ELSEVIER

Two statistical methods for the detection of environmental thresholds

Song S. Qian [a,*], Ryan S. King [b,1], Curtis J. Richardson [b]

[a] *The Cadmus Group, Inc., 6330 Quadrangle Drive, Suite 180, Chapel Hill, NC 27517, USA*
[b] *Duke University Wetland Center, Nicholas School of the Environment and Earth Sciences, Duke University, Durham, NC 27708, USA*

Abstract

A nonparametric method and a Bayesian hierarchical modeling method are proposed in this paper for the detection of environmental thresholds. The nonparametric method is based on the reduction of deviance, while the Bayesian method is based on the change in the response variable distribution parameters. Both methods are tested using macroinvertebrate composition data from a mesocosm experiment conducted in the Everglades wetlands, where phosphorus is the limiting nutrient. Using the percent of phosphorus tolerant species and a dissimilarity index as the response variables, both methods resulted in a similar and well-defined TP concentration threshold, with a distribution function that can be used to determine the probability of exceeding the threshold.
© 2003 Elsevier B.V. All rights reserved.

Keywords: Bray–Curtis dissimilarity; Changepoint; Deviance; Everglades; Gibbs sampling; Macroinvertebrate; Nonparametric; Total phosphorus

1. Introduction

Anthropogenic perturbations such as landscape fragmentation, cultural eutrophication, or introduction of toxic substances often cause changes in structure and function of aquatic and terrestrial ecosystems. For example, in the USA it is estimated that nearly 40% of all water bodies are ecologically impaired by various pollutants (USEPA, 1998a). Studies of how an ecosystem responds to such disturbances have important management implications, such as the es-

tablishment of water-quality or emission standards for particular geographic regions (Adams and Greeley, 2000). One common approach to establishing environmental criteria is to examine changes of selected population, community, or ecosystem attributes along a gradient of environmental conditions (e.g. Karr and Chu, 1997). Ecological attributes often show little change until a critical environmental value, or threshold, is reached (Fore et al., 1996; Richardson and Qian, 1999). Thus quantitative description of such exposure–response relationships can be very useful in support of the development of numerical environmental criteria (Suter, 1993; USEPA, 1998a).

Although the identification of environmental thresholds is deeply rooted in ecological risk assessment (Suter, 1993), surprisingly few statistical techniques are appropriate for their detection. Many traditional

---

* Corresponding author. Tel.: +1-919-403-5105;
fax: +1-919-401-5867.
E-mail address: sqian@cadmusgroup.com (S.S. Qian).
[1] Present address: Smithsonian Environmental Research Center, 647 Contees Wharf Road, Box 28, Edgewater, MD 21037, USA.

statistical techniques are not suitable for estimating such thresholds, nor are they adequate for estimating uncertainty in their predictions, both of which are critical components in ecological risk assessment (Bartell et al., 1992; Suter, 1993, 1996; Lemly and Richardson, 1997; USEPA, 1998b). Also, most methods require parametric assumptions, such as normality and linearity, that ecological data rarely meet (e.g. Clarke, 1993). Rather, ecological responses to environmental gradients often are nonlinear, non-normal, and heteroscadistic (Legendre and Legendre, 1998).

We propose two statistical methods for detecting a changepoint along an environmental gradient. The methods presented here can be seen as part of the larger effort of using mathematical and statistical models to study wetland ecology (e.g. Wang and Mitsch, 2000; Shukla, 1998; Buzzelli et al., 2000; Loiselle et al., 2000; Moreno-Grau et al., 1996) and to make risk assessment (e.g. Findlay and Zheng, 1999). The second method is also part of the increasing interest in the application of Bayesian statistics in the Ecol. Model. community (e.g. Omlin and Reichert, 1999; Prato, 2000; Reichert and Omlin, 1997; Steinberg et al., 1997; Borsuk et al., 2001; Aldenberg et al., 1995).

Our methods are, however, explicitly designed to detect ecological changes along an environmental gradient, and may be useful in criteria development because they (1) estimate discrete, numerical values of the predictor variable that lead to ecological changes, (2) provide an estimate of uncertainty by generating confidence limits, and (3) make very few assumptions regarding properties of the data.

The first method is a nonparametric approach that finds a changepoint that results in the largest reduction in the deviance of the response variable. The second method is based on a Bayesian hierarchical modeling approach. We present the statistical basis of these two methods and demonstrate the application using a data set collected from an experimental nutrient gradient in an Everglades wetland.

2. Method

Let $y_1, \ldots, y_n$ be the sequence of the ecological response variable observed along the ordered environmental gradient $x_1, \ldots, x_n$. A changepoint problem is

to find r ($1 \leq r \leq n$) that separates the response variable into two groups: $y_1, \ldots, y_r$ and $y_{r+1}, \ldots, y_n$, each with distinct characteristics such as the mean and the variance. The corresponding value of the environmental variable x_r is the threshold.

2.1. The nonparametric deviance reduction approach

The nonparametric deviance reduction approach is based on the idea that a structural change in an ecosystem may result in the change of both the mean and the variance of the ecological response variable used to indicate the change. When the observations from multiple sites are ordered along the gradient, the threshold or changepoint separates the observations into two groups. Each is relatively homogeneous. This idea was inspired by the tree-based modeling approach (Breiman et al., 1984), where binary splits were used to construct a predictive regression or classification model. Guisan and Zimmermann (2000) and Qian and Anderson (1999) are two examples of application of tree-based model in ecological studies. The concept of deviance reduction was used to develop the method for environmental threshold estimation and a bootstrap method was used to quantify uncertainty about the threshold.

