Hazard/Risk Assessment

ANALYSIS OF COMMUNITY-LEVEL MESOCOSM DATA BASED ON ECOLOGICALLY MEANINGFUL DISSIMILARITY MEASURES AND DATA TRANSFORMATION

CLEO TEBBY,*† SANDRINE JOACHIM,‡ PAUL J. VAN DEN BRINK,§‖ JEAN-MARC PORCHER,‡ and RÉMY BEAUDOUIN†
†Models for Ecotoxicology and Toxicology Unit, INERIS, Verneuil-en-Halatte, France
‡Ecotoxicology Unit, INERIS, Verneuil-en-Halatte, France
§Aquatic Ecology and Water Quality Management Group, Wageningen University, Wageningen, The Netherlands
‖Wageningen Environmental Research (Alterra), Wageningen, The Netherlands

Abstract: The principal response curve (PRC) method is a constrained ordination method developed specifically for the analysis of community data collected in mesocosm experiments, which provides easily understood summaries and graphical representations of community response to stress. It is a redundancy analysis method and is usually performed on log-transformed abundance data. The choice of a measure of dissimilarity between samples and the choice of the data transformation significantly affect the results of multivariate analysis. Dissimilarity measures that are more ecologically meaningful than the Euclidean distance can be incorporated into the PRC using distance-based redundancy analysis. The present study investigates the ordinations produced by a small selection of dissimilarity measures: the Euclidean distance using log-transformed and Hellinger-transformed data and the Bray-Curtis dissimilarity using raw and log-transformed data. It compares 2 data sets from experiments on the effect of the anti-inflammatory drug diclofenac and the insecticide chlorpyrifos on macroinvertebrate communities. The choice of dissimilarity measure can determine the outcome of a risk assessment. For the diclofenac data set, the PRCs were different depending on the dissimilarity measure: the community no-effect concentration was lowest for the Bray-Curtis on log-transformed data and Hellinger dissimilarity measures. For chlorpyrifos, however, the PRCs were similar for all dissimilarity measures. *Environ Toxicol Chem* 2017;36:1667–1679. © 2016 SETAC

INTRODUCTION

Mesocosm community data are usually analyzed using multivariate techniques, in particular ordination or gradient techniques, which provide useful summaries or visualization of the main information in the data [1–3]. This information can be constrained to represent treatment, time, and/or spatial variation.

The principal response curve (PRC) method [4] is a constrained ordination method developed specifically for the analysis of community data collected in microcosm and mesocosm experiments. It is designed to study the effects of chemicals (or other stressors) by showing the effect of different treatments on the community structure over time against an untreated control, resulting in an easily interpretable graphical representation.

The PRC is a partial redundancy analysis (RDA) method. The interaction between treatment and time is the constraining variable (explanatory variable), whereas time is a conditioning variable (covariable), so the principal effect of time is partialled out (removed) before analyzing the overall effect of the treatment × time interaction [4]. With PRC, the treatment is considered to be a categorical variable; no dose–response model is assumed.

The dissimilarity between samples can be calculated on the basis of community-level abundance counts of species using many different formulae. Because many dissimilarity measures have built-in transformations, the choice of a transformation depends closely on the dissimilarity measure. Whichever ordination method is used, the measure of similarity or dissimilarity between samples and the data transformation or scaling have a strong effect on the results of the analysis [5–8]. The choice of a dissimilarity measure depends on the aims of the research; the context of the descriptive study of which it is a part, including the type of data; and the analysis method to be used [9]. For a particular data set with a given objective, it is generally advisable to test several dissimilarity measures and data transformations to test a series of hypotheses that have been designed a priori, as advocated by Anderson et al. [10] regarding β diversity. Designing research hypotheses that are specific to one's data set may help to determine whether to include joint absence (or double zeros) of a taxon from 2 samples, abundance data, or simply presence–absence data [10], as well as the degree to which the similarity between samples is based only on the most abundant species or includes the less abundant species [6,11].

The question of joint absences is a critical aspect [2,9,12]. When a species is absent from 2 sites, does this mean that there is the same unfavorable condition at both sites? Or does this imply nothing about the similarity between the 2 sites? When a large number of low-abundance taxa are included in the samples resulting in sparse data, this question is all the more important because there is a higher frequency of double zeros [10]. As a result, very rare species are often removed from data sets prior to multivariate analysis [13,14] even though they might contribute significantly to the conclusions of the study [15]. This problem can be avoided by careful selection of the data transformation

and the type of data analysis. Numerous dissimilarity measures have been developed specifically for ecological abundance data to exclude joint absences [12] and overcome limitations intrinsic to the Euclidean distance used in RDA. These measures may, therefore, be more appropriate if adverse effects translate into a reduction in abundance rather than complete elimination of sensitive taxa. The Bray-Curtis dissimilarity [16], also known as the percentage difference, is commonly used to evaluate abundance data [17]. This measure is a form of standardized absolute deviation [6] and does not depend on the largest differences in species abundance between 2 samples [12] to the same extent as the Euclidean distance. The Hellinger transformation is worth considering because, as well as excluding double zeros [18], it can downweight rare species [19]. However, many ordination methods are not particularly sensitive to rare species simply because of data redundancy [20], rather than because of the Euclidean distance.

Several publications have compared dissimilarity measures using real or simulated data. Faith et al. [21] and Legendre and Gallagher [19] studied how well ecological patterns were retrieved from simulated data. Legendre and Gallagher [19] compared 6 dissimilarity measures, including the Hellinger distance obtained by transforming the data and computing the Euclidean distances between samples. Faith et al. [21] compared 10 different dissimilarity measures with 3 possible data scaling methods, in particular for robustness to differences in models of the community variation in the ecological space. Their analysis showed that the Bray-Curtis dissimilarity measure was one of the most effective measures, whereas the Euclidean distance was one of the least able to retrieve ecological distances. However, these studies provide general guidance based on simulated data to provide conclusions that are as general as possible, and the choice of measure for a new data set should be based on the data itself [22].

The PRC method is based on the Euclidean distance, which unfortunately is not recommended for the analysis of ecological data, in particular raw abundance data [12,23]. Abundance data are usually log-transformed prior to the PRC for 2 reasons. First, the logarithm transforms the multiplicative model for the counts into a linear model [4], which is in line with the hypothesis that, for a given magnitude of effect, the change in abundance is proportional to the total abundance and reduces the dominating influence of abundant taxa [5]. Secondly, log transformation is also intended to normalize data over replicates as, given the small number of replicates in mesocosm experiments (often between 2 and 4), checking normality using a statistical test would not be appropriate. If the data for each species at each sampling date were log-normally distributed and the standard deviation was proportional to the mean (quadratic mean–variance relationship) on individual sampling dates, log transformation would reduce the data for each species on each sampling date to equal variance and a normal distribution. The resulting transformed data would be particularly suited to the Euclidean distance [24].

In practice, the typical mean–variance relationship observed in abundance data is close to a quadratic relationship; but the proportionality between the standard deviation and the mean does not quite hold [24], largely because of the large proportion of null counts. Furthermore, when applied to count data with many zeros, the transformation is not strictly logarithmic because 1 is added to the data set prior to log transformation. This tends to distort the data set and may, therefore, counteract part of the beneficial effects of the log transformation by highlighting species with larger abundances slightly more than

desired. It has been found that parametric methods on log-transformed data do not perform as well as those obtained with specific error distributions such as Poisson or binomial with the raw data [25]. However, the PRC does not suffer from such limitations as it is an exploratory method with associated permutation-based (i.e., nonparametric) significance tests [26]. Despite these limitations, however, ter Braak and Smilauer [27] found that log transformation stabilizes the variance.

The widespread use of similarity analyses prompted the creators of the PRC to compare their method with a dissimilarity-based technique [8] for which 2 dissimilarity measures (including the Bray-Curtis dissimilarity on raw data) were used to produce a summary statistic of the change in community structure. The lack of interpretability down to the taxon level and the reduction down to a single dimension were major drawbacks of the dissimilarity-based method.

Although the PRC method is based on the Euclidean distance, it can accommodate any data transformation; but, so far as we are aware, it has mainly been used for log-transformed data. As suggested by Szöcs et al. [2], the underlying RDA in the PRC can also be based on metrics rather than the Euclidean distance by using distance-based redundancy analysis (db-RDA) [28] or canonical analysis of principal coordinates [29]. When non-Euclidean metric distances or nonmetric dissimilarities are computed between samples, the 2-dimensional graphs of the data cannot represent the actual distances because the graphs are based on a Euclidean space. Conventional multidimensional scaling (MDS), also known as principal coordinate analysis, is used in db-RDA to convert the dissimilarities into Euclidean distances while limiting the loss of information.

The present study illustrates how the properties of ecologically meaningful dissimilarity measures and data transformations affect multivariate analysis results. It studies the extent to which ecological risk assessments based on multivariate analysis results, in particular those obtained using the PRC method, can be affected by the choice of dissimilarity measure. The novelty of the present study lies in the use of ecologically meaningful dissimilarity measures and data transformations in PRC. It focuses on dissimilarity measures, which include information about abundance rather than only presence–absence data. It presents the results for a small selection of ecologically meaningful dissimilarity measures: the Euclidean distance on log-transformed data, which is normally used in PRC; the Bray-Curtis dissimilarity on raw and log-transformed data, which is often used in the field of ecology; and the Hellinger distance. Many ecological studies involve field data collection. An important characteristic of mesocosm semifield experiments is that the sites are under identical environmental conditions; therefore, the relevance of these dissimilarity measures must be reexamined to select the most appropriate dissimilarity measures. The value of this approach is illustrated using 2 sets of community-level macroinvertebrate data generated in a mesocosm setting. Several other dissimilarity measures were tested, including the quantitative version of the Jaccard coefficient, the Kulczynski dissimilarity [21], the Euclidean distance applied to raw data, raw data scaled to unit variance, and log-transformed data scaled to unit variance. Those results are not discussed because of their similarity to the measures selected.

