Estimation in the spatial linear model with a point source using MCMC

Graciela González-Farías[1], Felipe Peraza[1] and Jacqueline M. Hughes-Oliver[2]
[1]Centro de Investigación en Matemáticas, A.C.
[2]North Carolina State University.

November 26, 2001

Abstract

A point source is an entity that drives a nonstationary stochastic process either directly or indirectly . In this work, we consider the problem of estimation of the distribution of parameters for the spatial model, in the case of a fixed point source, using MCMC. The effect of the point source, is modelled supposing a deterministic trend function of the sites and the point source's coordinate, and the covariance function considered in Hughes-Oliver and Gonzáles-Farías, (1999) .

1 Introduction

One important problem in the study of environmental data, is the presence of point sources (factories, accidents, ...) that modify the original structure of the data. The effect on the trend and correlation behavior due to the point source can be modelled by choosing a function of the sites and the point source's coordinate.

A shock that modify a possible stationary data is consider a point source. Hughes-Oliver and González-Farías (1999) , (HOGF from now on), define a covariance function that combines the general exponential covariance $R_1(s,t) = \sigma^2(-\tau\|\mathbf{t}-\mathbf{s}\|^m)$ with the nonstationary Wiener process having covariance function $R_2(s,t) = \prod_{i=1}^d \min(t_i, s_i)$. Finally they obtain the parametric covariance function

$$R(\mathbf{s}_i, \mathbf{s}_j; \boldsymbol{\theta}) = \sigma^2 \exp\{-\tau [d(\mathbf{s}_i, \mathbf{s}_j)]^m\} \exp(\min\{[\delta + d(\mathbf{s}_i, c)]^a, [\delta + d(\mathbf{s}_j, c)]^a\}), \quad (1)$$

and estimate the parameter in (1) by maximum likelihood methods .

Some properties of the behavior of data with respect to the point source are described by the parameter a. The parameter a controls the nonstationarity behavior, in particular, $a = 0$ gives the stationary exponential covariance; when $a > 0$, the variance increases as you move away from the point source and decreases when $a < 0$. This parameter also controls the correlation between observations with respect to the point source. We give special attention to this parameter in this work.

Establishing the distribution function of a particular parameter or a test statistics is not and easy task. However, since we have a model based problem, it makes sense to use MCMC methods (Diggle, P.J., Tawn, J.A. and Moyeed, R.A, 1998) to determine the distribution of the parameters, based on (1).

In this work, we consider the estimation of the density function of the parameters using MCMC methods, for the covariance function considered in HOGF. We present a small simulation study to show the performance and an application to real data.

2 Methodology

Let $Y(\mathbf{s})$, represent measurements in $\mathbf{D} \subset R^2$ and $\mathbf{c} \in \mathbf{D}$ the coordinate of a point source, we model

$$Y(\mathbf{s}) = m(\mathbf{s}) + U(\mathbf{s}) \tag{2}$$

where

$$m(\mathbf{s}) = \sum_{j=1}^{k+1} f_{j-1}(\mathbf{s})\, \beta_{j-1} \tag{3}$$

and $\boldsymbol{\beta} = (\beta_0, ..., \beta_k)'$ is unknown, the functions f_0, f_1,..., f_k are known, and $u(\mathbf{s})$ is a random process with zero mean. The information about $\mathbf{c}$, could be included in $m(\mathbf{s})$, in the covariance function or both.

We write (2) as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{4}$$

where $\mathbf{y}$ is a $n \times 1$ vector of observations, $\mathbf{X}$ is a $n \times (k+1)$ matrix with elements (i, j) given by $f_{j-1}(s_{1i}, s_{2i})$, and $\mathbf{u} \sim N_n(0, \Omega_{\boldsymbol{\theta}})$,where

$$(\Omega_{\boldsymbol{\theta}})_{ij} = cov\,(u(\mathbf{s}_i), u(\mathbf{s}_j)) = R(\mathbf{s}_i, \mathbf{s}_j; \boldsymbol{\theta}).$$

and $\mathbf{s}_i, \mathbf{s}_j$ are two sites.

