Article

Bayesian Inference on the Memory Parameter for Gamma-Modulated Regression Models [†]

Plinio Andrade [1], Laura Rifo [2],*, Soledad Torres [3] and Francisco Torres-Avilés [4],‡

[1] Institute of Mathematics and Statistics, University of São Paulo, Rua do Matão 1010, 05508-090 São Paulo, Brazil; E-Mail: plinio@ime.usp.br

[2] Institute of Mathematics and Statistics, University of Campinas, Rua Sérgio Buarque de Holanda 651, 13083-859 Campinas, Brazil

[3] CIMFAV—Facultad de Ingeniería, Universidad de Valparaíso, General Cruz 222, Valparaíso 2362905, Chile; E-Mail: soledad.torres@uv.cl

[4] Departamento de Matemática y Ciencia de la Computación, Universidad de Santiago de Chile, Av. Libertador Bernardo O'Higgins 3363, Santiago 9170022, Chile

[†] This paper is dedicated to the memory of Professor Francisco Torres-Avilés.

[‡] Deceased.

* Author to whom correspondence should be addressed; E-Mail: lramos@ime.unicamp.br; Tel.: +55-19-3521-6052; Fax: +55-19-3289-5766.

Academic Editors: Carlos Alberto De Bragança Pereira and Adriano Polpo

Abstract: In this work, we propose a Bayesian methodology to make inferences for the memory parameter and other characteristics under non-standard assumptions for a class of stochastic processes. This class generalizes the Gamma-modulated process, with trajectories that exhibit long memory behavior, as well as decreasing variability as time increases. Different values of the memory parameter influence the speed of this decrease, making this heteroscedastic model very flexible. Its properties are used to implement an approximate Bayesian computation and MCMC scheme to obtain posterior estimates. We test and validate our method through simulations and real data from the big earthquake that occurred in 2010 in Chile.

1. Introduction

Diffusion processes have been a cornerstone of stochastic modeling of time series data, particularly in areas such as finance [1] and hydrology [2]. Many extensions to the classic diffusion model have been developed in recent years, addressing such diverse issues as asymmetry, kurtosis, heteroscedasticity and long memory; see, for instance, [3].

In the simplest case, the increments of a diffusion model are taken as independent Gaussian random variables, making the process a Brownian motion. In this work, by contrast, processes with long memory and non-Gaussian increments are considered.

The proposed model is a generalization of the Gamma-modulated (G-M) diffusion process, in terms of the memory parameter. This model was developed in [4] to address an asset market problem, extending the ideas of the Black–Scholes paradigm and using Bayesian procedures for model fitting. In that work, the memory parameter was assumed to be known and fixed, with some particular cases, such as the standard Brownian motion and the Student process. The latter one is a generalization of the Student process previously presented in [5], the marginals of which have a t-Student distribution with fixed degrees of freedom and a long memory structure.

Here, we enlarge the parameter space considering that the memory parameter is also unknown, provided we have a prior distribution on it.

This extension allows flexibility for the dependence structure of the process, where the Brownian motion and the G-M process become particular cases.

Typically, the trajectories generated by this process exhibit heteroscedasticity, with higher variability at the beginning, which we call "explosion at zero". In addition, as time increases, the variability decreases at a rate depending on the long memory parameter.

In particular, we will focus on estimation procedures for long-range memory stochastic models from a Bayesian perspective. Other parameters, such as location and dispersion, are also considered.

For the location and scale parameters, we can straightforwardly find natural conjugate prior distributions. However, the same does not occur for the memory parameter, as its marginal likelihood is not workable analytically. This implies that methods used for obtaining a posterior distribution, such as commonly-used likelihood-based solutions, are not suitable for this purpose.

In order to approximate the posterior distribution for the parameters involved, we propose a blended approximate Bayesian computation ABC-MCMC algorithm. This family of ABC algorithms and its very broad set of applications are well reviewed in [6]. In this work, the MCMC part is built for those components with full conditional posterior distributions that are able to be dealt with, and the ABC is implemented for the memory parameter. Grounded on previous results [7], for the ABC steps, an appropriate summary statistic was defined, based on the path properties and on the m-block variances. We obtain, by this proposal, very precise estimates for that parameter.

After generating samples from the posterior, a by-product is the e-value, an evidence measure for precise hypotheses, such as the Brownian motion and G-M cases. This measure was defined in [8] and used afterward, for instance, in [9–11].

We test and validate our method through simulations and illustrate it with data from the big earthquake that occurred in Chile in 2010.

The definition of the process and some properties are presented in Section 2. In Section 3, we describe the ABC-MCMC algorithm. The simulated and real data results are shown in Section 4, and finally, in Section 5, we give some final remarks.