The deviance (Venables and Ripley, 1994), a measure of homogeneity, is defined for a continuous variable, as:

$$D = \sum_{k=1}^{n} (y_k - \mu)^2 \tag{1}$$

where D is the deviance, n is the sample size, and μ is the mean of the n observations y_k. For a categorical variable, the deviance is defined as:

$$D = -2 \sum_{k=1}^{g} n_k \log(p_k) \tag{2}$$

where g is the number of classes, p_k is the proportion of observations and n_k is the number of observations in class k, respectively.

When the response data are divided into two groups, the sum of the deviance for the two subgroup is always less than or equal to the deviance of the entire data. Each possible changepoint is associated with a deviance reduction:

$$\Delta_i = D - (D_{\leq i} + D_{>i})$$

where D is the deviance of the entire data set $y_1, \ldots, y_n$, $D_{\leq i}$ is the deviance of the sequence $y_1, \ldots, y_i$, and $D_{>i}$ is the deviance of the sequence $y_{i+1}, \ldots, y_n$, where $i = 1, \ldots, n$. The changepoint r is the i value that maximizes $\Delta_i : r = \max_i \Delta_i$.

Uncertainty about the changepoint can be estimated by using a bootstrap simulation (Efron and Tibshi rani, 1993) and expressed as a 90% confidence interval. This uncertainty may be interpreted as a recognition that a changepoint may be best represented as a small range of values rather than one discrete value. A second consideration is that the deviance reduction approach will always find a changepoint no matter whether there is a real ecological change or not. Thus, we use the approximate χ^2 test to judge whether the resulting changepoint is statistically significant. The χ^2 test is based on the fact that the deviance reduction divided by the scale parameter is approximately χ^2 distributed (d.f. = 1) (Venables and Ripley, 1994). A large deviance reduction will result in a small *P*-value, thus the rejection of the null hypothesis of no changepoint.

Our method is consistent with the tree-based modeling approach. In fact, the changepoint is the first split of a tree model when x is used as the single predictor variable. As a result, the commonly available tree model software (e.g. `rpart` in S-Plus and R) can be used. In this paper, we wrote an S-Plus function to calculate the changepoint and used the S-Plus function `bootstrap` to evaluate the uncertainty.

2.2. The Bayesian hierarchical modeling approach

Under a Bayesian framework, we make specific probabilistic assumptions about the ecological response variable. Specifically, we assume that the response variable values, $y_1, \ldots, y_n$, collected from the n sites along the gradient of interest, are random samples from the sequence of random variables $Y_1, \ldots, Y_n$. In other words, we define a random variable for each site, and assume that these random variables belong to the same family of distributions with parameter θ.

The term "site" is used to refer a sampling location that has a distinct predictive variable value. Depending on the scale of the study, a site can be a $1 \, \text{m} \times 1 \, \text{m}$ sampling grid as in the Everglades example, or reaches of streams miles apart.

The random variables $Y_1, \ldots, Y_n$ have a changepoint r $(1 \leq r \leq n)$ if the parameter value changes at r:

$$Y_1, \ldots, Y_r \sim \pi(Y_i | \theta_1)$$
$$Y_{r+1}, \ldots, Y_n \sim \pi(Y_i | \theta_2) \tag{3}$$

Theoretical background on this type of changepoint analysis can be found in Smith (1975), Raftery and Akman (1986), Carlin et al. (1992). Extension to a multiple changepoint problem can be found in Stephens (1994). In this paper, we summarize the results of changepoint analysis presented in Smith (1975) and use the newly developed Gibbs sampling procedure for parameter estimation.

In our example, the response variables can be approximated by a normal distribution or a binomial distribution (see Section 3.2). Accordingly, we present the details of a changepoint problem for normal and binary response variables.

2.3. Normal distribution model

When the random variables $Y_1, \ldots, Y_n$ are from a normal distribution family, the changepoint problem is defined as follows:

$$Y_i \sim \begin{cases} N(\mu_1, \sigma_1^2), & i = 1, \ldots, r \\ N(\mu_2, \sigma_2^2), & i = r+1, \ldots, n \end{cases} \tag{4}$$

Let $\lambda_1 = 1/\sigma_1^2$ and $\lambda_2 = 1/\sigma_2^2$. As a result, model parameters are $\theta = (\mu_1, \lambda_1, \mu_2, \lambda_2)$. Assume the prior is of the form

$$\pi(\theta, r) \propto \pi(\lambda_1)\pi(\lambda_2)$$

In addition, let the prior distributions of λ_1 and λ_2 be from a gamma distribution family, i.e. $\lambda_1 \sim \gamma(\alpha_1', \beta_1')$ and $\lambda_2 \sim \gamma(\alpha_2', \beta_2')$. The proper prior distributions for λ_1 and λ_2 ensures a proper posterior distribution for r. In practice, values of the parameters (α_1', β_1') and (α_2', β_2') can be chosen to make the prior distributions nearly flat. We used 0.001 for all four parameters.

The joint distribution of data and parameters is proportional to the product of prior and likelihood:

$$\prod_{i=1}^{n} \pi(\theta, r)\pi(Y_i | r, \theta) \propto \lambda_1^{r/2 + \alpha_1' - 1}$$
$$\times e^{[-(1/2)r\lambda_1(\mu_1 - \bar{Y}_1)^2]} e^{(-\lambda_1 \delta_1)}$$
$$\times \lambda_2^{(n-r)/2 + \alpha_2' - 1} e^{[-(1/2)(n-r)\lambda_2(\mu_2 - \bar{Y}_2)^2]}$$
$$\times e^{(-\lambda_2 \delta_2)} \tag{5}$$

and the marginal distribution of r is:

