Environmental Pollution 234 (2018) 794–803

Contents lists available at ScienceDirect

Environmental Pollution

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

Spatiotemporal characterization and mapping of $PM_{2.5}$ concentrations in southern Jiangsu Province, China[☆]

Yong Yang [a, b], George Christakos [c, d, *], Xue Yang [a, b], Junyu He [c, d]

[a] College of Resources & Environment, Huazhong Agricultural University, Wuhan, China
[b] Key Laboratory of Arable Land Conservation (Middle & Lower Reaches of Yangtse River), Ministry of Agriculture, China
[c] Institute of Island & Coastal Ecosystems, Ocean College, Zhejiang University, Zhoushan, China
[d] Department of Geography, San Diego State University, San Diego, CA, USA

ARTICLE INFO

Article history:
Received 12 September 2017
Received in revised form
23 November 2017
Accepted 24 November 2017

Keywords:
Spatiotemporal variation
$PM_{2.5}$
Pollution
China
Kriging

ABSTRACT

As a result of rapid industrialization and urbanization, China is experiencing severe air pollution problems. Understanding the spatiotemporal variation and trends of air pollution is a key element of an improved understanding of the underlying physical mechanisms and the implementation of the most effective risk assessment and environmental policy in the region. The motivation behind the present work is that the study region of southern Jiangsu province of China is one of the most populated and developed regions in China. The daily concentrations of particulate matter with particle diameter smaller than 2.5 μm ($PM_{2.5}$) in southern Jiangsu province obtained during the year 2014 were used to derive the variogram model that provided a quantitative characterization of the spatiotemporal (ST) variation of $PM_{2.5}$ concentrations in the study region. A spatiotemporal ordinary kriging (STOK) technique was subsequently employed to generate informative maps of the ST pollutant distribution in southern Jiangsu province. The results generated by STOK showed that during 2014 about 29.3% of the area was $PM_{2.5}$ polluted (at various severity levels, according to the criteria established by the Chinese government), and that the number of days characterized as polluted varied from 59 to 164 at different parts of the study region. Nanjing, the capital of Jiangsu province, was the place with the highest $PM_{2.5}$ pollution (including 3 days of serious pollution). The $PM_{2.5}$ pollution exhibited a decreasing spatial trend from the western to the eastern part of southern Jiangsu. A similar temporal $PM_{2.5}$ pattern was found from the western to the eastern part of southern Jiangsu, which was characterized by 4 peaks and 3 troughs linked to different meteorological conditions and human factors.

1. Introduction

In recent years, as a result of industrialization and urbanization in China, atmospheric particulate matter (PM) pollution has become a serious environmental problem (Hu et al., 2014). Accordingly, environmental research in China is increasingly focusing on PM pollution assessment, including its spatial patterns, source apportionment, migration and transformation, and subsequent control measures (Zhao et al., 2014; Zhang et al., 2017). In this setting, the accurate characterization and mapping of the spatiotemporal (ST) pollutant distribution play a fundamental role. Since 2013, China's government has begun to release information about major ambient pollutants, including particulate matter with particle diameter smaller than 2.5 μm ($PM_{2.5}$), particulate matter with particle diameter smaller than 10 μm (PM_{10}), sulfur dioxide (SO_2), carbon monoxide (CO), and nitrogen dioxide (NO_2) at major cities and at an hourly rate. These data provide the basis for the modeling and evaluation of regional air pollution.

However, most of the PM correlation studies in China did not involve any rigorous ST modeling of atmospheric pollution, but they rather focused on a purely spatial representation of the situation, followed by an *ad hoc* quantitative analysis of the situation based on the results obtained during multiple time periods (Fang et al., 2016; Chen et al., 2016; Wang and Fang, 2016). Methodologically, this is like one is trying to fit the (spatiotemporal) reality to a (spatial) model, rather than fit a (spatiotemporal) model to the (spatiotemporal) reality. This is certainly paradoxical, given that

[☆] This paper has been recommended for acceptance by Dr. Hageman Kimberly Jill.

* Corresponding author. Hai Ke Building 421, No.1 Zheda Road, Lincheng Street, Dinghai District, Zhoushan, Zhejiang, 316021, China.
E-mail address: gchristakos@zju.edu.cn (G. Christakos).

https://doi.org/10.1016/j.envpol.2017.11.077

international studies have shown that air pollution characterization and mapping based on composite ST modeling offer a more accurate representation of PM distribution and allow a better understanding of the PM variation mechanisms than purely spatial data analysis (Christakos and Serre, 2000; Christakos et al., 2001; Snepvangers et al., 2003; Pang et al., 2009; Yang et al., 2014; Calculli et al., 2015; Fassò et al., 2016; Datta et al., 2016). Moreover, there exist physically significant cross space-time dependencies (visualized as diagonal correlations) in the $PM_{2.5}$ distribution. For example, depending on the meteorological conditions (wind direction, atmospheric pressure etc.), the high $PM_{2.5}$ concentration observed at location *A* today may affect the concentration at a different location *B* tomorrow to a larger degree than the $PM_{2.5}$ concentration at location *B* today affects the concentration at the same location *B* tomorrow. In the present study, in particular, the limited monitoring stations offer a relative sparse spatial coverage of the study region. Thus, the prediction (estimation) accuracy obtained from a spatial interpolation technique (e.g., spatial kriging) will be worse than that of spatiotemporal interpolation technique. Also, if the spatial kriging is used to predict spatial $PM_{2.5}$ concentrations in the study region separately for each time (e.g., day), it would be necessary to fit a separate spatial variogram model and run a spatial kriging algorithm for each individual day (a total of 365 models with the corresponding algorithms), in which case the workload would be very heavy and the space-time separable approach rather unrealistic. Therefore, ST analysis models are more adequate to explore the implicit information contained in the ST data and to help us reveal the mechanism of the ST variation for geographical attributes.

