Marine Pollution Bulletin 86 (2014) 68–75

Contents lists available at ScienceDirect

Marine Pollution Bulletin

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

Uncertainty analysis of total phosphorus spatial–temporal variations in the Yangtze River Estuary using different interpolation methods

Ruimin Liu *, Yaxin Chen, Chengchun Sun, Peipei Zhang, Jiawei Wang, Wenwen Yu, Zhenyao Shen

State Key Laboratory of Environment Simulation and Pollution Control, School of Environment, Beijing Normal University, Beijing 100875, China

ARTICLE INFO

Article history:
Available online 10 August 2014

Keywords:
Uncertainty
Spatial–temporal variations
Interpolation method
Total phosphorus
Yangtze River Estuary

ABSTRACT

Interpolation processes and results are generally accompanied by uncertainty which affects the spatial and temporal properties of pollutants. Based on the 4 period sample data of total phosphorus (TP) collected from the Yangtze River Estuary (YRE) in 2010 and 2011, the uncertainty of spatial–temporal variation was analyzed with interpolation methods of inverse distance weighted (IDW), local polynomial interpolation (LPI), ordinary kriging (OK) and disjunctive kriging (DK). The root mean square errors (RMSE) and the mean relative errors (MRE) were used to analyze the accuracy of different interpolation methods. The results showed that the uncertainty of DK was the lowest and the uncertainty of LPI was the highest among the 4 methods. The subtraction results between different interpolation methods showed that there was some distinct area of value in the disparate interval (not in [−0.05, 0.05] (mg/L)) in the 4 seasonal results, which was mainly distributed in the boundary region and around some sample sites. Both standard deviation (SD) and coefficient of variance (CV) in August 2010 were the highest in the 4 seasons and annual mean. The uncertainty may be caused by choice of interpolation methods, spatial data discrepancy and the lack of sample data.

1. Introduction

Yangtze River Estuary (YRE) is on the eastern coast of China, and it extends from Xuliujing in the west to the adjacent sea in the east. YRE is a special region which is a densely populated, important industrial and economic center of China. In YRE, there is average water discharge of approximately 0.9×10^{12} m^3/a and average suspended sediment discharge of approximately 4.2×10^8 t/a from the Yangtze River basin to the East China Sea (Edmond et al., 1985; Dai et al., 2011). In recent decades, owing to the growing anthropogenic influence in the Yangtze River basin and the increasing population near the YRE, the YRE has become an ecologically sensitive region and has received high loading of anthropogenic nutrients (Li et al., 2007; Feng et al., 2008). And the abundant nutrients were mainly from the upstream input and the land near YRE (Yang et al., 2007; Yu et al., 2012).

Any construction in the Yangtze River could have significant implications for the coastal area because 90% of the total river flow to the East China Sea is coming from Yangtze River outflow (Chen et al., 2012). Three-Gorges-Dam (TGD) in Yangtze River is the world's largest hydroelectric project. The response of coastal area to TGD has raised considerable interest and debate (Chai et al., 2009). For instance, completion of the TGD could change the metabolic status of the estuary by cutting off 70% of the downstream transport of organic carbon-containing particles (Chen et al., 2008a). The sediment of downstream has greatly reduced after completion of TGD, and the lower sediment concentration downstream of the dam may facilitate the intensity of riverbed erosion (Chen et al., 2008b).

With the high loading of anthropogenic nutrients and other pollution in YRE, eutrophication has become increasingly serious, and noxious algal blooms have become more frequent in the estuary (Chai et al., 2009). Nutrients, such as phosphorus, played an important role in the water ecological cycle; however, nutrients were also closely related to estuary pollution and may result in severe eutrophication (Smith, 2003; Chai et al., 2006). Research related to estuary nutrient pollution have shown that total phosphorus (TP) was one of the most important indicators of water quality and was the limiting nutrient that restricted microbial production in the estuary (Boynton et al., 1995; Turner et al., 2003). In order to control the eutrophication in the estuary, the characteristics of TP, such as spatial and temporal distribution, should be fully investigated.

Among all the pollutant research in estuaries, interpolation methods were basic tools in analyzing the spatial and temporal

* Corresponding author. Tel./fax: +86 10 5880 0829.
E-mail address: liurm@bnu.edu.cn (R. Liu).

http://dx.doi.org/10.1016/j.marpolbul.2014.07.041
0025-326X/

characteristics of pollutants because they could be used to estimate values of chemical constituents in locations where are not measured. IDW method was applied in YRE of China and the results demonstrated that TP and TN (total nitrogen) pollution in surface water had risen since the year 2003 (Chen et al., 2012). In the Chesapeake Bay area of the US, IDW and ordinary kriging (OK) interpolation methods were applied in the surface water and water quality evaluation was done in the research (Murphy et al., 2010). Different interpolation methods have also been applied in many other fields, such as water, soil and sediment studies (Gotway et al., 1996; Buttner et al., 1998; Zimmerman et al., 1999; Shi et al., 2007; Adhikary et al., 2011; Gu et al., 2012).

