Ecological Indicators 49 (2014) 162–168

Contents lists available at ScienceDirect

Ecological Indicators

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

Application of index number theory to the construction of a water quality index: Aggregated nutrient loadings related to the areal extent of hypoxia in the northern Gulf of Mexico

Gerald Whittaker[a], Bradley Barnhart[a,*], Rolf Färe[b], Shawna Grosskopf[b]

[a] National Forage Seed Production Research Center, Agricultural Research Service, USDA, 3450 Campus Way, Corvallis, OR 97330, USA
[b] Department of Economics, Oregon State University, 319D Ballard Extension Hall, Corvallis, OR 97331, USA

ARTICLE INFO

Article history:
Received 28 January 2014
Received in revised form 3 October 2014
Accepted 7 October 2014

Keywords:
Eutrophication
Surface water quality
Gulf hypoxia
Economic productivity
Distance functions

ABSTRACT

Numerous studies have linked individual nutrient loadings from the Mississippi and Atchafalaya Rivers to the growth of the hypoxic, or oxygen depleted, zone in the northern Gulf of Mexico. However, in the discussion of policy to remediate Gulf hypoxia, it is beneficial for stakeholders and policymakers to obtain a single measure for water quality that characterizes information from multiple water pollutants. This study aggregates loadings from six nutrients measured at the entrance to the Gulf of Mexico into a single time-varying index of water quality. The index is constructed using traditional index number theory originating from economic production theory, mainly, Shephard's distance functions calculated using data envelopment analysis (DEA). The methodology is an advance over other index construction schemes because the determined metric weights are endogenous, calculated from the data itself, and do not require external user input. To validate the index, May values of the index are used within a statistical regression model to model the areal extent of Gulf hypoxia using mid-July cruise measurements from 1985 to 2013, excluding 1989 when no cruise data were available. Regression results ($R^2_{adj} = 0.81$) suggest the index is successful at aggregating multiple pollutants into a single measure of water quality and may be useful for tracking their aggregated effect on the growth of the hypoxia area in the northern Gulf of Mexico. Calculation of the water quality index described here is automatic in the sense that no human intervention is required for variable selection, statistical analysis or assignment of weights. This is very useful for specifying a water quality objective in a multiple objective optimization for watershed management.

1. Introduction

Water quality indices have been developed and used to characterize a wide variety of phenomena including drinking water (Beamonte Córdoba et al., 2010), bioassessment (Aguiar et al. 2014; Blanchet et al., 2008; Kanno et al., 2010; Stoddard et al., 2008), fresh water habitat (Pinto et al., 2009; Simaika and Samways, 2011), effect of agriculture on stream water (Justus et al., 2010; Shiels, 2010), river water quality (Feio et al., 2009; Navarro-Llácer et al., 2010), ecological condition (Jordan et al., 2010; Marchini et al., 2009; Seilheimer et al., 2009; Tran et al., 2008) and variable reduction for selection of variables to monitor (Kantoussan et al., 2010). A recent survey of approaches for constructing indices (Bierman et al., 2011) cite the techniques of cluster analysis (Hargiss et al., 2008; Khalil et al., 2010; Styers et al., 2010), principal components analysis, factor analysis (Blocksom and Johnson, 2009; Leunda et al., 2009; Liou et al., 2004; Ou et al., 2010; Tran et al., 2010), discriminant analysis (Feio et al., 2009; Kane et al., 2009), and fuzzy logic (Ghosh and Mujumdar, 2010). This work has been extended to include errors in measurement of the constituent variables used to construct the water quality index (Beamonte Córdoba et al., 2010; Castoldi et al., 2009; Ghosh and Mujumdar, 2010; Sin et al., 2009; Taheriyoun et al., 2010).

In all of the studies cited above, the index is most commonly validated by a statistical comparison with observations of variables that characterize the purpose of the index. The underlying assumption of these water quality indices is that the calculated index is representative of the included components. How such an index should be constructed and what properties the index should have to ensure that it is representative of its included constituents have been the subject of extensive research over the past several hundred years. Original work on index theory began in 1707, where

* Corresponding author. Tel.: +1 541 738 4182.
E-mail address: bradleybarnhart@gmail.com (B. Barnhart).

http://dx.doi.org/10.1016/j.ecolind.2014.10.003
1470-160X/© 2014 Published by Elsevier Ltd.

a market basket of goods was utilized to compare prices among regions (see Diewert and Nakamura, 1993). Development of the consumer price index dates to 1823 (Lowe, 1823), and some of the properties necessary for the construction of a representative index were known by the late nineteenth century (see for example Pierson, 1895). For the theory of index numbers, see Diewert (1987). Nonetheless, few studies today utilize foundational index theory in order to construct water quality indices.

In this paper, we develop a water quality index that is based on the foundation of economic index theory. Our approach follows the work of Malmquist (1953) for constructing a consumer quantity index. In particular we apply Shephard's (1953, 1970) input distance functions to aggregate multiple water quality constituents into a single indicator of water quality. Caves et al. (1982) demonstrated that a Malmquist-type quantity index can be formulated using ratios of Shephard's (1953, 1970) distance functions. In practice, the distance functions are calculated using a linear programming method known as data envelopment analysis (DEA) (Charnes et al., 1978) or activity analysis (von Neumann, 1937). The method ensures that resulting indexes satisfy several mathematical properties including homogeneity, time-reversibility, transitivity, and dimensionality. Also, the weights applied to each metric in the construction of the water quality index are endogenous in our approach. This means they are not externally selected or subject to user bias, but are rather calculated from the data themselves. The technique provides a solution to common discussion on environmental performance indicators that often declare the need for index construction techniques that endogenously calculate metric weights and remove user bias (Bellenger and Herlihy, 2009, 2010; Tran et al., 2008; Tyteca, 1996; Zhou et al., 2007).