In view of the above considerations, the present work has three main goals: (a) to compute the empirical ST variograms and fit the theoretical variogram models that describe adequately the physical space-time variations of daily $PM_{2.5}$ concentrations in southern Jiangsu Province (China) during the year 2014; (b) to generate maps of the spatiotemporal $PM_{2.5}$ distribution using the spatiotemporal ordinary kriging technique (STOK, Christakos, 1992); and (c) based on the findings of (a) and (b), to provide a quantitative characterization of $PM_{2.5}$ pollution in the study region for risk assessment and population health management purposes.

2. Material and methods

2.1. Study area and data

The focus of this study is southern Jiangsu province, which is one of the most populated and developed area in China (Wang et al., 2016; Shen et al., 2014). Therefore, it is quite important to accurately assess the ST variation of the $PM_{2.5}$ concentrations in southern Jiangsu. The data used in this study were obtained at 53 monitoring sites at 5 cities of the southern Jiangsu region (Nanjing, Zhenjiang, Changzhou, Wuxi, and Suzhou, with a total of 22470 km^2) during the period January 1—December 31, 2014 (Fig. 1, top). The data used in the study were collected from the website http://106.37.208.233:20035/. The Ministry of Environmental Protection of the People's Republic of China (MEPPRC) releases real-time hourly concentrations of major ambient pollutant concentrations in this website. The ambient $PM_{2.5}$ concentrations were measured in-situ according to the Chinese Environmental Protection Standard HJ655-2013 (NHFPCC, 2013), and the daily $PM_{2.5}$ concentration were derived at each site by averaging the hourly data. The descriptive statistics and the histogram of the $PM_{2.5}$ concentration data used in this study are also shown in Fig. 1 (bottom). The mean $PM_{2.5}$ concentration was 65.63 $\mu g/m^3$, and the coefficient of variation (CV) was 0.57, indicating a medium level variability of the monitored $PM_{2.5}$ data (note that $0.1 < CV < 1$).

The daily $PM_{2.5}$ concentrations averaged over all monitoring sites available in the study region during the year 2014 are plotted in Fig. 2. This is a time series constructed on the basis of the original (raw) $PM_{2.5}$ data, which exhibits a seasonal pattern with elevated concentrations during spring and winter due to seasonal fluctuations of both the emissions and meteorological conditions (for a relevant discussion, see Hu et al., 2014). We notice that the stations were sorted/labeled according to their S_2 values (Fig. 1 top), and for every group of 5 stations one was sequentially selected as the validation station, so that in the end a total number of 3611 samples from 10 stations (almost 20% of all monitoring stations) served as the validation dataset.

2.2. Spatiotemporal ordinary Kriging (STOK)

Air pollutants, such as $PM_{2.5}$ or PM_{10} concentrations in the atmosphere, are physical attributes that develop simultaneously in space and time under conditions of uncertainty. The simultaneous space and time variation of $PM_{2.5}$ concentrations in southern Jiangsu is best represented mathematically in terms of the spatiotemporal random field model (S/TRF)

$$PM_{2.5}(\boldsymbol{p}) = \{PM_{2.5} : \boldsymbol{p} = (\boldsymbol{s}, t), \boldsymbol{s} = (s_1, s_2) \in S, t \in T\},$$

which is defined on a geographical domain $S \in R^2$ and a time period $T \in R^1$ in a way that distinguishes between the physically different arguments of space and time. In-situ uncertainty is represented in the S/TRF model by a collection of possible ST $PM_{2.5}$ realizations, where the likelihood of each one of these realizations occurring is computed by the $PM_{2.5}$ probability law. The ST distance between the points $\boldsymbol{p}_i$ and $\boldsymbol{p}_j$ separated by the space distance $\boldsymbol{h}_{ij}$ and the time difference τ_{ij} is denoted by the vector

$$\boldsymbol{p}_j - \boldsymbol{p}_i = \Delta \boldsymbol{p}_{ij} = (\boldsymbol{h}_{ij}, \tau_{ij}),$$

$i, j = 1, 2, \ldots, n$ (for a detailed technical presentation of the S/TRF model, see Christakos, 1992, 2017).

The ST pollutant variation (which is due to physical processes and mechanisms) is quantitatively expressed in terms of the correlation between $PM_{2.5}$ concentrations at any pair of points $\boldsymbol{p}_i$ and $\boldsymbol{p}_j$. This correlation is represented by the variogram $\gamma_{PM_{2.5}}(\boldsymbol{h}_{ij}, \tau_{ij})$, which is a function of the space lag $\boldsymbol{h}_{ij}$ and the time separation τ_{ij} between $\boldsymbol{p}_i$ and $\boldsymbol{p}_j$. Two types of variogram functions are considered in the S/TRF modeling setting: (*a*) the empirical (experimental) variogram that is computed only between the $PM_{2.5}$ sampling points (say *N* points), and (*b*) the theoretical variogram (fitted to the empirical variogram) that provides $PM_{2.5}$ correlation values between any pair of points on the ST mapping grid (say $n > N$ grid points). In particular, the empirical ST variogram was calculated in this study based on the $PM_{2.5}$ data available at points $\boldsymbol{p}_i$ $(i = 1, 2, \ldots, N)$ using the formula

$$\widehat{\gamma}_{PM_{2.5}}(h_S, h_T) = \frac{1}{2N(h_S, h_T)} \sum_{i=1}^{N(h_S, h_T)} [PM_{2.5}(s_i, t_i) - PM_{2.5}(s_i + h_S, t_i + h_T)]^2 \quad (1)$$