Assuming the covariance function in HOGF, denote $\boldsymbol{\theta} = (\sigma^2, \tau, m, \delta, a)$, that is,

$$R(\mathbf{s}_i, \mathbf{s}_j; \boldsymbol{\theta}) = \sigma^2 \exp\{-\tau [d(\mathbf{s}_i, \mathbf{s}_j)]^m\} \exp(\min\{[\delta + d(\mathbf{s}_i, c)]^a, [\delta + d(\mathbf{s}_j, c)]^a\}),$$

where $d(\cdot,\cdot)$ is the Euclidean distance and $\mathbf{c}$, is the point source.

In the following work we take m and δ fixed ($m = 1$ and $\delta = 0.5$, see HOGF for details). Let $\boldsymbol{\theta} = (\sigma^2, \tau, a)$, where $\sigma^2 > 0$, $\tau > 0$, and $-\infty < a < \infty$.

We will use a Gibbs sampling scheme to find the posterior distribution of $\boldsymbol{\beta}$ and σ^2 and a Langevin-Hasting type algorithm, to find the posterior distribution of $\boldsymbol{\tau}$ and a.

The likelihood function for the model is given by,

$$f(\mathbf{y}|\boldsymbol{\theta}, \boldsymbol{\beta}) \propto |\Omega_{\boldsymbol{\theta}}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^t\, \Omega_{\boldsymbol{\theta}}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})\right\} \tag{5}$$

We use the inverse gamma as the conjugate prior for σ^2 , the scale parameter (Carlin and Louis, 1998), a prior gamma distribution for τ and a standard normal distribution for a, as follows:

$$\sigma^2 \sim IGa(\alpha_1, \gamma_1),$$

$$f(\sigma^2) \propto \frac{e^{-1/(\gamma_1 \sigma^2)}}{(\sigma^2)^{\alpha_1 + 1}},\ \sigma^2 > 0$$

$$\tau_1 = \ln(\tau) \sim N(0, 1),$$

and

$$a \sim N(0,1).$$

For the parameter $\boldsymbol{\beta}$, we consider the following prior:

$$\boldsymbol{\beta} \sim N(\mathbf{b}_0, \Sigma_0)$$

where $\alpha_1, \gamma_1, \alpha_2, \gamma_2, \mathbf{b}_0$ and Σ_0, are given.

The full conditional distribution is

$$\begin{aligned} f(\boldsymbol{\theta},\boldsymbol{\beta}|\mathbf{y}) \quad &\propto \quad |\Omega_{\boldsymbol{\theta}}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^t\,\Omega_{\boldsymbol{\theta}}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\right\}\frac{e^{-1/\left(\gamma_1\sigma^2\right)}}{(\sigma^2)^{\alpha_1+1}}e^{-\tau_1^2/2}e^{-a^2/2} \\ &\times \exp\left\{-(\boldsymbol{\beta}-\boldsymbol{\beta}_0)^t\,\Sigma_0^{-1}(\boldsymbol{\beta}-\boldsymbol{\beta}_0)\right\}. \end{aligned} \tag{6}$$

Thus the posterior distribution of $\boldsymbol{\beta}$, given $\boldsymbol{\theta}$ is:

$$\boldsymbol{\beta}|\boldsymbol{\theta},\mathbf{y} \sim N(\mathbf{Dd},\mathbf{D}), \tag{7}$$

where

$$\mathbf{D}^{-1} = \mathbf{X}^T\Omega_{\boldsymbol{\theta}}^{-1}\mathbf{X} + \Sigma_0^{-1}$$

and

$$\mathbf{d} = \mathbf{X}^T\Omega_{\boldsymbol{\theta}}^{-1}\mathbf{y} + \Sigma_0^{-1}\mathbf{b}_0.$$

