

Soil total phosphorus threshold in the Everglades: a Bayesian changepoint analysis for multinomial response data

Song S. Qian^{a,*}, Yangdong Pan^b, Ryan S. King^c

^a The Cadmus Group Inc., 6330 Quadrangle Drive, Suite 180, Chapel Hill, NC 27517, USA

^b Environmental Sciences and Resources, Portland State University, Portland, OR 97207, USA

^c Smithsonian Environmental Research Center, 647 Contee's Wharf Road, Box 28, Edgewater, MD 21037, USA

Abstract

The Everglades is a historically phosphorus limited freshwater wetland ecosystem. Increasing the supply of phosphorus will lead to changes in the ecosystem. As a cumulative indicator of phosphorus loading to the wetland system, the top 10 cm soil TP concentration is often used to indicate eutrophication. The current study used multivariate species composition data to find the TP threshold, above which species compositional changes are expected. A Bayesian changepoint model was developed in this study. The statistical model can directly use species assemblage data as a measure of structural change. Using diatom species and macroinvertebrate species assemblage data collected from the field as well as from a mesocosm experiment, the soil TP threshold was estimated to be around 400 mg/kg.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Everglades; Species composition data; Soil

1. Introduction

Because shifts in species composition are often caused by changing environmental conditions, ecologists have long considered relative abundance of different species an important measure for evaluating ecosystems (Odum, 1985; Schindler, 1990; Stoemer and Smol, 1999). For example, an anthropogenic disturbance, such as a slight to moderate increase in nutrient supply, can alter the relative abundance of periphyton (Pan et al., 2000). Under intense human-related stress, significant compositional and functional changes in biotic communities may occur, which may result in the interruption or loss of some or all ecological functions in the ecosystem. Odum (1985) sug-

gested and Schindler (1990) experimentally demonstrated that many measures of ecological function such as nutrient cycling, energy flow, and succession can be severely disrupted as anthropogenic stress intensifies. For example, excessive nutrient enrichments may result in complete replacement of sensitive and native algal species (Schindler, 1974; Pan et al., 2000), and algal species assemblages, dominated by exotic and non-palatable nuisance species (e.g., toxic blue-green algae), may alter the energy transfer path from primary producers to consumers at higher trophic levels and reduce “food-chain” length in aquatic ecosystems (Schindler, 1990). As a result, shifts in species composition are often used as indicators or early warning signs of worsening environmental conditions.

Studies of how an ecosystem responds to human disturbances can have important management implications, such as the establishment of water-quality or emission standards for particular geographic regions

* Corresponding author. Tel.: +1-919-403-5105;

fax: +1-919-401-5867.

E-mail address: sqian@cadmusgroup.com (S.S. Qian).

(Adams and Greeley, 2000). One common approach to estimating environmental criteria is to examine the spatial changes of selected population, community, or ecosystem attributes along a gradient of environmental conditions (e.g., Karr and Chu, 1997). Using this approach, ecologists can quantify the environmental threshold, which is an important environmental variable given that ecological attributes often show little change until a critical environmental value, or threshold, is reached (Fore et al., 1996; Richardson and Qian, 1999). In this way, quantitative descriptions of exposure-response relationships can be very useful to support the development of numerical environmental criteria (Suter, 1993; USEPA, 1998a).

Statistical analysis of compositional data is a difficult subject. Aitchison (1982, 1986) developed the logit normal distribution as a framework for compositional data analysis. The logit normal distribution is, however, difficult to interpret in ecological terms. The Bayesian hierarchical extension by Billheimer et al. (2001) used a conditional multinomial distribution on top of the logit normal distribution, such that internal parameter structure changes can be transmitted to changes in relative abundance of species. Both the logit normal distribution and its Bayesian hierarchical extension can be used to describe differences in species composition at different locations. The changepoint method we proposed directly links the species composition changes to a dominant environmental gradient, with an objective for estimating the environmental threshold.

