C Program for Computing Sampson-Guttorp Spatial Deformations

User's Manual

September, 2004

Contents

1 A Brief Introduction to the Manual

This document is meant to provide guidance in using the C program through which we compute spatial deformation models for environmental processes and obtain samples from the posterior distribution of the model parameters. The program implements a Metropolis-Hastings

(MH) algorithm. Throughout this manual it is assumed that the user is familiar with the MH general structure.

The next section describes the model, including the likelihood, the thin-plate spline modelling of the spatial deformation, and the model of the temporal variance as a random field. Section 3 lists the prior distributions currently implemented in the program. Section 4 explains the input required to run the program and section 5 describes the output.

2 The Model

2.1 Model Assumptions

We assume that temporally independent samples $Z_{it} = Z(x_i, t)$ are available at each of N geographic locations at the same T points in time: $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$. We consider the following model for the underlying process:

$$Z(x,t) = m(x) + \nu(x)^{1/2} H_t(x) + \varepsilon(x,t) \tag{1.1}$$

where x denotes location, t denotes time and $m(x)$ represents the spatial mean.

$H_t(x)$ is a zero mean, variance one, Gaussian spatial process with a Sampson-Guttorp correlation function. (The subscript indexes replicates in time and will be dropped in the following notation unless needed explicitly, usually in a summation.). Thus,

$$\mathrm{Corr}(H(x), H(y)) = \rho_\theta(f(x) - f(y)) . \tag{1.2}$$

$\nu(x)$ represents the temporal variability of the process observed at location x. For details on how this is modelled see section 4. $\varepsilon(x,t)$ is a white-noise process (i.e. it has mean zero, constant variance and zero correlation), independent of $H(x)$. Its variance, usually called the *nugget* effect, will be denoted as σ_ε^2.

Note that the assumptions of constant temporal mean and temporal independence often do not strictly hold in practice, and thus some preprocessing or filtering of data may be needed.

2.2 Likelihood

We denote by $Z_1,\ldots,Z_T$ the T independent N-dimensional vectors of observations, by $\bar{Z}$ the vector of site means, $\bar{Z}=\frac{1}{T}\sum_{t=1}^{T}Z_t$, and by $\mathbf{S}$ the sample covariance matrix, $\mathbf{S}=\frac{1}{T-1}\sum_{t=1}^{T}(Z_t-\bar{Z})(Z_t-\bar{Z})^T$. ($\mathbf{S}$ is non-singular if $T>N$). Since by assumption the mean $m(x,t)$ does not vary in time, we denote by m the N-dimensional vector of means at the geographic locations.

The variance at the gauged locations will be denoted by $\mathbf{v}$ when it refers to an N-dimensional vector (when the variance is modeled as a spatial process) and by v when it refers to a scalar (when the variance is assumed fixed throughout space).

The Z_t are assumed to be a sample from a population with theoretical $N\times N$ covariance matrix $\mathbf{\Sigma}$ with elements

$$\sigma_{ij}=\begin{cases}\sqrt{v_i v_j}\,\rho_\theta\left(\left\|\xi_i-\xi_j\right\|\right) & i\neq j\\ v_j+\sigma_\varepsilon^2 & i=j\end{cases} \tag{1.3}$$

Here we write $\xi_i=f(x_i)$. We refer to the x_i as the geographic or "*G*-plane" coordinates while the ξ_i represent the "dispersion" or "*D*-plane" coordinates of the sites. The parameters θ, f, v, and σ_ε^2 appear in the likelihood only through the matrix $\mathbf{\Sigma}$. Therefore the likelihood, which is the multivariate normal density of the observations conditional on the parameters m, θ, f, v, σ_ε^2 can be expressed as a function of m and $\mathbf{\Sigma}$. It is a function of the observations only through the sample mean and covariance, $\bar{Z}$ and $\mathbf{S}$:

$$\begin{aligned}\left[Z_1,\ldots,Z_N\left|m,\theta,f,v,\sigma_\varepsilon^2\right.\right]&=\left[\bar{Z},\mathbf{S}\left|m,\mathbf{\Sigma}\right.\right]\\ &=\left|2\pi\mathbf{\Sigma}\right|^{-T/2}\exp\left\{-\frac{T}{2}\operatorname{tr}\mathbf{\Sigma}^{-1}\mathbf{S}-\frac{T}{2}\left(\bar{Z}-m\right)^T\mathbf{\Sigma}^{-1}\left(\bar{Z}-m\right)\right\}.\end{aligned} \tag{1.4}$$