MATERIALS AND METHODS

Two macroinvertebrate community data sets were used. The first tested the effects of a continuous exposure to diclofenac, and the second was the well-known single-exposure chlorpyrifos data set, which has been used as an example in several

statistical packages and was used in the article introducing the PRC [4].

Diclofenac data set

Study design. The data were obtained using an outdoor facility composed of 12 lotic mesocosms, each measuring 20 m by 1 m and divided into a shallow upper section (1–9 m from the inlet, 0.25 m deep) and a lower section (10–20 m from the inlet, 0.7 m deep). The experimental setup was detailed in de Kermoysan et al. [30] for a previous experiment. Every year, the mesocosm substrate (sand and clay) is renewed and the communities are reconstructed by introducing macrophytes, zooplankton, macroinvertebrates, and fish. The macroinvertebrates introduced were *Gammarus pulex*, *Asellus aquaticus*, *Potamopyrgus antipodarum*, *Planorbis*, *Radix balthica*, *Notonecta* sp., *Glossiphonia complananta*, and *Erpobdella octoculata*.

Three concentrations of diclofenac (0.1 μg/L, 1 μg/L, and 10 μg/L) and controls were tested, with 3 mesocosms for each concentration. The exposure was continuous from 15 April 2013 until the end of the summer (150 d). The concentrations of diclofenac in each of the 9 mesocosms were monitored every month at 0 m, 5 m, and 19 m from the water inlet to check the actual exposure concentrations.

The mean effective concentrations were calculated for each treatment using the mean values of the 3 replicates at 5 m and 19 m. The mean concentrations in the mesocosms over the whole experiment were lower than the nominal concentrations: 0.05 ± 0.01 μg/L for the 0.1 μg/L concentration, 0.45 ± 0.02 μg/L for the 1 μg/L concentration, and 4.10 ± 0.41 μg/L for the 10 μg/L concentration.

Sampling methodology. Macroinvertebrates were sampled approximately every 4 wk from the end of February until mid-September (8 sampling dates in total) using different types of artificial substrate (tubes and tiles). Each tube substrate comprised 7 tubes strapped together (2 cm wide and 20 cm long). The terra-cotta tiles were 11 cm long, 16 cm wide, and 1.5 cm thick. Ten tubes and 10 tiles were placed horizontally at the bottom of each mesocosm. In the upper section, the substrates were arranged, in series, starting with tubes (tubes 0.5 m, 2 m, 4 m, 6 m, 8 m from the water inlet; tiles 1 m, 3 m, 5 m, 7 m, and 9 m from the water inlet). The same pattern was used in the lower section (tubes: 10 m, 12 m, 14 m, 16 m, and 18 m from the water inlet; tiles: 11 m, 13 m, 15 m, 17 m, and 19 m from the water inlet).

On each sampling date, the artificial substrates were retrieved from each mesocosm using a landing net to prevent any loss of organisms. Each substrate was then washed and scrubbed clean in a container to remove the invertebrates trapped inside. The rinsing water was then passed through a sieve with a mesh size of 50 μm and placed in plastic dish half-filled with water.

Each macroinvertebrate taxon was identified and enumerated immediately before being returned to each mesocosm. The tiles and tubes were also replaced in each mesocosm. Macroinvertebrates were identified to the lowest practical taxonomic level (genus, except for Chironomidae, which were identified down to the family level) based on Tachet and Richoux [31].

In addition to the use of artificial substrates, macroinvertebrates on the walls of each mesocosm were sampled using a landing net in each mesocosm at 1 m, 7 m, 13 m, and 19 m from the water inlet. The organisms were scraped off the walls and placed in a plastic dish half-filled with water. The macroinvertebrates were identified and enumerated as described

in the previous paragraph. All data for each mesocosm were pooled for further analysis.

Chlorpyrifos data set

The same methodology was applied to the chlorpyrifos macroinvertebrate data set, which contains log-transformed, $\log[10 \times (x+1)]$, abundances of 121 macroinvertebrate taxa [32]. Twelve mesocosms were sampled on 11 occasions from week –4 pretreatment to week 24 posttreatment. Four mesocosms served as controls, and the remaining 8 were treated once on 8 May 1990 with the insecticide chlorpyrifos, 2 mesocosms each receiving a nominal dose of 0.1 μg/L, 0.9 μg/L, 6 μg/L, and 44 μg/L. This data set is part of the combined zooplankton and macroinvertebrate data set used by Van den Brink and ter Braak [4] in their seminal paper on principal response curves. One species (*Chaoborus obscuripes*) was removed because of differences between the raw data and the transformed data provided in the "vegan" R package [33]. The combined data set is also used as an example in CANOCO 4 and in the vegan R package.

Dissimilarity measures and transformations tested

The present study focused on the Euclidean distance using log-transformed data, the Euclidean distance using Hellinger-transformed data (also known as the Hellinger distance), and the Bray-Curtis dissimilarity using raw and log-transformed data. Both the Bray-Curtis and the Hellinger dissimilarities exclude double zeros: 2 sites are not more similar if both lack certain species, a common hypothesis in ecology. The Euclidean distance, on the other hand, takes account of double zeros; this can be more meaningful in the context of species disappearance. Mesocosms are an experimental setting where environmental factors are identical between replicates and between treatments at any given time point. The relevance of the disappearance of a species and, hence, whether to include double zeros depend on how well established the communities are as well as on the abundance of this species. Also, different experimental setups can vary in how they are affected by random variations as a result of species colonizing the mesocosms from external sources.

In the following equations, $x_{i,j}$ is the vector of abundances (raw or transformed) for species i at site j, and $d_{j,k}$ represents the dissimilarity between sites j and k.

The Euclidean distance is the "ordinary distance" between 2 points in a Euclidean space

$$d_{j,k} = \sqrt{\sum_i \left(x_{i,j} - x_{i,k}\right)^2} \qquad (1)$$

The Hellinger distance [34] is equivalent to the Euclidean distance using Hellinger-transformed data and is frequently used in the analysis of ecological abundance data [19]. The Hellinger distance can be calculated by substituting $x_{i,j}$ by $x'_{i,j}$ in Equation 1

$$x'_{i,j} = \sqrt{\frac{x_{i,j}}{\sum_i x_{i,j}}} \qquad (2)$$

The Hellinger distance is intended to be used for 2-way contingency tables such as species × sites data matrices; it is not intended to be calculated using log-transformed data. The

scaling to equal sample sums used in the Hellinger distance excludes double zeros [18]; this measure does not consider the joint absence of a taxon from 2 samples as a criterion of resemblance. Furthermore, if the total abundance increases with time, the Hellinger transformation removes this trend. Information about the increase in total abundances in a particular treatment is also lost; the analysis considers only the abundance profile—that is, the relative abundance of each species in each sample.

The Bray-Curtis dissimilarity [16] sums the absolute differences between raw abundances in 2 samples and relates this sum to the sum of the abundances in both samples

$$d_{j,k} = \left(\frac{\sum_i \left| x_{i,j} - x_{i,k} \right|}{\sum_i \left(x_{i,j} + x_{i,k} \right)} \right) \tag{3}$$

The Bray-Curtis dissimilarity is nonmetric. It does not satisfy the triangle inequality, and therefore, the dissimilarity matrix cannot be projected into a Euclidean space without being transformed. By definition, the Bray-Curtis dissimilarity measure excludes double zeros (Equation 3).

Some authors recommend using a $\log(x + 1)$ or a square, cubic, or other root transformation to count data prior to using the Bray-Curtis dissimilarity [11]. This gives less weight to the dominant species and thus more weight to the qualitative aspects (i.e., presence–absence) of the data [35]. Other authors argue that Bray-Curtis is well suited to raw abundance data [12]. This measure does not emphasize the largest differences in species abundance between 2 samples to the same extent as the Euclidean distance. Furthermore, absolute differences in abundance for less abundant species contribute as much as absolute differences for the more abundant species [12]. The Bray-Curtis dissimilarity was therefore tested using both raw and log-transformed data.

Exploratory analysis: Unconstrained ordination

As a first step toward a distance-based PRC, the data sets were analyzed using unconstrained ordination methods to examine the data set properties, pairwise dissimilarities, and impact of logarithmic and Hellinger transformations on the overall pattern of dispersion of points [29].

As usually recommended for PRC [36,37], the data were transformed with a $\log(Ax + 1)$ transformation, such that $Ax = 2$ when x is the smallest non-null abundance. Therefore, A was set to 2 in the diclofenac data set and 10 in the chlorpyrifos data set.

A first unconstrained ordination was obtained with metric MDS on the Bray-Curtis dissimilarity matrix and with principle component analysis on the log-transformed and Hellinger-transformed data sets (this is equivalent to MDS on the Euclidean distance matrices). The non-Euclidean dissimilarity matrix input to the MDS was adjusted to Euclidean space by adding a constant to the whole matrix [28] using the method described in Gower and Legendre [9]. This procedure modifies the total inertia and thus the test statistics but does not affect the p values obtained by permutation in the case of 1-way analysis of variance with either Euclidean distances or Bray-Curtis dissimilarities [38].