$$\text{pr}(r|Y) \propto$$

$$\begin{cases} \dfrac{1}{r^{1/2}} \dfrac{1}{(n-r)^{1/2}} \dfrac{\Gamma(\gamma_1)}{\delta_1^{\gamma_1}} \dfrac{\Gamma(\gamma_2)}{\delta_2^{\gamma_2}}, & \text{for } r < n \\[2ex] \dfrac{1}{n^{1/2}} \dfrac{\Gamma(\gamma_n)}{\delta_n^{\gamma_n}} \dfrac{\Gamma(\alpha'_2)}{\beta_2'^{\alpha'_2}}, & \text{for } r = n \end{cases} \qquad (6)$$

where

$$\bar{Y}_1 = \frac{1}{r}\sum_{i=1}^{r} Y_i, \qquad \bar{Y}_2 = \frac{1}{n-r}\sum_{i=r+1}^{n} Y_i,$$

$$\gamma_1 = \frac{r-1}{2} + \alpha'_1, \qquad \delta_1 = \frac{1}{2}\left[\sum_{i=1}^{r} Y_i^2 - r\bar{Y}_1^2\right] + \beta'_1,$$

$$\gamma_2 = \frac{n-r-1}{2} + \alpha'_2, \quad \delta_2 = \frac{1}{2}\left[\sum_{i=r+1}^{n} Y_i^2 - (n-r)\bar{Y}_2^2\right] + \beta'_2,$$

$$\gamma_n = \frac{n-1}{2} + \alpha'_1, \qquad \delta_n = \frac{1}{2}\left[\sum_{i=1}^{n} Y_i^2 - n\bar{Y}_1^2\right] + \beta'_1$$

and $\Gamma(\cdot)$ represents the Gamma function.

This is a discrete probability distribution. Because the order of the response variable is the same as the environmental gradient variable, the probability of x_i being the threshold is also defined by Eq. (6). We may choose to use the mode of the distribution as the estimate of the changepoint, or the expected value of the corresponding environmental gradient variable.

The posterior conditional distributions of parameters θ are:

$$\mu_1|\mu_2, \lambda_1, \lambda_2, r \sim N(\bar{Y}_1, r\lambda_1),$$

$$\lambda_1|\mu_1, \mu_2, \lambda_2, r \sim \gamma(\gamma_1, \delta_1),$$

$$\mu_2|\mu_1, \lambda_1, \lambda_2, r \sim N(\bar{Y}_2, (n-r)\lambda_1), \text{ and}$$

$$\lambda_2|\mu_1, \mu_2, \lambda_1, r \sim \gamma(\gamma_2, \delta_2)$$

2.4. Binomial distribution model

When the random variables $Y_1, \ldots, Y_n$ are from binomial distributions, the changepoint problem becomes

$$Y_i \sim \begin{cases} \text{binomial}(\theta_1, N_i), & i = 1, \ldots, r \\ \text{binomial}(\theta_2, N_i), & i = r+1, \ldots, n \end{cases} \qquad (7)$$

where θ_1 and θ_2 are probabilities of success before and after the change, respectively. Assuming a uniform

prior on r, θ_1, and θ_2, the joint distribution of data and parameter is proportional to:

$$\pi(\theta_1, \theta_2, r) L(Y; \theta_1, \theta_2, r) \propto \theta_1^{\sum_{i=1}^{r} Y_i}$$

$$\times (1-\theta_1)^{\sum_{i=1}^{r}(N_i-Y_i)} \theta_2^{\sum_{i=r+1}^{n} Y_i}$$

$$\times (1-\theta_2)^{\sum_{i=r+1}^{n}(N_i-Y_i)} = \theta_1^{S_{11}}$$

$$\times (1-\theta_1)^{S_{12}} \theta_2^{S_{21}} (1-\theta_2)^{S_{22}} \qquad (8)$$

where $S_{11} = \sum_{i=1}^{r} Y_i$, $S_{12} = \sum_{i=1}^{r}(N_i - Y_i)$, $S_{21} = \sum_{i=r+1}^{n} Y_i$, and $S_{22} = \sum_{i=r+1}^{n}(N_i - Y_i)$.

Integrating out θ_1 and θ_2 from the joint distribution (8), the posterior marginal distribution of r is

$$\pi(r|Y) \propto \int \theta_1^{S_{11}} (1-\theta_1)^{S_{12}} \theta_2^{S_{21}} (1-\theta_2)^{S_{22}} \, d\theta_1 \, d\theta_2$$

$$\propto \frac{\Gamma(S_{11}+1)\Gamma(S_{12}+1)}{\Gamma(S_{11}+S_{12}+2)} \frac{\Gamma(S_{21}+1)\Gamma(S_{22}+1)}{\Gamma(S_{21}+S_{22}+2)} \qquad (9)$$

The conditional posterior distributions of θ_1 and θ_2 are available:

$$\pi(\theta_i|r, Y) = \text{beta}(S_{j1}+1, S_{j2}+1), \quad \text{for } j = 1, 2 \qquad (10)$$

Inference on whether a changepoint exists can be made by the probability of no changepoint $\text{pr}(n|Y)$.

Inference on the posterior distributions of r, μ_1, μ_2, λ_1, and λ_2 for the normal model, and r, θ_1, and θ_2 for the binomial model are made by using the Gibbs sampler, a Markov chain Monte Carlo simulation (MCMC) method (Gelfand and Smith, 1990; Gelfand et al., 1990; Smith and Roberts, 1993). Casella and George (1992) provides an intuitive exposition of the Gibbs sampler. Examples of environmental ap-

plication of the Gibbs sample are found in Qian and Reckhow (1998), Qian and Richardson (1997), and Qian et al. (2000).

We note that multiple samples from each site can be (and should be) directly used in both methods' computation procedures.

For the binomial model, we wrote S-Plus functions to calculate the changepoint distribution (Eq. (9)) and to sample the posterior distributions of the parameters (Eq. (10)). The calculation can be done using any software that evaluates a Γ-function and generates random samples from a beta distribution. For the normal model, we used WinBUGS (Spiegelhalter et al., 2000), a freely available software for Bayesian analysis using Gibbs sampler.