The posterior distribution of σ^2 is,

$$\begin{aligned} f\left(\sigma^2|\mathbf{y},\boldsymbol{\beta},\tau,a\right) &= f(\mathbf{y}|\boldsymbol{\theta},\boldsymbol{\beta})\,f\left(\sigma^2\right) \\ &\propto |\Omega_{\boldsymbol{\theta}}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T\,\Omega_{\boldsymbol{\theta}}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\right\}\frac{e^{-1/\left(\gamma_1\sigma^2\right)}}{(\sigma^2)^{\alpha_1+1}} \\ &= \left(\sigma^2\right)^{-n/2}\left|\Omega_{\boldsymbol{\theta}_{-1}}^{-1}\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T\,\Omega_{\boldsymbol{\theta}_{-1}}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\right\}\frac{e^{-1/\left(\gamma_1\sigma^2\right)}}{(\sigma^2)^{\alpha_1+1}} \\ &\propto \left(\sigma^2\right)^{-n/2}\left(\sigma^2\right)^{-\alpha_1-1} \exp\left\{-\frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T\,\Omega_{\boldsymbol{\theta}_{-1}}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) - \frac{1}{\gamma_1\sigma^2}\right\} \\ &= \left(\sigma^2\right)^{-(n/2+\alpha_1+1)} \exp\left\{-\frac{1}{\sigma^2}\left[\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T\,\Omega_{\boldsymbol{\theta}_{-1}}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) + \frac{1}{\gamma_1}\right]\right\} \\ &= \exp\left\{-\frac{1}{\sigma^2\left[(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T\,\Omega_{\boldsymbol{\theta}_{-1}}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) + \frac{1}{\gamma_1}\right]^{-1}}\right\}\left(\sigma^2\right)^{-(n/2+\alpha_1)-1} \end{aligned}$$

thus

$$\sigma^2|\mathbf{y},\boldsymbol{\beta},\tau,a \sim IGa\left(\frac{1}{2}n+\alpha_1, \left[\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T\,\Omega_{\boldsymbol{\theta}_{-1}}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) + \frac{1}{\gamma_1}\right]^{-1}\right) \tag{8}$$

where $\Omega_{\boldsymbol{\theta}} = \sigma^2\Omega_{\boldsymbol{\theta}_{-1}}$. Thus, we can obtain a sample of $\sigma^2|\mathbf{y},\boldsymbol{\beta},\tau,a$ directly from the inverse gamma (8).

The posterior distribution of θ_2, θ_3 is,

$$\begin{aligned} f(\tau, a|\mathbf{y}, \boldsymbol{\beta}) &= f(\mathbf{y}|\boldsymbol{\theta}, \boldsymbol{\beta}) f(\tau) f(a) \\ &\propto |\Omega_{\boldsymbol{\theta}}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T \Omega_{\boldsymbol{\theta}}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\right\} e^{-\tau_1^2/2} e^{-a^2/2} \\ &\propto \left|\Omega_{\boldsymbol{\theta}_{-1}}\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T \Omega_{\boldsymbol{\theta}}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) - \frac{\tau_1^2}{2} - \frac{a^2}{2}\right\} \end{aligned} \tag{9}$$

note that the posterior distribution of (τ, a) is not a known distribution. So, we will use a Metropolis-Langevin [3] type algorithm as described next.

Let, $\boldsymbol{\gamma} = (\tau, a)$, the Langevin transition is based on the proposal

$$q(\cdot|\boldsymbol{\gamma}) = N(\xi(\boldsymbol{\gamma}), hI), \tag{10}$$

where

$$\xi(\boldsymbol{\gamma}) = \boldsymbol{\gamma} + \tfrac{h}{2}\nabla \log \pi(\boldsymbol{\gamma}),$$

and $\pi(\boldsymbol{\gamma})$ is the posterior of the target distribution, in this case given by the posterior (9), $\nabla \log \pi(\boldsymbol{\gamma})$ is the gradient of the logarithm of $\pi(\boldsymbol{\gamma})$, and h is the step variance. Some guidelines for choosing h are given in Roberts and Rosenthal, (1998).