A second unconstrained ordination was obtained by parceling out the time variation, which is the first step of the partial RDA underlying the PRC. When the Euclidean distance was used, the mean transformed abundances for each sampling

date for each species were subtracted from the transformed data for analysis. When the Bray-Curtis dissimilarity was used, the mean MDS coordinates at each sampling date for each species were subtracted from the MDS coordinates.

The resemblance between dissimilarity measures was quantified using the RV coefficient [39], a multivariate version of Pearson's correlation coefficient, which can be used to compare 2 dissimilarity matrices. The RV coefficient between 2 column-wise centered matrices X and Y is defined by the following equation, where $COVV(X,Y)$ is the scalar-valued covariance between X and Y, $VAV(X)$ is the scalar-valued variance of X, and $VAV(Y)$ is the scalar valued variance of Y

$$RV_{X,Y} = \frac{COVV(X,Y)}{\sqrt{VAV(X)VAV(Y)}} \tag{4}$$

Constrained ordination: Treatment and time effect

A PRC was performed using the Euclidean distance on log-transformed and Hellinger-transformed data sets. In the analysis with the Bray-Curtis dissimilarity index, a distance-based PRC was obtained simply by replacing the underlying partial RDA by distance-based partial RDA. This was implemented with the function dbrda() in the vegan package, using direct decomposition. One key aspect of the PRC is that it allows interpretation down to the species level [8]. Unfortunately, when the PRC is based on a dissimilarity matrix rather than on species abundance, the information about species contributions to the axes is lost; the species scores cannot be represented on the diagram, but the significance of the PRC axes and of the underlying RDA can be assessed. The points representing the replicates (individual mesocosms) were added to the PRC diagram by centering the species-based sample scores; the means of these scores are the mean points usually represented on the PRC diagram.

Significance testing

The significance of the RDA in general and of each PRC axis was assessed by performing 9999 permutations of time series residuals from the reduced model [40]. A large number of permutations was preferred, as the aim was to compare the p values rather than simply assess significance at a 5% threshold.

The final result expected from the analysis of mesocosm data is often a no-observed-effect concentration (NOEC). Community-level NOECs were calculated in 2 steps. First, the significance of the effect of treatment on the partial RDA axes for each sampling date was assessed with 9999 permutations of the residuals on all axes. Second, for significant dates, the first principal component analysis or MDS component was calculated, and a 2-sided Dunnett's test was used to compare the mean of each treatment with the control and determine an NOEC for each date. An overall community NOEC was then derived when the same NOEC was calculated for 2 consecutive dates.

All calculations were carried out with R 3.2.3 [41], vegan package 2.4-1 [33], FactoMineR package for ordination plots and RV coefficients [42], and multcomp for Dunnett's test [43]. The R code for the chlorpyrifos data set analysis is available in the Supplemental Data.

RESULTS

The ordination graphs for the Bray-Curtis dissimilarity using log-transformed data were very similar to those obtained with

the Euclidean distance using log-transformed data. They are, therefore, given only in the Supplemental Data.

Brief description of the data sets

For the diclofenac data set, 21 taxa were collected throughout the experiment (Supplemental Data, Table S1) belonging to Annelida, Gastropoda, Ephemeroptera, Crustacea, Turbellaria, Diptera, and other insects. The abundances for each dose are shown on a log scale in Supplemental Data, Figure S1, and with the Hellinger transformation in Supplemental Data, Figure S2; the most abundant taxa are mostly the same for each treatment. The mean–variance relationship for raw data was near quadratic, but with log-transformed data and Hellinger-transformed data the variance did not vary as strongly with the mean (Supplemental Data, Figure S3).

The abundances for each dose are shown with the logarithmic transformation in Supplemental Data, Figure S4, and with the Hellinger transformation in Supplemental Data, Figure S5. As with the diclofenac data set, the mean–variance relationship for raw data was near quadratic, whereas with log-transformed data and Hellinger-transformed data the variance did not vary as strongly with the mean (Supplemental Data, Figure S6).

A comparison of the total abundances in each mesocosm throughout the experiment provides additional insight into the differences between the 2 data sets (Figure 1). In the diclofenac data set, the mean abundance for each dose increased 10-fold in the first months of the experiment. The mean abundances were roughly equal for each dose on each sampling date except during the last 3 sampling mo, when it was 10-fold higher for the highest dose than for the other doses. For both data sets, an analysis of variance of the effect of dose, time, and their interaction on the logarithm of the total abundances in each mesocosm was performed. This confirmed that, in the chlorpyrifos data set, the mean abundances for each dose were more even throughout the experiment and that the time-dependent treatment effects were slightly smaller. The residual variability between mesocosms for each dose was also slightly lower in the chlorpyrifos data set.

Unconstrained ordination

With the diclofenac data set, the unconstrained ordination based on all dissimilarity measures, and in particular on the Bray-Curtis dissimilarity on raw data, clearly showed an artifact called a "horseshoe effect" owing to the strong time-related gradient (Supplemental Data, Figure S7); the samples from the beginning of the experiment were drastically different from those toward the end of experiment. This occurred even though the lengths of gradients calculated with detrended correspondence analysis were not very large (1.91 standard deviation [SD] for the first axis with raw data, 1.66 SD with log-transformed data, and 1.74 SD with Hellinger-transformed data). The horseshoe was more pronounced with the Bray-Curtis dissimilarity on raw data because of the increase in total abundance throughout the experiment (Supplemental Data, Figures S7 and S8). With the Bray-Curtis dissimilarity on raw data, 62% of the total inertia was attributed to time (Supplemental Data, Table S2). With the Hellinger-transformed data, there was hardly any variation in the sample sums; but a horseshoe can still be observed, showing that there was a marked change in relative abundance of species throughout the experiment, and 67% of the total inertia was attributed to time.

With the chlorpyrifos data set, the time gradient was not as strong; although it was visible as a main axis of variability with the log-transformed data (with the same magnitude as the
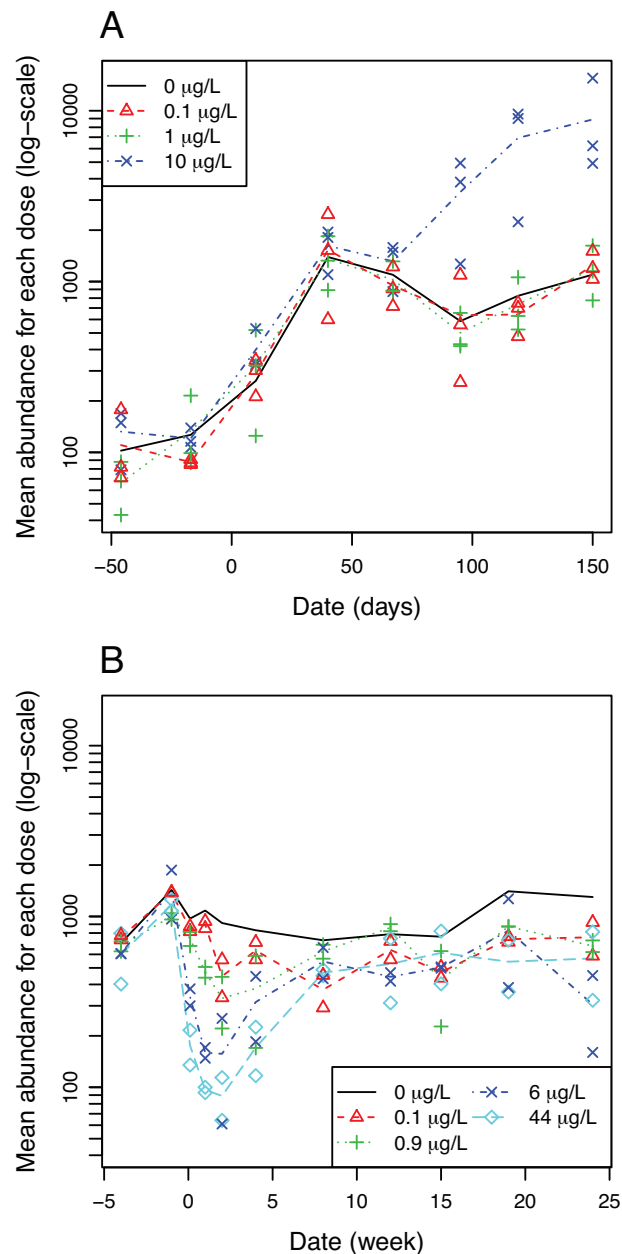

Figure 1. Mean total abundance in each treatment level in (**A**) the diclofenac data set and (**B**) the chlorpyrifos data set.

dose-related gradient), it was not sufficient to form a horseshoe (Supplemental Data, Figure S9). In the log-transformed data, 21% of the total variance could be attributed to time. With the Bray-Curtis dissimilarity on raw data, 25% of the total inertia was attributed to time (Supplemental Data, Table S3).

When the time variation had been partialled out, the unconstrained ordination revealed differences in taxa abundances between doses on the first axes of the diagrams, although the differences were only slight with the Hellinger distance (Figure 2B).