In a previous study, we presented a Bayesian changepoint analysis for estimating environmental thresholds using univariate response variables such as percentage of pollution sensitive species (Qian et al., 2003). Because the method in our previous study was limited to univariate response variable, species composition data must be processed to produce an index of some sort. By reducing the dimension of species composition data, loss of information is inevitable. In this paper, we present a Bayesian changepoint analysis method that uses species composition as the response variable. Four data sets were used to illustrate the method: (1) macroinvertebrate species compositions measured along a phosphorus gradient created in a mesocosm experiment conducted in an Everglades wetland; (2) macroinvertebrate species compositions measured along a phosphorus gradient

in the northern Everglades; (3) periphytic algae (algae attached on aquatic plants) species composition data collected near the transition zone of the phosphorus gradient in the northern Everglades; and (4) periphytic algae species composition data with sampling sites covering the entire phosphorus gradient in the northern Everglades. The rest of the paper presents the study area and data sets, the statistical methods, the results, and a discussion.

2. Method

2.1. Study area and data

The Florida Everglades is a phosphorus limited ecosystem (e.g., Richardson and Qian, 1999; Steward and Ornes, 1975a,b; Swift and Nicholas, 1987; Flora et al., 1988). Excessive inputs of phosphorus to the Everglades led to the 1994 Everglades Forever Act (EFA, Florida Administrative Code, 17-302.530), which mandated that a numerically defensible water-column total phosphorus (TP) standard be established to ensure a balance of flora and fauna.

Four data sets, collected from the Water Conservation Area-2A (WCA2A) in the northern Everglades, are used in this paper. WCA2A is a 547 km² diked marsh with a mosaic of sawgrass (*Cladium jamaicense*) prairies and open water sloughs. It has been receiving agricultural runoff from the north for decades. As a result, a north to south phosphorus gradient has developed (Craft and Richardson, 1993; Reddy et al., 1993; Urban et al., 1993).

The first data set consists of macroinvertebrate species composition data collected from a mesocosm dosing study conducted by the Duke University Wetland Center in the southern end of WCA2A (Richardson et al., 2000; Pan et al., 2000). Qian et al. (2003) estimated TP threshold for the Everglades using the same data set with a single response variable model using a derived dissimilarity index and the percentage of phosphorus tolerant species.

The second data set includes macroinvertebrate species composition data collected along three north to south transects directly south of the three major water inlets in WCA2A. Details of data collection and initial analysis can be found in King and Richardson (2002).

The third and fourth data sets are diatom species composition data collected from WCA2A separately by Pan et al. (2000) and Cooper and Goman (2002). The Pan et al. (2000) data include algae and sediment total phosphorus from 32 slough sites near the southern end of the north–south phosphorus gradient in WCA2A sampled from 6 to 8 April 1995 (see Pan et al., 2000 for details of data collection and analysis). A total of 45 diatom species were identified and recorded in the Pan et al. (2000) data. The Pan et al. (2000) data were collected in the transitional zone of the phosphorus enriched and unenriched areas of WCA2A. The Cooper and Goman (2002) data were collected in 1996 from 31 sites covering the entire WCA2A. There were 90 diatom species that can be positively identified in the data set. The number of diatom species included in the Cooper and Goman (2002) study is twice as many as in the Pan et al. (2000) study. However, the site locations were more carefully chosen in the Pan et al. (2000) study to focus on the impact of phosphorus.

2.2. Statistical method

We discuss the discrete parameter changepoint problem. Let $y_1, \dots, y_n$ be the sequence of the ecological response variables observed along the ordered environmental gradient $x_1, \dots, x_n$. A changepoint problem is to find r ($1 \leq r \leq n$) that separates the response variable into two groups: $y_1, \dots, y_r$ and $y_{r+1}, \dots, y_n$, each with distinct characteristics such as the relative species abundance. The corresponding value of the environmental variable x_r is the threshold. Another type of changepoint problem is discussed in Qian and Richardson (1997).

Under a Bayesian framework, we make specific probabilistic assumptions about the ecological response variable. Specifically, we assume that the response variable values, $y_1, \dots, y_n$, collected from the n sites along the gradient of interest, are random samples from the sequence of random variables $Y_1, \dots, Y_n$. 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 $Y_1, \dots, Y_n$ have a change point at r ($1 \leq r \leq n$) if

$$\begin{aligned} Y_1, \dots, Y_r &\sim \pi(Y_i|\theta_1) \\ Y_{r+1}, \dots, Y_n &\sim \pi(Y_i|\theta_2) \end{aligned} \quad (1)$$

where π represent a generic probability density function. The general setup of a Bayesian analysis can be described as follow.