3. The Everglades example

3.1. Introduction

Numerous studies have shown that the Everglades is a phosphorus limited ecosystem (e.g. Richardson et al., 1999; Steward and Ornes, 1975a,b; Swift and Nicholas, 1987; Flora et al., 1988). One of the most publicized sources of perturbation to the Everglades ecosystem has been excessive inputs of phosphorus. Thus, as part of the 1994 Everglades Forever Act (EFA, Florida Administrative Code, 17-302.530), the state of Florida mandated that a numerically defensible water-column total phosphorus (TP) standard be established by 2001. Moreover, as a Florida Class III waterbody, legislation has specifically mandated that the TP concentration selected as a standard must not result in an imbalance of flora or fauna.

In accordance with this legislation, a phosphorus dosing study was established in the Water Conservation Area-2A (WCA-2A) in the northern Everglades to experimentally examine changes in biological attributes in response to TP concentrations. Two dosing facilities, each with five walled mesocosms and one unwalled control, were constructed in adjacent sloughs in southern part of WCA-2A where anthropogenic impact is negligible. Mesocosms were 2-m wide and 8-m long flumes, continuously dosed with P from the north ends. Each walled mesocosm was assigned a soluble reactive phosphate (SRP) treatment in one facility and replicated in the other. These treatments ranged from background concentrations (walled control) to approximately 125 μg/l in the highest treatment. Water-column SRP and TP were measured biweekly throughout the duration of the study at multiple stations down the length of each mesocosm. The dosing study was inaugurated on 30 November 1992 and terminated on 21 September 1998. Greater detail of the dosing study experimental design and operation can be found in Richardson et al. (2000) and Pan et al. (2000).

3.2. Data

While several levels of biological organization were studied to develop a phosphorus threshold for the Everglades, here we only consider the relationship between macroinvertebrate assemblages and TP as a test for the two statistical methods. Macroinvertebrates are one of the most widely used biological indicators in aquatic systems (Rosenberg and Resh, 1993; Barbour et al., 1999), thus were expected to be useful monitors of ecological condition in the Everglades. We initiated the macroinvertebrate component of the phosphorus dosing study in 1996, 4 years after dosing had begun. Samples were collected at the 2, 4, and 6-m stations within each mesocosm (including unwalled control) on four dates, two wet season (2 September 1996, 21 September 1998) and two dry season (8 January 1997, 4 February 1998). Thus, there were 36 observations on each date. Greater detail on sampling design and methods are described elsewhere (Richardson et al., 2000; King and Richardson, in press).

To assess dose–response relationships between macroinvertebrate communities and TP, 2 metrics were calculated using the species abundance data: (1) Bray–Curtis dissimilarity (BCD) and (2) percent tolerant individuals.

BCD was selected as a metric because it has been shown to be one of the most robust and ecologically interpretable measures for species abundance data (Bray and Curtis, 1957; Faith et al., 1987; Clarke, 1993; Legendre and Legendre, 1998; Legendre and Anderson, 1999). Before calculation, a $\log_{10}(x + 1)$ transformation was applied to taxon abundances to increase the relative contribution of the uncommon and rare taxa (e.g. Gauch et al., 1982; Efron and Tibshirani, 1989; Cao et al., 1998). Since BCD is based on pairwise comparisons between all sample

pairs, samples were ordinated using nonmetric multi-dimensional scaling (nMDS), rotated using varimax rotation to maximize variation along nMDS Axis 1, and extracted as univariate scores along nMDS Axis 1 (McCune et al., 1997; Legendre and Legendre, 1998). Because the standardized log-abundance variables are approximately normal, nMDS Axis 1, a linear combination of these nearly normal variables, is approximately normal as well. The objective in the use of nMDS was to recover a multivariate assemblage pattern that could potentially be attributed to a gradient in TP concentration, and to reduce dimensionality to allow for univariate changepoint analysis.

Percent tolerant individuals was calculated using a list of taxa shown to be highly abundant in high phosphorus, eutrophic areas in the Everglades but uncommon in low phosphorus areas (King, 2001; King and Richardson, in press). This metric was recorded in the form of two counts: number of tolerant species and total number of individuals. Because the nature of the data is binary (a subject is either phosphorus tolerant, success, or non-tolerant, failure) and we are interested in the proportion of tolerant species (or the probability of success), a binomial distribution is appropriate.

In accordance with the Everglades Forever Act (Florida Administrative Code 17-373.4592), long-term geometric mean values of TP were used as predictors in this analysis. We define the "long-term" as the approximate life span of most long-lived taxa present at the dosing study, which is about 6 months. Each geometric mean corresponded to the precise location of each macroinvertebrate sample collected from the mesocosms.

Table 1
Changepoint estimation results for macroinvertebrate responses in a phosphorus dosing mesocosm in the Everglades

Sample no.	TP threshold (μg/l)		Intervals	
	Nonparametric	Bayesian	Nonparametric	Bayesian
	Response variable: BCD			
1	12.25 ($P = 0.00123$)	10.23 ($P < 0.00001$)	10.05, 18.38	10.05, 10.55
2	11.60 ($P < 0.00001$)	11.81 ($P < 0.00001$)	11.12, 12.76	11.27, 12.76
3	10.53 ($P = 0.00073$)	10.68 ($P < 0.00001$)	10.07, 10.68	10.61, 11.59
4	10.81 ($P = 0.00073$)	13.94 ($P < 0.00001$)	8.31, 13.94	10.55, 13.94
	Response variable: percent phosphorus tolerant species			
1	12.11 ($P < 0.00001$)	10.54 ($P < 0.00001$)	10.05, 18.05	10.05, 10.55
2	14.04 ($P < 0.00001$)	13.47 ($P < 0.00001$)	11.27, 16.43	12.72, 15.21
3	10.67 ($P < 0.00001$)	10.68 ($P < 0.00001$)	9.07, 11.99	10.61, 10.68
4	10.69 ($P < 0.00001$)	11.80 ($P < 0.00001$)	7.12, 14.40	8.31, 12.38

