International Journal of Applied Earth Observation and Geoinformation 30 (2014) 128–138

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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

Bayesian area-to-point kriging using expert knowledge as informative priors

Phuong N. Truong[a,*], Gerard B.M. Heuvelink[a], Edzer Pebesma[b]

[a] Soil Geography and Landscape Group, Wageningen University, PO Box 47, 6700 AA Wageningen, The Netherlands
[b] Institute for Geoinformatics (IFGI), University of Münster, Heisenbertstr 2, 48149 Münster, Germany

A R T I C L E I N F O

Article history:
Received 12 August 2013
Accepted 31 January 2014
Available online 28 February 2014

Keywords:
Spatial disaggregation
Area-to-point kriging
Informative Bayesian area-to-point estimator
Statistical expert elicitation
Expert knowledge
Area-to-point conditional simulation

A B S T R A C T

Area-to-point (ATP) kriging is a common geostatistical framework to address the problem of spatial disaggregation or downscaling from block support observations (BSO) to point support (PoS) predictions for continuous variables. This approach requires that the PoS variogram is known. Without PoS observations, the parameters of the PoS variogram cannot be deterministically estimated from BSO, and as a result, the PoS variogram parameters are uncertain. In this research, we used Bayesian ATP conditional simulation to estimate the PoS variogram parameters from expert knowledge and BSO, and quantify uncertainty of the PoS variogram parameters and disaggregation outcomes. We first clarified that the nugget parameter of the PoS variogram cannot be estimated from only BSO. Next, we used statistical expert elicitation techniques to elicit the PoS variogram parameters from expert knowledge. These were used as informative priors in a Bayesian inference of the PoS variogram from BSO and implemented using a Markov chain Monte Carlo algorithm. ATP conditional simulation was done to obtain stochastic simulations at point support. MODIS (Moderate Resolution Imaging Spectroradiometer) atmospheric temperature profile data were used in an illustrative example. The outcomes from the Bayesian ATP inference for the Matérn variogram model parameters confirmed that the posterior distribution of the nugget parameter was effectively the same as its prior distribution; for the other parameters, the uncertainty was substantially decreased when BSO were introduced to the Bayesian ATP estimator. This confirmed that expert knowledge brought new information to infer the nugget effect at PoS while BSO only brought new information to infer the other parameters. Bayesian ATP conditional simulations provided a satisfactory way to quantify parameters and model uncertainty propagation through spatial disaggregation.

1. Introduction

Spatial disaggregation (downscaling) is becoming more important in a world where the demand for data transformation from global to local scales is rapidly increasing. In climate research, for example, regional or local climate models may require data of spatial climate attributes (e.g. precipitation, air temperature or atmospheric vapour) at finer resolution than those measured using remote sensing (RS) instruments or predicted using global climate models. Here, spatial resolution or pixel size stands for the spatial support, i.e. the geometrical size, shape and spatial orientation of a spatial unit of an observation or a prediction. Changing the spatial support of a variable changes its statistical and spatial properties (Schabenberger and Gotway, 2005). This is the well-known change of support problem (COSP) (Cressie, 1996; Gotway and Young, 2002; Schabenberger and Gotway, 2005).

Spatial support and COSP have been acknowledged as an important source of uncertainty in RS analyses due to aggregation and zoning effects (Marceau and Hay, 1999; Dungan, 2006). Spatial disaggregation of remotely sensed imagery through interpolation shows an important application of geostatistics to RS analysis (Van der Meer, 2012). Well-known geostatistical techniques for downscaling remotely sensed imagery of continuous variables are Area-to-Point (ATP) kriging and multivariate ATP kriging (Atkinson, 2013).

In this study, we focused on ATP kriging (Kyriakidis, 2004) for spatial disaggregation of a Gaussian random field. ATP kriging follows the principle of classical kriging and makes predictions of an attribute at point support (PoS) from block support observations (BSO) of the same attribute. It also quantifies the uncertainty about

Abbreviations: ATP, area-to-point; BSO, block support observations; COSP, change of support problem; ESA, European Space Agency; KNMI, Royal Netherlands Meteorological Institute; MCMC, Markov chain Monte Carlo; MODIS, Moderate Resolution Imaging Spectroradiometer; PoS, point support; pdf, probability density function; RS, remote sensing; SEE, statistical expert elicitation.