Interpolation accuracy is related to the precise definition of the polluted area and its boundaries. However, interpolation processes and results are generally accompanied by uncertainty which affects the spatial and temporal properties of pollutants (Morgan et al., 1990; Goovaerts, 1999; Hijmans et al., 2005). The uncertainty of different interpolation methods can lead to different results. It was observed an increasing awareness of the importance of assessing the uncertainty about the value of target properties at unsampled locations (Yamamoto et al., 2012). In the soil heavy metal study, the uncertainty of different interpolation methods were assessed and compared using particular evaluation parameters and raster subtractions (Xie et al., 2011). Uncertainty was modeled using the E-type mean of total phosphorus (TP) in Florida, USA (Grunwald et al., 2004). However, the uncertainty of different interpolation methods has never been previously studied in YRE.

In this research, based the 4 period sample data in 2010 and 2011, the spatial and temporal variations of TP were interpolated using inverse distance weighted (IDW), local polynomial interpolation (LPI), ordinary kriging (OK) and disjunctive kriging (DK), and the uncertainty associated with the 4 different interpolation methods was analyzed. The results of this study would contribute to further research on the characteristics of total phosphorous in the estuary.

2. Materials and methods

2.1. Study area and sampling

YRE is a large and partially mixed estuary on the eastern coast of China. According to the water discharge characteristics and the marine DEM, an area of 11,000 km^2 was selected as the study area which limited to 120.95°E–122.7°E and 30.8°N–31.8°N. The area can be separated into North Branch and South Branch by Chongming Island, and the South Branch can be separated into North Channel and South Channel by Changxing Island and Hengsha Island.

In this research, 30 sample sites in total were designed in the study area. TP were measured from surface water samples. The sample periods and measurement process were detailed in previous papers (Chen et al., 2012).

2.2. Methodology

2.2.1. Interpolation methods

The interpolation methods used in this study were Inverse distance weighted (IDW) interpolation, local polynomial interpolation (LPI), ordinary kriging (OK) and disjunctive kriging (DK). The ESRI ArcGIS was used as the analysis tool.

IDW interpolation is a commonly used spatial interpolation method and is an accurate and global interpolator (Watson and Philip, 1985). The assigned values to interpolated points are calculated with a weighted average (the inverse of the distance to each sample point) of the values available at the sample points.

$$z(x) = \left[\sum_{i=1}^{n} \frac{z_i}{d_i^u}\right] \Big/ \left[\sum_{i=1}^{n} \frac{1}{d_i^u}\right] \quad (1)$$

where $z(x)$ is the predicted value at an interpolated point, z_i is the i-th sample point, n is the total number of sample points, d_i is the distance between the i-th sample point and the interpolated point, u is the weighting power which may decide the weight affected by distance.

Polynomial interpolation method is a process of finding a formula (often a polynomial) whose graph will pass through a given set of sample points. LPI is a moderately quick deterministic and smooth interpolator, and can be seen as a combination of global polynomial methods and the moving average procedure (Xie et al., 2010; Gribov and Krivoruchko, 2011). There are no assumptions required for the sample data. LPI fits the specified order (zero, first, second, third, etc.) polynomial using all points only within the defined neighborhood (ESRI, 2008).

Kriging is a common interpolation method based on the assumption that the interpolated parameter can be treated as a regionalized variable. The estimator is given by a linear combination of the observed values with weights. Depending on the stochastic properties of random fields, there are different types of kriging, and the type of kriging determines the linear constraint on the weights implied by the unbiased condition (Cressie, 1993; Li and Heap, 2011). The weights of ordinary kriging (OK) are derived from kriging equations using a semivariance function. The parameters of the semivariance function and the nugget effect can be estimated by an empirical semivariance function (Webster and Oliver, 2007).

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2 \quad (2)$$

where $\gamma(h)$ is the semivariance value at distance interval h; $N(h)$ is the number of sample pairs within the distance interval h; and $z(x_i + h)$ and $z(x_i)$ are sample values at two points separated by the distance interval h (Webster and Oliver, 2007).

Unlike OK, disjunctive kriging (DK) is a newly-developed and commonly-applied kriging method used for spatial nonlinear interpolation. It can be suitable for abnormal distribution spatial data. This interpolation tool applies a measure of spatial correlation between pairs of locations describing variance over distance (Von Steiger et al., 1996; Liu et al., 2006; Eldeiry and Garcia, 2012).

$$\begin{cases} \sum_{i=1}^{n} \lambda_i(u)\gamma(u_\alpha - u_i) + \mu = \gamma(u_\alpha - u) \\ \sum_{i=1}^{n} \lambda_i(u) = 1 \end{cases} \quad (3)$$

where λ is the weighting coefficient, μ is the Lagrange parameter, $\alpha = 1, 2 \ldots n$, and $\gamma(u_\alpha - u_i)$ and $\gamma(u_\alpha - u)$ is the semivariogram value of the vector $(u_\alpha - u_i)$ and $(u_\alpha - u)$.

2.2.2. Accuracy analysis methods

The root mean square errors (RMSE) and the mean relative errors (MRE) of each sample sites can be used to evaluate the accuracy of different interpolation methods. Both of them can be calculated as a measure of the accuracy of numerical indicators. Smaller RMSE and MRE values indicate less error (Xie et al., 2011).

$$\mathrm{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} [z^*(x_i) - z(x_i)]^2} \quad (4)$$

$$\mathrm{MRE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{z^*(x_i) - z(x_i)}{z(x_i)} \right| \quad (5)$$