Computers & Geosciences 37 (2011) 320–330

Contents lists available at ScienceDirect

Computers & Geosciences

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

A multigrid method for the estimation of geometric anisotropy in environmental data from sensor networks

I. Spiliopoulos [a], D.T. Hristopulos [a,*], M.P. Petrakis [a], A. Chorti [b]

[a] Geostatistics Research Unit, Department of Mineral Resources Engineering, Technical University of Crete, Chania 73100, Greece
[b] Department of Computer Communications, Middlesex University, T110, The Burroughs, London NW4 4BT, UK

ARTICLE INFO

Article history:
Received 2 October 2009
Received in revised form
12 April 2010
Accepted 3 June 2010
Available online 5 November 2010

Keywords:
Segmentation
Clustering
Spatial interpolation
Clustered CHI
Non-parametric
Kriging
Extreme values

ABSTRACT

This paper addresses the estimation of geometric anisotropy parameters from scattered spatial data that are obtained from environmental surveillance networks. Estimates of geometric anisotropy improve the accuracy of spatial interpolation procedures that aim to generate smooth maps for visualization of the data and for decision making purposes. The anisotropy parameters involve the orientation angle of the principal anisotropy axes and the anisotropy ratio (i.e., the ratio of the principal correlation lengths). The approach that we employ is based on the covariance Hessian identity (CHI) method, which links the mean gradient tensor with the Hessian matrix of the covariance function. We extend CHI to clustered CHI for application in data sets that include patches of extreme values and clusters of varying sampling density. We investigate the impact of CHI anisotropy estimation on the performance of spatial interpolation by ordinary kriging using a data set that involves both real background radioactivity measurements and a simulated release of a radioactive plume.

1. Introduction

Geoinformatics aims at the development of information science infrastructures for handling problems in the geosciences, as well as in geotechnical, environmental and mining engineering. It involves the development and management of data structures and databases, the utilization of networking and communication technologies for the transfer of the data, as well as the development and application of statistical methodologies for the processing of the data. Terrestrial environmental monitoring networks involve irregular distribution of measurement stations in space. The local density and characteristics of the networks are influenced by various factors, including national environmental policies, terrain topography and proximity to urban centers. However, in order to visualize and analyze the information provided by the network for decision making purposes, smooth maps of the monitored process are required. To generate smooth maps, a spatial model is needed for interpolation of the measurements on regular map grids. Remote sensing measurements are also affected by the problem of missing data (Rossi et al., 1994; Foster and Evans, 2008).

Geostatistical methods are important for the statistical processing of the information provided by environmental networks, since they permit characterization and quantification of spatial dependencies from scattered data for subsequent use in spatial interpolation and map generation. An accurate spatial model used in a geostatistical setting should incorporate estimates of anisotropy. For use in a near-automatic system, anisotropy estimation should require at most a few free parameters, to allow utilization by non-experts in Geostatistics. In addition, the anisotropy estimation approach must be flexible and computationally fast, in order to handle sensor networks with dynamic geometry.

Anisotropy in spatial data appears in two forms: trend anisotropy and statistical anisotropy. For a polynomial trend function $m_z(\mathbf{s}) = \sum_{k=1}^{n} \sum_{i=1}^{k} a_{i,k-i} x^i y^{k-i}$, where $\mathbf{s}=(x,y)$, the coefficients $a_{i,k-i}$ are determined by spatial regression and anisotropy implies that $a_{i,k-i} \neq a_{k-i,i}$. Statistical anisotropy involves geometric and zonal anisotropy. The former involves different correlation lengths in different principal directions, while the latter implies different variogram sills in different directions. Zonal anisotropy is typically observed when the data involve two processes, one of which evolves in a lower-dimensional space. Here we focus on the estimation of statistical geometric anisotropy. In two spatial dimensions, geometric anisotropy is determined from two parameters: the orientation angle of the principal axes and the ratio of the correlation lengths along the principal directions.

Estimation of the anisotropy parameters is typically based on empirical methods such as fitting of experimental directional

* Corresponding author.
E-mail addresses: spiliopoulos@gmail.com (I. Spiliopoulos), dionisi@mred.tuc.gr (D.T. Hristopulos), petrakis@mred.tuc.gr (M.P. Petrakis), ersi.chorti@gmail.com (A. Chorti).
URL: http://www.mred.tuc.gr/home/hristopoulos/dionisi.htm (D.T. Hristopulos).

0098-3004/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cageo.2010.06.007

variograms or on the rigorous, but computationally demanding, maximum likelihood method. The estimation of empirical variograms involves many *ad hoc* decisions, such as number of lag classes, lag tolerances, maximum lag, the minimum number of pairs per lag class and the type of estimator (e.g., classical Matheron or robust estimator), thus introducing considerable uncertainty in variogram estimation (Marchant and Lark, 2004). Anisotropy estimates require fitting the empirical variograms to theoretical model expressions and selecting an optimal model. The latter depends on the suite of models considered and the fitting method used (e.g., weighted least squares, generalized least squares) (Genton, 1998).

Maximum likelihood (ML) and restricted maximum likelihood (REML) estimation are elegant methods with good asymptotic properties. However, they are computationally demanding, especially for large data sets (e.g., more than 1000 points). The computational complexity of these methods scales as $O(N^3)$, where N is the number of sampling points. The total number of steps, and thus the total computational time, depends on the initial guesses for the variogram model parameters. Every step in the optimization procedure requires the inversion of a covariance matrix, which implies an $O(N^3)$ computational cost. In addition, the computation needs to be repeated for every variogram model tested. The methods based on the empirical variogram and the likelihood-based methods require specification of a variogram model to generate anisotropy estimates.