For the chlorpyrifos data set, the shift in communities with increasing doses was itself sufficiently large to produce a very slight horseshoe with the Bray-Curtis dissimilarity (Figure 3). This suggests that the change in community structure is gradual from one dose to another and that the control and highest-dose communities are considerably different. All dissimilarity measures appeared to discriminate between doses in these
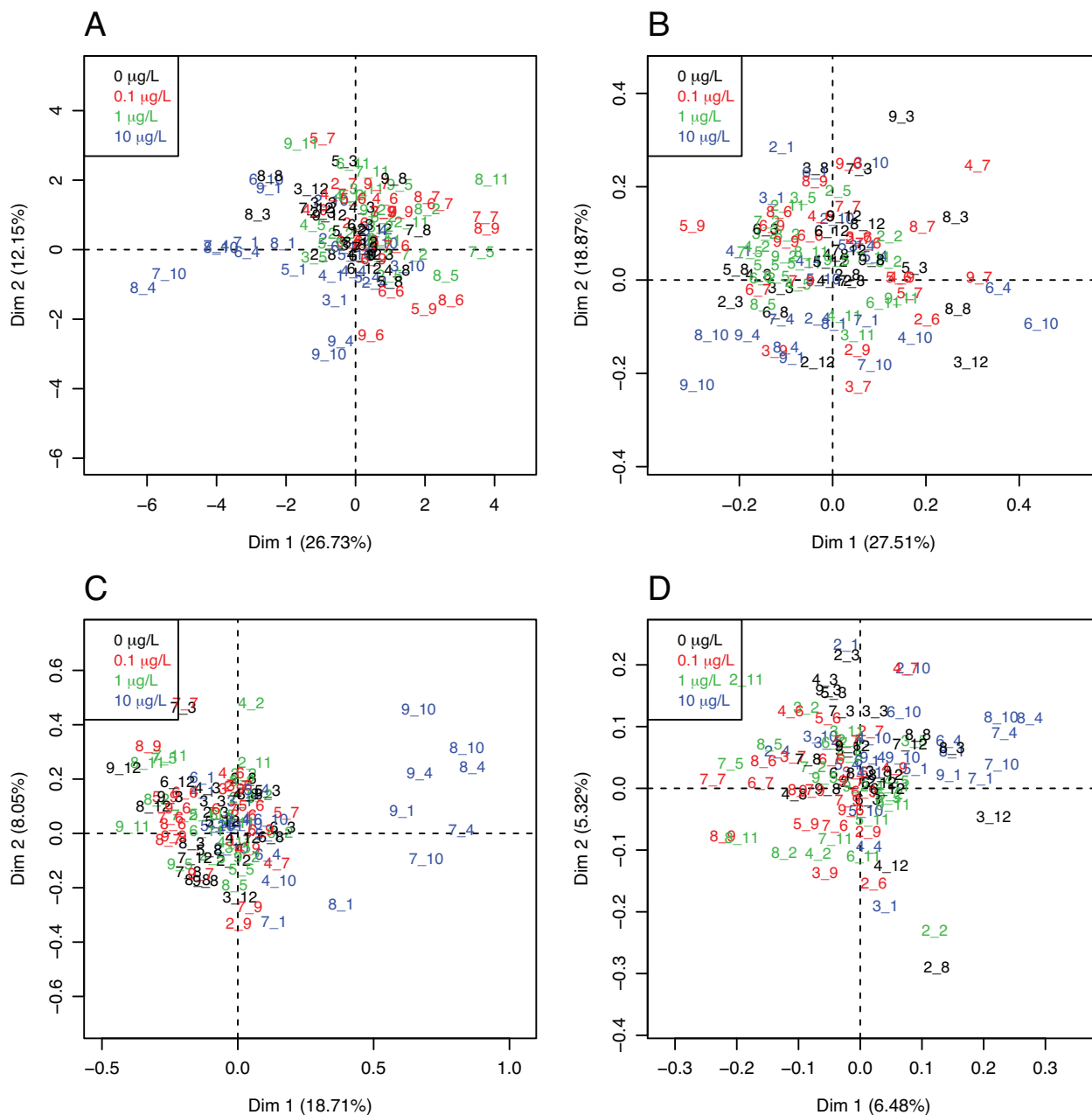
Figure 2. Multidimensional scaling using the Euclidean distance on log-transformed (**A**) and Hellinger-transformed (**B**) data and using the Bray-Curtis dissimilarity on raw (**C**) and log-transformed (**D**) diclofenac data. The variation that could be attributed to time was partialled out. Labels indicate the sampling month (2 = February, 3 = March, etc.), followed by the mesocosm identifier (1–12). Colors indicate the dose. Dim = dimension.

unconstrained analyses, in particular by opposing the 2 highest doses to the others.

The RV coefficients between each pair of dissimilarity matrices ranged between 0.47 and 0.69 for the diclofenac data set and between 0.65 and 0.92 for the chlorpyrifos data set, the 4 dissimilarity measures being only partially redundant (Supplemental Data, Tables S4 and S5). In both data sets, the most similar results were obtained by applying the Bray-Curtis and Euclidean dissimilarity measures to log-transformed data, in particular for the chlorpyrifos data set (RV = 0.92). On the other hand, the most dissimilar results were not obtained by the same set of dissimilarity measures for both data sets.

Distance-based principal response curves

The first PRC (the first axis in this constrained analysis) and the RDA of the whole of the diclofenac data set showed a

significant ($p < 0.05$) effect of dose for all dissimilarity measures (Table 1) and represented a large part of the treatment-related variance in particular with the Bray-Curtis dissimilarity on raw data (Supplemental Data, Table S6). The subsequent PRCs did not show any significant effect (Supplemental Data, Table S7). The first PRC was similar when using the Bray-Curtis dissimilarity and the Euclidean distance for log-transformed data (Figure 4A and D). The diagram with the Euclidean distance for log-transformed data mainly reflected the dynamics of *Spongilla lacustris*, which was confirmed by the response curves for this particular species (Supplemental Data, Figure S10A). The second most abundant taxon was *R. balthica*, whose response curves did not resemble the PRC at a first glance except for the difference between the highest dose and the other doses at the end of the experiment (Supplemental Data, Figure S10B). The weight of the third most abundant
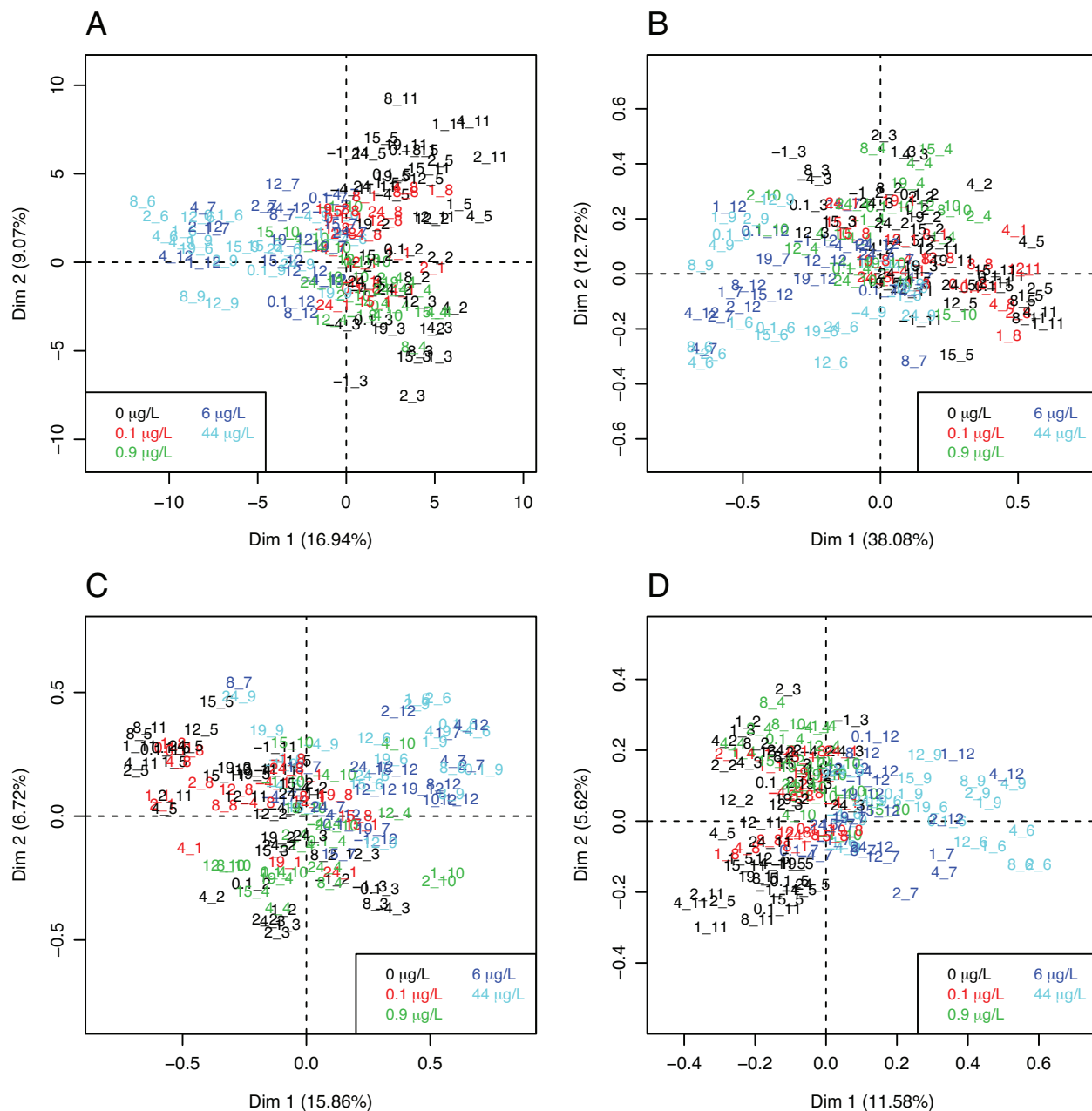
Figure 3. Multidimensional scaling using the Euclidean distance on the log-transformed (**A**) and Hellinger-transformed (**B**) data and using the Bray-Curtis dissimilarity on raw (**C**) and log-transformed (**D**) chlorpyrifos data. The variation that could be attributed to time was partialled out. Labels indicate the sampling week (0 = beginning of treatment), followed by the mesocosm identifier (1–12). Colors indicate the dose. Dim = dimension.

taxon, *Dugesia*, was almost equal to that of *R. balthica*, and other taxa were almost as abundant. The score for *S. lacustris* was approximately twice the score for *R. balthica* and *Dugesia*.