2. Generalized Gamma-Modulated Process

Let us consider the standard Brownian motion, $\{B_t\}_{t>0}$, and a Gamma process, $\{\gamma_t\}_{t>0}$, as defined in [4].

A Gamma process is a pure-jump increasing Lévy process with independent and stationary Gamma increments for non-overlapping intervals. For this process, the intensity measure is given by $\kappa(x) = ax^{-1}\exp(-bx)$, for any positive x. That is, jumps whose size lies in the interval $[x, x + dx]$ occur as a Poisson process with intensity $\kappa(x)dx$. The parameters involved in the intensity measure are a, which controls the rate of jump arrivals, and b, the scaling parameter, which controls the jump sizes.

The marginal distribution of a Gamma process at time t is a Gamma distribution with mean at/b and variance at/b^2, allowing also the parametrization in terms of the mean, μ, and variance, σ^2, per unit time, that is, $a = \mu^2/\sigma^2$ and $b = \mu/\sigma^2$.

For the one-dimensional distributions, we have that $\alpha\gamma_t(a, b) = \gamma_t(a, b/\alpha)$ in distribution; $E(\gamma_t^n) = b^{-n}\Gamma(at + n)/\Gamma(at)$, $n \geq 0$, where $\Gamma(z)$ is the Gamma function; $E(\exp(\theta\gamma_t)) = (1 - \theta/b)^{-at}$, for $\theta < b$, and its characteristic function, $\phi_{\gamma_t}(u) = E(\exp(iu\gamma_t))$, is:

$$\phi_{\gamma_t}(u) = \left(1 - \frac{u}{b}i\right)^{-at}.$$

Given times $s < t$, $Corr(\gamma_s, \gamma_t) = \sqrt{s/t}$, and given $h > 0$, the density, $f_h(y)$, of the increment $\gamma_{t+h} - \gamma_t$ is given by the Gamma density function with mean ah/b and variance ah/b^2,

$$f_h(y) = \frac{y^{ah-1}b^{ah}e^{-by}}{\Gamma(ah)}.$$

In this work, we will consider $a = b = 1/2$.

Given a real value $\alpha \in [-1, 0]$, we define the generalized Gamma-modulated (G-M) process by:

$$^\alpha X_t = B_t\,\gamma_t^\alpha, \quad \text{for } t > 0. \tag{1}$$

Figure 1 shows typical realizations of the generalized G-M process for different values of the parameter α. In particular, the value $\alpha = 0$ corresponds to the Brownian motion and $\alpha = -0.5$ to the Student process studied in [4].

The next subsections present some useful path characteristics of the process that could lead us to choose this model as an appropriate one for a given problem and to help us make inferences about its parameters, such as the presence of long memory, the variability profile and the variance of the increment process.
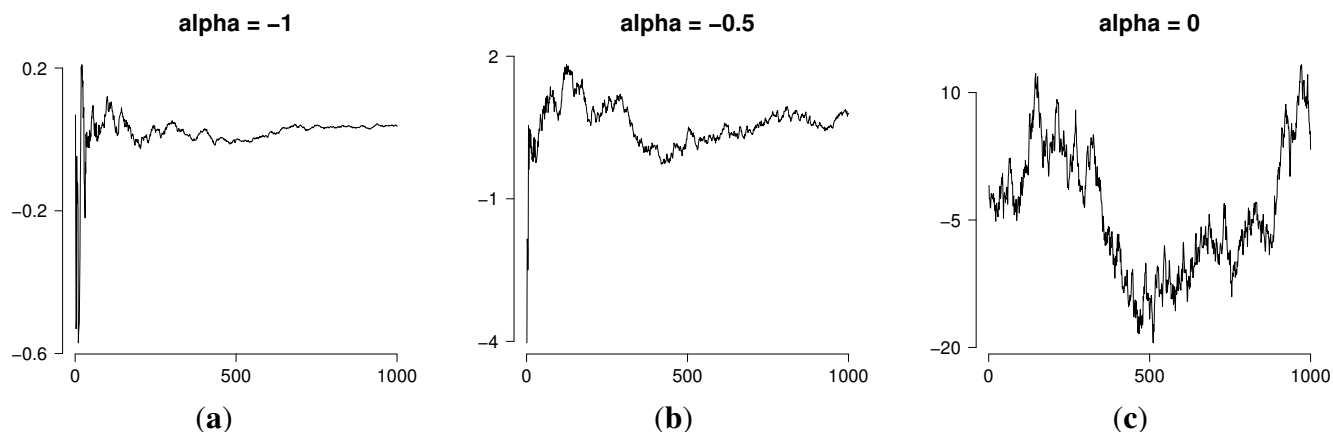
Figure 1. Realizations of the generalized Gamma-modulated (G-M) process for different values of α. (**a**) $\alpha = -1$; (**b**) $\alpha = -0.5$; (**c**) $\alpha = 0$.

