and 12.5 g and 25 g of carbon. Such situations can be expected when there are no consistent gradients in the factors, i.e. when each species has its own location in the plot and its own gradients.

The figure shows that there is a marked difference between the abundances of species when there is no or little disturbance or very high disturbance. Very high disturbances need to have high levels of enrichment to be a viable environment for benthic species, but there are only very few such species.

The species discussed in Widdicombe & Austin (2001) are marked with black numbers in Fig. 7. They are generally close to the origin, partly due to their high abundances. Apart from nr. 26 (*Scoletoma fragilis*), they lie in the top part of Fig. 7 with relatively low enrichment and limited disturbance. Note that there are no species located in the South-

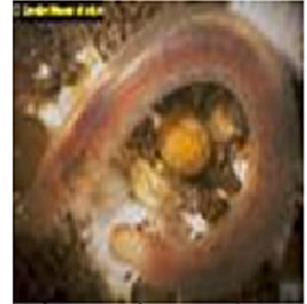
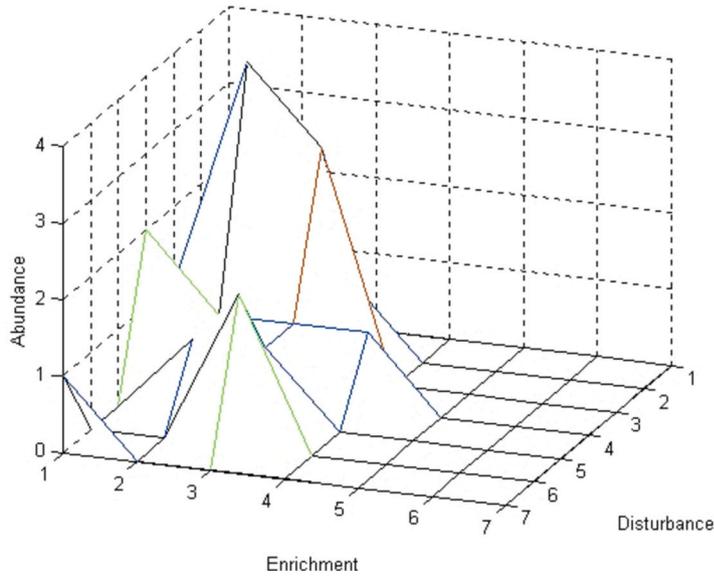


Fig. 5. *Ophelina modesta* (Appendix nr. 50). Abundance for all enrichment- × -disturbance combinations. No individuals are found above enrichment level 5. Varying abundances for enrichments below level 5. So this confirms that the species has a high abundance with low enrichment and low disturbance, as Fig. 3 shows.