Interestingly, the second PRC with the Hellinger distance resembled the first PRC for Euclidean and Bray-Curtis

Table 1. *p* values resulting from the overall significance test to assess the effect of diclofenac on the macroinvertebrate community

Dissimilarity measure	First principal response curve	All principal response curves
Euclidean on log abundances	1.00E-04	1.00E-04
Bray-Curtis on raw data	1.00E-04	2.00E-04
Bray-Curtis on log abundances	1.00E-04	0.01
Hellinger on raw data	2.00E-04	1.00E-04

dissimilarities in that it showed mainly the abundance of *S. lacustris* (Supplemental Data, Figure S11). This PRC was, however, not significant.

The doses were significant at a larger number of dates for the Hellinger-transformed data (4 consecutive dates at the end of the experiment; see Supplemental Data, Table S8). The significance at each sampling date was consistent regardless of whether only the first axis or all axes of the RDA were used at each date.

With Dunnett's test, neither the Euclidean distance nor the Bray-Curtis dissimilarity based on the log-transformed data provided 2 identical NOECs at consecutive dates, although the dose–response pattern was similar. At 95 d and 119 d after the start of the treatment, the overall effect of the dose was nonmonotonic using these dissimilarity measures. The highest dose appeared to have opposite effects from the 2 lowest doses.
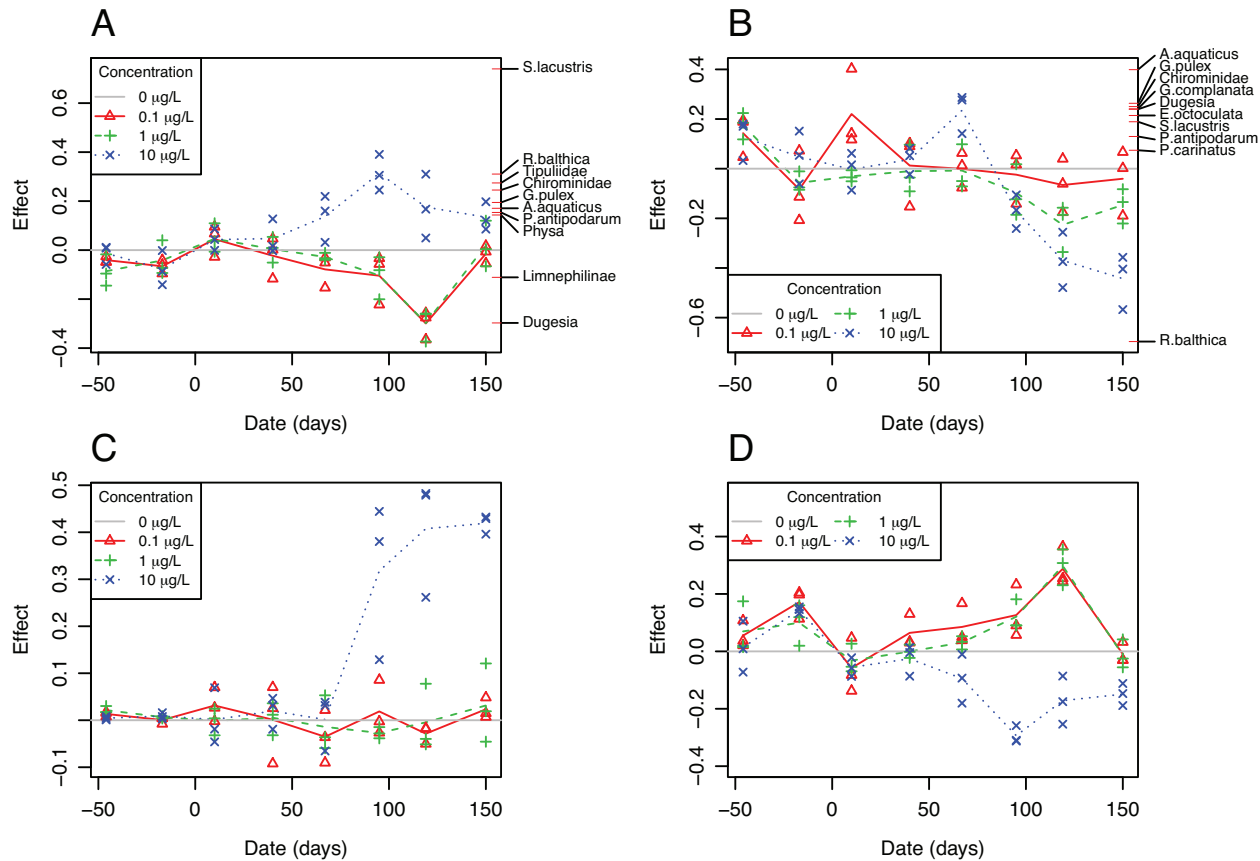
Figure 4. First principal response curves with the Euclidean distance on log-transformed (**A**) and Hellinger-transformed (**B**) data and with the Bray-Curtis dissimilarity on raw (**C**) and log-transformed (**D**) diclofenac data. Only the 10 species with the highest scores are represented (species: *Asellus aquaticus*, *Erpobdella octoculata*, *Gammarus pulex*, *Glossiphonia complananta*, *Pareas carinatus*, *Potamopyrgus antipodarum*, *Radix balthica*, *Spongilla lacustris*).

Slight differences in magnitude of the effect implied that at 95 d after treatment only the highest dose was significantly different from the control (NOEC = 1 µg/L), and at 119 d after treatment the lowest dose was significantly different from the control (NOEC = 0 µg/L; Supplemental Data, Figure S12). The Bray-Curtis and Hellinger dissimilarity measures both provided community NOECs of 1 µg/L. The differences between the control and the lowest dose were not significant with these measures, but the differences between the control and the highest dose were significant on 3 and 4 consecutive dates, respectively. Although these measures were possibly less sensitive to some differences with the control, they provided an overall community NOEC (Table 2).

The first PRC of the chlorpyrifos data set showed a significant part of the dose variance (Supplemental Data, Table S9), and the overall test showed a significant ($p < 0.05$) effect of the dose for all dissimilarity measures (Table 3). The number of significant PRCs varied from 1 dissimilarity measure

to another (Supplemental Data, Table S10). For the Euclidean distance and Bray-Curtis dissimilarity using log-transformed data, the first and third PRCs were significant and the second PRC was only just over the selected threshold ($p = 0.052$ and $p = 0.054$). For the Bray-Curtis dissimilarity using raw data and for the Hellinger distance, the 3 first PRCs showed significant effects. The subsequent PRCs were not significant.

The first PRC for chlorpyrifos was similar for all dissimilarity measures (Figure 5) and clearly showed a dose–response pattern and a gradual recovery after treatment. The diagram for the Euclidean and Hellinger distances reflects the dynamics of several taxa (in particular *Caenis horaria*, *Cloeon dipterum*, and Oligochaeta in the case of the Hellinger distance). This was confirmed for *C. horaria* and *C. dipterum* by the response curves for these species (Supplemental Data, Figure S13). On the other hand, the resemblance between the first PRC with the Hellinger distance and the dynamics for Oligochaeta was not so clear, the difference between the

Table 2. Community no-observed-effect concentrations determined by Dunnett's test ($p < 0.05$) at each sampling date: Diclofenac data set

	Days after treatment							
Dissimilarity measure	−46	−17	10	40	67	95	119	150
Euclidean on log abundances	>	>	0	>	>	1	0	>
Bray-Curtis on raw data	>	>	>	>	>	1	1	1
Bray-Curtis on log abundances	>	>	>	>	>	1	0	>
Hellinger on raw data	>	>	>	>	1	1	1	1

Table 3. p values resulting from the overall significance test to assess the effect of chlorpyrifos on the macroinvertebrate community