The recently proposed *covariance Hessian identity* (*CHI*) or covariance tensor identity (CTI) method (Hristopulos, 2002; Chorti and Hristopulos, 2008) provides a computationally fast approach for the estimation of geometric anisotropy parameters. Unlike methods based on maximum likelihood or the empirical variogram, CHI does not require a parametric covariance model. The tradeoff is that CHI assumes a differentiable covariance model. A deterministic interpolation method (e.g., bilinear or bicubic, minimum curvature) or Savitzky–Golay polynomial filtering is used to estimate the sample derivatives of scattered data (Chorti and Hristopulos, 2008). Here we opt for the former approach, which necessitates the use of *anisotropy estimation grids*. Several of these grids may be used (see Section 3.1), if the sampling density varies significantly over the study area.

The differentiability assumption is not a significant restriction on the choice of the variogram model: In addition to the classical Gaussian model, both Matérn and Spartan (Hristopulos, 2003; Hristopulos and Elogne, 2007) families provide covariance functions with controllable differentiability properties. A more significant issue is the presence of uncorrelated random noise (nugget) in the data, which destroys the differentiability of the observed realization. Nonetheless, the estimates of the derivatives are based on finite differences and not on a mathematical limit, and thus they admit finite values at the scale of the anisotropy estimation grid. The CHI method uses these values to extract anisotropy parameters corresponding to a differentiable random field. If the nugget is non-zero, as is typically the case for geospatial data, CHI tends to underestimate the true anisotropy. The impact of white Gaussian and lognormal noise addition is investigated in Chorti and Hristopulos (2008).

The remainder of this manuscript is structured as follows: Section 2 briefly describes the radioactivity data, which are used to illustrate the proposed methodological developments. The clustered CHI method for anisotropy analysis is presented in Section 3. Clustered CHI involves segmentation of the data into domains to handle extreme values separately and partitioning of the domains into clusters based on sampling density variations for efficient estimation of anisotropy. In Section 4 we use the clustered CHI method to determine the anisotropy of the radioactivity data. We also use cross-validation measures to compare the performance of kriging-based spatial interpolation with and without domain partitioning and anisotropic corrections. In addition, we present kriged maps of the radioactivity distribution and compare the different modeling assumptions. Finally, in Section 5 we present our conclusions. From the methodological viewpoint, this paper extends the scope of CHI by combining it with image segmentation techniques for clustering and by developing coarse-grained anisotropy parameters over different clusters. From a practical perspective, it investigates the impact of anisotropy estimation on interpolation performance for a data set characterized by sampling density irregularities and extreme values.

Table 1
Statistics of GDR data used in case study. Symbols used denote: $z_{\min}$, minimum value; q_1, first quartile; q_2, median; m_z, mean value; q_3, third quartile; $z_{\max}$, maximum; σ_z, standard deviation; μ_z, skewness; k_z, kurtosis.

$z_{\min}$	q_1	q_2	m_z	q_3	$z_{\max}$	σ_z	μ_z	k_z
29.0	85.8	131.0	2442.0	3082.0	26990.0	4371.36	2.29	5.25

2. Description of the data

The data set involves scattered measurements of gamma dose rates (GDR) in Europe. Our aim is to analyze the anisotropy of the data and to investigate the impact of anisotropy modeling in the performance of spatial interpolation and map generation. The GDR data were provided by the German Federal Office for Radiation Protection (Bundesamt für Strahlenschutz) in the framework of the INTAMAP project.[1] The sampling network is represented by the sites of the *European Radiological Exchange Platform* (EURDEP). A total of $N=3626$ sampling sites are used with their positions expressed in the INSPIRE coordinate system.[2] GDR is measured in *nanoSievert per hour* (nSv/h). The network involves both densely sampled areas (e.g., Germany and Austria) and sparsely sampled ones (e.g. South Europe).

Real background radioactivity measurements are combined with simulations that include systematic errors, local peaks due to washout effects caused by heavy rainfall, single peaks due to lighting strikes, and areas of extreme values resulting from the dispersion of a radioactive plume caused by a simulated reactor accident in central Europe. The simulations are generated with the RODOS system (Ehrhardt, 1997) using meteorological information from the German weather service. The time of the simulated accident was 23:40 on January 6, 2008. Forecasts of the plume dispersion were produced at +18, +30, +42, and +54 h from the starting time, for an area of 2500×2500 km^2 centered at the city of Offenbach. We used the +42 h time slice for the spatial analysis. The statistics of the GDR data, given in Table 1, exhibit large variability and strong deviations from Gaussianity.

Below, we estimate the anisotropy of this data set, and we analyze its impact on interpolation (ordinary kriging) performance as well as on the patterns of the generated maps.

3. The clustered CHI method for anisotropy estimation

This section presents in detail the methodologies proposed for anisotropy estimation. The procedures described below focus on the following tasks: (i) segmentation of the sensor network in domains of normal and extreme values, (ii) the subsequent

[1] Interoperability and Automated Mapping. 6th Framework Programme, ICT for Environmental Risk Management. URL: http://www.intamap.org/

[2] Infrastructure for Spatial Information in the European Community. EU 7th Framework Programme. URL: http://inspire.jrc.ec.europa.eu/