2.1. Explosion at Zero

The graphs in Figure 1a,b, with $\alpha < 0$, show that the process is highly variable near $t = 0$, but as $t \to \infty$, its variability decreases. We call this path property "explosion at zero" and prove it in the next result.

Proposition 1. *Let $^{\alpha}X_t$ be the generalized Gamma-modulated process as defined in Equation ([1]). Then, for all $M > 0$, we have:*

$$\lim_{s \to 0} \mathbb{P}\left(|^{\alpha}X_s| > M\right) = \begin{cases} 1, & \text{if } \alpha < 0; \\ 0, & \text{if } \alpha = 0. \end{cases}$$

Proof. Let us consider $M > 0$ and $\alpha < 0$. Conditioning in γ_s, we obtain:

$$
\begin{aligned}
\mathbb{P}\left(|B_s| > M\gamma_s^{-\alpha}\right) &= \int_0^{\infty} \frac{1}{\Gamma(s/2)} \frac{x^{s/2-1}}{2^{s/2}} e^{-x/2} \mathbb{P}\left(|B_s| > Mx^{-\alpha}\right) dx \\
&= \int_0^{\infty} \frac{1}{\Gamma(s/2)} \frac{x^{s/2-1}}{2^{s/2}} e^{-x/2} \int_{Mx^{-\alpha}}^{\infty} 2 \frac{e^{-u^2/2s}}{\sqrt{2\pi s}} du\, dx \\
&\geq e^{-As^{-\alpha}/2} \frac{1}{2^{s/2}} \int_0^{As^{-\alpha}} \frac{x^{s/2-1}}{\Gamma(s/2)} \int_{MA}^{\infty} 2 \frac{e^{-v^2/2}}{\sqrt{2\pi}} dv\, dx \\
&\geq \left(\frac{MAe^{-(MA)^2/2}}{1 + (MA)^2}\right) \frac{e^{-As^{-\alpha}/2} A^{s/2} s^{-\alpha s/2}}{2^{s/2}\Gamma(s/2)} \frac{2}{s},
\end{aligned}
$$

with $A > 0$. The last quantity tends to one as $s \to 0$ for $\alpha < 0$.

Let us consider now the case $\alpha = 0$,

$$\mathbb{P}\left(|B_s| > M\right) = \int_{M/\sqrt{s}}^{\infty} 2 \frac{e^{-v^2/2}}{\sqrt{2\pi}} dv \leq e^{-M^2/2s} \frac{\sqrt{s}}{M},$$

that tends to zero as $s \to 0$. $\square$

2.2. The Increment Process

Let us now consider the increment process, $\Delta(^\alpha X_t) = B_t \gamma_t^\alpha - B_{t-1}\gamma_{t-1}^\alpha$. The next results describe the asymptotic behavior of the variance-covariance structure for these differences and, hence, for the process itself, since $^\alpha X_t$ has zero expectation.

Proposition 2. *For the increment process,* $\Delta(^\alpha X_t)$,

$$Var(\Delta(^\alpha X_t)) \sim t^{2\alpha}.$$

Proof. Let us observe that the increment process, $\Delta(^\alpha X_t) = B_t \gamma_t^\alpha - B_{t-1}\gamma_{t-1}^\alpha$, can be written as $\Delta(^\alpha X_t) = W_t + V_t$, where:

$$W_t = B_t(\gamma_t^\alpha - \gamma_{t-1}^\alpha), \quad V_t = (B_t - B_{t-1})\gamma_{t-1}^\alpha.$$

Working out each term, we obtain:

$$Var(V_t) = E((B_t - B_{t-1})^2 \, \gamma_{t-1}^{2\alpha}) = E(\gamma_{t-1}^{2\alpha}) \sim t^{2\alpha},$$

$$Var(W_t) = E(B_t^2(\gamma_t^\alpha - \gamma_{t-1}^\alpha)^2) = t \, E((\gamma_t^\alpha - \gamma_{t-1}^\alpha)^2) \sim t^{2\alpha-1},$$

and:

$$Cov(V_t, W_t) \sim (t-1)^\alpha(t^\alpha - (t-1)^\alpha) \sim t^{2\alpha-1}.$$

$\square$

This property leads us to consider the variance of the observed increment process, $Var(\Delta(^\alpha X_t)) \sim t^{2\alpha}$, as informative for the parameter α and, therefore, helpful to implement the approximated simulation for its marginal posterior.