Thus, we accept the new value $\boldsymbol{\gamma}'$ with probability

$$\alpha(\boldsymbol{\gamma}, \boldsymbol{\gamma}') = 1 \wedge \tfrac{\pi(\boldsymbol{\gamma}') q(\boldsymbol{\gamma}', \boldsymbol{\gamma})}{\pi(\boldsymbol{\gamma}) q(\boldsymbol{\gamma}, \boldsymbol{\gamma}')} \tag{11}$$

and using (10) we get,

$$\begin{aligned} q(\boldsymbol{\gamma}, \boldsymbol{\gamma}') &= \frac{1}{\sqrt{2\pi h}} \exp\left\{-\|\boldsymbol{\gamma}' - \xi(\boldsymbol{\gamma})\|^2 \diagup 2h\right\} \\ &= \frac{1}{\sqrt{2\pi h}} \exp\left\{-\left\|\boldsymbol{\gamma}' - \boldsymbol{\gamma} - \tfrac{h}{2}\nabla \log \pi(\boldsymbol{\gamma})\right\|^2 \diagup 2h\right\} \end{aligned}$$

then substituting in (11), we obtain

$$\begin{aligned} \alpha(\boldsymbol{\gamma}, \boldsymbol{\gamma}') &= 1 \wedge 1 \wedge \tfrac{\pi(\boldsymbol{\gamma}') q(\boldsymbol{\gamma}', \boldsymbol{\gamma})}{\pi(\boldsymbol{\gamma}) q(\boldsymbol{\gamma}, \boldsymbol{\gamma}')} \\ &= 1 \wedge \frac{\left|\Omega_{\theta'_{-1}}\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\mathbf{X}\beta)^T \Omega_{\theta'}^{-1}(\mathbf{y}-\mathbf{X}\beta) - \frac{(\tau_1')^2}{2} - \frac{(a')^2}{2}\right\} \exp\left\{-\frac{1}{2h}\left\|\boldsymbol{\theta} - \xi(\boldsymbol{\theta}')\right\|^2\right\}}{\left|\Omega_{\theta_{-1}}\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\mathbf{X}\beta)^T \Omega_{\theta}^{-1}(\mathbf{y}-\mathbf{X}\beta) - \frac{\tau_1^2}{2} - \frac{a^2}{2}\right\} \exp\left\{-\frac{1}{2h}\|\boldsymbol{\theta}' - \xi(\boldsymbol{\theta})\|^2\right\}}. \end{aligned}$$

3 Simulation Results

We use the same configuration for the covariance parameters as it was suggested by HOGF. That is, $\sigma^2 = 1$, $\tau = .05, 2.5$ and $a = -1.2, 0, 1.2$, under model (1) in an 11×11 regular grid with a drift function

$$m(\mathbf{s}) = \beta_0 + \beta_1 \{\ln(dc) - mean[\ln(dc)]\}$$

using the values 1 and 2 for β_0 and β_1, respectively.

We simulate chains of size 20,000, deleting the first 10,000 values. The value h in the Langevin-Hasting step is chosen to obtain an acceptance probability around 0.5 to ensure convergence (Roberts, G.O. and Rosenthal, J. S. , 1998). The results are summarize in Table 1.

The parameter σ^2, is underestimated in all cases when the parameter $\tau = 0.05$, and behaves better for $\tau = 2.5$ and $a = 1.2$, (the median is 1.075931). Except for $\tau = 0.05$, $a = -1.2$, $\tau = 0.05$, $a = 0$, all the 95% probability intervals contain the true value $\sigma^2 = 1$.

The parameter τ is reasonably estimated when the true value is 0.05 ; underestimated when $a = -1.2$ and $\tau = 2.5$; overestimated when $a = 0$ and $\tau = 2.5$. Except for the case $\tau = 0.05, a = 0$, all 95% probability intervals contain the true value.