### First versus Second Component of Mode 2 (Species )

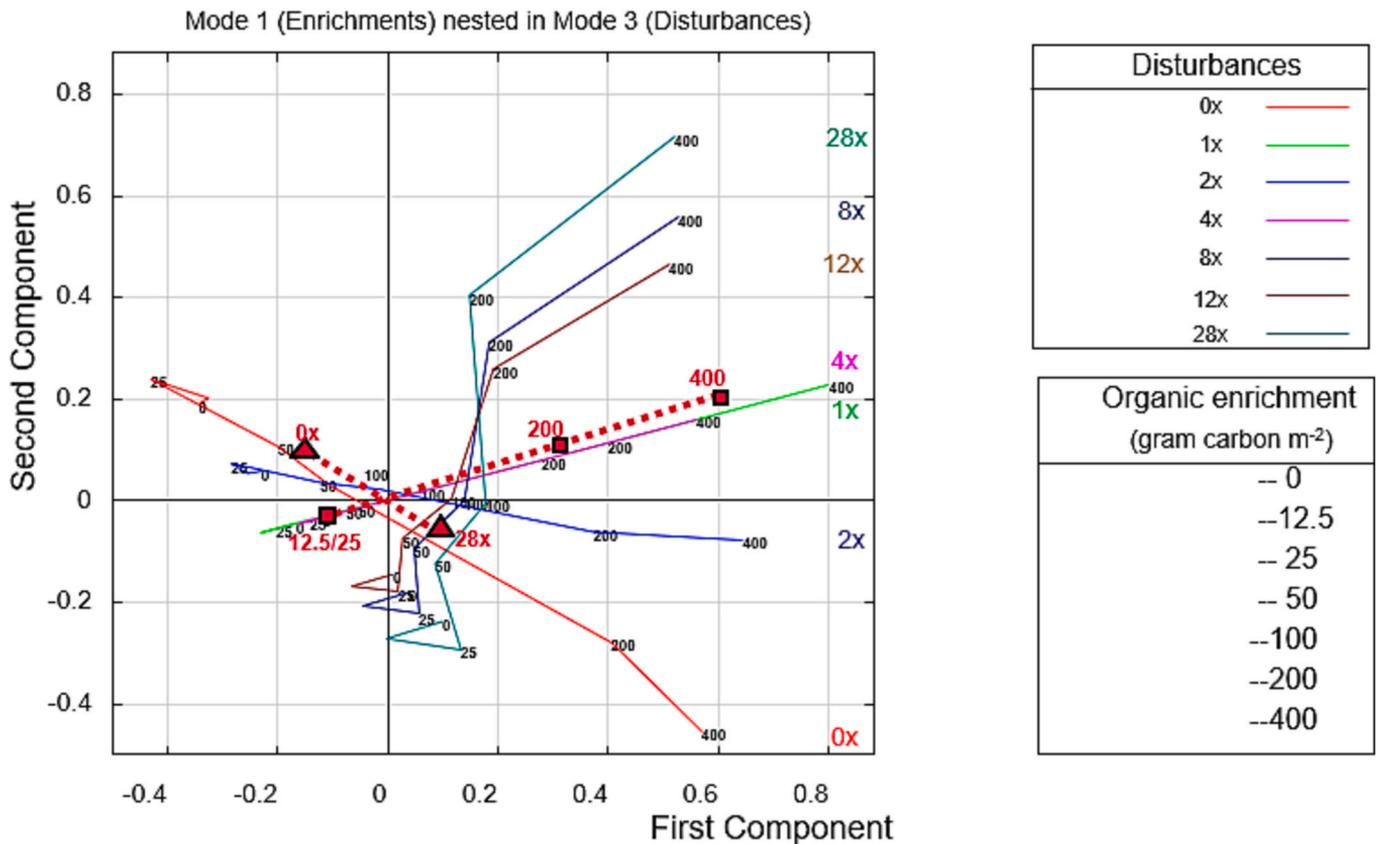


Fig. 6. Interactive biplot showing the three-way interaction. Lines of increasing disturbance are marked with the levels of enrichment. Thus the continuous red line shows the enrichments marked on the zero disturbance line, and the other disturbance lines are marked similarly. No disturbance is intrinsically different from even light disturbance with respect to the abundance of certain species. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



influenced by an organism's body *surface to volume ratio*. Their main conclusions are repeated here to provide a biological explanation of the species' location in the plot. Moreover, the locations of the less abundant species with respect to the more abundant ones can assist in explaining the influence of the factors on their abundance.

#### Biological explanations for sensitivities to disturbance and enrichment

- Tube building species such as *Pseudopolydora paucibranchiata* (nr. 36 -  $n = 492$ ) and *Anobothrus gracilis* (nr. 56 -  $n = 140$ ) were shown to be extremely sensitive to increases in physical disturbance. This sensitivity is assumed to be a result of either damage to individuals or the failure of the organism to regain/maintain its position within the sediment during or after disturbance.
- *Large mobile species*, such as *Goniada maculata* (nr. 14 -  $n = 363$ ) and *Scoletoma fragilis* (nr. 26 -  $n = 130$ ) showed no such intolerance to physical disturbance. However, their abundance was significantly reduced in areas receiving the highest levels of organic enrichment. It is likely that low levels of oxygen, characteristic of extremely enriched environments, prevented them from persisting in these areas due to relatively small body surface to volume ratios. The other species of *Scoletoma*, *Scoletomaumberineris tetraura* (nr. 27 -  $n = 131$ ) was generally smaller and this size difference might explain why its abundance was not significantly reduced in increased organic enrichment.
- *Species of limited mobility and without branchial structures* among the species, such as *Nemertea* indet. (6–291), *Pholoe minuta* (nr. 11 -  $n = 159$ ), and *Cossura longocirrata* (nr. 42 -  $n = 966$ ), were both affected by disturbance and enrichment.
- *Species with branchial filaments*, such as *Chaetozone setosa* (nr. 40 -  $n = 3121$ ) and *Diplocirrus glaucus* (nr. 43 -  $n = 265$ ) were also affected by both factors, but also showed an interaction which was characterised by higher abundances than expected on the basis of the main effects model, especially for the higher levels of the factors. Possibly the branchial filaments allowed them to get more oxygen even in very enriched circumstances if combined with the increase of oxygen due to the higher levels of disturbance.