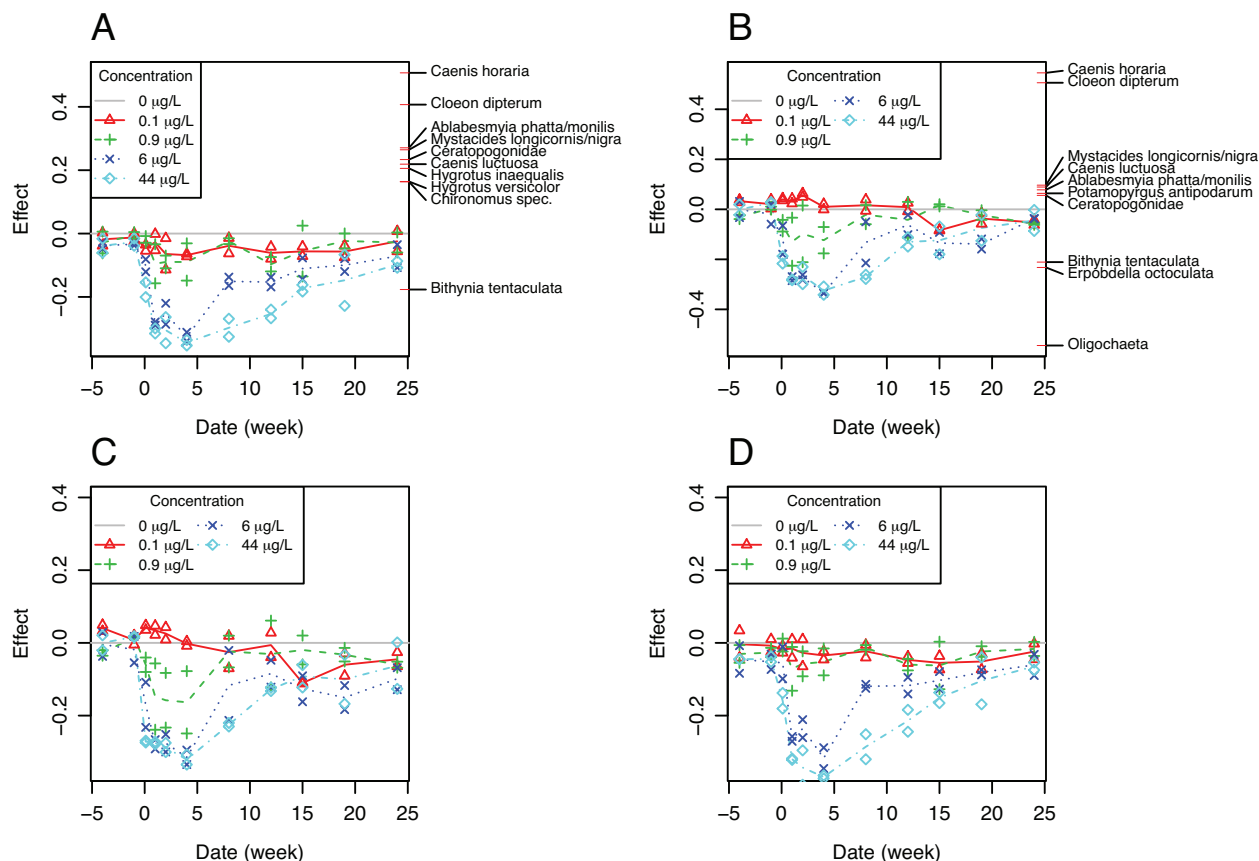| Dissimilarity measure	First principal response curve	All principal response curves
Euclidean on log abundances	2.00E-04	0.0041
Bray-Curtis on raw data	2.00E-04	0.001
Bray-Curtis on log abundances	1.00E-04	2.00E-04
Hellinger on raw data	3.00E-04	0.001

Figure 5. First principal response curves using the Euclidean distance on log-transformed (**A**) and Hellinger-transformed (**B**) data and using the Bray-Curtis dissimilarity on raw (**C**) and log-transformed (**D**) chlorpyrifos data. Only the 10 species with the highest scores are represented.

2 highest doses and the other doses being more pronounced in the PRC than for Oligochaeta. However, this taxon was slightly less abundant for the 0.9 μg/L dose than for the other doses 14 wk after exposure. This difference was visible on the PRC with the Hellinger distance but not on the PRC obtained with the Euclidean distance using log-transformed data.

The interpretation of the second PRC was not as straightforward and depended on the dissimilarity measure (Supplemental Data, Figure S14). The second PRC based on the Hellinger distance shows the differences between the dynamics of the 2 main taxa (*C. horaria* and *C. dipterum*) in terms of recovery at the highest dose. The Euclidean distance and Bray-Curtis dissimilarity on log-transformed data highlighted the differences between the 0.9 μg/L and 44 μg/L doses compared to the other doses, whereas the second PRC obtained for the Bray-Curtis dissimilarity measure using raw data did not show any consistent time-dependent differences between doses.

The significance of the effect of the dose at each sampling date varied from 1 dissimilarity measure to another (Supplemental Data, Table S11), although the effect was always significant at 1 wk, 2 wk, and 4 wk after treatment. Immediately after treatment (0.1 wk) and 8 wk, 12 wk, and 19 wk after treatment also showed significant effects depending on the dissimilarity measure. The dissimilarity measures based on log-transformed data produced the largest number of consecutive significant dates.

The community-level NOECs were 0.9 μg/L of chlorpyrifos for all dissimilarity measures, as this concentration was the NOEC on at least 3 consecutive dates (Table 4).

Relationship between unconstrained analysis and constrained analysis

The effects of the dosage were sufficiently large to be observable on the first axes of the unconstrained analyses for both data sets. The correlations between the sample scores on

Table 4. Community no-observed-effect concentrations determined by Dunnett's test ($p < 0.05$) at each sampling date: Chlorpyrifos data set

Dissimilarity measure	Days after treatment										
	–4	–1	0.1	1	2	4	8	12	15	19	24
Euclidean on log abundances	>	>	>	0.9	0.9	0.9	0.9	0.9	>	>	>
Bray-Curtis on raw data	>	>	6	0.9	0.9	0.9	>	>	>	6	>
Bray-Curtis on log abundances	>	>	>	0.9	0.9	0.9	6	0.9	>	>	>
Hellinger on raw data	>	>	>	0.9	0.9	0.9	6	>	>	6	>

the first axis produced by the partial RDA and by the MDS on data with time partialled out were relatively high, and the plots of the scores showed a relatively even distribution without outliers for any of the dissimilarity measures. With the diclofenac data set, the first axes of both analyses were strongly correlated (Supplemental Data, Table S12). The second axes were not as strongly correlated, in particular for the Bray-Curtis dissimilarity using log-transformed data and for the Hellinger distance. Both the first and the second axes of the unconstrained and constrained analyses of the chlorpyrifos data set were strongly correlated with each other except for the Bray-Curtis dissimilarity using the raw data (Supplemental Data, Table S13). This confirms that the effects of the dosage were the major axes of variability in the data once the effect of time had been partialled out.

DISCUSSION

This analysis of 2 data sets using several exploratory techniques illustrates how the choice of a dissimilarity measure and data transformation can reveal different features of the data sets and how this can affect the final results of the ecological risk assessment. Ideally, the data analysis should not be restricted to 1 dissimilarity measure; a set of contrasting measures each associated with different hypotheses about both the experiment and data characteristics should be evaluated [10]. Examining the data from the different angles offered by the variety of dissimilarity measures can increase the understanding of the final results produced by constrained analysis and their associated *p* values. Expertise in community ecology is then required to assess which changes in community structure or abundance are ecologically significant. In some cases, among a set of dissimilarity measures that are relevant given the experiment, the aim of the study, and the data characteristics, observation of the various ordination plots could support the choice of a particular dissimilarity measure for that data set.

The present analysis also illustrates how PRC results can be complemented by less complex data representations and analyses. The results may be complemented by univariate analyses on individual taxa, paying special attention to the species with the highest weights in the significant PRCs and by comparing the actual response curves relative to the control for the species with the highest contribution to the PRCs [8,44,45].

The choice of a measure of dissimilarity

As illustrated in the present study, the sample and species scores in both the unconstrained and constrained ordinations can sometimes provide arguments in favor of a particular measure of dissimilarity. This can identify important features such as sample outliers, which can have a large weight in the resulting PRC, or horseshoe, arch, or bow [46] effects attributable either to the factor under investigation or to other experimental conditions. Horseshoes can be a result of the fact that the samples at either ends of an ecological gradient way have nothing in common. These artifacts may go unnoticed if the sample scores are not plotted. The sample and species scores can help us to understand which part of the variation each PRC represents and, in particular, to identify individual samples or mesocosms that may outweigh others, thus avoiding drawing conclusions about treatment-level conclusions where only mesocosm-level conclusions are appropriate. The horseshoe obtained in ordinations with the Bray-Curtis dissimilarity on raw data has been reported previously using data ranging from low-abundance sites to high-abundance sites [47]. This illustrates the need to test dissimilarity measures on each new data set before selecting the measures to be used.

The choice of a dissimilarity measure will depend on the characteristics of the study. With mesocosm data, an important question is whether the disappearance or appearance of a species is relevant and how likely it is given the experimental setting. Attention also must be paid to the dispersion of the abundance data and the scarcity of the data matrix—the proportion of 0 counts and rare species. Table 5 formulates working hypotheses for the analysis of both of the mesocosm data sets and identifies which of the dissimilarity measures we studied can be recommended. This table was elaborated based on the knowledge of the details of the experimental setups, on the features observed in the summary description of the 2 data sets, and on the way these features showed through the results of the unconstrained ordinations. Given a new data set from these

Table 5. Examples of working hypotheses and recommendations for the choice of dissimilarity measures applied to the 2 data sets

Hypothesis	Clues	Recommendation	Strategy	
			Diclofenac data set	Chlorpyrifos data set
Community structure is stable at the beginning of experiment; treatment is likely to be the main factor influencing disappearance or appearance of a species, once effect of time has been removed.	Age of communities at the beginning of experiment	Include double zeros (Euclidean distance)	No	Yes
Differences in proportion of species in community as a result of treatment are expected.	Mode of action of the treatment (differences in species sensitivity)	Hellinger's distance (raw data)	Yes	Yes
Total abundance is relevant—there is evidence that the abundance per species is equivalent across replicates.	Check total abundances and species abundance profiles	Bray-Curtis dissimilarity or Euclidean distance	Yes	Yes
There are large differences in abundances between species; the less abundant species may be affected.	Check species abundance profiles and histograms per species	Log-transform with $\log(Ax + 1)$	Yes	Yes
Very few zeros in data	Check for skewness with histograms per species	Scale species data to equal variance	No	No
Summary: Recommended distances			Bray-Curtis dissimilarity on log-transformed data Hellinger distance	Euclidean distance on log-transformed data Hellinger distance

facilities, the choice of a dissimilarity measure could be narrowed down earlier.