We simplify the likelihood by putting a flat (improper) prior distribution on m, $\left[m\right]\propto 1$, and integrating it out of the likelihood:

$$\left[\mathbf{S}\middle|\boldsymbol{\Sigma}\right]=\int\left[\bar{Z},\mathbf{S}\middle|m,\boldsymbol{\Sigma}\right]\left[m\right]\mathrm{d}m \;\propto\; \left|\boldsymbol{\Sigma}\right|^{-(T-1)/2}\exp\left\{-\frac{T}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\mathbf{S}\right\}. \tag{1.5}$$

In the computer program the right-hand side of eq. (5) is called the likelihood (except that the power of $|\boldsymbol{\Sigma}|$ is T, not $T-1$).

2.3 Modeling f as a Pair of Thin-Plate Splines (PTPS)

2.3.1 Motivation and Basic Properties

We denote by $\mathbf{X}$ the $N\times 2$ matrix of gauged locations (each row represents one location, as in the program's input), and start by assuming that the images of the gauged locations are known. We denote these by the $N\times 2$ matrix $\boldsymbol{\Xi}$.

Consider all the interpolating functions Ψ from $\mathbb{R}^2\to\mathbb{R}^2$ that map $\mathbf{X}$ onto $\boldsymbol{\Xi}$, that are twice differentiable. To each Ψ we attribute a scalar, named the total bending energy functional:

$$J(\Psi)=\sum_{j=1}^{2}\iint\limits_{\mathbb{R}^2}\left(\frac{\partial^2\psi_j}{\partial s_1^2}\right)^2+2\left(\frac{\partial^2\psi_j}{\partial s_1 s_2}\right)^2+\left(\frac{\partial^2\psi_j}{\partial s_2^2}\right)^2\mathrm{d}s_1\mathrm{d}s_2 \tag{1.6}$$

The negative of $J(\Psi)$ can be viewed as a measure of smoothness.

It can be shown that of all the interpolating functions that are twice differentiable with integrable second derivatives, the PTPS minimizes the total bending energy functional. It is an interpolating function $\Phi(s)=\left(\Phi_1(s),\Phi_2(s)\right)^T$: $\mathbb{R}^2\to\mathbb{R}^2$ of the following form:

$$\Phi(s)=c+\mathbf{A}s+\mathbf{W}^T\tilde{\sigma}(s). \tag{1.7}$$

Here, $c+\mathbf{A}s$ is the linear part of the deformation, with c a 2×1 vector and $\mathbf{A}$ a 2×2 matrix. $\mathbf{W}^T\tilde{\sigma}(s)$ is the non-linear part, with $\tilde{\sigma}(s)=\left(\sigma(s-x_1),\ldots,\sigma(s-x_N)\right)^T$ an $N\times 1$ vector and $\mathbf{W}$ an $N\times 2$ matrix of coefficients. The function $\sigma:\mathbb{R}^2\to\mathbb{R}^2$ is defined as

$$\sigma(h)=\begin{cases}\|h\|^2\log(\|h\|) & \|h\|>0\\ 0 & \|h\|=0\end{cases} \tag{1.8}$$

