Ecological Indicators 57 (2015) 249–258

Contents lists available at ScienceDirect

Ecological Indicators

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

A dynamic water quality index model based on functional data analysis

Yan Feng[a], Liu Ling[a,*], Li Yanfeng[b], Zhang You[a], Chen Musong[a], Xing Xigang[a]

[a] State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China
[b] Wentian College, Hohai University, Maanshan 243031, China

ARTICLE INFO

Article history:
Received 7 October 2014
Received in revised form 29 April 2015
Accepted 1 May 2015

Keywords:
Dynamic water quality index
Dynamic weighting method
Functional data analysis
Water quality assessment

ABSTRACT

A water quality index (WQI) incorporates two shortcomings in the dynamic assessment of water quality, namely: (1) the sampling time series must be identical for each indicator and no missing data should occur, and (2) stationary weights cannot represent the changes in the pollutant importance. To solve these problems, the present study introduces the functional data analysis method into WQI research and establishes a dynamic WQI (D-WQI) model. D-WQI is a generalization of the conventional WQI. In the D-WQI model, the changes of water quality and pollutant importance are represented in the form of dynamic functional curves. The generation methods of the concentration curves, sub-index curves, dynamic weight curves, and WQI curves are discussed. As an illustration, the D-WQI model is applied in the water quality assessment of the Changjiang River in Sanjiangying in 2012. Result shows that the river can be classified as II (good) throughout the year, which can satisfy the requirement of the Chinese South-to-North Water Diversion Project.

1. Introduction

Water is essential to life, and access to freshwater with a suitable quality is crucial for the sustainable development of societies (Srebotnjak et al., 2012). Therefore, water quality assessment is an important aspect of water resource management (Banerjee and Srivastava, 2009). A water quality index (WQI) is a common tool for the quantitative assessment of water quality. It converts pollutant concentration data into sub-index values and then combines them into a single score (Dobbie and Dail, 2013; Fox, 2014). The major advantage of WQI is its summary and intuitive communication capability. WQI provides a summary of an entire water environment system by integrating the information of various indicators (Dobbie and Dail, 2013), which reflect the management objectives and the environmental characteristics of the research area. Moreover, WQI represents conflicting water quality conditions with a numerical score, which makes it easy for the public and the policy makers to understand the condition of an aquatic environment (Massoud, 2012). Therefore, WQI has wide applications in water quality assessment since its introduction in 1848 (Terrado et al., 2010). For example, Cude (2001) established a WQI model for the water quality evaluation of the Tualatin River. The Canadian Council of Ministers of the Environment (CCME, 2001) released the Canadian Water Quality Index for Canadian jurisdictions to report water quality information to both the management and the public. Lermontov et al. (2009) established a fuzzy WQI method for the environmental analysis of the Iguape River. In addition, water is an indispensable factor in the ecosphere, and thus, WQI is frequently used as a component of the ecological health index. For example, WQI, along with the phytoplankton indices of biotic integrity, benthic indices of biotic integrity, and submerged aquatic vegetation indices, was introduced into the establishment of the Bay Health Index in the health assessment of the Chesapeake Bay (Williams et al., 2009).

Current WQI calculation methods are frequently restricted by two shortcomings in dynamic assessment. (1) The importance of a pollutant frequently varies with time; however, this variation cannot be reflected with a stationary weight. (2) The sampling time series must be identical for each indicator and no missing data should occur; otherwise, aggregation calculation in WQI will be invalid. Sometimes the researchers solve the data missing problem by linear interpolation or depicting the remaining indicators without generating the comprehensive index, however, the former method cannot reflect the nonlinear trend and the latter method cannot assess the entire water environment quantitatively.

* Corresponding author. Tel.: +86 15005182351.
E-mail addresses: yfmilan@163.com (F. Yan), liulinghhu@163.com (L. Liu).

http://dx.doi.org/10.1016/j.ecolind.2015.05.005
1470-160X/

Functional data analysis (FDA), a new statistical method based on the concept that information is represented in the form of functions (Ullah and Finch, 2013), provides a new means to solve the aforementioned problems. FDA is particularly useful in dynamic analysis and highly flexible because the sampling time series for each indicator is not necessary to be identical, and cases with missing data are acceptable. The concept of FDA was introduced by Ramsay in 1997; since then, this method has been widely used in the fields of medicine (Newell et al., 2006; Koh et al., 2007; Barati et al., 2013), ecology (Sánchez et al., 2007; Embling et al., 2012; Ruiz-Medina et al., 2014), and economy (Ramsay et al., 2002; Wang et al., 2008; Müller et al., 2011). In water quality trend analysis, FDA has also demonstrated preliminary applications. For example, Champely and Doledec (1997) used FDA to separate long-term trends from periodic variations of pollutants in the Seine River. Henderson (2006) explored the differences in nutrient and sediment trends in three dams in South East Queensland using FDA. Haggarty et al. (2012) conducted a clustering research on 21 Scottish lakes using FDA. To eliminate the effects of missing values and different sampling times on the dynamic fuzzy analysis of water quality, Feng and Ling (2014) generalized discrete membership vectors into continuous membership curves by introducing FDA into conventional fuzzy synthetic evaluation.