In a data exploratory phase, we could examine the graph of the empirical variances of $\Delta(^\alpha X_t)$ from t to the end of the process, as a function of t in logarithmic scale. As t increases, that graph should become linear with slope 2α. Figure 2 exhibit this result for some values of α. Observe that the asymptotic result can be visualized from $\log t \approx 3$, that is from $t \approx 20$, for $\alpha < 0$.
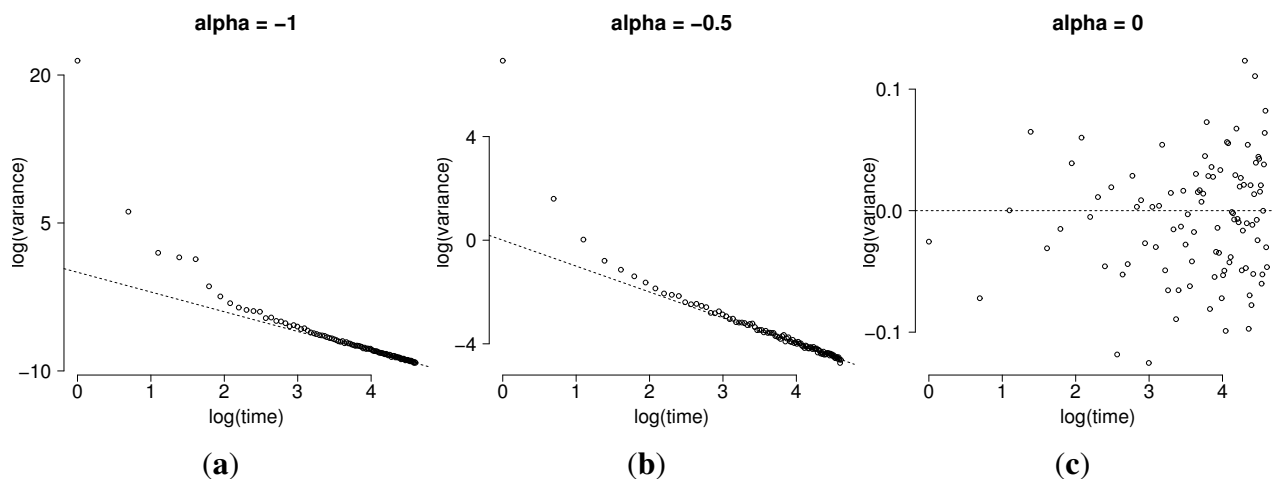
Figure 2. Sample distribution of $Var(\Delta(^\alpha X_t))$, for 1000 simulated trajectories of the generalized G-M process, for different values of α. (**a**) $\alpha = -1$; (**b**) $\alpha = -0.5$; (**c**) $\alpha = 0$.

Proposition 3. *The auto-covariance for lag n of the increment process, denoted by $M_\alpha(n)$, is of order $n^{\alpha-1}$, if $-1 < \alpha < 0$, and it is zero, if $\alpha = 0$, as n increases.*

Proof. Let us consider $n \geq 3$, because of the explosion at zero. Then:

$$
\begin{aligned}
M_\alpha(n) &= E\left(\Delta(^\alpha X_3)\,\Delta(^\alpha X_{n+3})\right)\\
M_\alpha(n) &= E\left((B_3\gamma_3^\alpha - B_2\gamma_2^\alpha)(B_{n+3}\gamma_{n+3}^\alpha - B_{n+2}\gamma_{n+2}^\alpha)\right)\\
&= E(B_3\gamma_3^\alpha B_{n+3}\gamma_{n+3}^\alpha) - E(B_3\gamma_3^\alpha B_{n+2}\gamma_{n+2}^\alpha)\\
&\quad - E(B_2\gamma_2^\alpha B_{n+3}\gamma_{n+3}^\alpha) + E(B_2\gamma_2^\alpha B_{n+2}\gamma_{n+2}^\alpha)\\
&= E(B_3 B_{n+3})E(\gamma_3^\alpha\gamma_{n+3}^\alpha) - E(B_3 B_{n+2})E(\gamma_3^\alpha\gamma_{n+2}^\alpha)\\
&\quad - E(B_2 B_{n+3})E(\gamma_2^\alpha\gamma_{n+3}^\alpha) + E(B_2 B_{n+2})E(\gamma_2^\alpha\gamma_{n+2}^\alpha)\\
&= 3E(\gamma_3^\alpha\gamma_{n+3}^\alpha - \gamma_3^\alpha\gamma_{n+2}^\alpha) - 2E(\gamma_2^\alpha\gamma_{n+3}^\alpha - \gamma_2^\alpha\gamma_{n+2}^\alpha)\,.
\end{aligned}
$$