Table 2
Mean response variable values

| Sample no.	BCD (left)		BCD (right)	
	Nonparametric	Bayesian	Nonparametric	Bayesian
	Response variable: BCD			
1	-0.80 ± 0.33	-0.82 ± 0.36	0.35 ± 0.63	0.34 ± 0.66
2	-0.95 ± 0.41	-0.92 ± 0.48	0.48 ± 0.40	0.49 ± 0.41
3	-0.76 ± 0.74	-0.74 ± 0.78	0.55 ± 0.44	0.56 ± 0.45
4	-0.65 ± 0.71	-0.53 ± 0.74	0.46 ± 0.42	0.52 ± 0.41
	Response variable: percent phosphorus tolerant species			
1	0.044 ± 0.032	0.041 ± 0.008	0.231 ± 0.109	0.221 ± 0.012
2	0.070 ± 0.041	0.071 ± 0.013	0.224 ± 0.086	0.216 ± 0.015
3	0.026 ± 0.027	0.033 ± 0.007	0.182 ± 0.147	0.178 ± 0.011
4	0.056 ± 0.043	0.079 ± 0.016	0.194 ± 0.096	0.197 ± 0.017
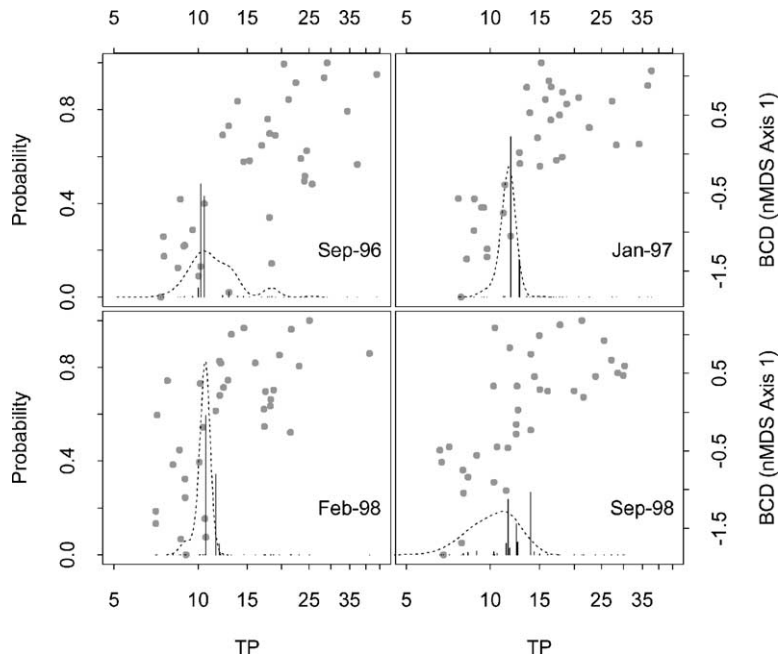
Fig. 1. Changepoint distributions estimated for BCD using the nonparametric (dashed lines) and the Bayesian (the solid lines) methods. Data are shown as shaded dots.
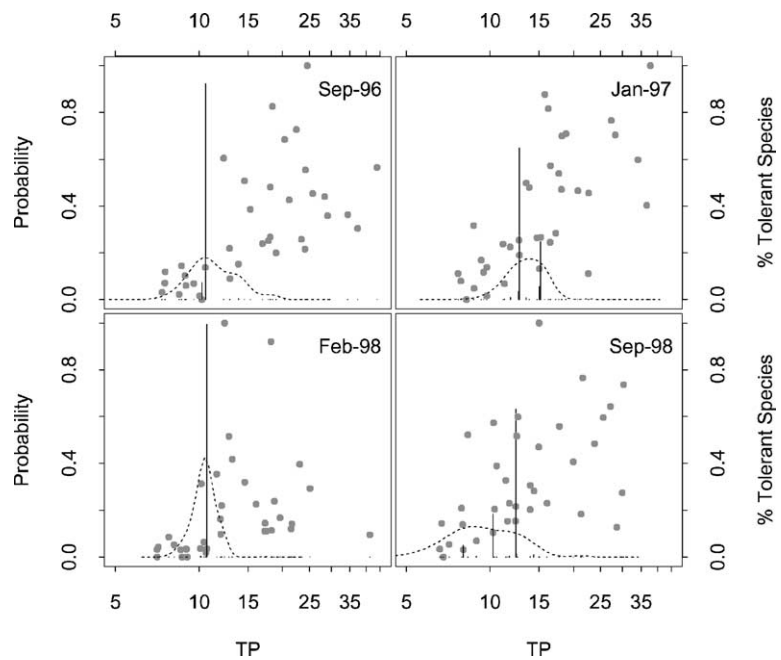

Fig. 2. Changepoint distributions estimated for percent phosphorus tolerant species using the nonparametric (dashed lines) and the Bayesian (the solid lines) methods. Data are shown as shaded dots.

3.3. Results

For each of the four sampling events, we present the changepoint estimated using both the nonparametric and the Bayesian methods (Table 1), as well as the estimated mean and standard deviation of the BCD and percent phosphorus tolerant species on both sides of the changepoint (Table 2). Uncertainty in the estimated changepoint are presented by using (1) the range of the middle 90% of the 1000 bootstrap simulation replicates, and (2) the 90% credible intervals for the Bayesian estimates (under the column *Intervals* in Table 1). The values of *P* under the two *TP threshold* columns in Table 1 are the *P*-value for testing whether a changepoint exists estimated based on the approximate χ^2 test discussed in the method section for the nonparametric method, and is the probability of no changepoint, $pr(n|Y)$, for the Bayesian method.