First, the likelihood function $L(Y; r)$ under the model of Eq. (1) is:

$$L(Y; r) = \prod_{i=1}^r \pi(Y_i|\theta_1) \prod_{i=r+1}^n \pi(Y_i|\theta_2) \quad (2)$$

With the prior density of $r: \pi(r)$, the posterior density of $r|Y$ becomes:

$$\pi(r|Y) = \frac{L(Y; r)\pi(r)}{\sum_{r=1}^n L(Y; r)\pi(r)} \quad (3)$$

Features of the posterior distribution (3) can be easily estimated, and the posterior odds of no change ($P(r = n|Y)/\{1 - P(r = n|Y)\}$) is often used to determine whether or not a change has occurred.

For the species composition data, we assume a multinomial distribution. That is, $y_i = \{y_{i1}, \dots, y_{ik}\}$ is a vector of observed counts of individual species from site i . The parameter of interest is the relative abundance. When the environmental variable is less than the threshold ($i \leq r$), the relative abundance vector is $\theta_1 = \{p_{11}, \dots, p_{1k}\}$ and the relative abundance vector is $\theta_2 = \{p_{21}, \dots, p_{2k}\}$ for $i > r$. The number of species included (the dimension of the multinomial distribution) is k . We further assume that θ_1 is a priori independent of θ_2 , and the prior distribution of parameters θ_1, θ_2, r is:

$$\pi(\theta_1, \theta_2, r) \propto \pi(\theta_1)\pi(\theta_2). \quad (4)$$

The prior distribution on r is assumed to be a uniform distribution over all possible values of i . Ideally, the prior distribution of r should be chosen based on ecological knowledge. In addition, TP threshold is a real value, not an index to the ordered values. A continuous prior distribution on the threshold would be closer to the reality than the uniform distribution on the index. We choose the the uniform distribution on the index for two reasons. First, there is no consensus on what

constitutes a change in the species assemblage. Different interpretation may lead to different prior belief. Consequently, a uniform distribution would be a compromise. Second, a discrete distribution of r makes computation convenient. Simulation studies we conducted showed no discernable difference in the posteriors of r estimated from models using continuous and discrete priors.

We use the Dirichlet distribution to describe the prior distribution of θ_1 and θ_2 :

$$\pi(\theta_1) \propto \prod_{j=1}^k p_{1j}^{\alpha_{1j}-1},$$

and

$$\pi(\theta_2) \propto \prod_{j=1}^k p_{2j}^{\alpha_{2j}-1}.$$

The likelihood of observing data $Y = \{y_{ij}\}$ is

$$\begin{aligned} L(Y, \theta_1, \theta_2, r) &\propto \prod_{i=1}^r \left(\prod_{j=1}^k p_{1j}^{y_{ij}} \right) \times \prod_{i=r+1}^n \left(\prod_{j=1}^k p_{2j}^{y_{ij}} \right) \\ &= \prod_{j=1}^k p_{1j}^{\sum_{i=1}^r y_{ij}} \times \prod_{j=1}^k p_{2j}^{\sum_{i=r+1}^n y_{ij}} \quad (5) \end{aligned}$$

The joint distribution of θ_1, θ_2, r is the product of the prior (Eq. (4)) and the likelihood (Eq. (5)):

$$\begin{aligned} \pi(\theta_1, \theta_2, r|Y) &\propto \prod_{j=1}^k p_{1j}^{\sum_{i=1}^r y_{ij} + \alpha_{1j} - 1} \times \prod_{j=1}^k p_{2j}^{\sum_{i=r+1}^n y_{ij} + \alpha_{2j} - 1} \quad (6) \end{aligned}$$