When several measures have been tested for the data set, the choice of a dissimilarity measure must not be based on the outcome of the entire analysis (e.g., the NOEC in risk assessment). The criteria for preferring 1 dissimilarity measure over another may not be lower *p* values when testing for the effect of treatment in individual PRCs or in the RDA as a whole, a larger number of consecutive significant dates, or a large proportion of explained variance. For example, in the case of the Bray-Curtis dissimilarity and the Euclidean distance using raw diclofenac data, the first PRC represents a large proportion of explained inertia and several nonconsecutive sampling dates displayed significant effects, yet the first PRC was largely dominated by a single species, *R. balthica*. The response observed for that particular species could be considered to be a principal response for the community if it was liable to have an effect on other taxa or communities in the ecosystem, the strongest statistical signal not being necessarily the most ecologically significant [14]. The main questions then are whether the response is a community response and, if not, whether it causes cascading community responses, hence the need to explore subsequent PRCs and test their significance [8,48]. Also, according to Moser et al. [44], the species with the highest weight determined by PRC is not necessarily the species with the largest response; the general response may in fact hide a larger response from a very small number of species.

Unconstrained multivariate analysis

Unconstrained ordinations with various dissimilarity measures on the raw abundance data provided insight into the main features of the data sets and revealed a strong time gradient in the diclofenac data set in particular for Bray-Curtis dissimilarity using raw data. This could be attributed both to the increase in total abundance in all mesocosms and to the overall variation in relative abundances throughout the experiment. This variation can be explained both by the experimental setup and by seasonal variation. The mesocosm ecosystems were established only a few months before the beginning of the experiment; they were set up over autumn and winter, and a relatively small number of individuals were introduced according to a defined protocol. The samples were collected between the end of February and the beginning of October while the various species introduced colonized the mesocosms to varying extents. Some taxa, such as Ephemeroptera, Diptera, and Odonata, colonized the mesocosms naturally because of outdoor conditions, and other Gastropoda and Oligochaeta were introduced via eggs on the plants collected from previous experiments. These results tend to indicate that Bray-Curtis dissimilarity on raw data is not the best measure for newly established communities as they are likely to show large increases in abundance. In this case, it would be advisable to transform the data to reduce the weight of large abundances. Also, because of these important changes in community structure throughout the experiment, the disappearance or appearance of species is more likely to occur unrelated to the treatment. Thus, it may be advisable to use a dissimilarity measure that excludes double zeros.

In the chlorpyrifos data set, in contrast, the time gradient was not as marked. This could possibly be explained by the fact that the communities had been established a few years before the experiment started, so only seasonal differences were present and not ecosystem development [49]. In this data set, the slight

differences observed between the results with the various dissimilarity measures in the unconstrained analysis did not provide any arguments for or against the use of a particular dissimilarity measure. The stability of well-established communities set in identical environmental conditions could actually call for a distance measure that includes double zeros (such as the Euclidean distance), unlike studies based on data collection, because the disappearance of a species in a treatment level would be a concern.

Effect of total abundance on dissimilarities between samples

The dissimilarity measures selected varied considerably in the extent to which the total abundance was taken into account. Both data sets illustrated that both the Bray-Curtis dissimilarity and the Euclidean distance using raw data highlighted differences in total abundances that were evident in the graphical representation of the total abundances in each sample (Figure 1). The Bray-Curtis dissimilarity measure is designed to take absolute differences in abundance for a species, regardless of the abundance of that species [12]. Assuming that the differences in abundance follow a multiplicative model, the most abundant species would still have a higher weighting because the absolute differences in abundance would be larger for these species.

The overall trend in the total abundance was strikingly different between the 2 data sets. In the diclofenac data, the mean abundances were roughly equal for each dose at each sampling date except in the last 3 sampling mo, where it was 10-fold higher for the highest dose than for the other doses (Figure 1). For this data set, the Bray-Curtis dissimilarity appeared to focus on the most abundant species, as did the Euclidean distance, as stated by Clarke and Green [11], who recommend using logarithmic or power transformations, such as the square-root transformation [50]. The Euclidean distance, which highlights differences in total abundances, clearly opposed the highest dose to the other treatments and focused on the differences at the end of the experiment owing to increasing abundances throughout the experiment. In the diclofenac data set, *R. balthica*, the most abundant species, was affected by the dosage but in a particular way. Using the Euclidean distance on raw data, the effect of diclofenac appeared to be considerable but only on the most abundant species. On the other hand, in the chlorpyrifos data set, the mean abundances for each dose were more even throughout the experiment, and the time-dependent effects of the dosage were slightly smaller (Figure 1). The choice of a dissimilarity measure had less effect on the results, possibly also because there were more species sampled and 1 species did not outweigh all the others. The magnitude of the overall effect of chlorpyrifos on the total abundance was comparable to that of diclofenac, but the dose–response pattern was more complex. Total abundances were lower in the mesocosms that were dosed than in the control mesocosms throughout the experiment, in particular for the highest doses immediately after exposure (Figure 1). In both unconstrained and constrained analyses, dissimilarity measures that highlighted differences in total abundances (such as the Euclidean distance on raw data; Supplemental Data, Figure S15) therefore tended to oppose the control to the other mesocosms as the focus was on the total abundances.

The present results illustrate that the log transformation downweights high-abundance samples and abundant species [5]. In the diclofenac data set, when the log abundances were used, *R. balthica* was still among the most contributing species but did

not outweigh them to the same extent. On the other hand, the Hellinger transformation removes the information on total abundances. As in correspondence analysis, ordinations based on the Hellinger distance are therefore based only on relative abundances. The disadvantage is that this is only appropriate assuming that the community being studied is sufficiently diverse for ecotoxicological effects to vary between subgroups of the community because no effect is observed if all species are reduced in abundance. For both data sets, the Hellinger distance was able to detect differences in community structure. It is worth noting that, in the chlorpyrifos data set, the pattern formed by the sharp decrease in total abundances after treatment is also reflected in terms of community structure as the dose–response type pattern is also observed in the first PRC with the Hellinger distance.

Effect of abundance mean and variance on dissimilarities between samples

The results of the PRC are constrained to represent treatment effects by extracting the most dominant response pattern from the data set. The mean–variance relationship in the data and the way it is taken into account by the dissimilarity measure are important for understanding the results in terms of the magnitude of the effect and residual variability. With the typical near-quadratic relationship between mean and variance [24], a large variance is observed for abundant species. The Bray-Curtis dissimilarity using raw data suffers less from this than the Euclidean distance, which tends to show large differences in variance when there are differences between mean abundances [24]. The logarithm transforms the multiplicative model for the counts into a linear model and would make the variance independent from the mean if the mean–variance relationship was quadratic. In practice, with our data, with the logarithmic transformation $\ln(Ax + 1)$, the variance tended to decrease for large abundances. For both data sets the Hellinger transformation appeared to be more effective at removing the dependence between mean and variance (Supplemental Data, Figures S3 and S6), thus providing a better basis for the Euclidean distance.

Other dissimilarity measures or data transformations

Several other dissimilarity measures were tested, but the results were not discussed because of their similarity to the measures selected. The quantitative version of the Jaccard coefficient and the Kulczynski dissimilarity [21] produced ordinations similar to those obtained with the Bray-Curtis distance. The Euclidean distance was also applied to raw data, raw data scaled to unit variance, and log-transformed data scaled to unit variance. With standardization to unit variance, rare taxa presented outstandingly large abundances in a small number of samples owing to the large number of zeros. As a consequence, when several low-abundance taxa were found in the same sample, the sample became an outlier with a strong contribution to the ordination. This side effect was more pronounced with log-transformed data than with raw data.

Graphical interpretation of the PRCs

Comparisons of the actual response curves to the PRCs showed that a 2-fold difference in taxa scores in a PRC can be sufficient for the most abundant taxon to completely outweigh the other taxa in the interpretation of the PRC, as illustrated in the diclofenac data set. This can be explained by the fact that the second, third, and fourth largest contributions had similar weights and therefore contributed equally in small amounts, with the principal response reflecting the abundance of the taxon which contributed the largest effect.

The usual PRC representation, with means for each dose at each sampling date, together with statistical significance tests provide the user with a simple graphical representation of time-dependent dose–response relationships. In some cases, large differences in mean abundance over replicates could be nonsignificant. In such cases, we believe that providing visual information by representing individual points for replicates on the PRC diagram can help by showing whether particular data points can explain the results or whether there is simply large variability.

Impact on the community NOEC

The choice of dissimilarity measure only had an impact on the overall community NOEC of the diclofenac data set (Table 2), not of the chlorpyrifos data set (Table 4). In the diclofenac data set the Euclidean and Bray-Curtis dissimilarity measures using log-transformed data did not provide 2 consecutive sampling dates with the same community NOEC, although they tended to highlight differences with the control at lower concentrations, partly a result of the nonmonotonicity of the dose–response relationship. Selecting the dissimilarity measure that highlights the most critical changes in community structure and abundance not only is important for understanding and visualizing the data; it can determine the outcome of a risk assessment.