Figs. 1 and 2 present the threshold distributions along with the BCD and percent tolerant species data, respectively, for all four sampling events. In the figures, the bootstrap simulation results for the nonparametric method were shown by the dashed lines (the estimated probability density function of the 1000 bootstrap simulation replicates); the Bayesian method results are shown by the vertical lines with the height representing the probability of the corresponding TP value being the threshold; the shaded circles are the data used for the analysis.

The results from the nonparametric method are comparable to those of the Bayesian method. This is expected since the probability distribution assumptions (normal and binomial) made on the response variables under the Bayesian method are appropriate. Because the Bayesian method uses the distributional information, it resulted in narrower 90% intervals for the changepoint. When the proper response variable probability distribution cannot be ascertained, the nonparametric method should be used.

4. Discussions

We presented two statistical methods for quantifying environmental thresholds using data from an experiment conducted in the Everglades as an example. Both methods are designed specifically for the detection of change in the selected ecological response variable along a gradient. The nonparametric method does not make probabilistic assumptions about the response data; it is therefore more robust. Computation of the nonparametric method is straightforward and can be done without special software. However, it is necessary to identify the type of response data (e.g. continuous, counts, or categorical) in order to calculate the deviance properly. Because the nonparametric method does not make use of information about the probabilistic distribution of the response variable, it is less efficient than the Bayesian method when such information exist. The Bayesian method requires specific information on the distribution of the response variable. This information is, however, readily available for most ecological data, either from past experience (e.g. log-normal is a good approximation for concentration variables), or the nature of data (e.g. binomial distribution is appropriate for binary response variables and Poisson distribution is often used for counts data). Computation of the Bayesian method is more intense and complicated than the computation of the nonparametric method. If only the changepoint is of interest, any software that is capable of evaluating the Gamma function is sufficient (Eqs. (6) and (9)). We recommend that both methods be used in order to fully explore all possible outcomes.

From the probability density (or distribution) functions (Figs. 1 and 2) one can determine the most likely threshold values and their uncertainty. From a risk assessment view point, a cumulative density or distribution function (CDF) can then be used to directly read out the probability of exceeding the threshold (Fig. 3). The CDFs for the percent tolerant species presents the risk exceeding the threshold at various TP concentrations.

It is more often the case that a threshold is not well defined. As a result, the change in the selected response variable may be gradual. If the data cover the change well, i.e. there are enough data points to described both before and after the change, our methods will work well, resulting in a flatter changepoint distribution.

The EFA requires the TP threshold to be set to prevent flora and fauna imbalance. However, there was no definition about what constitutes an imbalance. In the example, we selected the percent phosphorus tolerant macroinvertebrate species and the Bray–Curtis dissimilarity index as the response variables. A change
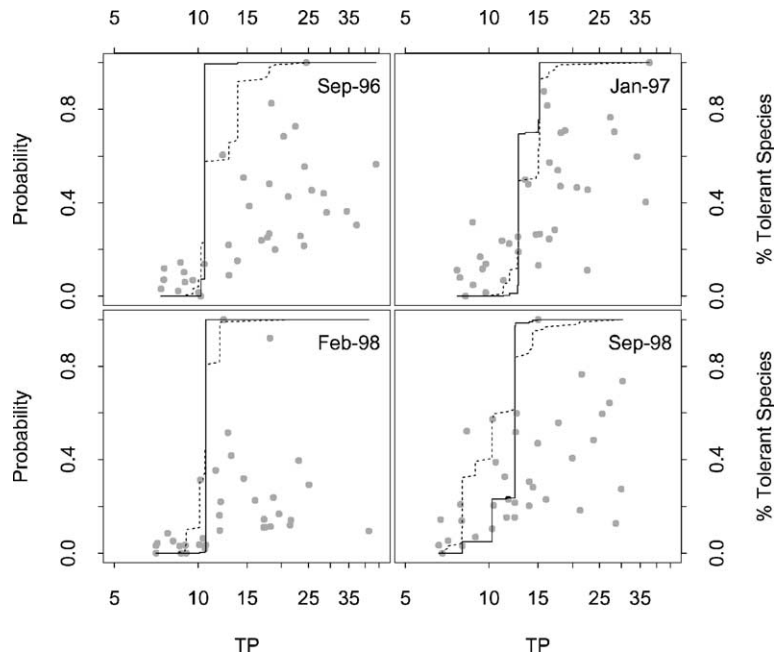
Fig. 3. Cumulative changepoint distributions estimated for percent phosphorus tolerant species using the nonparametric (dashed lines) and the Bayesian (the solid lines) methods, indicating the risk of exceeding the TP threshold.

in the two response variables we used only represents an "imbalance" in the macroinvertebrate composition metrics we tested, which may or may not be the same imbalance as defined in the EFA. The results presented in Table 1 indicated a TP threshold value of slightly above 11 μg/l. Separately using data from the same mesocosm experiment, we estimated the TP threshold values using additional variables representing various ecological trophic levels. We found that, in general, the TP thresholds were different for response variables representing different trophic levels or the community level, which indicates that an environmental threshold estimated using our methods should not be isolated from the intended response variable.

Dimension reduction is a common practice in quantitative ecology (Legendre and Legendre, 1998). In our case, we used Bray–Curtis dissimilarity index and percent tolerant individuals. Reducing the dimension of the data will inevitably lead to loss of information. Therefore, using the full species composition is advisable. We will report our work on using species composition data for estimating threshold in a separate

paper (Qian et al., 2003). In the mean time, we suggest that multiple metrics (e.g. abundance, number of taxa) be used for estimating the same threshold to better understand what aspect of the ecosystem changes at which TP concentration. For the same reason, the definite total phosphorus threshold for the Everglades ecosystem should be estimated using response variables from multiple trophic levels.