The analyses in this paper supplied further insight about the effects observed by [Widdicombe and Austen \(2001\)](#) by examining in more detail the relationships between the three ways. It was shown how to

combine the effect of both factors on the abundance of the species in a single framework. It would require further investigations beyond the scope of this paper to go into the biological details of each species. Especially the location in the interactive biplot of less abundant species with respect to the more abundant species could be used to explain their sensitivity with respect to disturbance and enrichment and their interaction.

#### CRedit authorship contribution statement

**Pieter M. Kroonenberg:** Formal analysis, Methodology, Software. **Stephen Widdicombe:** Conceptualization, Funding acquisition, Methodology, Writing – original draft. **Paul Somerfield:** Conceptualization, Writing – original draft.

#### Declaration of competing interest

None.

#### Data availability

Data will be made available on request.

#### Acknowledgements

This study was funded in part by the EU P5 project COST-IMPACT and by UK DEFRA (project no. AE1137). The authors would like to thank the following for their contributions and advice during the course of this study: Dr. K. R. Clarke (experimental design and statistical analysis of the original mesocosm study), Oddbjørn Pettersen, Per Ivar Johhannsen and Sigurd Øksnevad (technical support at the Solbergstrand Marine Station); and the crew of RVs 'Trygve Braarud' and 'Bjørn Føyn' (collection of sediment and bioturbators).

Thanks also go to Charles Franssen, Naturalis, Leiden, The Netherlands for his comments and assistance with some ecological issues. The reviewers deserve a special thanks for their meticulous reading of the text, and for their many suggestions for improvement and correction. Finally, we would like to pay tribute to our colleague and co-author Professor Paul Somerfield who sadly passed away during the final stages of this paper.

#### Appendix A. Oslofjord species

Id	Species	Abundance	Family	Ecological explanation
04	<i>Golfingia</i> indet	47	Golfingiidae	
06	<i>Nemertea</i> indet.	291	n/a	limited mobility without branchial structure
07	<i>Priapulius caudatus</i>	70	Priapulidae	close to 3, 13, 35, 72
09	<i>Phascolion strombus</i>	16	Golfingiidae	close to 54, 56
11	<i>Pholoe minuta</i>	159	Sigalionidae	limited mobility without branchial structure
12	<i>Eteone flava</i>	88	Phyllococidae	
13	<i>Glyceria lapidum</i>	53	Glyceridae	close to 7, 14, 35, 72
14	<i>Goniada maculata</i>	362	Goniadidae	large mobile; close to 7, 13, 35, 72
18	cf. <i>Syllis cornuta</i>	166	Syllidae	; close to 75
20	cf. <i>Syllis</i> sp. B	61	Syllidae	
24	<i>Nephtys ciliata</i>	26	Nephtydidae	caea; close to 26, 60
26	<i>Scoletoma fragilis</i>	130	Lumbrineridae	large mobile; close 24, 60
27	<i>Scoletoma tetraura</i>	131	Lumbrineridae	smaller mobile
28	<i>Parougia caeca</i>	147	Dorvilleidae	
30	<i>Scoloplos armiger</i>	14	Orbiniidae	
32	<i>Paraonis fulgens</i>	1407	Paraonidae	
33	<i>Prionospio cirrifera</i>	20	Spionidae	
34	<i>Prionospio fallax</i>	76	Spionidae	
35	<i>Prionospio steenstrupi</i>	43	Spionidae	close to 7, 13, 14, 72
36	<i>Pseudopolydora paucibranchiata</i>	492	Spionidae	tube building
40	<i>Chaetozone setosa</i>	3121	Cirratulidae	branchial filaments

