Ecological Indicators 56 (2015) 125–133

Contents lists available at ScienceDirect

Ecological Indicators

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

Prediction of soil heavy metal distribution using Spatiotemporal Kriging with trend model

Yong Yang[a,b,c], Jiaping Wu[d], George Christakos[d,*]

[a] Department of Resources and Environmental Information, College of Resources and Environment, Huazhong Agricultural University, Wuhan, China
[b] Key Laboratory of Arable Land Conservation (Middle and Lower Reaches of Yangtse River), Ministry of Agriculture, China
[c] Department of Geography, San Diego State University, San Diego, CA, USA[1]
[d] Ocean College, Zhejiang University, Hangzhou, Zhejiang, China

ARTICLE INFO

Article history:
Received 18 December 2014
Received in revised form 9 February 2015
Accepted 25 March 2015

Keywords:
Soil heavy metal
Spatiotemporal
Geostatistics
Kriging
Trend

ABSTRACT

Soil heavy metal concentrations exhibit significant space-time trends due to their accumulation along the time axis and the varying distances from the pollution sources. Thus, concentration trends cannot be ignored when performing spatiotemporal soil heavy metal predictions in an area. In this work, datasets were used of soil cadmium (Cd) concentrations in the Qingshan district (Wuhan City, Hubei Province, China) sampled during the period 2010–2014. Spatiotemporal Kriging with four Trend models (STKT) and non-separable space-time correlation was implemented to assimilate multi-temporal data in the mapping of Cd distribution within the contaminated soil area. Soil Cd trends were represented by four different space-time polynomial functions, and a non-separable power function-exponential variogram model of Cd distribution was assumed. Plots of the predicted space-time Cd distributions revealed a marked tendency of the Cd concentrations over time to spread from the southwest part to the entire study area (higher soil Cd concentrations are found in the southwest part of the Qingshan area, whereas the temporal Cd trend is characterized by a constant increase from 2010 to 2014). Thus, the maps indicate that the entire study area is contaminated by Cd, a situation that seems to be stable over time. STKT can reduce prediction errors in practically and statistically significant ways. A numerical comparison of the STKT technique vs. the mainstream Spatiotemporal Ordinary Kriging (STOK) technique showed that STKT can perform better than STOK when the trend model's goodness of fit to the Cd data was satisfactory (producing minimal data fit error statistics), implying that adequate trend modeling is a key issue for space-time prediction accuracy purposes. In particular, quantitative results obtained at the Qingshan region showed that, by incorporating local Cd values and distance-based dependence structures the STKT techniques produced the best prediction error statistics, resulting in considerable prediction error reductions (the level of which depend on the trend model specification; e.g., in the case of STKT with trend model 3 the improvement comparing to STOK was almost 30%). Future studies of Cd contamination in the region (sampling design optimization) can benefit from the results of the geostatistical analysis of the present paper (variogram and trend modeling, etc.).

1. Introduction

Soil heavy metals have been identified as crucial components of the food chain and as important co-factors for many diseases, such as teratogenic, toxic, and cardiovascular diseases (De Vivo et al., 2008; Giaccio et al., 2012; Morton-Bermea et al., 2009). With industrialization and urbanization, the concentrations of heavy metals in soils became increasingly higher (Li et al., 2009; D'Emilio et al., 2013; Fernandez et al., 2007). Thus, the adequate characterization of the composite space-time variability of soil heavy metals is extremely important in a variety of agronomic and environmental activities. Spatiotemporal (ST) Geostatistics offers a variety of modeling and prediction techniques, such as Spatiotemporal Kriging in its various forms (STK, Christakos, 1991, 1992), and spatiotemporal Bayesian Maximum Entropy (BME, Christakos, 1998, 2000). During the last few decades, spatiotemporal techniques have been applied in a variety of earth, environmental, and health studies,

* Corresponding author at: Zhejiang University, Hangzhou 310058, China. Tel.: +86 571 88981701/139 68048183.
E-mail addresses: gchrista@mail.sdsu.edu, gchristakos@zju.edu.cn (G. Christakos).

[1] On Leave.

http://dx.doi.org/10.1016/j.ecolind.2015.03.034
1470-160X/

such as springwater ion processes (Christakos and Bogaert, 1996), thermometric data (Bogaert and Christakos, 1997), soil water content (Snepvangers et al., 2003; Jost et al., 2005), soil salinity content (Douaik et al., 2005), radioactive soil contamination (Savelieva et al., 2005), El Niño effects (Choi et al., 2006), air pollution (Pang et al., 2009; Yu et al., 2009; Kolovos et al., 2010; Akita et al., 2012; Kloog et al., 2014), and infectious disease mapping (Choi et al., 2008; Angulo et al., 2013). However, there has been relatively little concern in the literature as regards the ST modeling and mapping of soil heavy metals. For example, Modis et al. (2013) studied the spatiotemporal distribution of soil heavy metals in an industrial area assuming a spatially homogeneous/temporally stationary variation of heavy metal concentrations. However, due to reasons like their varying distances from the pollution source and their accumulation along the time axis, heavy metal concentrations can exhibit significant space-time trends, so that their variation cannot be represented by a fixed mean value over the study area during the period of interest (i.e., the homogeneity/stationarity assumption is not valid). In view of the above considerations, the analysis and modeling of soil heavy metal trends should be an important component of accurate and physically meaningful space-time prediction.