Observe that:

$$
\begin{aligned}
E(\gamma_t^\alpha\gamma_s^\alpha) &= \frac{1}{2^{t/2}}\frac{1}{\Gamma((t-s)/2)\Gamma(s/2)}\int_0^\infty\int_0^{\gamma_t}\gamma_t^\alpha\gamma_s^{\alpha+s/2-1}(\gamma_t-\gamma_s)^{(t-s)/2-1}e^{-\gamma_t/2}d\gamma_s d\gamma_t\\
&= \frac{1}{2^{t/2}}\frac{1}{\Gamma(\frac{t-s}{2})\Gamma(s/2)}\int_0^\infty\int_0^{\gamma_t}\gamma_s^\alpha(\gamma_t-\gamma_s)^{(t-s)/2-1}\gamma_s^{s/2-1}d\gamma_s\gamma_t^\alpha e^{-\gamma_t/2}d\gamma_t\\
&= \frac{1}{2^{t/2}}\frac{1}{\Gamma(\frac{t-s}{2})\Gamma(s/2)}\int_0^\infty\int_0^1 z^{\alpha+s/2-1}(1-z)^{(t-s)/2-1}dz\gamma_t^{2\alpha+t/2-1}e^{-\gamma_t/2}d\gamma_t\\
&= \frac{1}{2^{t/2}}\frac{1}{\Gamma(\frac{t-s}{2})\Gamma(s/2)}\frac{\Gamma(\alpha+s/2)\Gamma((t-s)/2)}{\Gamma(\alpha+t/2)}\int_0^\infty\gamma_t^{2\alpha+t/2-1}e^{-\gamma_t/2}d\gamma_t\\
&= \frac{1}{2^{t/2}}\frac{1}{\Gamma(\frac{t-s}{2})\Gamma(s/2)}\frac{\Gamma(\alpha+s/2)\Gamma((t-s)/2)}{\Gamma(\alpha+t/2)}\frac{\Gamma(t/2+2\alpha)}{(1/2)^{t/2+2\alpha}}\\
&= 4^\alpha\frac{\Gamma(\alpha+s/2)\Gamma(2\alpha+t/2)}{\Gamma(s/2)\Gamma(\alpha+t/2)}\,.
\end{aligned}
$$

If $-1 < \alpha < 0$, we can apply the Gautschi inequality [12], obtaining:

$$
E(\gamma_t^\alpha\gamma_s^\alpha) \sim (st)^\alpha\,,
$$

and then:

$$
M_\alpha(n) = E(\gamma_1^\alpha\gamma_{n+1}^\alpha) - E(\gamma_1^\alpha\gamma_n^\alpha)) \sim n^{\alpha-1}\,.
$$

On the other hand, for $\alpha = 0$, $E(\gamma_t^\alpha\gamma_s^\alpha) = 1$, and therefore, $M_0(n) = 0$. $\square$

In other words, for $\alpha = 0$, the process has no memory, and we recover the Brownian motion case as already mentioned; for $\alpha < 0$, the process is called anti-persistent. Recalling the Hurst parameter H associated with the fractional Brownian motion, the relationship between H and α is $\alpha = 2H - 1$; see [4], for instance.

3. ABC-MCMC Study

When dealing with non-standard posterior distributions, the usual procedure is to use Markov chain Monte Carlo simulation to produce an approximate sample from this posterior. However, when the likelihood function is intractable, MCMC methods cannot be implemented. The class of likelihood-free methods termed, approximate Bayesian computation (ABC), can deal with this problem as long as we are able to simulate from the probabilistic model and a suitable set of summary statistics is available.

The ABC idea was proposed by Pritchard *et al.* [13] and developed further in the last decade. In particular, with respect to the choice of the summary statistics from among diverse options, in [14], the authors consider a sequential test procedure for deciding whether the inclusion of a new statistic improves the estimation. In [15], the discussion refers to the choice of informative statistics related to the algorithmic properties.

In our case, in order to perform the ABC algorithm for the memory parameter, α, we have to be able to generate a realization of the target process and we have to know which are the important observable characteristics of the process that lead us to increase our information about α.

Observe that we already know, by Equation (1), how to simulate from the generalized G-M for each value of the parameter. After obtaining a simulated trajectory, we can compare it then to the observed trajectory through adequate statistics. Intuitively, if they are similar enough, the chosen value of the parameter can be thought of as an appropriate one.