In addition to the $2N$ interpolating constraints,

$$\Phi_j(x_i) = \xi_{ij},\ 1 \le i \le N;\ j = 1,2, \tag{1.9}$$

six additional constraints are imposed on the coefficients of the non-linear part of the deformation, $\mathbf{W}$, in order for the matrix $\mathbf{\Gamma}^{11}$ in eq. (15) to be defined; namely

$$1^T \mathbf{W} = 0,\ \ \mathbf{X}^T \mathbf{W} = 0. \tag{1.10}$$

In other words, the columns W_1 and W_2 of $\mathbf{W}$ are vectors in the subspace of dimension $N-3$ of $\mathbb{R}^N$ orthogonal to the subspace spanned by $\left\{1, X_1, X_2\right\}$. This subspace will be denoted by $\mathbb{V}$:

$$\mathbb{V} = \left\{ v \in \mathbb{R}^N : v^T 1 = 0,\ v^T X_1 = 0,\ v^T X_2 = 0 \right\}. \tag{1.11}$$

The PTPS can therefore be parameterized either in terms of the $2N$ images $\mathbf{\Xi}$ or in terms of the $2N+6$ parameters c, $\mathbf{A}$, $\mathbf{W}$, which, together with the 6 constraints in eq. (10), allow for $2N$ free parameters. This link is summarized in the following equation, combining eqs. (7), (9) and (10):

$$\begin{bmatrix} \mathbf{\Xi} \\ 0 \\ 0 \end{bmatrix} = \underbrace{\begin{bmatrix} \tilde{\mathbf{S}} & 1 & \mathbf{X} \\ 1^T & 0 & 0 \\ \mathbf{X}^T & 0 & 0 \end{bmatrix}}_{\mathbf{\Gamma}} \begin{bmatrix} \mathbf{W} \\ c^T \\ \mathbf{A}^T \end{bmatrix} \tag{1.12}$$

where $\tilde{\mathbf{S}}$ is an $N \times N$ matrix with elements $\tilde{\mathbf{S}}_{ij} = \sigma(x_i = x_j)$.

Equivalently, starting from a matrix of images $\mathbf{\Xi}$, the coefficients of the corresponding spline are obtained from the dual equation to eq. (12):

$$\begin{bmatrix} \mathbf{W} \\ c^T \\ \mathbf{A}^T \end{bmatrix} = \mathbf{\Gamma}^{-1} \begin{bmatrix} \mathbf{\Xi} \\ 0 \\ 0 \end{bmatrix} \tag{1.13}$$

The minimized total bending energy, $J(\Phi)$, has two equivalent forms:

$$J(\Phi) = \mathrm{tr}\left(\mathbf{W}^T\tilde{\mathbf{S}}\mathbf{W}\right) = \mathbf{W}_1^T\tilde{\mathbf{S}}\mathbf{W}_1 + \mathbf{W}_2^T\tilde{\mathbf{S}}\mathbf{W}_2 \tag{1.14}$$

and

$$J(\Phi) = \mathrm{tr}\left(\mathbf{\Xi}^T\mathbf{\Gamma}^{11}\mathbf{\Xi}\right) = \mathbf{\Xi}_1^T\mathbf{\Gamma}^{11}\mathbf{\Xi}_1 + \mathbf{\Xi}_2^T\mathbf{\Gamma}^{11}\mathbf{\Xi}_2 \tag{1.15}$$

where $\mathbf{\Gamma}^{11}$ is the N × N upper corner of $\mathbf{\Gamma}^{-1}$.

2.3.2 Identifiability Constraints

The likelihood (eq. (5)) is invariant to translations, rotations, changes in scale and reflections about a line. In order to solve this identifiability problem, we choose a reference frame by holding two image (*D*-plane) points, ξ_{i_1} and ξ_{i_2} fixed equal to their original geographic coordinates, x_{i_1} and x_{i_2}. The program allows the user to specify any two sites to have *D*-plane coordinates set equal to the *G*-plane coordinates. (We usually hold fixed two relatively distant gauged locations in order to control the overall scale of the resulting configuration and minimize possible numerical instabilities.)