The marginal posterior distribution of r is

$$\pi(r|Y) = \iint \pi(\theta_1, \theta_2, r|Y) d\theta_1 d\theta_2 \propto \begin{cases} \frac{\prod_{j=1}^k \Gamma(\alpha_{1j} + \sum_{i=1}^r y_{ij})}{\Gamma\{\sum_{j=1}^k (\alpha_{1j} + \sum_{i=1}^r y_{ij})\}} \times \frac{\prod_{j=1}^k \Gamma(\alpha_{2j} + \sum_{i=r+1}^n y_{ij})}{\Gamma\{\sum_{j=1}^k (\alpha_{2j} + \sum_{i=r+1}^n y_{ij})\}}, & \text{for } r < n \\ \frac{\prod_{j=1}^k \Gamma(\alpha_{1j} + \sum_{i=1}^n y_{ij})}{\Gamma\{\sum_{j=1}^k (\alpha_{1j} + \sum_{i=1}^n y_{ij})\}} \times \frac{\prod_{j=1}^k \Gamma(\alpha_{2j})}{\Gamma\{\sum_{j=1}^k (\alpha_{2j})\}}, & \text{for } r = n \end{cases}$$

This integral is evaluated using the fact that the two parts in the posterior distribution resemble the density function of a Dirichlet distribution.

The marginal posterior distributions of θ_1 and θ_2 can be estimated using the Gibbs sampler since their conditional posterior distributions are easy to obtain:

$$\pi(\theta_1|\theta_2, r) \propto \prod_{j=1}^k p_{1j}^{\sum_{i=1}^r y_{ij} + \alpha_{1j} - 1},$$

which is a Dirichlet distribution; and

$$\pi(\theta_2|\theta_1, r) \propto \prod_{j=1}^k p_{2j}^{\sum_{i=r+1}^n y_{ij} + \alpha_{2j} - 1},$$

which is also a Dirichlet distribution.

Parameters of the Dirichlet prior distributions can be selected based on our knowledge of the species distribution or be selected to make the prior near non-informative. There are three options for selecting a noninformative Dirichlet prior distribution. The most obvious one is to set $\alpha_{ij} = 0$, an improper prior distribution. The posterior distributions of r, θ_1 , and θ_2 will be proper distributions only when all species appeared in the data at least once. The second option is to set $\alpha_{ij} = 0.5$ (the Jeffreys' prior, [Jeffreys, 1961](#)) and the third option is to set $\alpha_{ij} = 1$ (the "natural" uniform density used by Bayes and Laplace, [Berger, 1985](#)). Good arguments can be made for any of the three options ([Ripley, 1996](#)). When the sample size is large, the choice of priors will be less important. We used the Jeffrey's prior.

3. Results

The estimated changepoint distributions from the data sets show a clear and distinct change point along the TP gradient ([Fig. 1](#)). The mean change point values ranging from 367 to 452 mg/kg ([Table 1](#)). Comparing the 95% credible intervals for the four data sets ([Fig. 2](#)), we note that the interval for the mesocosm macroinvertebrate data is wider (more uncertain) than the same for the gradient data. The credible interval for the [Pan et al. \(2000\)](#) data is narrower than the same for the [Cooper and Goman \(2002\)](#) data, even though the latter had twice as many diatom species. The field macroinvertebrate species data also produce

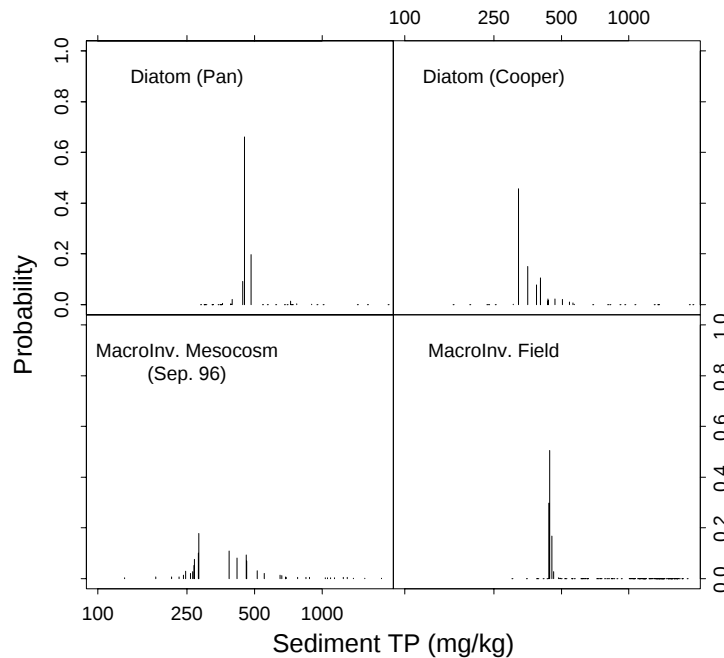


Fig. 1. The probability distributions of soil TP thresholds. The heights of the lines represents the probability the corresponding soil TP value being the threshold.

very narrow credible interval for the threshold, as the samples were collected from a well established phosphorus gradient shared by many research groups.