More concretely, suppose that we take the observed increments $d_t = \Delta(x_t)$ from a sample x_t and that for each $\alpha \in [-1, 0]$, we generate a realization from a Brownian motion b_t and a realization from a Gamma process γ_t, obtaining:

$$\Delta(^\alpha x_t) = b_t \gamma_t^\alpha - b_{t-1} \gamma_{t-1}^\alpha .$$

For the ease of notation, set $^\alpha y_t = \Delta(^\alpha x_t)$. To compute the proximity between d_t and $^\alpha y_t$, we will determine the distance between some statistics for each sample. The usual choices for the memory parameter are, among others, the rescaled range R/S or the rescaled variance V/S, the periodogram, quadratic variations, aggregated variances, Whittle estimator and functions, as the logarithm, inverse or square root, of these ones [15,16]. In [7], we used those statistics to obtain approximate posterior samples for the memory parameter for fractional self-similar processes.

We tried all of them in this work; however, their performance was not good enough, because of the non-stationarity and the inherent heteroscedasticity of the process. To solve this situation, we considered the slope of the regression of the sampling variance of m-size blocks on the time, regarding the results in Section 2.2, and this solution worked much better than the former ones.

Let T^* denote the following statistic. Giving n observations and an integer $m << n$, we take consecutive blocks of size m and obtain the sampling variance for each block, $s^2(k)$, for $k = 1, \ldots, \lfloor n/m \rfloor$, where $\lfloor \xi \rfloor$ is the integer part of ξ. Those values are used as estimates for the variances at times $m, 2m, \ldots, \lfloor n/m \rfloor m$. In $\log - \log$ scale, the slope of the regression line obtained through the points $(km, s^2(k))$, $k = 1, \ldots, \lfloor n/m \rfloor$, should be of order of 2α, as time increases. We define T^* as this slope.

The ABC steps for α are then given by the following algorithm.

Algorithm 1 ABC procedure for α.

---

1. Select α from a prior distribution $\pi(\alpha) \sim Unif[-1, 0]$;
2. Generate a trajectory $^{\alpha}y_t$ for that value α;
3. Given $\varepsilon > 0$, if $|T^*(x_t) - T^*(^{\alpha}y_t)| < \varepsilon$, then select α for the sample S_α, otherwise, reject it;
4. Repeat Steps 1–3 until reaching an adequate sample size for the parameter α.

The sample S_α obtained in the last step is an approximate sample from the posterior distribution for α, the goodness of which strongly depends on the choice of the statistics and the threshold ε in Step 3.

Note that if ε is too small, then the rejection rate is high, and the algorithm becomes too slow; on the other hand, if ε is too big, the algorithm accepts too many values of α, giving a less precise approximate sample. In general, what is done is to take a small percentile of the simulated distances, that is we select those α's giving the closest values to the observed one. In this work, after some trial, we used the first percentile, as proposed by [17].

Let us consider now the more general model for the increments given by:

$$Y_t = \mu + \frac{1}{\sqrt{\tau}} \Delta(^{\alpha}X_t) \,, \tag{2}$$

for $\mu, \tau \in \mathbb{R}$ and $\tau > 0$, representing the precision of the fluctuations. As: $\Delta(^{\alpha}X_t) = \left(B_t \gamma_t^\alpha - B_{t-1} \gamma_{t-1}^\alpha \right)$, with $B_t - B_{t-1} \sim N(0, 1)$, $\gamma_t - \gamma_{t-1} \sim \chi^2(1)$ and B_t independent of γ_t, for all $t > 0$, and following the theory associated with the Gamma-modulated process [4], it is possible to assume a hierarchical representation.

We can rewrite the above model in a multivariate way as:

$$(Y_1, \ldots, Y_n) \mid \mu, \tau, \gamma, \alpha \sim N\left(\mu\mathbf{1}, \tau^{-1}\Sigma(\gamma, \alpha) \right) \,,$$

where, for $i, j = 1, \ldots, n$,

$$\begin{aligned}
\Sigma_{ii} &= i \left(\gamma_i^\alpha - \gamma_{i+1}^\alpha \right)^2 + \gamma_{i+1}^{2\alpha} \,, \\
\Sigma_{ij} &= \left(\gamma_j^\alpha - \gamma_{j+1}^\alpha \right) \left(i\gamma_i^\alpha - (i+1)\gamma_{i+1}^\alpha \right) \,, \text{ for } i < j.
\end{aligned}$$