After choosing the two fixed points, four out of the six linear coefficients, c and $\mathbf{A}$, are determined. The additional two coefficients are free, and we put a proper prior distribution on them. The prior distribution of $\mathbf{W}$, when restricted to $\mathbb{V}\times\mathbb{V}$ (and not the whole of $\mathbb{R}^N\times\mathbb{R}^N$) is proper as well. The posterior distribution of the parameters is then proper as well.

In the remainder of this subsection we show how to constrain f to keep two gauged locations fixed. We denote these two locations by x_1 and x_2, but emphasize that they can be any two locations, and moreover that the equations below are easily modified to hold fixed any two points on the plane. Fixing x_1 we obtain

$$\begin{aligned} & f(x_1) = x_1 \\ & \quad \Rightarrow x_1 = c + \mathbf{A}x_1 + \mathbf{W}^T\tilde{\sigma}(x_1) \\ & \quad \Rightarrow c = (\mathbf{I}-\mathbf{A})x_1 - \mathbf{W}\tilde{\sigma}(x_1). \end{aligned} \tag{1.16}$$

Fixing also x_2 we obtain:

$$\begin{aligned} &f(x_2) = x_2 \\ &\qquad \Rightarrow x_2 = c + \mathbf{A}x_2 + \mathbf{W}^T\tilde{\sigma}(x_2) \\ &\qquad \Rightarrow \mathbf{A}\left(x_2 - x_1\right) = \left(x_2 - x_1\right) - \mathbf{W}^T\left(\tilde{\sigma}(x_2) - \tilde{\sigma}(x_1)\right). \end{aligned} \tag{1.17}$$

This equation determines two of the entries of the matrix $\mathbf{A}$. We denote $\mathbf{z} = x_2 - x_1$ and $\left(x_2 - x_1\right) - \mathbf{W}^T\left(\tilde{\sigma}(x_2) - \tilde{\sigma}(x_1)\right)$ and assume without losing generality that $z_1 \neq 0$. (Otherwise, necessarily $z_2 \neq 0$, because $x_1 \neq x_2$, and instead of solving for the first column of A we can solve for its second column). It follows that:

$$\left.\begin{aligned} a_{11}z_1 + a_{12}z_2 = k_1 \\ a_{21}z_1 + a_{22}z_2 = k_2 \end{aligned}\right\} \Rightarrow \begin{aligned} a_{11} &= \frac{k_1 = a_{12}z_2}{z_1} \\ a_{21} &= \frac{k_2 - a_{22}z_2}{z_1} \end{aligned} \tag{1.18}$$

Therefore, only a_{12} and a_{22}, the entries in the second column of A, are free.

Note that in the particular case in which $a_{12} = 0,\ a_{22} = 1$ and $\mathbf{W}$ is the zero matrix, we obtain that $a_{11} = (x_{21} - x_{11})/(x_{21} - x_{11}) = 1$ and $a_{21} = (x_{22} - x_{12} - x_{22} - x_{22})/(x_{21} - x_{11}) = 0$; thus $\mathbf{A}$ is the identity matrix, and from eq. (16) c is the zero vector, and these parameters correspond to f being the identity transformation.

2.3.3 Prior Distribution of *f*

Since the planar deformation f is modelled as a PTPS, it is completely specified by its parameters, $a_{12} = 0,\ a_{22} = 1$ and $\mathbf{W}$. The prior distribution on these is a product between the priors of the linear free parameters and the non-linear parameters. a_{12} and a_{22} have independent normal distributions with means 0 and 1 respectively, and variances that allow for diffuse densities. In practice we have typically set both of these prior variances to 10.