When comparing the relative proportion of each species before and after the changepoint (i.e., the soil TP less than or larger than the threshold), only a handful of species are found to have substantial changes (Fig. 3). For example, the relative abundances of the three species on the right end of the x -axis (the dominant species in P-limited Everglades habitats or “native species”) are much smaller when the soil TP concentration is above the threshold level,

while the relative abundance of *Nitzschia frustulum* (a well known “eutrophic” tolerant diatom species) increased dramatically. As a result, it is conceivable that a limited number of “indicator” species may be used for indicating eutrophication. However, as many have pointed out (e.g., King and Richardson, 2002), eliminating “rare” species may result in loss of information. We tested our method for its sensitivity to the number of species by removing rare species (defined as occurring less than 1, 5, and 10% in each observation). For all four data sets, removing species occurring with less than 1% in each observation resulted in

Table 1
Data sets used in the paper

Study ID	Reference	Date collected	Species	Location
1.1	King and Richardson (2002)	January 1997	Macroinvertebrates	Mesocosm
1.2		February 1998		
1.3		September 1996		
1.4		September 1998		
2	Pan et al. (2000)	1995	Diatom	Field
3	Cooper and Goman (2002)	1996	Diatom	
4	King and Richardson (2002)	1998	Macroinvertebrates	

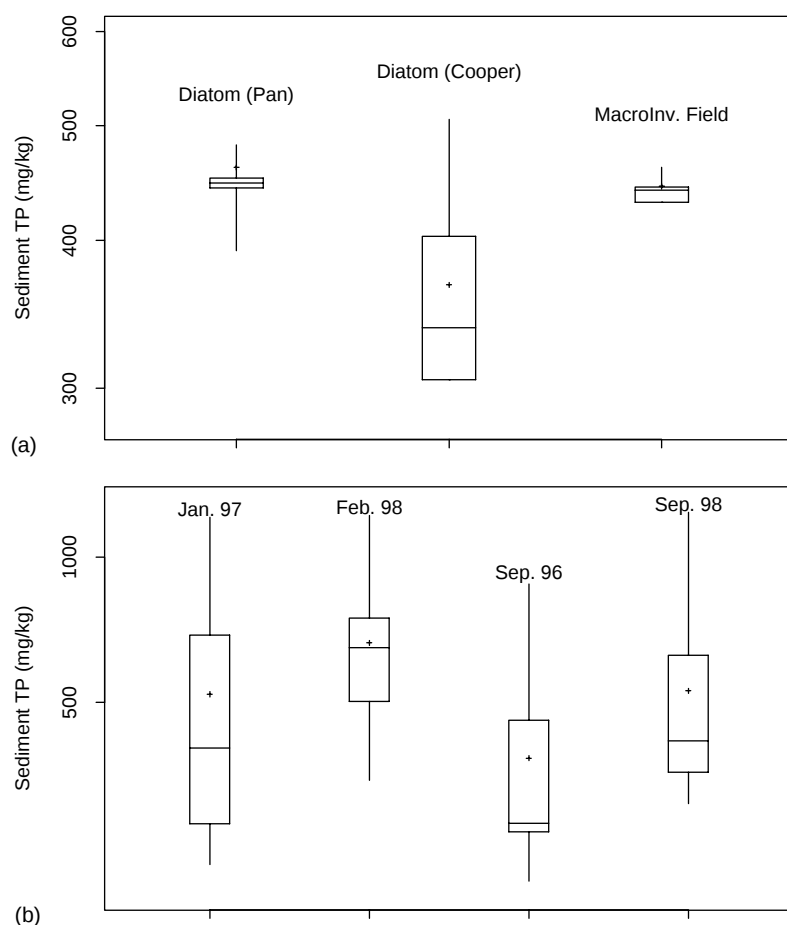


Fig. 2. Summary statistics plots of the soil TP threshold distributions: (a) the soil TP threshold distributions from the three field studies; (b) the soil TP threshold distributions from the mesocosm study data collected in four dates.

insignificant or no changes. When removing species that occurred with values less than 5% in each observation, the TP threshold for the Pan et al. (2000) data changed significantly while the other three showed slight changes. All show significant changes in the estimated thresholds when removing species occurring less than 10% in each observation.