As a final step, let us assume prior distributions given by:

$$\begin{aligned}
\mu \mid \tau &\sim N(b_0, g(\tau)v_0) \\
\tau &\sim \Gamma(a_0, d_0),
\end{aligned}$$

and prior knowledge for α, $\pi(\alpha)$. Then, given $0 < \gamma_1 < \gamma_2 < \ldots < \gamma_{n+1}$, auxiliary random effects distributed as:

$$\pi(\gamma_1, \ldots, \gamma_n) \propto \prod_{i=1}^{n+1} (\gamma_i - \gamma_{i-1})^{-1/2} e^{-\gamma_{n+1}/2} \,,$$

and an observed trajectory $\mathbf{y} = (y_1, \ldots, y_n)$, the posterior distribution can be computed from:

$$\begin{aligned}
&\pi(\mu, \tau, \gamma_1, \ldots, \gamma_n, \alpha \mid \mathbf{y}) \\
&\propto \pi(\gamma_1, \ldots, \gamma_n)\, \pi(\tau)\, \pi(\alpha)\tau^{n/2} \left| \Sigma(\gamma, \alpha) \right|^{-1/2} \times \exp\left\{ -\frac{\tau}{2} (\mathbf{y} - \mu\mathbf{1})'\Sigma(\gamma, \alpha)^{-1}(\mathbf{y} - \mu\mathbf{1}) \right\}.
\end{aligned}$$

After straightforward computations and assuming that $g(\tau) = 1$, the full conditional posterior distributions for μ, τ are:

$$
\begin{aligned}
\mu \mid \mathbf{y}, \tau, \gamma_1, \ldots, \gamma_n, \alpha &\sim N\left(b_1, v_1\right), \\
\tau \mid \mathbf{y}, \mu, \gamma_1, \ldots, \gamma_n, \alpha &\sim \Gamma\left(a_1, d_1\right),
\end{aligned}
\tag{3}
$$

where:

$$
b_1 = \left(v_0 + \tau \mathbf{1}'\Sigma(\gamma, \alpha)^{-1}\mathbf{1}\right)^{-1}\left(\tau \mathbf{1}'\Sigma(\gamma, \alpha)^{-1}\mathbf{y} + v_0 b_0\right), \; v_1 = \left(v_0 + \tau \mathbf{1}'\Sigma(\gamma, \alpha)^{-1}\mathbf{1}\right),
$$

$$
a_1 = n/2 + a_0, \; d_1 = d_0 + 0.5(\mathbf{y} - \mu\mathbf{1})'\Sigma(\gamma, \alpha)^{-1}(\mathbf{y} - \mu\mathbf{1}).
$$

With this notation, for parameters (μ, τ, α), we propose the following ABC-MCMC procedure.

Algorithm 2 ABC-MCMC procedure for (μ, τ, α).

---

1. Apply the ABC proposal given by Algorithm 1, to obtain a posterior sample for α, S_α;
2. Sample uniformly from $\alpha \in S_\alpha$;
3. Generate a trajectory $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_n)$ from the Gamma process;
4. Sample (μ, τ) from the conditional posterior $\pi(\mu, \tau \mid \alpha, \boldsymbol{\gamma}, \mathbf{y})$;
5. Repeat Steps 3–4 until the convergence is reached, which can be checked by the usual graphical criterion for (μ, τ), and take their sampling mean;
6. Repeat Steps 2–5, for obtaining a complete posterior sample for (μ, τ, α).

The whole procedure is to then use the ABC algorithm for α and the MCMC scheme for μ, τ to perform the Bayesian inference of our proposal.

3.1. Posterior Evidence for Sharp Hypotheses

As a by-product of the previous algorithm, we are able to compute approximately the so-called e-value, an evidence measure defined in [8] for precise hypothesis testing, which we describe briefly below.

Let us consider a hypothesis of interest $H : \alpha \in \Omega_0$, and define the tangential set T_0 to Ω_0 as the set:

$$
T_0 = \{(\mu, \tau, \alpha) \in \Omega : \pi(\mu, \tau, \alpha \mid \mathbf{y}) > \pi_0\}, \qquad \text{where } \pi_0 = \sup_{\Omega_0} \pi(\mu, \tau, \alpha \mid \mathbf{y}).
$$

In other words, the tangential set to Ω_0 considers all points "more probable" than those in Ω_0, according to the posterior law.

The evidence measure e-value for the hypothesis H is defined as:

$$
ev(\Omega_0) = 1 - \int_{T_0} \pi(\mu, \tau, \alpha \mid \mathbf{y}) \, d\mu \, d\tau \, d\alpha.
\tag{4}
$$

Therefore, if the tangential set has high posterior probability, the evidence in favor of H is small; if it has low posterior probability, the evidence against H is small. For instance, suppose that we are considering the sharp hypothesis $H : \alpha = \alpha_0$. If the subset $\alpha = \alpha_0$ has high density, that is it lies

near the mode of $\pi(\mu, \tau, \alpha \mid \mathbf{y})$), then the e-value must be large, giving high evidence for that sharp hypothesis. Some interesting hypotheses are the precise ones defining the Brownian motion case, when $\alpha = 0$, and the G-M process, when $\alpha = -0.5$.