The prior of $\mathbf{W}$ is multivariate normal with a kernel derived from the bending energy functional, and it has the form

$$[\mathbf{W}] \propto \left\{ -\frac{1}{2\tau}\left(\mathbf{W}_1^T \tilde{\mathbf{S}}\, \mathbf{W}_1\right) + \left(\mathbf{W}_2^T \tilde{\mathbf{S}}\, \mathbf{W}_2\right) \right\} \mathbf{I}_{\{\mathbf{W} \in \mathbb{V} \times \mathbb{V}\}}, \tag{1.19}$$

where the matrix $\tilde{\mathbf{S}}$ is defined above following equation eq. (12), and is completely specified by the gauged locations. The only parameter in this distribution that needs to be specified by the user is τ – the smaller its value, the more the prior mass is concentrated around linear deformations. It should be noted that this parameter depends on the scale or units of the coordinate system. We find it convenient to scale the coordinates so that the range of the coordinates is on the order of [-1,1]. Values of τ that define diffuse priors, and yet do not provide substantial probability on inappropriate "folding" (not onto) deformations depend also on the particular geographic configuration. In fact, we recommend sampling from priors for various values of τ in order to appreciate this prior. With scaling of coordinates as suggested here, values of τ between .02 and .001 have worked well.

2.4 Model for the Temporal Variance

This section describes our model for the temporal variance when this is assumed to vary spatially. $\mathbf{v}$, the vector of variances at the gauged locations, is modelled as resulting from a hidden covariance non-stationary random process, $v(x)$, that is assumed to vary smoothly over the region of interest. The user may choose to model the temporal variance as constant or varying in space.

2.4.1 Details of the Variance Model

Let $\eta(x)$ be a real Gaussian process defined for $x \in \mathbb{R}^2$, with constant spatial mean μ, constant spatial variance $\tilde{\sigma}^2$, and SG non-stationary correlation $\mathrm{Corr}(\eta(x),\eta(y)) = \rho_{\tilde{\theta}}\left(\|f(x) - f(y)\|\right)$, where $\rho(\bullet)$ belongs to a parametric family with parameter(s) $\tilde{\theta}$. The function f is taken to be the same planar deformation as the one used in modeling the non-stationarity in $Z(\bullet)$. Then, at the gauged sites, $\boldsymbol{\eta} = \left(\eta(x_1),\ldots,\eta(x_N)\right)$ has density of the form

$$[\eta] \propto \left|\tilde{\Sigma}\right|^{-1/2} \exp\left\{-\frac{1}{2}(\eta-\mu\cdot 1)^T \tilde{\Sigma}^{-1}(\eta-\mu\cdot 1)\right\} \tag{1.20}$$

for $\tilde{\mathbf{\Sigma}}$ an $N \times N$ matrix with elements $\tilde{\sigma}_{ij} = \tilde{\sigma}^2 \rho_{\tilde{\theta}}\left(\left\|f(x_i)-f(x_j)\right\|\right)$.

We define the variance process, $\nu(x)$, as $\mu(x) = \exp(\eta(x))$ and we readily obtain from eq. (20) the density of $\nu(\bullet)$ and its moments at the gauged sites:

$$[\mathbf{\nu}] \propto \frac{1}{\prod_{i=1}^{N} \nu_i} \left|\tilde{\Sigma}\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\log(\mathbf{\nu})-\mu\cdot 1)^T \tilde{\Sigma}^{-1}(\log(\mathbf{\nu})-\mu\cdot 1)\right\} \tag{1.21}$$

where $\log(\mathbf{\nu}) = \left(\log(\nu_1),\ldots,\log(\nu_N)\right)$. The process $\nu(x)$ has constant spatial mean and variance given by:

$$\mathrm{E}\left(\nu(x)\right) = \exp\left(\mu + \frac{1}{2}\tilde{\sigma}^2\right) \tag{1.22}$$

and:

$$\mathrm{Var}\left(\nu(x)\right) = \exp(2\mu+\tilde{\sigma}^2)(\exp(\tilde{\sigma}^2)-1) \tag{1.23}$$