4. Discussion

In a separate study, Qian et al. (2002) reported the historical soil TP in the Everglades to be 472 mg/kg with a 95% credible interval of 250, 790 mg/kg, inferred by using the fossil diatom species composition

collected by Cooper and Goman (2002) from both the enriched and unenriched areas in WCA2A. The historical soil TP concentrations were about the same in both the phosphorus enriched and unenriched areas. The TP thresholds estimated in the current study are similar to the historical soil TP concentrations (see Table 2). This similarity indicates that site-specific fossil diatom species responses to historical phosphorus enrichment are similar to the responses of diatom species along the current TP gradient in the Everglades. Based on diatom species responses along the current TP gradient, a diatom species transfer function model can be developed and used for reconstructing site-specific historical changes of TP in the Everglades.

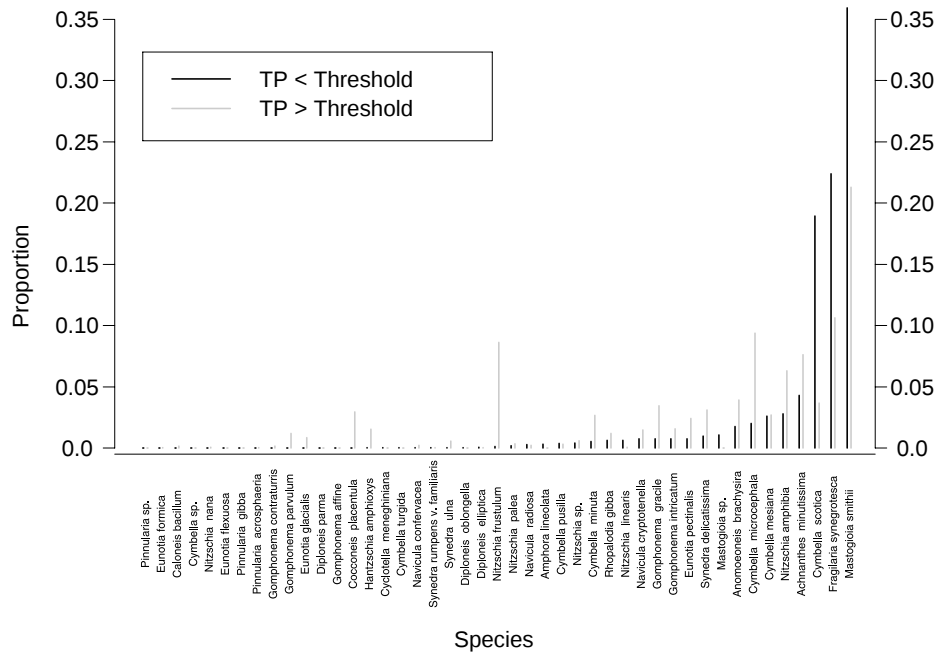


Fig. 3. The average relative abundances of the identified diatom species associated with soil TP concentrations below and above the estimated TP threshold values. Results shown are derived from the Pan et al. (2000) data.

The statistical method presented here assumes a multinomial distribution of the species counts. As a result, the actual counts are needed for model fitting. The conventional practice in many ecological studies is to report only the relative abundance (percentage) of each species. Although the percentage is an unbiased estimate of a species relative abundance, the variance of the relative abundance is determined by the total number counted. Therefore, the reporting of only the percentage could lead to loss of information, unless the total numbers of species counted in all samples are the same and reported.