(continued on next page)

(continued)

Id	Species	Abundance	Family	Ecological explanation
41	Tharyx sp.	67	Cirratulidae	
42	<i>Cossura longocirrata</i>	966	Cossuridae	limited mobility, no branchial structure
43	<i>Diplocirrus glaucus</i>	265	Flabelligeridae	branchial filaments
45	Capitella sp.	33	Capitellidae	
47	<i>Heteromastus filiformis</i>	22,608	Capitellidae	
49	<i>Maldane sarsi</i>	16	Maldanidae	
50	<i>Ophelina modesta</i>	26	Ophellidae	
51	<i>Polyphysia crassa</i>	10	Scalibregmatidae	
52	<i>Galathowenia oculata</i>	15	Oweniidae	
54	<i>Melinna cristata</i>	19	Melinnidae	close to 9, 56
55	<i>Ampharete baltica</i>	67	Ampharetidae	close to 29,71
56	<i>Anobothrus gracilis</i>	140	Ampharetidae	tube building; close to 9,54
60	<i>Polycirrus medusa</i>	34	Terebellidae	(hair?); close to 24, 26
62	Tubificoides indet.	87	Naididae	
71	<i>Ennucula tenuis</i>	1105	Nuculidae	bivalve; close to 29, 55
72	<i>Thyasira flexuosa</i>	59	Thyasiridae/	close to 7,13, 14, 35
75	<i>Abra nitida</i>	20	Semelidae	close to 18
78	Nudibranch indet. A.	10	n/a	

Note: All species in green occur >88 times in the experiment (Widdicombe and Austen, 2001); those in yellow 10 or more times. Not included are species with <10 occurrences; all class/order are polychaeta, except those indicated by a second name.