Its non-stationary covariance is given by:

$$\mathrm{Cov}\left(\nu(x),\nu(y)\right) = \exp\left(2\mu+\tilde{\sigma}^2\right)\left(\exp\left\{\tilde{\sigma}^2 \rho_{\tilde{\theta}}\left(\left\|f(x)-f(y)\right\|\right)\right\}-1\right), \tag{1.24}$$

and its correlation by:

$$\mathrm{Corr}\left(\nu(x),\nu(y)\right) = \frac{\left(\exp\left\{\tilde{\sigma}^2 \rho_{\tilde{\theta}}\left(\left\|f(x)-f(y)\right\|\right)\right\}-1\right)}{\exp(\tilde{\sigma}^2)-1}, \tag{1.25}$$

2.4.2 Prior Distributions of the Hyper-Parameters μ and $\tilde{\sigma}^2$

We impose a normal prior distribution on μ, $\mu \sim N(\alpha,\beta)$, , which results in a normal (full) conditional distribution, $\mu \big| \mathbf{\nu},\theta,\sigma_\varepsilon^2,a,\mathbf{W},\tilde{\theta},\tilde{\sigma}^2,\mathbf{S}$. Similarly, the prior of $1/\tilde{\sigma}^2$ is $\Gamma(\gamma,\delta)$, implying a Gamma full conditional distribution, $\left(1/\tilde{\sigma}^2\right) \big| \mathbf{\nu},\theta,\sigma_\varepsilon^2,a,\mathbf{W},\tilde{\theta},\tilde{\sigma}^2,\mathbf{S}$. These two facts facilitate the sampling of μ and $\tilde{\sigma}^2$ in the program; these are proposed in the algorithm through Gibbs steps and always accepted.

It remains to specify the parameters of the normal and Gamma hyperprior distributions. If we have no other prior knowledge, we can make use of the moments of the sample variances $S_{ii}, 1 \le i \le N$ as follows. For simplicity we assume that there is no nugget, $\sigma_\varepsilon^2 = 0$, and obtain that

$$\mathrm{E}(S_{ii}) = \mathrm{E}(E(S_{ii} \big| \nu_i)) = \mathrm{E}(\nu_i) = \exp\left\{\mu + \frac{1}{2}\tilde{\sigma}^2\right\} \tag{1.26}$$

and

$$\begin{aligned} Var(S_{ii}) &= Var(E(S_{ii} | \nu_{ii})) + E(Var(S_{ii} | \nu_{ii})) \\ &= Var(\nu_i) + E\left(S_{ii} | \nu_{ii}\, \nu_i^2\right) \\ &= \exp\left\{2\mu + \tilde{\sigma}^2\right\}\left(\exp(\tilde{\sigma}^2)\left(1 + S_{ii} | \nu_{ii}\right) - 1\right) \\ &\approx \exp\left\{2\mu + \tilde{\sigma}^2\right\}\left(\exp(\tilde{\sigma}^2) - 1\right) \end{aligned} \tag{1.27}$$

By equating the sample moments of S_{ii} with the theoretical moments,

$$\overline{S} = \frac{1}{N}\sum_{i=1}^{N} S_{ii} = \exp\left\{\mu + \tfrac{1}{2}\tilde{\sigma}^2\right\} \tag{1.28}$$

and

$$\tfrac{1}{N}\sum_{i=1}^{N}(S_{ii} - \overline{S})^2 = \exp\left(2\mu + \tilde{\sigma}^2\right)\left(\exp(\tilde{\sigma}^2) - 1\right) \tag{1.29}$$

we can solve for μ and $1/\tilde{\sigma}^2$ to get:

$$\mu = \log(\overline{S}) - \tfrac{1}{2}\log\left(\frac{\tfrac{1}{N}\sum_{i=1}^{N}(S_{ii} - \overline{S})^2 + \overline{S}^2}{\overline{S}^2}\right) \tag{1.30}$$

and

$$\frac{1}{\tilde{\sigma}^2} = \frac{1}{\log\left(\dfrac{\frac{1}{N}\sum_{i=1}^{N}(S_{ii}-\bar{S})^2+\bar{S}^2}{\bar{S}^2}\right)} \tag{1.31}$$

The expressions in the right hand sides of eq. (30) and eq. (31) can be used as means in the prior distributions of μ and $1/\tilde{\sigma}^2$. They are sample-dependent and need to be computed in advance by the user. As variances in the prior distributions of μ and $1/\tilde{\sigma}^2$ one can use multiples of the means that render the priors conveniently diffuse.