The behavior of parameter a is good in all cases except when $\tau = 2.5, a = -1.2$, however the 95% probability interval includes the true value.

It is interesting to observe the posterior distribution for the parameter a, in all cases (independently of the parameter τ) , shows a right skewed distribution and becomes almost symmetric when a is positive (Fig. 1).

4 Example: electromagnetism data

The measurements of electromagnetism are taken at sites falling on a regular grid, as shown in Figure 2, where the sites are one meter apart in both vertical and horizontal directions. Electromagnetism is expected to be a fairly constant pattern in the field, but an existing metal pole affects the measuring device so that, we observed circular rings around the pole and expect a bigger variance near the pole. In this sense the metal pole is a point source.

Let $Y(s)$ the electromagnetism data previously transformed by

$$y = \ln \left[-\ln \left(\frac{data - 46300}{-46300} \right) \right]$$

Let dc represents the distance from the metal pole to site $\mathbf{s}$, and consider the drift function

$$m(\mathbf{s}) = \beta_0 + \beta_1 \{ \ln(dc) - mean[\ln(dc)] \} .$$

We implement maximum likelihood and restricted maximum likelihood methods in order to estimate the covariance parameters (Table 3). The maximum likelihood estimators are $\left(\hat{\sigma}^2, \hat{\tau}, \hat{a}\right)_{MV} =$ $(0.004076, 0.0283753, -1.7671504)$ and the restricted maximum likelihood estimators are $\left(\hat{\sigma}^2, \hat{\tau}, \hat{a}\right)_{REML} =$ $(0.004814, 0.024077, -1.8196177)$. There is not too much difference since we have only two parameters in the mean function.

We generate a chain of size $20,000$, with an acceptation probability of 0.64, we observe a strong autocorrelation and since *iid* samples are needed, we took a subsample of the chain to ensure convergence (Robert and Mengersen, 1999). Choosing a batch of 10, we observe a small autocorrelation in all cases, so that the following results are obtained under this subsample scheme.

In Figure 3, we can observe the convergence of the chains, which are satisfactory and the distributions look as it is expected from the simulation results. Table 3, summarize the results of the MCMC. For illustration, we compare the estimations from the likelihood methods (Table 2) with those obtain by the MCMC procedure. The maximum likelihood estimators are not too close to the median of distribution obtained by MCMC, however, the 90% probability intervals contain the estimators found above. It is not always the case for REML estimators, except for a , β_0 and β_1.

Note that the median for the distribution of τ is small (0.047), it means that there is a strong spatial dependence between observations (for equally spaced observations it behaves almost as a unit root). The range of distribution for a is always negative (as physics expected!!), that is, the variance decreases as we move away from the point source. This is reasonable because the device is less affected far from the metal pole.

An advantage with respect to the results in HOGF, is the fact that we obtain the posterior distributions for all the parameters, in particular a, and we can construct probability intervals to test the hypothesis $H_o : a = 0$, which is of particular interest in this case. Finally, the function drift is estimated in median sense by

$$m(\mathbf{s}) = 1.52099 + .37100 \{\ln(dc) - mean[\ln(dc)]\}.$$

The drift is an increasing nonnegative function of dc, thus, the median value of measurements also increases when we move away from the metal pole.

5 Conclusions

The presence of a point source in spatial data can modify the structure of the data. So, in pollution data, just to mention one case, will be particularly important to estimate the effect of the point sources given the fact that, the prediction values and mainly , the errors associated with those prediction values maybe be greatly affected forcing us to take erroneous decisions or actions.

Further work is needed to establish a more stable behavior of the parameters involved given different factors as the strong correlation and the explosive variances around the point source, as well as not having enough data in that region. Also it will be important to study the predictive functions and compare with the results obtain using likelihood methods.

In the case of several point sources known or unknown, the models will be more complex and the MCMC methods will be a good instrument to estimate the parameters involved as well as the predictive functions.

