On Reading "*Model-Based Geostatistics*" by Diggle, Tawn, and Moyeed, JRSSC 1998

Whitney Huang

CLEMSON UNIVERSITY

wkhuang@clemson.edu

Clemson ENVR Group, October 28, 2020

Appl. Statist. (1998)
47, *Part* 3, *pp.* 299–350

Model-based geostatistics

P. J. Diggle† and J. A. Tawn

Lancaster University, UK

and R. A. Moyeed

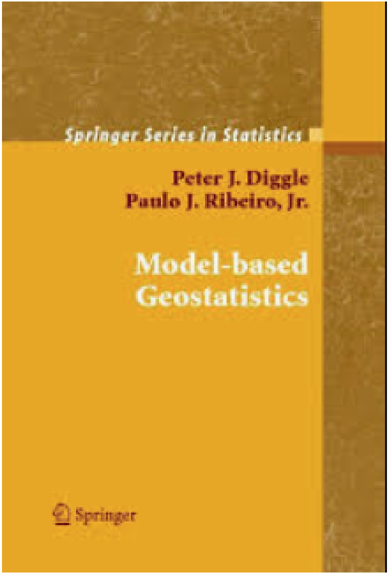
Plymouth University, UK

[*Read before* The Royal Statistical Society *on Wednesday, November 12th, 1997, the President*, Professor R. N. Curnow, *in the Chair*]

Summary. Conventional geostatistical methodology solves the problem of predicting the realized value of a linear functional of a Gaussian spatial stochastic process $S(\mathbf{x})$ based on observations $Y_i = S(\mathbf{x}_i) + Z_i$ at sampling locations $\mathbf{x}_i$, where the Z_i are mutually independent, zero-mean Gaussian random variables. We describe two spatial applications for which Gaussian distributional assumptions are clearly inappropriate. The first concerns the assessment of residual contamination from nuclear weapons testing on a South Pacific island, in which the sampling method generates spatially indexed Poisson counts conditional on an unobserved spatially varying intensity of radioactivity; we conclude that a conventional geostatistical analysis oversmooths the data and underestimates the spatial extremes of the intensity. The second application provides a description of spatial variation in the risk of campylobacter infections relative to other enteric infections in part of north Lancashire and south Cumbria. For this application, we treat the data as binomial counts at unit postcode locations, conditionally on an unobserved relative risk surface which we estimate. The theoretical framework for our extension of geostatistical methods is that, conditionally on the unobserved process $S(\mathbf{x})$, observations at sample locations $\mathbf{x}_i$ form a generalized linear model with the corresponding values of $S(\mathbf{x}_i)$ appearing as an offset term in the linear predictor. We use a Bayesian inferential framework, implemented via the Markov chain Monte Carlo method, to solve the prediction problem for non-linear functionals of $S(\mathbf{x})$, making a proper allowance for the uncertainty in the estimation of any model parameters.

Keywords: Generalized linear mixed model; Geostatistics; Kriging; Markov chain Monte Carlo method; Spatial prediction

Book by Diggle and Ribeiro, 2007



Gaussian Process (GP) Spatial Model

We assume the observed data $\{y(\boldsymbol{x}_i)\}_{i=1}^{n}$ is one partial realization of a spatial GP $\{Y(\boldsymbol{x})\}_{\boldsymbol{x}\in\mathcal{X}}$.

Model:

$$Y(\boldsymbol{x}) = m(\boldsymbol{x}) + S(\boldsymbol{x}) + \epsilon, \qquad \boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^d$$

where

- Mean function:

$$m(\boldsymbol{x}) = \mathbb{E}\left[Y(\boldsymbol{x})\right] = \boldsymbol{Z}^T(\boldsymbol{x})\boldsymbol{\beta}$$

- Covariance function:

$$\{S(\boldsymbol{x})\}_{\boldsymbol{x}\in\mathcal{X}} \sim \mathrm{GP}\left(0, K\left(\cdot,\cdot\right)\right), \quad K(\boldsymbol{x}_1, \boldsymbol{x}_2) = \mathrm{Cov}\left(S(\boldsymbol{x}_1), S(\boldsymbol{x}_2)\right)$$

An Equivalent Representation

To simply the presentation let's assume $m(\boldsymbol{x}) = \mu$. The data model

$$Y_i = \mu + S(\boldsymbol{x}_i) + \epsilon_i, i = 1, \cdots, n,$$

can be presented as

$$Y_i | S(\boldsymbol{x}_i) \sim \mathrm{N}(\mu + S(\boldsymbol{x}_i), \tau^2),$$

where $\epsilon_i \sim \mathrm{N}(0, \tau^2)) \perp S(\cdot)$.

Parameter estimation can be done via likelihood-based method. "Plug-in" spatial prediction is typically used.