3 Summary of Prior Distributions

The following is a list of the prior distributions set in the program. Parameters are specified by the alphanumeric names used in the template of the parameter input file provided below.

Deformation Parameters

$$a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} s1a & 0 \\ 0 & s2a \end{bmatrix}\right) \tag{1.32}$$

$$[\mathbf{W}] \propto \exp\left\{-\tfrac{1}{2tau}\left(W_1^T\tilde{\mathbf{S}}W_1 + W_2^T\tilde{\mathbf{S}}W_2\right)\right\}\mathbf{I}_{\{\mathbf{W}\in\mathbb{V}\times\mathbb{V}\}}. \tag{1.33}$$

Correlation Parameters

$$\theta(\theta_1) \sim \exp(ptheta) \tag{1.34}$$

$$\theta_2 \sim U(0, ptheta2) \tag{1.35}$$

Variance Parameters

$$\sigma_\varepsilon^2 \sim \exp(pnugget) \tag{1.36}$$

$$[\nu] \propto \frac{1}{\prod_{i=1}^{N} \nu_i} \left|\tilde{\Sigma}\right|^{-\frac{1}{2}} \exp\left\{-\tfrac{1}{2}\left(\log(\nu)-\mu\cdot 1)^T\right)\tilde{\Sigma}^{-1}\left(\log(\nu)-\mu\cdot 1)^T\right)\right\} \tag{1.37}$$

$\tilde{\Sigma}$ is an $N \times N$ matrix with elements $\tilde{\sigma}_{ij} = \tilde{\sigma}^2 \rho_{\tilde{\theta}}\left(\left\|f(x_i)-f(x_j)\right\|\right)$, and in this program, $\rho_{\tilde{\theta}}(d) = \exp\left(-\tilde{\theta}d\right)$

$$\tilde{\sigma}^2 \sim \Gamma(palpha, pbeta) \tag{1.38}$$

$$\tilde{\theta} \sim \exp(pthetat) \tag{1.39}$$

$$\mu \sim N(pmu, psigma2) \tag{1.40}$$

4 Program Input

Upon execution, the program will prompt the user for a number things, including the names of 4 or 5 input files:

1. A name for a folder that will be created to hold the various output files
2. The name of the input parameter file described in detail below
3. Whether to use the geographic coordinates as initial d-plane coordinates in the MCMC ("y"/"n"). Note that these coordinates determine initial estimates for the deformation parameters in $\mathbf{W}$ and $\mathbf{A}$.
4. (If the answer to 3. is "n", the user will be prompted for a text file with the initial D-plane coordinates (N rows, 2 coordinates/row). This if often used when restarting the MCMC program.)
5. The name of the file containing the geographic coordinates of the N sites (N rows, 2 coordinates/row).
6. The name of the file containing the $N \times N$ sample covariance matrix (N rows, N covariances/row).
7. The name of the file containing the vector of N initial values for the temporal variances.

The following text provides as an example template for the input parameter file, #2 above (excluding the lines in a red, underlined font which should not be included). The program reads only the first field on each line; the parameter names and explanatory definitions may be provided, as here, for user labeling.