References

[1] Carlin B.P. and Louis, T.A.(1998) *Bayes and Empirical Bayes Methods for Data Analysis.* Chapman and Hall/CRC: London.

[2] Diggle, P.J., Tawn, J.A. and Moyeed, R.A. (1998). Model-based geostatistics (with discussion). *Applied Statistics* **47**, 299-350.

[3] Grenadder, U. and Miller, M. (1994). Representations of knowledge in complex systems (with discussion). *J. Roy. Satist. Soc. Ser. B* **56**, 549-603.

[4] Hughes-Oliver, J. M. and González-Farías, G. (1999). Parametric covariance models for shock-induced stochastic processes. *Journal of Statistical Planning and Inference* **77**, 51-72.

[5] Robert, C.P. and Mengersen, K.L. (1999) . Reparametrization issues in mixture estimation and their bearings on the Gibbs sampler. *Comput. Statist. Date Ana.* **29**, *325-342.*

[6] Roberts, G.O. and Rosenthal, J. S. (1998). Optimal scaling of discrete approximations to Langevin diffusions. *J. Roy. Satist. Soc. Ser. B* **60**, 255-268.

6 Tables and figures

$\tau = 0.05, a = -1.2$					
	mean	median	minimum	maximum	95% PI
σ^2	0.427	0.397	0.114	1.447	$(0.210, 0.805)$
τ	0.087	0.087	0.025	0.408	$(0.045, 0.167)$
a	-1.142	-1.095	-3.041	-0.1406	$(-2.099, -0.449)$
β_0	-0.047	-0.051	-2.814	2.967	$(-1.161, 1.086)$
β_1	2.259	2.264	0.862	3.627	$(1.596, 2.911)$
$\tau = 0.05, a = 0$					
	mean	median	minimum	maximum	95% PI
σ^2	0.440	.410	0.140	2.326	$(0.204, 0.895)$
τ	0.122	0.123	0.036	0.431	$(0.060, 0.231)$
a	-0.040	0.011	-1.414	0.438	$(-0.920, 0.285)$
β_0	0.193	0.193	-3.309	4.224	$(-1.301, 1.700)$
β_1	2.896	2.900	0.957	4.530	$(2.118, 3.637)$
$\tau = 0.05, a = 1.2$					
	mean	median	minimum	maximum	95% PI
σ^2	0.604	0.558	0.229	2.320	$(0.322, 1.146)$
τ	0.072	0.076	0.028	0.118	$(0.035, 0.105)$
a	1.214	1.214	1.135	1.285	$(1.173, 1.253)$
β_0	0.206	0.201	-24.092	19.929	$(-9.368, 9.801)$
β_1	1.327	1.317	-15.359	15.981	$(-5.34, 8.047)$
$\tau = 2.5, a = -1.2$					
	mean	median	minimum	maximum	95% PI
σ^2	0.795	0.783	0.218	2.465	$(0.348, 1.354)$
τ	1.529	1.529	2.406	5.040	$(0.814, 2.932)$
a	-0.645	-0.525	-3.06	0.299	$(-2.112, 0.157)$
β_0	0.970	0.974	-0.192	2.063	$(0.596, 1.332)$
β_1	1.996	1.990	0.182	3.935	$(1.258, 2.746)$
$\tau = 2.5, a = 0$					
	mean	median	minimum	maximum	95% PI
σ^2	0.882	0.775	0.234	3.552	$(0.394, 2.416)$
τ	3.312	3.108	0.998	32.492	$(1.597, 11.752)$
a	0.044	0.123	-1.501	0.570	$(-1.888, 0.412)$
β_0	0.810	0.809	0.122	2.031	$(0.480, 1.150)$
β_1	2.524	2.521	0.264	4.132	$(1.829, 3.248)$
$\tau = 2.5, a = 1.2$					
	mean	median	minimum	maximum	95% PI
σ^2	1.076	1.030	0.380	2.975	$(0.605, 1.824)$
τ	2.928	2.807	1.217	7.784	$(1.562, 6.522)$
a	1.198	1.198	1.101	1.287	$(1.149, 1.245)$
β_0	.240	0.262	-7.47	7.791	$(-3.512, 3.948)$
β_1	-0.715	-0.692	-7.849	6.949	$(-4.304, 2.841)$

Table 1. Statistics of simulation.