Three most commonly seen types of data are: (1) continuous data whose distribution can be approximated by a normal distribution, (2) counts data that can be approximated by a Poisson distribution, and (3) binomial or multinomial data (e.g. presence/absence of a particular species and counts of several species found in a site). Both the nonparametric and the Bayesian methods can be applied to all three types of data. We presented the normal and binomial response variable cases in the Everglades example. When the response variable is counts, a log transformation should be performed before applying the nonparametric method. The Bayesian changepoint method for counts data can be found in Raftery and Akman (1986). The application of Bayesian hierarchical modeling approach to

multinomial or categorical data can be found in Qian et al. (2003).

Because TP is the limiting nutrient in the Everglades ecosystem, it is expected that the ecosystem will respond to the elevated TP concentration. Because BCD summarizes the community pattern of the macroinvertebrate assemblage data, the variable is a ranking of some sort. A clear and well-defined changepoint along the TP gradient (Figs. 1 and 2 and Table 2) indicates the existence of a TP threshold and supports the theory of a phosphorus assimilative capacity for wetlands as proposed by Richardson and Qian (1999).

Acknowledgements

References

Adams, S.M., Greeley, M.S., 2000. Ecotoxicological indicators of water quality: using multi-response indicators to assess the health of aquatic ecosystems. Water Air Soil Pollut. 123, 103–115.

Aldenberg, T., Janes, J.H., Kramer, P.R.G., 1995. Fitting the dynamic model PClake to a multi-lake survey through Bayesian statistics. Ecol. Model. 78, 83–99.

Barbour, M.T., Gerritsen, J., Snyder, B.D., Stribling, J.B., 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates, and fish. U.S. Environmental Protection Agency, Office of Water, Washington, DC. EPA 841-0B-99-002.

Bartell, S.M., Gardner, R.H., O'Neill, R.V. (Eds.), 1992. Ecological Risk Estimation. Lewis Publishers, Chelsea, MI.

Borsuk, M.E., Higdon, D., Stow, C.A., Reckhow, K.H., 2001. A Bayesian hierarchical model to predict benthic oxygen demand from organic matter loading in estuaries and coastal zones. Ecol. Model. 143, 165–181.

Bray, J.R., Curtis, J.T., 1957. An ordination of the upland forest communities of southern Wisconsin. Ecol. Monogr. 27, 325–349.

Breiman, L., Friedman, J.H., Olshen, R., Stone, C.J., 1984. Classification and Regression Trees. Wadsworth International Group, Belmont, CA.

Buzzelli, C.P., Childers, D.L., Dong, Q., Jones, R.D., 2000. Simulation of periphyton phosphorus dynamics in Everglades National Park. Ecol. Model. 134, 103–115.

Cao, Y., Williams, D.D., Williams, N.E., 1998. How important are rare species in aquatic community ecology and bioassessment? Limnol. Oceanogr. 43, 1403–1409.

Carlin, B.P., Gelfand, A.E., Smith, A.F.M., 1992. Hierarchical Bayesian analysis of changepoint problems. Appl. Statist. 41, 389–405.

Casella, G., George, E., 1992. Explaining the Gibbs sampler. The Am. Statist. 46, 167–174.

Clarke, K.R., 1993. Nonparametric multivariate analyses of changes in community structure. Aust. J. Ecol. 18, 117–143.

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman and Hall, London.

Faith, D.P., Norris, R.H., 1989. Correlation of environmental variables with patterns of distribution and abundance of common and rare freshwater macroinvertebrates. Biol. Conserv. 50, 77–98.

Faith, D.P., Minchin, P.R., Belbin, L., 1987. Compositional dissimilarity as a robust measure of ecological distance. Vegetatio 69, 57–68.

Findlay, C.S., Zheng, L., 1999. Estimating ecosystem risks using cross-validated multiple regression and cross-validated holographic neural networks. Ecol. Model. 119, 57–72.

Flora, M.D., Walker, D.R., Scheidt, D.J., Rice, R.G., Landers, D.H., 1988. The response of the Everglades marsh to increased nitrogen and phosphorus loading: Part I. Nutrient dosing, water chemistry, and periphyton productivity. National Park Service, South Florida Research Center, Everglades National Park, Homestead, FL.

Fore, L.S., Karr, J.R., Wisseman, R.W., 1996. Assessing invertebrate responses to human activities: evaluating alternative approaches. J. North Am. Benthol. Soc. 15, 212–231.

Gauch, H.G., 1982. Multivariate Analysis in Community Ecology. Cambridge University Press, Cambridge, UK.

Gelfand, A.E., Smith, A.F.M., 1990. Sampling-based approaches to calculating marginal densities. J. Am. Statist. Assoc. 85 (410), 398–409.

Gelfand, A.E., Hills, S.E., Racine-Poon, A., Smith, A.F.M., 1990. Illustration of Bayesian inference in normal data models using Gibbs sampling. J. Am. Statist. Assoc. 85 (412), 972–985.

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecol. Model. 135, 147–186.

Karr, J.R., Chu, E.W., 1997. Biological monitoring and assessment: using multimetric indexes effectively. University of Washington, Seattle, WA. EPA 235-R97-001.

King, R.S., 2001. Dimensions of invertebrate assemblage organization in a phosphorus-limited Everglades landscape. Ph.D. disseration. Duke University, Durham, NC.

King, R.S., Richardson, C.J., in press. Macroinvertebrate and fish responses to experimental P additions in Everglades sloughs. In: Richardson, C.J. (Ed.), The Everglades Experiments: Lessons for Ecosystem Restoration. Springer-Verlag, New York.

Legendre, P., Anderson, M.J., 1999. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. Ecol. Monogr. 69, 1–24.