Table 2
Estimated soil TP thresholds (mg/kg) from the four data sets

Study ID	2.5%	Median	Mean	97.5%
1.1	230.7	401.8	519.7	1210.4
1.2	345.0	649.7	664.6	1219.5
1.3	212.9	280.8	383.2	880.3
1.4	308.5	416.4	528.7	1240.2
2	392.0	447.2	460.9	481.5
3	305.0	337.5	366.8	506.0
4	430.9	441.0	444.4	461.1

The method developed in this paper assumes an abrupt change in the species composition along an environmental gradient. Ecological changes are not always abrupt (Muradian, 2001). Because the change-point is estimated as a probability distribution, our method can be used to show a zone of change reflected by, e.g., the 95% credible interval. Our experience (from simulation studies) show that (1) when the underlying pattern is an abrupt change, the model will result in a posterior changepoint distribution that is narrowly concentrated around the threshold; and (2) a wide-spreading posterior changepoint distribution may indicate a gradual change or other noise (e.g., secondary gradient). Care must be taken when making inference about a wide-spreading posterior changepoint distribution. No statistics method is capable of discerning whether the cause of a wide-spreading posterior changepoint distribution is a gradual change or other noise. In our study, we see highly concentrated posterior changepoint distributions from the Pan et al. (2000) data and the field macroinvertebrate data, which indicate that both diatom and macroinvertebrate species compositions change abruptly as soil

TP increases. We also see widely distributed posterior changepoint distributions from the mesocosm macroinvertebrate data and the Cooper and Goman (2002) data (Fig. 1), which can only suggest additional source of uncertainty in light of the other two data sets.

The ranges of the estimated credible intervals reflect the level of uncertainty about the changepoint. There are many sources of uncertainty, but the following two may be the most influential. First is the source of phosphorus. In the mesocosm, orthophosphate is added as the source of phosphorus; while in the gradient, the source of phosphorus is mainly from the agriculture runoff. When the same mesocosm macroinvertebrate species data were modelled against the water column TP gradient, the estimated credible interval (not shown) is very narrow (another evidence of abrupt change). This shows that the mesocosm soil TP may have not reached an equilibrium with TP loading and thus may not be indicative of phosphorus availability. Second, a secondary gradient (e.g., the SO_4 gradient near the southern end of WCA2A) may affect the result as we move further into the southern end of WCA2A, which may explain why the credible interval for the Pan et al. (2000) data is narrower than the same for the Cooper and Goman (2002) data, even though the latter had twice as many diatom species.

Two additional factors further complicated our study: (1) the difference between mesocosm and natural gradient; and (2) impact of species richness. Mesocosms are commonly used for hypotheses testing in ecological and environmental studies. However, the contribution of mesocosm studies to our understanding of complex ecological systems remains unclear (Daehler and Strong, 1996). Our data showed that the credible intervals for the mesocosm macroinvertebrate data are wider than the same for the gradient data. This may be because that duration of the mesocosm study is relatively short compared to the time needed for the soil TP to reach equilibrium. Pan et al. (2000) compared changes in periphyton assemblages along the experimental phosphorus gradient in the mesocosm and observed phosphorus gradient in the Everglades. They noted that several changes in algal assemblages differed between the mesocosm and Everglades phosphorus gradient. The discrepancies were attributed largely to the difference between observed and experimental phosphorus gradients in habitat complexity, spatial and temporal scales of im-

portant ecological processes, and the frequency and magnitude of environmental changes.

Species richness often increases as a function of sampling area (Rosenzweig, 1996). Increased species richness is typically contributed by so-called rare species. Contribution of rare species to ecological community analysis and bio-assessment remains unclear (see the review by Cao et al., 2001). Gauch (1982) argued that rare species may only add noise to the community analysis and thus should be excluded in the analysis. Other authors indicated that rare species may contain significant information on communities and their responses to environmental disturbances (Faith and Norris, 1989). However, conventional sampling and lab analytic procedures often cannot quantify rare species and their population with high certainty. On the one hand, we see significant changes in our results when species occurring less than 5% in each sample were removed. On the other hand, the Pan et al. (2000) data with only half as many species as the Cooper and Goman (2002) data are most informative (much narrower credible interval). A well-targeted sampling design may be most important.