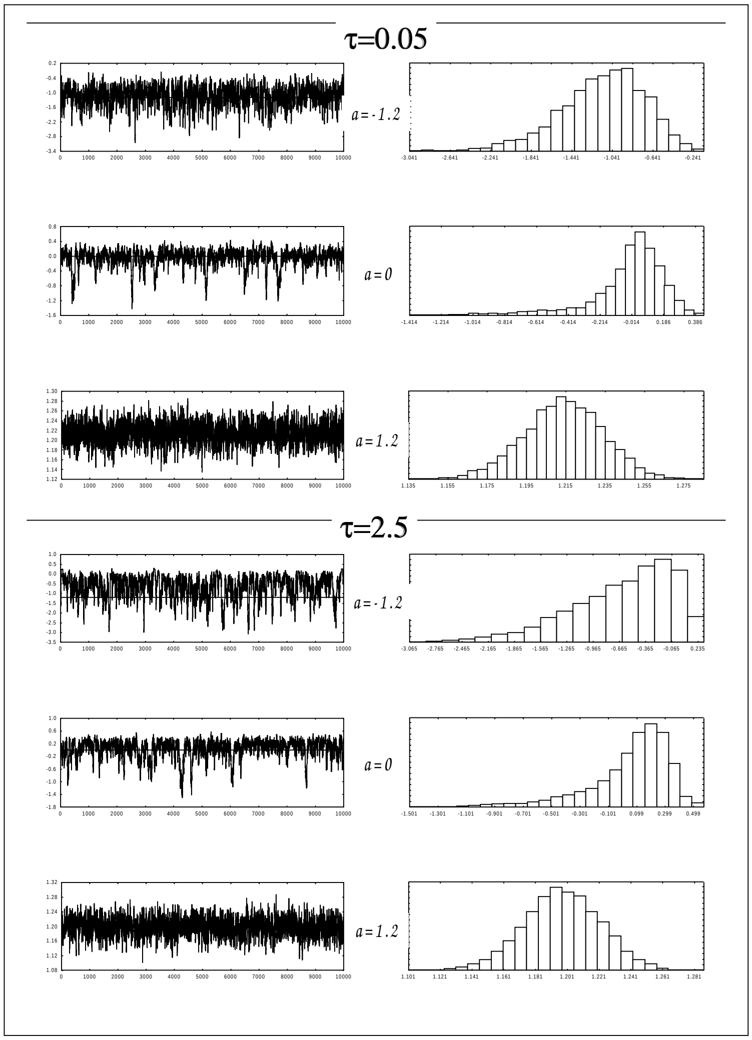

Figure 1: Trace and histogram for covariance parameters on simulation.

	$\hat{\sigma}^2$	$\hat{\tau}$	$\hat{a}$	$\hat{\beta}_0$	$\hat{\beta}_1$
MV	0.00408	0.02837	−1.76715	1.51244	0.36496
REML	0.00481	0.02408	−1.81962	1.51246	0.36490

Table 2. Likelihood estimators for magnetism data

	mean	median	minimum	maximum	90% PI
σ^2	0.00275	0.00257	0.00086	0.00800	$(0.00153, 0.00408)$
τ	0.04996	0.04714	0.01460	0.15606	$(0.02534, 0.08368)$
a	−1.54589	−1.52644	−2.87905	−0.58132	$(-2.05652, -1.09841)$
β_0	1.52059	1.52099	1.35350	1.72285	$(1.44670, 1.59323)$
β_1	0.37111	0.37100	0.30070	0.445030	$(0.33988, 0.40311)$