Main Novelty of this Paper

Extend the GP spatial model to model non-Gaussian spatial data

- Incorporating spatial structure via GP model within the framework of generalized linear model (GLM)
- Inference/Prediction is carry out under Bayesian framework via Markov chain Monte Carlo (MCMC)
- Demonstrates the modeling approach with applications where Poisson and binomial distributional assumptions are more tenable

Spatial Generalized Linear Models

- Data Level: Y_i **conditionally** follow a distribution within the exponential family (e.g., Poisson; binomial) where

$$\mathbb{E}[Y_i|S(\boldsymbol{x}_i)] = \mu(\boldsymbol{x}_i), \quad i = 1, \cdots, n.$$

- Linear "fixed effects" plus spatial "random effects":

$$\eta(\boldsymbol{x}_i) = g(\mu(\boldsymbol{x}_i)) = \boldsymbol{Z}(\boldsymbol{x}_i)^{\mathrm{T}}\boldsymbol{\beta} + S(\boldsymbol{x}_i)$$

- Latent spatial process:

$$S(\boldsymbol{x}) \sim GP(0, K(\cdot,\cdot)), \qquad \boldsymbol{x} \in \mathcal{X}$$

Example: Spatial Binary Data

- $Y_i | p(\boldsymbol{x}_i) \sim \text{Bernoulli}(p(\boldsymbol{x}_i))$
- $\log(\frac{p(\boldsymbol{x}_i)}{1-p(\boldsymbol{x}_i)}) = \boldsymbol{Z}(\boldsymbol{x}_i)^{\mathrm{T}}\boldsymbol{\beta} + S(\boldsymbol{x}_i)$
- $S(\boldsymbol{x}) \sim GP(0, K(\cdot,\cdot)), \qquad \boldsymbol{x} \in \mathcal{X}$

Example: Spatial Count Data

- $Y_i|\lambda(\boldsymbol{x}_i) \sim \text{Poisson}(\lambda(\boldsymbol{x}_i))$
- $\log(\lambda(\boldsymbol{x}_i)) = \boldsymbol{Z}(\boldsymbol{x}_i)^{\mathrm{T}}\boldsymbol{\beta} + S(\boldsymbol{x}_i)$
- $S(\boldsymbol{x}) \sim GP(0, K(\cdot,\cdot)), \qquad \boldsymbol{x} \in \mathcal{X}$

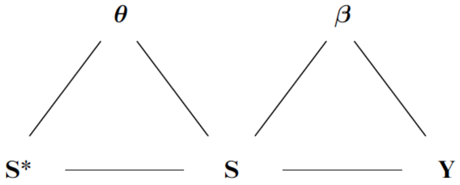
Inference and Prediction: MCMC

Goal: To obtain the posterior distribution of

$$[(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{S})|\boldsymbol{Y}],$$

where $\boldsymbol{\theta}$ are the parameters for GP; $\boldsymbol{\beta}$ consist of the regression parameters; $\boldsymbol{S}$ and $\boldsymbol{Y}$ are the values of S and Y at $\{\boldsymbol{x}_i\}_{i=1}^n$

To implement MCMC we would need to sample $\pi(\boldsymbol{\theta}|\boldsymbol{Y}, \boldsymbol{S}, \boldsymbol{\beta})$, $\pi(S_i|\boldsymbol{S}_{-i}, \boldsymbol{Y}, \boldsymbol{\theta}, \boldsymbol{\beta})$, and $\pi(\boldsymbol{\beta}|\boldsymbol{Y}, \boldsymbol{S}, \boldsymbol{\theta})$



Source: Diggle et al. (13)

Inference and Prediction: MCMC

- $\pi(\boldsymbol{\theta}|\boldsymbol{Y}, \boldsymbol{S}, \boldsymbol{\beta}) = \pi(\boldsymbol{\theta}|\boldsymbol{S}) \propto f(\boldsymbol{S}|\boldsymbol{\theta}) f(\boldsymbol{\theta})$

- $\pi(S_i|\boldsymbol{S}_{-i}, \boldsymbol{Y}, \boldsymbol{\theta}, \boldsymbol{\beta}) \propto f(\boldsymbol{Y}|\boldsymbol{S}, \boldsymbol{\beta}) f(S_i|\boldsymbol{S}_{-i}, \boldsymbol{\theta}) = \left\{\prod_{j=1}^{n} f(y_j|s_j, \boldsymbol{\beta})\right\} f(S_i|\boldsymbol{S}_{-i}, \boldsymbol{\theta})$

- $\pi(\boldsymbol{\beta}|\boldsymbol{Y}, \boldsymbol{S}, \boldsymbol{\theta}) = \pi(\boldsymbol{\beta}|\boldsymbol{Y}, \boldsymbol{S}) \propto f(\boldsymbol{Y}|\boldsymbol{S}, \boldsymbol{\beta}) f(\boldsymbol{\beta}) = \left\{\prod_{j=1}^{n} f(y_j|s_j, \boldsymbol{\beta})\right\} f(\boldsymbol{\beta})$