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

Lemly, A.D., Richardson, C.J., 1997. Guidelines for risk assessment in wetlands. Environ. Monit. Assess. 47, 117–134.

Loiselle, S., Carpaneto, G.M., Hull, V., Waller, T., Rossi, C., 2000. Feedback analysis in reserve management: studying local myths using qualitative models. Ecol. Model. 129, 25–37.

McCune, B., Dey, J.P., Peck, J.E., Cassell, D., Heiman, K., Will-Wolf, S., Neitlich, P.N., 1997. Repeatability of community data: species richness versus gradient scores in large-scale lichen studies. The Bryologist 100, 40–46.

Moreno-Grau, S., Garcia-Sanchez, A., Moreno-Clavel, J., Serrano-Aniorte, J., Moreno-Grau, M.D., 1996. A mathematical model for waste water stabilization ponds with macrophytes and microphytes. Ecol. Model. 91, 91–99.

Omlin, M., Reichert, P., 1999. A comparison of techniques for the estimation of model prediction uncertainty. Ecol. Model. 115, 45–59.

Pan, Y., Stevenson, R.J., Vaithiyanathan, P., Slate, J., Richardson, C.J., 2000. Changes in algal assemblages along observed and experimental phosphorus gradients in a subtropical wetland, USA. Freshwater Biol. 44, 339–353.

Prato, T., 2000. Multiple attribute Bayesian analysis of adaptive ecosystem management. Ecol. Model. 113, 181–193.

Qian, S.S., Anderson, C.W., 1999. Exploring factors controlling the variability of pesticide concentrations in the Willamette River Basin using tree-based models. Environ. Sci. Technol. 33, 3332–3340.

Qian, S.S., Richardson, C.J., 1997. Estimating the long-term phosphorus accretion rate in the Everglades: a Bayesian approach with risk assessment. Water Resour. Res. 33 (7), 1681–1688.

Qian, S.S., Lavine, M.L., Stow, C.A., 2000. Bayesian noparametric binary response regression models with application in environmental management. Environ. Ecol. Statist. 7, 75–89.

Qian, S.S., Pan, Y., King, R.S., 2003. Soil total phosphorus threshold in the Everglades: a Bayesian changepoint analysis for multinomial response data. Ecological Indicators (to appear).

Raftery, A.E., Akman, V.E., 1986. Bayesian analysis of a Poisson process with a change-point. Biometrika 73, 85–89.

Reichert, P., Omlin, M., 1997. On the usefulness of overparameterized ecological models. Ecol. Model. 95, 289–299.

Richardson, C.J., Qian, S.S., 1999. Long-term phosphorus assimilative capacity in freshwater wetlands: a new paradigm for sustaining ecosystem structure and function. Environ. Sci. Technol. 33 (10), 1545–1551.

Richardson, C.J., Ferrell, G.M., Vaithiyanathan, P., 1999. Nutrient effect on stand structure, resorption efficiency, and secondary compounds in Everglades sawgrass. Ecology 80, 2182–2192.

Richardson, C.J., Vaithiyanathan, P., Stevenson, R.J., King, R.S., Stow, C.A., Qualls, R.G., Qian, S.S., 2000. The ecological basis for a phosphorus (P) threshold in the Everglades: directions for sustaining ecosystem structure and function. Duke University, Durham, NC. Duke Wetland Center Publication 00-02.

Rosenberg, D.M., Resh, V.H. (Eds.), 1993. Freshwater Biomonitoring and Benthic Macroinvertebrates. Chapman and Hall, New York.

Shukla, V.P., 1998. Modeling the dynamics of wetland macrophytes: Keoladeo National Park Wetland, India. Ecol. Model. 109, 99–114.

Smith, A.F.M., 1975. A Bayesian approach to inference about a change-point in a sequence of random variables. Biometrika 62, 407–416.

Smith, A.F.M., Roberts, G.O., 1993. Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. J. Roy. Stat. Soc. B 55 (1), 3–23.

Spiegelhalter, D., Thomas, A., Best, N., 2000. WinBUGS: Version 1.3 User Manual. http://www.mrc-bsu.cam.ac.uk/bugs.

Steinberg, L.J., Reckhow, K.H., Wolpert, R.L., 1997. Characterization of parameters in mechanistic models: a case study of a PCB fate and transport model. Ecol. Model. 97, 35–46.

Stephens, D.A., 1994. Bayesian retrospective multiple-changepoint identification. Appl. Statist. 43, 159–178.

Steward, K.K., Ornes, W.H., 1975a. Assessing a marsh environment for wastewater renovation. J. Water Pollut. Control Federation 47, 1880–1891.

Steward, K.K., Ornes, W.H., 1975b. The autecology of sawgrass in the Florida Everglades. Ecology 56, 162–171.

Suter, G.W. 1993. Ecological Risk Assessment. Lewis Publishers, Chelsea, MI.

Suter, G.W., 1996. Abuse of hypothesis testing statistics in ecological risk assessment. Hum. Ecol. Risk Assess. 2, 331–347.

Swift, D.R., Nicholas, R.B., 1987. Periphyton and water quality relationships in the Everglades Water Conservation Areas, 1978–1982. South Florida Water Management, West Palm Beach, FL.

Venables, W.N., Ripley, B.D., 1994. Modern Applied Statistics with S-Plus. Springer, New York.

USEPA (United States Environmental Protection Agency), 1998a. National strategy for the development of regional nutrient criteria. Office of Water, Washington, DC. EPA 822-R-98-002.

USEPA (United States Environmental Protection Agency), 1998b. Guidelines for ecological risk assessment. Risk Assessment Forum, Washington, DC.

Wang, N., Mitsch, W.J., 2000. A detailed ecosystem model of phosphorus dynamics in created riparian wetlands. Ecol. Model. 126, 101–130.