Only a small number of studies have used Shephard's (1953, 1970) input distance functions or similar approaches in order to aggregate multiple constituents into an indicator, or to produce environmental quality indexes using production theory (Bellenger and Herlihy, 2009, 2010; Färe et al., 2004; Zhou et al., 2008). Prompted by an environmental indicator review by Tyteca (1996), Färe et al. (2004) demonstrated a DEA approach to calculate an environmental performance index that simultaneously accounted for resources used, good outputs produced and bad outputs emitted. Most recently, Bellenger and Herlihy (2009, 2010) used a distance function approach to construct environmental performance indicators using six macroinvertebrate metrics and compared them to an existing Environmental Protection Agency index of biotic integrity in the Appalachian mountains.

This paper reiterates the utility of distance function techniques to aggregate multiple constituents into a single indicator. We apply these approaches to construct a water quality index utilizing six nutrient loadings supplied to the Gulf of Mexico from the Mississippi and Atchafalaya Rivers. The time-varying index, representing the aggregated effect of nutrient loadings, will then be validated by using it to model the areal extent of Gulf hypoxia.

The area of the summer hypoxic (oxygen $\leq 2\,\mathrm{mg\,l^{-1}}$) region in the northern Gulf of Mexico has been widely studied and publicized for nearly two and a half decades (Forrest et al., 2011; Obenour et al., 2012; Rabalais et al., 2001; Scavia et al., 2013). Since 1997, numerous sources have demonstrated high correlations of the Gulf hypoxic area with nutrient loadings delivered from the Mississippi and Atchafalaya Rivers (Forrest et al., 2011; Greene et al., 2009; Scavia et al., 2003; Turner et al., 2006; Wiseman et al., 1997). To track the variability of Gulf hypoxia, and perhaps the progress in reducing its size, shelfwide measurement cruises have been conducted annually in late July since 1985 (Rabalais et al., 2007). This data set has enabled researchers to model the response of Gulf hypoxia to a variety of forcing variables, including nitrogen and phosphorus concentrations, discharge, and wind speed (Forrest et al., 2011; Scavia and Donnelly, 2007; Scavia et al., 2003, 2004). May nitrogen loadings have been found to have the highest correlation with July hypoxia areas (Forrest et al. (2011) state $R^2 = 0.24$ using 24 measurements). Therefore, the May nitrogen loadings have been used to predict hypoxia areas in the months before each annual July cruise (Forrest et al., 2011; Scavia and Donnelly, 2007; Scavia et al., 2003, 2004, 2013). However, no single model for Gulf hypoxia incorporates more than two water pollutants (total N and total P) simultaneously, despite research that other nutrients may greatly affect eutrophication and hypoxia area, including silica and individual types of N and P (Correll, 1998; Howarth et al., 2011). Also, we are not aware of any previous study that has aggregated multiple nutrient loadings into a composite index of water quality for this region. This paper will construct a monthly water quality index from observations of six nutrient loadings measured at the mouth of the Gulf of Mexico, including dissolved nitrite plus nitrate, total organic nitrogen plus ammonia nitrogen (total Kjeldahl nitrogen), dissolved ammonia, total phosphorous, dissolved orthophosphate, and dissolved silica. The index will be validated by use in a statistical regression in order to model Gulf hypoxia area. Model results will be compared to results from the latest biophysical scenario and forecast model for Gulf hypoxia (Scavia et al., 2013).

In Section 2, the theory and methods necessary to construct the index will be discussed. Section 3 will discuss the data used as well as motivation for the formulation of the statistical regression model. Section 4 will display the resulting index and regression model results while Section 5 will include discussion and conclusions.

2. Theory

In general, an index attempts to compare two vectors x^0 and x^1, where each vector may contain N inputs so that $x^0, x^1 \in \Re^N_+$. The idea of Malmquist (1953) was to introduce a benchmark curve (an indifference curve, in his case) and measure the distance of an input vector from the curve. In this study, we assume that the water characteristics are undesirable, and should be minimized. A two-dimensional case is illustrated in Fig. 1, where we wish to compare vector x^0 of distance OC with vector x^1 of distance OB. The benchmark that we use for comparison is the best practice benchmark (II), where we are "indifferent" as to the choice among points that lie on the curve. Each distance DC and AB is the distance that can be proportionally reduced for each vector to reach the benchmark (II). We conclude for this example that x^1 has better water quality than x^0 because it requires a smaller contraction to reach II than x^0 does.

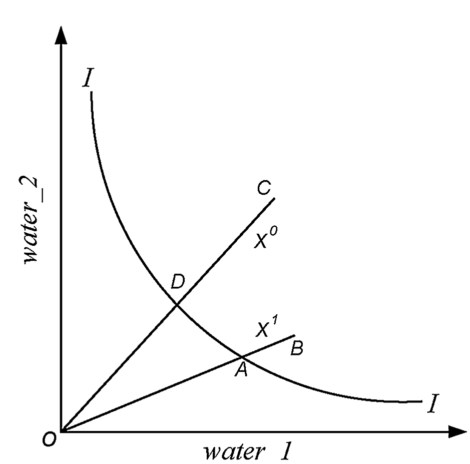

Fig. 1. Demonstration of index construction with two constituents of water quality.