Model structure & program control parameters

62	N:	number of sites
350	T:	number of temporal replications
n	V:	variance field, constant ("c") or nonconstant random field ("n")
p	C:	spatial correlation model, exponential ("e") or power exponential ("p")
	i,j:	indices of the two sites to be held fixed in the D-plane
100	I:	how many MCMC iterations to run before printing/saving results

1000	J:	the number of samples to print/save (total number of iterations is I*J)

Spatial deformation parameters

.01	tau:	the scale (variance) parameter for the multivariate normal prior for the parameters, W , of the nonlinear component of the thin-plane spline
.00001	w_mcs:	MCMC sampling s.d. for the spline nonlinear parameters W
10	s1a:	prior variances of the two spline linear parameters
10	s2a:	
.002	mcsa:	MCMC sampling s.d. for the two spline linear parameters

D-plane correlation parameters

.01	theta:	initial value for parameter theta of the D-plane [power] exponential correlation function
1	ptheta:	parameter of the exponential prior on theta
.02	theta_mcs:	MCMC sampling s.d. for theta (sampling is from a gamma distribution with mean being the current MCMC estimate)
1.5	theta2:	initial value for parameter theta2 of the D-plane power exponential correlation function
2	ptheta2:	upper limit of the uniform prior on theta2 [this is reasonably left fixed at the value 2 as the domain of theta2 is (0,2).]
.02	theta2_mcs:	MCMC sampling s.d. for theta2 (sampling from a gamma distribution with mean being the current MCMC estimates)
1	nugget:	initial value of the nugget parameter of the D-plane correlation function
1	pnugget:	parameter of the exponential prior on the nugget
.02	nugget_mcs:	MCMC sampling s.d. for nugget

Variance field:

0	imu:	initial value of the mean of the log variance field
6	isigma2:	initial value of the variance of the log variance field
2	pmu:	mean of the normal prior for the variance field nu
8	psigma2:	variance of the normal prior for the variance field nu
20	palpha	parameter alpha of the gamma prior for the variance of the log variance field
5	pbeta	parameter beta of the gamma prior for the variance of the log variance field [Note that there are no MCMC sampling parameters for mu and sigma2 because these are updated using Gibbs steps, rather than Hastings. But there is a Hastings step for the values of the variance field at the monitoring sites.]
98	nu_lb:	MCMC uniform proposal distribution for the elements of the variance field at the monitoring sites, nu_i, is [(nu_lb/100)*current value,(100/nu_lb)*current value].
1	thetat:	initial value of the parameter thetat of the exponential spatial correlation function for the variance field
1	pthetat:	parameter of the exponential prior on thetat
	thetat_mcs:	MCMC sampling s.d. for thetat

5 Output

The following output files are created in the subdirectory specified by the user as noted above. Each contains J samples from posterior distributions.

A.new: $(2J \times 2)$-dim matrix **A** (Each (2×2)-dim submatrix results from one run).

c.new: $(J \times 2)$-dim matrix **c** (Each row results from one run).

debug.new: numeric output for debugging purposes only

last_coord.txt: the last $N \times 2$ D-plane coordinate matrix- generated by the MCMC algorithm, typically used as the initial coordinates when restarting the program to run further interations.

last_varvalue.txt: the last N-vector of temporal variances generated by the MCMC algorithm, typically used to provide initial values when restarting the program to run further iterations.

mu: J-dim vector μ.

newcoord.txt: $(NJ \times 2)$-dim matrix of D-plane coordinates (each $(N \times 2)$-dim submatrix results from one run).

nu.new: J-dim vector v if the temporal variance is constant.

$J \times N$ matrix $\boldsymbol{v}$ if the temporal variance is non-constant (each row results from one run).

nugget.new: J-dim vector

results.new: Summary of the run: input, running time, acceptance rates

rev_input.txt: A new input template containing as initial parameter values, the results from the final MCMC iteration just run. To be used to restart the program.

Newcoord.txt: $(NJ \times 2)$-dim matrix of D-plane coordinates (each $(N \times 2)$-dim submatrix results from one run).

Sigma2.new: J-dim vector $\tilde{\sigma}2$.

theta: J-dim vector – θ if the correlation model is exponential and $\theta 1$ if the correlation model is power-exponential.

theta2: J-dim vector $\theta 2$ if the correlation model is power-exponential.

thetat: J-dim vector $\tilde{\theta}$.

W.new: ($NJ \times 2$)-dim matrix **W** (Each ($N \times 2$)-dim submatrix results from one run).