REFERENCES

1. Kedwards TJ, Maund SJ, Chapman PF. 1999. Community level analysis of ecotoxicological field studies: II. Replicated-design studies. *Environ Toxicol Chem* 18:158–166.
2. Szöcs E, Van den Brink PJ, Lagadic L, Caquet T, Roucaute M, Auber A, Bayona Y, Liess M, Ebke P, Ippolito A, ter Braak CJF, Brock TCM, Schafer RB. 2015. Analysing chemical-induced changes in macro-invertebrate communities in aquatic mesocosm experiments: A comparison of methods. *Ecotoxicology* 24:760–769.
3. Maund S, Chapman P, Kedwards T, Tattersfield L, Matthiessen P, Warwick R, Smith E. 1999. Application of multivariate statistics to ecotoxicological field studies. *Environ Toxicol Chem* 18:111–112.
4. Van den Brink PJ, Ter Braak CJF. 1999. Principal response curves: Analysis of time-dependent multivariate responses of biological community to stress. *Environ Toxicol Chem* 18:138–148.
5. Wong DCL, Whittle D, Maltby L, Warren P. 2003. Multivariate analyses of invertebrate community responses to a C(12-15)AE-3S anionic surfactant in stream mesocosms. *Aquat Toxicol* 62:105–117.
6. Clarke KR. 1999. Nonmetric multivariate analysis in community-level ecotoxicology. *Environ Toxicol Chem* 18:118–127.
7. Sparks TH, Scott WA, Clarke RT. 1999. Traditional multivariate techniques: Potential for use in ecotoxicology. *Environ Toxicol Chem* 18:128–137.
8. Van den Brink PJ, Ter Braak CJF. 1998. Multivariate analysis of stress in experimental ecosystems by principal response curves and similarity analysis. *Aquat Ecol* 32:163–178.
9. Gower JC, Legendre P. 1986. Metric and Euclidean properties of dissimilarity coefficients. *J Classif* 3:5–48.
10. Anderson MJ, Crist TO, Chase JM, Vellend M, Inouye BD, Freestone AL, Sanders NJ, Cornell HV, Comita LS, Davies KF, Harrison SP, Kraft NJB, Stegen JC, Swenson NG. 2011. Navigating the

multiple meanings of beta diversity: A roadmap for the practicing ecologist. *Ecol Lett* 14:19–28.

11. Clarke KR, Green RH. 1988. Statistical design and analysis for a biological effects study. *Mar Ecol Prog Ser* 46:213–226.

12. Legendre P, Legendre L. 1998. *Numerical Ecology*, 2nd ed. Elsevier, Amsterdam, The Netherlands.

13. Borcard D, Legendre P, Drapeau P. 1992. Partialling out the spatial component of ecological variation. *Ecology* 73:1045–1055.

14. Cao Y, Williams DD, Williams NE. 1998. How important are rare species in aquatic community ecology and bioassessment? *Limnol Oceanogr* 43:1403–1409.

15. Poos MS, Jackson DA. 2012. Addressing the removal of rare species in multivariate bioassessments: The impact of methodological choices. *Ecol Indic* 18:82–90.

16. Bray JR, Curtis JT. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecol Monogr* 27:326–349.

17. Clarke KR, Warwick RM. 2001. *Change in Marine Communities: An Approach to Statistical Analysis and Interpretation*, 2nd ed. Primer-e, Plymouth, UK.

18. Leps J, Smilauer P. 2014. *Multivariate Analysis of Ecological Data using CANOCO 5*, 2nd ed. Cambridge University Press, Cambridge, UK.

19. Legendre P, Gallagher ED. 2001. Ecologically meaningful transformations for ordination of species data. *Oecologia* 129:271–280.

20. Marchant R. 1999. How important are rare species in aquatic community ecology and bioassessment? A comment on the conclusions of Cao et al. *Limnol Oceanogr* 44:1840–1841.

21. Faith DP, Minchin PR, Belbin L. 1987. Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio* 69:57–68.

22. Hui FKC, Taskinen S, Pledger S, Foster SD, Warton DI. 2015. Model-based approaches to unconstrained ordination. *Methods Ecol Evol* 6:399–411.

23. Orloci L. 1978. *Multivariate Analysis in Vegetation Research*, 2nd ed. Dr. W. Junk, The Hague, The Netherlands.

24. Warton DI, Wright ST, Wang Y. 2012. Distance-based multivariate analyses confound location and dispersion effects. *Methods Ecol Evol* 3:89–101.

25. O'Hara RB, Kotze DJ. 2010. Do not log-transform count data. *Methods Ecol Evol* 1:118–122.

26. Legendre P, Oksanen J, ter Braak CJF. 2011. Testing the significance of canonical axes in redundancy analysis. *Methods Ecol Evol* 2:269–277.

27. ter Braak CJF, Smilauer P. 2015. Topics in constrained and unconstrained ordination. *Plant Ecol* 216:683–696.

28. Legendre P, Anderson MJ. 1999. Distance-based redundancy analysis: Testing multispecies responses in multifactorial ecological experiments. *Ecol Monogr* 69:1–24.

29. Anderson MJ, Willis TJ. 2003. Canonical analysis of principal coordinates: A useful method of constrained ordination for ecology. *Ecology* 84:511–525.

30. de Kermoysan G, Joachim S, Baudoin P, Lonjaret M, Tebby C, Lesaulnier F, Lestremau F, Chatellier C, Akrour Z, Pheron E, Porcher JM, Pery ARR, Beaudouin R. 2013. Effects of bisphenol A on different trophic levels in a lotic experimental ecosystem. *Aquat Toxicol* 144:186–198.

31. Tachet H, Richoux P. 2000. *Invertébrés d' Eau Douce: Systématique, Biologie, Ecologie*. CNRS Editions, Paris, France.

32. Van den Brink PJ, Van Wijngaarden RPA, Lucassen WGH, Brock TCM, Leeuwangh P. 1996. Effects of the insecticide Dursban® 4E (active ingredient chlorpyrifos) in outdoor experimental ditches. 2.

33. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2016. *vegan: Community Ecology Package*, R package Ver 2.4-1 [cited 2016 November 15]. Available from: https://CRAN.R-project.org/package=vegan

34. Rao CR. 1995. A review of canonical coordinates and an alternative to correspondence analysis using Hellinger distance. *Qüestiiò* 19:23–63.

35. Jongman RHG, Ter Braak CJF, van Tongeren OFR. 1995. *Data Analysis in Community and Landscape Ecology*. Cambridge University Press, Cambridge, UK.

36. Van den Brink PJ, Van Donk E, Gylstra R, Crum SJH, Brock TCM. 1995. Effects of chronic low concentrations of the pesticides chlorpyrifos and atrazine in indoor fresh-water microcosms. *Chemosphere* 31:3181–3200.

37. Van den Brink PJ, Hattink J, Bransen F, Van Donk E, Brock TCM. 2000. Impact of the fungicide carbendazim in freshwater microcosms. II. Zooplankton, primary producers and final conclusions. *Aquat Toxicol* 48:251–264.

38. McArdle BH, Anderson MJ. 2001. Fitting multivariate models to community data: A comment on distance-based redundancy analysis. *Ecology* 82:290–297.

39. Robert P, Escoufier Y. 1976. A unifying tool for linear multivariate statistical methods—The RV-coefficient. *J R Stat Soc Series C Appl Stat* 25:257–265.

40. Freedman D, Lane D. 1983. A nonstochastic interpretation of reported significance levels. *J Bus Econ Stat* 1:292–298.

41. R Core Team. 2014. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

42. Husson F, Josse J, Le S, Mazet J. 2014. *FactoMineR: Multivariate Exploratory Data Analysis and Data Mining with R*, Ver 1.26 [cited 2016 November 15]. Available from: https://CRAN.R-project.org/package=FactoMineR

43. Hothorn T, Bretz F, Westfall P. 2008. Simultaneous inference in general parametric models. *Biom J* 50:346–363.

44. Moser T, Rombke J, Schallnass HJ, Van Gestel CAM. 2007. The use of the multivariate principal response curve (PRC) for community level analysis: A case study on the effects of carbendazim on enchytraeids in terrestrial model ecosystems (TME). *Ecotoxicology* 16:573–583.

45. Cuppen JGM, Van den Brink PJ, Camps E, Uil KF, Brock TCM. 2000. Impact of the fungicide carbendazim in freshwater microcosms. I. Water quality, breakdown of particulate organic matter and responses of macroinvertebrates. *Aquat Toxicol* 48:233–250.

46. Swan JMA. 1970. An examination of some ordination problems by use of simulated vegetational data. *Ecology* 51:89–102.

47. Podani J, Miklos I. 2002. Resemblance coefficients and the horseshoe effect in principal coordinates analysis. *Ecology* 83:3331–3343.

48. Van den Brink PJ, Van den Brink N, Ter Braak CJF. 2003. Multivariate analysis of ecotoxicological data using ordination: Demonstrations of utility on the basis of various examples. *Australasian Journal of Ecotoxicology* 9:141–156.

49. Drent J, Kersting K. 1993. Experimental ditches for research under natural conditions. *Water Res* 27:1497–1500.

50. Willis TJ, Anderson MJ. 2003. Structure of cryptic reef fish assemblages: Relationships with habitat characteristics and predator density. *Mar Ecol Prog Ser* 257:209–221.

Invertebrate community responses and recovery. *Environ Toxicol Chem* 15:1143–1153.