* Corresponding author. Tel.: +31 3174 82008; fax: +31 3174 19000.
E-mail addresses: phuong.truong@wur.nl (P.N. Truong), gerard.heuvelink@wur.nl (G.B.M. Heuvelink), edzer.pebesma@uni-muenster.de (E. Pebesma).

0303-2434/$ – see front matter
http://dx.doi.org/10.1016/j.jag.2014.01.019

the disaggregated predictions by means of the ATP kriging variance. ATP kriging satisfies the condition that the arithmetic average of the predictions (and simulations) at all point locations within a block equals the value of this block when the same number of BSO is used as conditioning data (Goovaerts, 2008). Hence, to use ATP kriging, BSO must be (assumed to be) the arithmetic average of PoS data within the blocks.

Let z be the variable of interest that is assumed to be a realisation of a second-order stationary Gaussian random function Z and let $\bar{z}(B_i) = 1/\left|B_i\right| \int_{s \in B_i} z(s)\,ds$ be the value of z at block support, where $z(\boldsymbol{s})$ is the value of z at point location $\boldsymbol{s}$ and $\left|B_i\right|$ is the area of a block B indexed by i. Because the arithmetic averaging is linear in its argument, the random process at block support is also a Gaussian process.

Let $\boldsymbol{Z_p} = (Z(\boldsymbol{s_1}), \ldots, Z(\boldsymbol{s_M}))^T$ and $\bar{\boldsymbol{Z}}_B = (\bar{Z}(B_1), \ldots, \bar{Z}(B_N))^T$ denote vectors of Z at point and block support, then their joint probability distribution is jointly Gaussian:

$$\begin{bmatrix} \boldsymbol{Z_p} \\ \boldsymbol{Z_B} \end{bmatrix} \sim N\left(\mu \begin{bmatrix} \boldsymbol{1_M} \\ \boldsymbol{1_N} \end{bmatrix}, \begin{bmatrix} \boldsymbol{C_{pp}} & \boldsymbol{C_{pB}} \\ \boldsymbol{C_{Bp}} & \boldsymbol{C_{BB}} \end{bmatrix}\right) \quad (1)$$

where μ is the constant spatial mean of Z, $\boldsymbol{1}_M$ and $\boldsymbol{1}_N$ are M and N vectors of ones, $\boldsymbol{C_{pp}}$ is the $M \times M$ variance – covariance matrix of $\boldsymbol{Z_p}$, $\boldsymbol{C_{BB}}$ is the $N \times N$ variance-covariance matrix of $\bar{\boldsymbol{Z}}_B$, $\boldsymbol{C_{pB}}$ and $\boldsymbol{C_{Bp}}$ are the variance-covariance matrix between $\boldsymbol{Z_p}$ and $\bar{\boldsymbol{Z}}_B$ and vice versa.

Because their joint distribution is normal, the optimal predictor of $\boldsymbol{Z_p}$ given $\bar{\boldsymbol{Z}}_B$ is a linear combination of the BSO (Chilès and Delfiner, 1999, Section 3.3.4):

$$\hat{Z}_p = \mu\boldsymbol{1}_M + \boldsymbol{C_{pB}}\boldsymbol{C_{BB}^{-1}}(\bar{\boldsymbol{Z}}_{\boldsymbol{B}} - \mu\boldsymbol{1}_N) \quad (2)$$

The variance-covariance matrix of the prediction error, called $\boldsymbol{C}(Z_p - \hat{Z}_p)$, is given by:

$$\boldsymbol{C}(Z_p - \hat{Z}_p) = \boldsymbol{C_{pp}} - \boldsymbol{C_{pB}}\boldsymbol{C_{BB}^{-1}}\boldsymbol{C_{pB}^{T}} \quad (3)$$