The present work focuses on soil heavy metal predictions by means of the ST Kriging with a Trend (STKT) technique (Kyriakidis and Journel, 1999; Snepvangers et al., 2003). The dataset used in this study consists of soil cadmium (Cd) concentrations obtained at the Qingshan district (Wuhan City, Hubei Province, China) during the time period 2010–2014. The Cd trend models are determined in terms of space-time polynomial functions (linear, quadratic, etc.) of the relevant coordinates, and a non-separable variogram model of an exponential-power function form is assumed to represent the space-time dependency structure of the Cd distribution. Furthermore, soil Cd concentrations are predicted using ST Ordinary Kriging (STOK, Christakos, 1992; Bogaert, 1996), which ignores the presence of any space-time trends. A numerical comparison of the computationally less complex STOK vs. STKT is presented, and the pros and cons of STKT compared to STOK are discussed in some detail. This kind of comparison between Kriging techniques of different levels of complexity (ordinary, trend-dependent, etc.) is of considerable interest in geostatistical modeling practice, although the existing comparisons are restricted mostly to purely spatial domains (e.g., Ahmed and De Marsily, 1987; Journel and Rossi, 1989; Zimmerman et al., 1999; Eldeiry and Garcia, 2010).

2. Materials and methods

2.1. Study area

The study area is located in the east of the Qingshan district (latitude 30°37′ N, longitude 114°26′ E), which is one of the seven districts of Wuhan city, the capital of Hubei Province (China). Wuhan is the largest city in the middle reach of the Yangtze River in China. There are some very large, heavy industry enterprises, such as the Wuhan Iron & Steel Corporation, the China First Metallurgical Construction Co. Ltd., and the Wuhan Heavy Casting & Forging plant. These factories were built in the center of the district. In the east of the Qingshan district, close to those factories, the land was used to plant crops and vegetables, such as rice, eggplant, cabbage, and cayenne pepper. The regional planting history is about 30–40 years.

2.2. Sample collection and analysis

An elaborate soil study was performed in the area of interest during the month of October, 2010. In this study, 124 topsoil samples (0–20 cm depth) were collected in the area of interest (Fig. 1). We found that there was a serious soil contamination situation. In order to monitor the degree of contamination, we collected topsoil samples from the study area every year, from 2011 until 2014. The numbers of soil samples collected during 2011, 2012, 2013 and 2014 were 45, 48, 55, and 48, respectively (Fig. 1). All sampling points were selected randomly. At each sampling point, 5 sub-samples were randomly collected and then mixed to obtain a composite soil sample. Approximately 1 kg of each soil sample was collected using a wood spade and stored in self-sealing plastic bags. The spade was washed with de-ionized water and wiped dry with paper towels between each use. Any foreign debris in the soil samples was manually removed during sample collection. The coordinates of each sample location were recorded with a help of GPS. All soil samples were air-dried at room temperature, and passed through a 0.15 mm sieve. The prepared soil samples were then stored in polyethylene bottles for analysis.

The soil cadmium (Cd) concentrations of soil samples (digested in Teflon beakers with a mixture of nitric acid, HNO_3, and perchloric acid, $HClO_4$, using a hot plane) were determined by means of inductively coupled plasma mass spectrometry (ICP-MS, TMO, USA). Quality assurance and quality control of Cd in soil samples were based on the determination of the metal contents in blank and duplicate samples and standard reference materials (GSS-3) obtained from the Center of National Standard Reference Material of China. A subset of samples consisting of 25% of the total number of samples available every year during the period 2010–2014 were randomly selected and served as the validation set assessing the performance of the different space-time prediction technique, including STKT with different trend models, and STOK. The remaining 75% of the total number of samples served as training points (see Section 2.3).

2.3. Spatiotemporal Kriging with Trend model (STKT)

2.3.1. The STRF formulation

Spatiotemporal attributes, such as soil Cd concentrations, are physical attributes that develop simultaneously in space and time. The aim of this study is the composite space-time prediction of soil Cd concentrations, $z(\mathbf{p}) : \mathbf{p} = (\mathbf{s}, t)$, $\mathbf{s} = (s_1, s_2) \in S$, $t \in T$, occurring at a space-time point $\boldsymbol{p}$, characterized by the geographical domain $S \subset R^2$ during a time period $T \subset R$. Specifically, Cd predictions $z(\mathbf{p}_0)$ are sought at unsampled space-time points $\mathbf{p}_0 = (\mathbf{s}_0, t_0)$ based on n measurements $z(\mathbf{p}_i)$ obtained at points $\mathbf{p}_i = (\mathbf{s}_i, t_i)$, $i = 1, \ldots, n$. It is assumed that $z(\mathbf{p}_i)$ is a realization of a spatiotemporal random field $Z(\mathbf{p})$ (STRF, Christakos, 1992). In this study the STRF is assumed to have the decomposition form

$$Z(\mathbf{p}) = m(\mathbf{p}) + R(\mathbf{p}), \tag{1}$$

i.e., the STRF has two distinct components: (a) a mean component $m(\mathbf{p})$ representing the space-time Cd trend, and (b) a residual component $R(\mathbf{p})$ representing fluctuations of Cd concentrations around that trend function. Using the STRF above, the present study aims at the realistic reconstruction of the spatiotemporal Cd distribution at the Qingshan district using all space-time data efficiently.

2.3.2. The trend component

In the STRF framework, the soil Cd distribution trend can be represented quantitatively in terms of the space-time deterministic series (Christakos, 1991)

$$m(s, t) = \sum\nolimits_{\rho=0}^{\mu} \sum\nolimits_{\zeta=0}^{\nu} b_{\rho\zeta} f_{\rho\zeta}(s, t) \tag{2}$$

where $f_{\rho\zeta}(s, t)$ are $\mu \times \nu$ known basis functions expressing the mean variation of the observed Cd dataset across space and time, $b_{\rho\zeta}$ are unknown coefficients to be determined by data fitting, and