In the present study, a generalization of WQI and FDA theory is first presented. Then, a dynamic WQI (D-WQI) model is established by introducing the FDA method into conventional WQI theory. Lastly, the D-WQI model is applied in the water quality assessment of the Changjiang River in Sanjiangying, which is the water source area of the Eastern Route Project of the South-to-North Water Diversion Project in China.

2. Materials and methods

2.1. Study area

Sanjiangying is a town located in the north bank of the Changjiang River in Eastern China with the geographical coordinates 119°42′53″ E, 32°18′45″ N (Fig. 1). Sanjiangying is the origin of the Eastern Route of the Chinese South-to-North Water Diversion Project. Approximately 8.9 billion m^3 of water are diverted from the Changjiang River in Sanjiangying, through the Beijing–Hangzhou Grand Canal every year, and supplied to the eastern Huang-Huai-Hai Plain (Office of the South-to-North Water Diversion Project Commission of the State Council, 2001). Therefore, the water quality assessment of the Changjiang River in Sanjiangying is significant to the water resource management of the Chinese South-to-North Water Diversion Project.

From the perspective of the entire basin of the Changjiang River, the main pollutant sources are industrial waste and agricultural nonpoint source (NPS) pollution. From the perspective of the study area, the effect of agricultural waste products on water quality is more significant than that of industrial pollution near Sanjiangying, which is being strictly controlled through the South-to-North Water Diversion Project. According to Fan and Zhang (2008), a senior engineer at the Chinese Changjiang Water Resources Commission, the major pollutants in the Changjiang River are permanganate index (COD_{Mn}) and ammonia nitrogen (NH_3-N). Furthermore, the high temperature during summer frequently leads to low concentration of dissolved oxygen (DO), which is harmful to aquatic organisms during this season. Thus, we chose {DO, COD_{Mn}, NH_3-N} as the indicator vector.

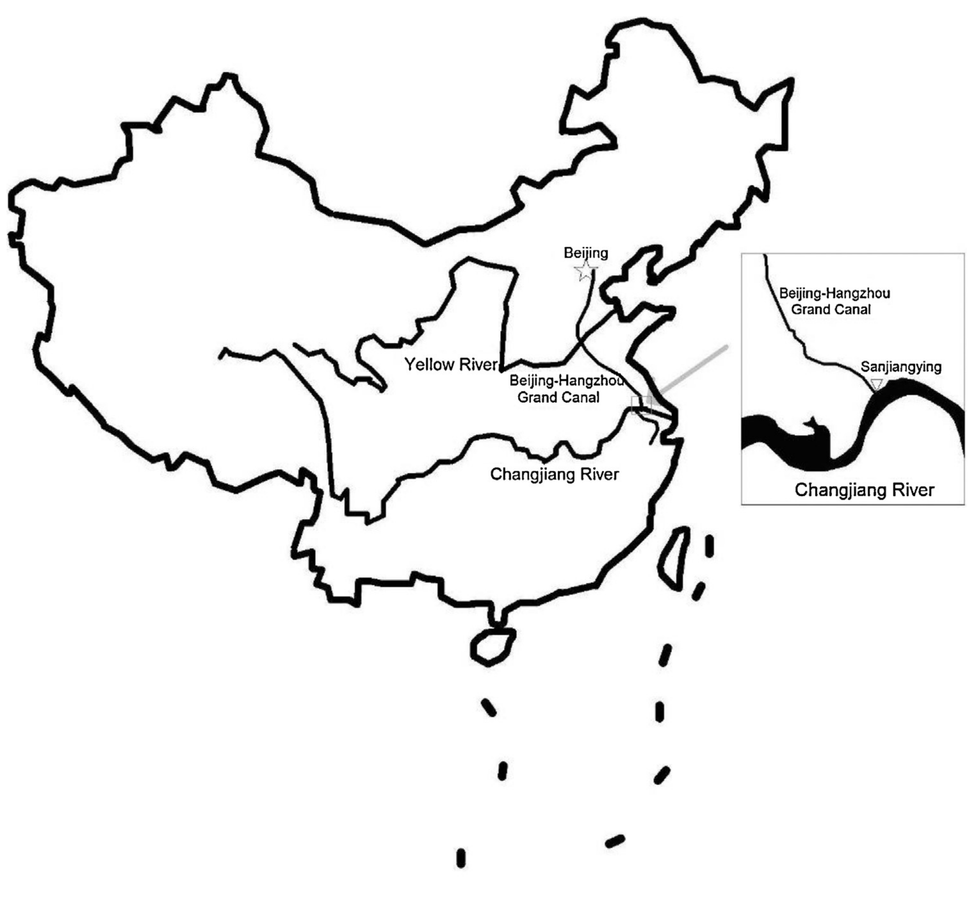

Fig. 1. Map of the study area.