Table 3. Statistics for electromagnetism data

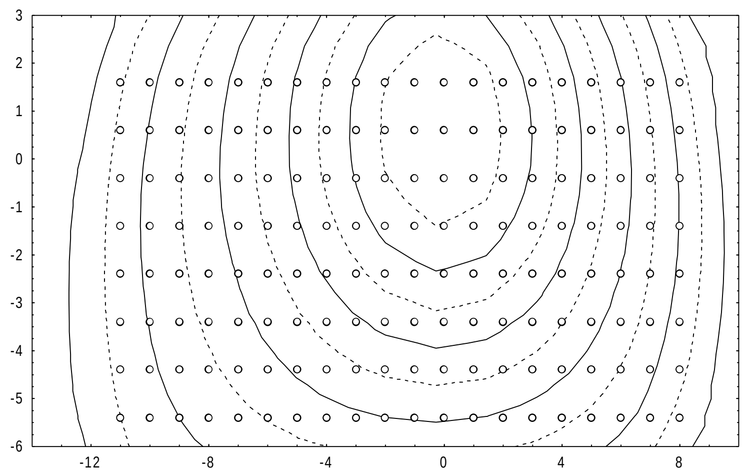

Figure 2: Electromagnetism data

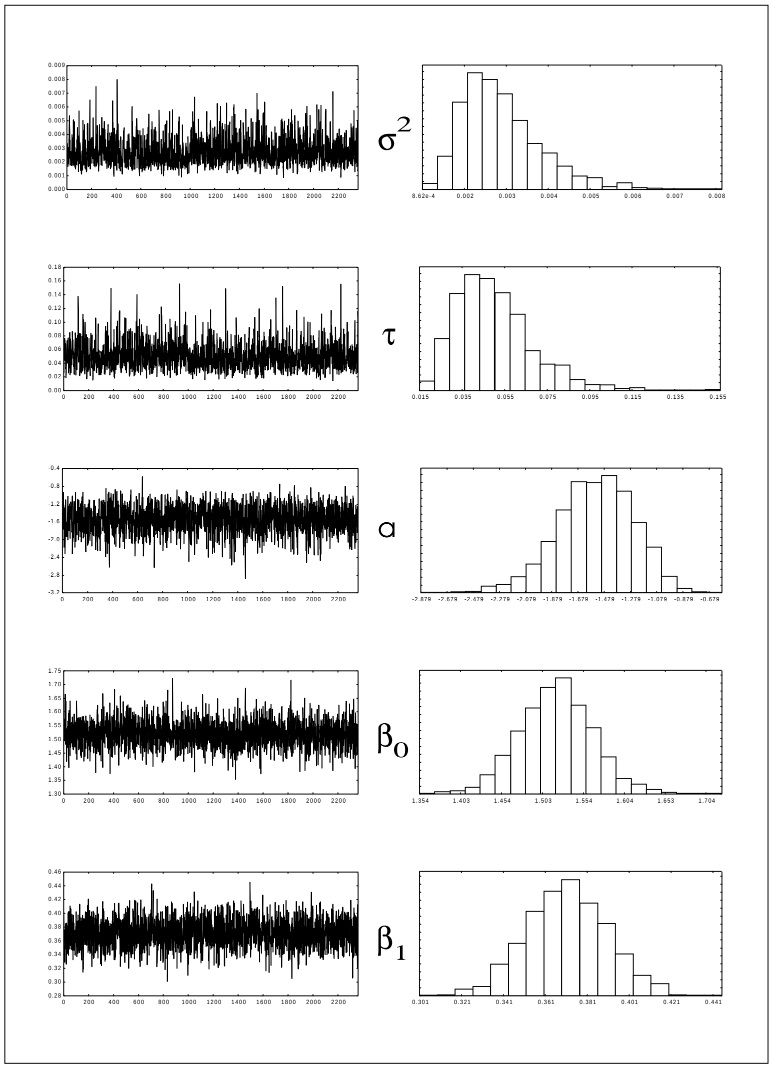

Figure 3: Trace and histogram for electromagnetism data