This shows that ATP kriging is straightforward and very similar to common kriging, but its main difficulty is that it requires the PoS variogram (Kyriakidis, 2004) to calculate the point-point, point-block and block-block variance-covariance matrices in Eq. (1), where the latter two require a regularisation (Journel and Huijbregts, 1978, Section II.D.4). Estimation of the PoS variogram from BSO is usually done using deregularisation or deconvolution (Journel and Huijbregts, 1978, Section II.D.4). Pardo-Igúzquiza and Atkinson (2007) introduced an iterative numerical deconvolution method to derive the PoS variogram from regular BSO (i.e. satellite imagery). In their study, the types of models included in the nested PoS variogram model were defined based on the nested variogram model fitted to the BSO. The optimisation condition was that the derived PoS variogram was the one minimising the difference between the theoretically regularised variogram model and the model fitted to the BSO. Goovaerts (2008) extended the method of Pardo-Igúzquiza and Atkinson (2007) to derive the PoS variogram from both regular and irregular (i.e. different size and shape) BSO. Gotway and Young (2007) presented an iterative generalised estimation approach to estimate the parameters of the PoS covariance function and the trend surface using irregular BSO. Nagle et al. (2011) used maximum likelihood estimation for the PoS covariance function using BSO. Gelfand et al. (2001) addressed Bayesian estimation of PoS variogram parameters from BSO of a spatial–temporal process. Their study focused on developing objective Bayesian inference methods, where the priors of the PoS variogram model parameters were given as noninformative priors. This is one of few studies that addressed PoS variogram estimation from BSO using a Bayesian approach.

In all aforementioned methods for deriving the PoS variogram, the nugget component of the PoS variogram was dismissed and assumed to be zero. There was surprisingly little attention on resolving the issue of inferring the nugget parameter from BSO, despite the material impact of the nugget variance on the ATP prediction and associated uncertainty (Kyriakidis, 2004). From the performance assessment of the iterative numerical deconvolution method using irregular BSO, Goovaerts (2008) concluded that the behaviour at the origin of the PoS variogram model (i.e. the nugget effect and within-block semivariance) could not be characterised with only BSO. Recently, Nagle et al. (2011) pointed out that the BSO retain little information to infer the nugget component of the PoS variogram and recommended using prior knowledge to overcome this problem.

The advantage of using a Bayesian approach is that the Bayesian estimator can quantify the uncertainty about the inference of the PoS variogram parameters. It is also the only formalised method to combine prior knowledge with BSO. However, extracting expert knowledge as informative priors is a delicate process in order to obtain reliable information. Much research has been done recently on using statistical expert elicitation (SEE) to extract expert knowledge to use as informative priors for Bayesian statistical models, e.g. in Bayesian environmental and ecological modelling (Choy et al., 2009; Kuhnert et al., 2010; Kuhnert, 2011) and Bayesian geological modelling (Wood and Curtis, 2004). Formal SEE (Garthwaite et al., 2005; O'Hagan et al., 2006) provides transparent and reliable techniques to elicit from expert knowledge the probability distributions of the PoS variogram parameters to use as informative priors (Truong and Heuvelink, 2013; Truong et al., 2013). The SEE procedure comprises several structured stages: starting from defining the issues that require expert knowledge, finding experts, choosing an elicitation approach and doing the real elicitation task with experts to post-processing and using the SEE outcomes. There is increasing literature presenting detailed guidelines of developing and using SEE methods, e.g. Hahn (2006), Knol et al. (2010), Kuhnert et al. (2010), O'Hagan (2012) to name a few. This promises to be a sufficient solution for the issue of lacking information from BSO to infer the nugget component of the PoS variogram.

Our aim in this study was twofold. Firstly, we wanted to resolve the issue of poor estimation of the nugget effect from BSO by using a Bayesian approach that incorporates knowledge of multiple experts. Secondly, we wanted to quantify the propagation of PoS variogram parameters and ATP kriging model uncertainty to the disaggregated outcomes using Bayesian ATP conditional simulation. We illustrate the method with an example on disaggregating MODIS air temperature data measured on a coarse grid of 5 km resolution to a finer grid of 1 km resolution. To this end, the remainder of this paper has three main sections. Section 2 presents the statistical methods and a description of the example. Section 3 presents the main results of the study and a discussion. Section 4 provides the conclusions and recommendations for further research.

2. Materials and methods

Fig. 1 shows the three main steps of the method.

2.1. Data

Spaceborne thermal imagery is becoming important in climate modelling, soil moisture assessment, irrigation management, etc. (Kuenzer et al., 2013a,b; Ha et al., 2013). Products of daily spaceborne thermal imagery often have lower spatial resolution (e.g., MODIS: 1–5 km, NOAA-AVHRR: 1 km, Sentinel 3-ESA future mission: 1 km), whereas higher spatial resolution at several tenths of metres is often required, e.g. in precision agriculture or irrigation management at field level or in assessing urban heat effect (Kuenzer et al., 2013a,b). For these reasons and for illustration