## References

- Abdi, H., Valentin, D., 2007. Multiple correspondence analysis. In: Salkind, N.J. (Ed.), *Encyclopedia of Measurement and Statistics*. SAGE Publications, Thousand Oaks, CA, pp. 1–13.
- Beh, E.J., Lombardo, R., 2014. Correspondence Analysis. Theory, Practice, and New Strategies. Wiley, Chichester, UK. <https://doi.org/10.1002/9781118762875>.
- Beh, E.J., Lombardo, R., Albert, G., 2018. Correspondence analysis and the Freeman-Tukey statistic: a study of archaeological data. *Comput. Stat. Data Anal.* 128, 73–86. <https://doi.org/10.1016/j.csda.2018.06.012>.
- Carlier, A., Kroonenberg, P.M., 1996. Decompositions and biplots in three-way correspondence analysis. *Psychometrika* 61, 355–373. <https://doi.org/10.1007/BF02294344>.
- Carlier, A., Kroonenberg, P.M., 1998. The case of the French cantons: An application of three-way correspondence analysis. In: Blasius, J., Greenacre, M.J. (Eds.), *Visualization of Categorical Data*. Academic Press, pp. 253–275. <https://doi.org/10.1016/B978-012299045-8/50021-8>.
- Choulakian, V., 1988. Analyse factorielle des correspondances de tableaux multiples [Factorial correspondence analysis of multiway tables]. *Rev. Stat. Appl.* 36, 33–41.
- Clarke, K.R., 1993. Non-parametric multivariate analyses of changes in community structure. *Aust. Ecol.* 18, 117–143. <https://doi.org/10.1111/j.1442-9993.1993.tb00438.x>.
- Darroch, J.N., 1974. Multiplicative and additive interaction in contingency tables. *Biometrika* 61, 207–214. <https://doi.org/10.1093/biomet/61.1.207>.
- Dequier, A., 1974. Notes sur les tables de contingence entre trois caractères. [Notes on three-way contingency tables] (Unpublished Document, Nr. 1033). C.I.R.O., Arcueil, France.
- Field, J.G., Clarke, K.R., Warwick, R.M., 1982. A practical strategy for analysing multispecies distribution patterns. *Mar. Ecol. Prog. Ser.* 8, 37–52. <https://doi.org/10.3354/meps008037>.
- Goria, M.N., 1992. On the fourth root transformation of chi-square. *Aust. J. Stat.* 34, 55–64. <https://doi.org/10.1111/j.1467-842X.1992.tb01043.x>.
- Greenacre, M.J., 1993. Biplots in correspondence analysis. *J. App. Stat.* 20, 251–269. <https://doi.org/10.1080/02664769300000021>.
- Greenacre, M.J., 2013. The contributions of rare objects in correspondence analysis. *Ecology* 94, 241–249. <https://doi.org/10.1890/11-1730.1>.
- Kroonenberg, P.M., 1983. Three-mode principal component analysis: Theory and applications. DSWO Press. Errata, 1989. <https://three-mode.leidenuniv.nl>.
- Kroonenberg, P.M., 1989. Singular value decompositions of interactions in three-way contingency tables. In: Coppi, R., Bolasco, S. (Eds.), *Multiway Data Analysis*. Elsevier, pp. 169–184 (ISBN: 0 444 87410 0).
- Kroonenberg, P.M., 2008. Applied Multiway Data Analysis. Wiley. <https://doi.org/10.1002/9780470238004>. ISBN: 978-0-470-16497-6.
- Kroonenberg, P.M., Anderson, C.J., 2006. Additive and multiplicative models for three-way contingency tables: Darroch (1974) Revisited (pp. 455–486). In: Greenacre, M. J., Blasius, J. (Eds.), *Multiple Correspondence Analysis and Related Methods*. Chapman and Hall/CRC. ISBN: 978-1-58488-628-0.
- Lancaster, H., 1951. Complex contingency tables treated by the partition of the chi-square. *J. R. Stat. Soc. Ser. B* 13, 242–249. <https://doi.org/10.1111/j.2517-6161.1951.tb00089.x>.
- Legendre, P., Legendre, L., 1998. *Numerical Ecology* (2nd English Ed.). Elsevier. ISBN: 978-0-444-53868-0.
- Lombardo, R., Beh, E.J., Kroonenberg, P.M., 2016. Modelling trends in ordered correspondence analysis using orthogonal polynomials. *Psychometrika* 81, 325–349. <https://doi.org/10.1007/s11336-015-9448-y>.
- Lombardo, R., Beh, E.J., Kroonenberg, P.M., 2021. Symmetrical and non-symmetrical variants of three-way correspondence analysis for ordered variables. *Stat. Sci.* 36, 542–561. <https://doi.org/10.1214/20-STS814>.
- Lombardo, R., Van de Velden, M., Beh, E.J., 2023. Three-way correspondence analysis in R. *R J.* 15, 237–262. <https://doi.org/10.32614/RJ-2023-049>.
- Ter Braak, C.J.F., 1985. Correspondence analysis of incidence and abundance data: properties in terms of a unimodal response model. *Biometrics* 41, 859–873. <https://doi.org/10.2307/2530959>.
- Tucker, L.R., 1966. Some mathematical notes on three-mode factor analysis. *Psychometrika* 31, 279–311. <https://doi.org/10.1007/BF02289464>.
- Widdicombe, S., Austen, M.C., 2001. The interaction between physical disturbance and organic enrichment: an important element in structuring benthic communities. *Limnol. Oceanogr.* 46, 1720–1733. <https://doi.org/10.4319/lo.2001.46.7.1720>.
- Widdicombe, S., Austen, M.C., Kendall, M.A., Olsford, F., Schaanning, M.T., Dashfield, S. L., Needham, H.R., 2004. Importance of bioturbators for biodiversity maintenance: indirect effects of fishing disturbance. *Mar. Ecol. Prog. Ser.* 275, 1–10. <https://doi.org/10.3354/meps275001>.