Acknowledgements

Qian and Pan's work were supported by a US EPA STAR grant (R827898). King's work was supported by the Duke University Wetland Center. E. Conrad Lamon, Craig A. Stow, and Karen Wesson provided constructive comments and suggestions on an earlier version of the manuscript. Comments and suggestions from two referees and the associate editor are greatly appreciated. The data collected by Cooper and Goman (2002) were provided by the Duke University Wetland Center.

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.
- Aitchison, J., 1982. The statistical analysis of compositional data (with discussion). *J. R. Stat. Soc. B* 44, 139–177.
- Aitchison, J., 1986. *The Statistical Analysis of Compositional Data*. Chapman & Hall, New York.

- Berger, J.O., 1985. *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York.
- Billheimer, D., Guttorp, P., Fagan, W.F., 2001. Statistical interpretation of species composition. *J. Am. Stat. Assoc.* 96, 1205–1214.
- Cao, Y., Larsen, D.P., Thorne, R.S.T.-J., 2001. Rare species in multivariate analysis for bioassessment: some considerations. *J. North Am. Benthol. Soc.* 20, 144–153.
- Cooper, S., Goman, M., 2002. Historical changes in water quality and vegetation in WCA2A determined by paleoecological analyses. In: Richardson, C.J. (Ed.), *The Everglades Experiments: Lessons for Ecosystem Restoration*. Springer-Verlag, New York, in press.
- Craft, C.B., Richardson, C.J., 1993. Peat accretion and phosphorus accumulation along a eutrophication gradient in the northern Everglades. *Biogeochemistry* 22, 133–156.
- Daehler, C.C., Strong, D.R., 1996. Can you bottle nature? The roles of microcosms in ecological research. *Ecology* 77, 663–664.
- 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.
- 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.
- Jeffreys, H., 1961. *Theory of Probability*, third ed. Oxford University Press, London.
- Karr, J.R., Chu, E.W., 1997. *Biological monitoring and assessment: using multimetric indexes effectively*. University of Washington, Seattle. EPA 235-R97-001.
- King, R.S., Richardson, C.J., 2002. Evaluating subsampling approaches and macroinvertebrate taxonomic resolution for wetland bioassessment. *J. North Am. Benthol. Soc.* 21 (1), 150–171.
- Muradian, R., 2001. Ecological thresholds: a survey. *Ecol. Econ.* 38, 7–24.
- Odum, E.P., 1985. Trends expected in stressed ecosystems. *BioScience* 35, 419–422.
- Pan, Y., Stevenson, R.J., Vaithyanathan, 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.
- Qian, S.S., King, R.S., Richardson, C.J., 2003. Two statistical methods for the detection of environmental thresholds. *Ecol. Model.* 166, 87–97.
- Qian, S.S., Pan, Y., Cooper, S., 2002. Unpublished manuscript presented at the Annual Meetings of the American Geophysical Union, Spring 2002, Washington, DC.
- Qian, S.S., Richardson, C.J., 1997. Estimating the long-term phosphorus accretion rate in the Everglades: a Bayesian approach with risk assessment. *Water Resources Res.* 33, 1681–1688.
- 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., Vaithyanathan, 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, North Carolina. Duke Wetland Center Publication 00–02.
- Ripley, B.D., 1996. *Pattern Recognition and Neural Networks*. Cambridge University Press, Cambridge, UK.
- Rosenzweig, M.L., 1996. *Species Diversity in Space and Time*. Cambridge University Press, Cambridge, UK.
- Schindler, D.W., 1974. Eutrophication and recovery in experimental lakes: implications for lake management. *Science* 184, 897–899.
- Schindler, D.W., 1990. Experimental perturbations of whole lakes as tests of hypotheses concerning ecosystem structure and function. *Oikos* 57, 25–41.
- 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.
- Stoermer, E.F., Smol, J.P., 1999. *The Diatoms: Applications for the Environmental and Earth Sciences*. Cambridge University Press, Cambridge, UK.
- Suter, G.W., 1993. *Ecological Risk Assessment*. Lewis Publishers, Chelsea, MI, USA.
- 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, USA.
- USEPA (United States Environmental Protection Agency), 1998a. National strategy for the development of regional nutrient criteria. Office of Water, Washington, DC, USA. EPA 822-R-98-002.