In this work, we approximate empirically the integral in the e-value using the posterior sample obtained by the previous algorithms. As usual, in the Bayesian paradigm, the predictive distribution for the next steps can be computed using this ABC-MCMC sample and the model Equation (2).

4. Numerical Results

In this section, we present the main results after implementing our algorithms in the R-gui software [18], using simulated and real data. Hence, we will show the performance of our proposal and its use in practice.

4.1. Simulated Results

To illustrate the performance of our procedure, we simulate 500 replicates of length $n = 1000$, for a grid with $\alpha \in [-1, 0]$ by 0.1, when $\mu = 0$ and $\tau = 1$ are fixed.

Figure 3 shows three trajectories for $\alpha = -0.8, -0.5, -0.2$ and their respective approximate posterior densities obtained by Algorithm 1, with posterior means (and 95% credible intervals) -0.835 $(-0.939, -0.735)$, -0.495 $(-0.579, -0.421)$, -0.217 $(-0.278, -0.157)$.
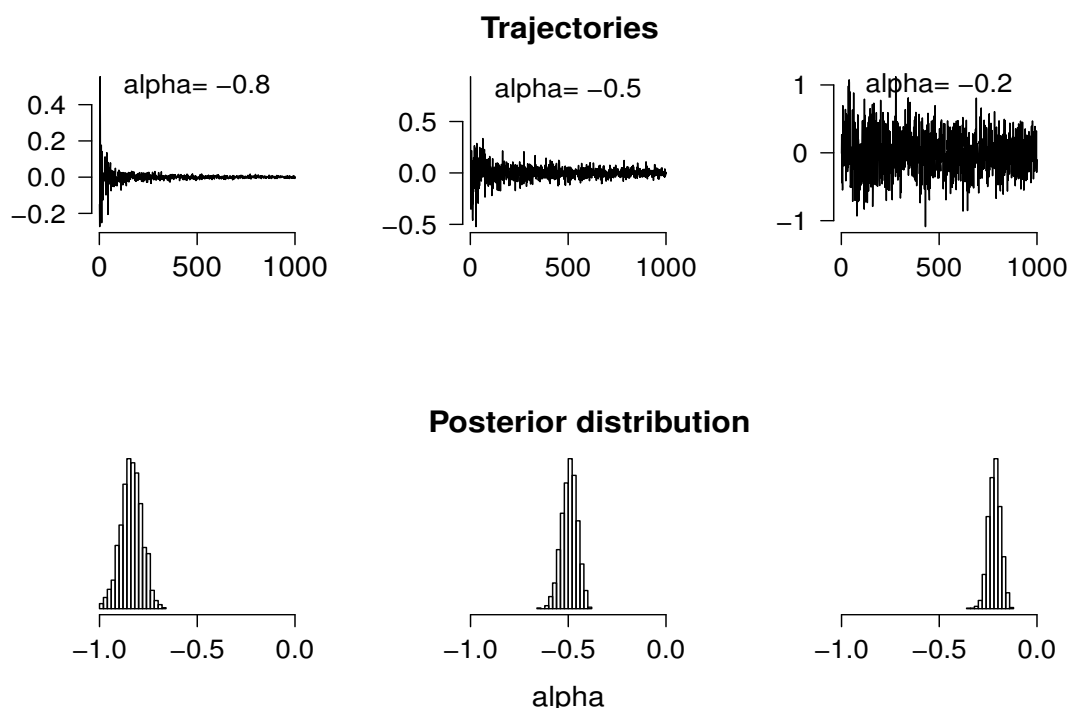
Figure 3. Realizations of the increments of the generalized Gamma-modulated process for different values of α and the posterior density obtained by the approximate Bayesian computation (ABC) algorithm for α.

In Figure 4, we see the sampling distribution of the posterior mean for α. Observe that the estimates we obtained are fairly precise for all values in the interval $[-1, 0]$.

In order to confirm that precision, we obtained the e-values associated with the hypothesis $\alpha_0 = -0.5$, when the nominal value for α varies in $[-1, 0]$, as shown in Figure 5a. The boxplots show the sampling distribution for these e-values from the 500 replicates for each $\alpha \in \{-1, -0.9, \ldots, -0.1, 0\}$. Observe that the obtained e-values are coherent with the previous estimates, giving high evidence to $\alpha_0 = -0.5$ when the nominal value is close to it and low evidence otherwise.
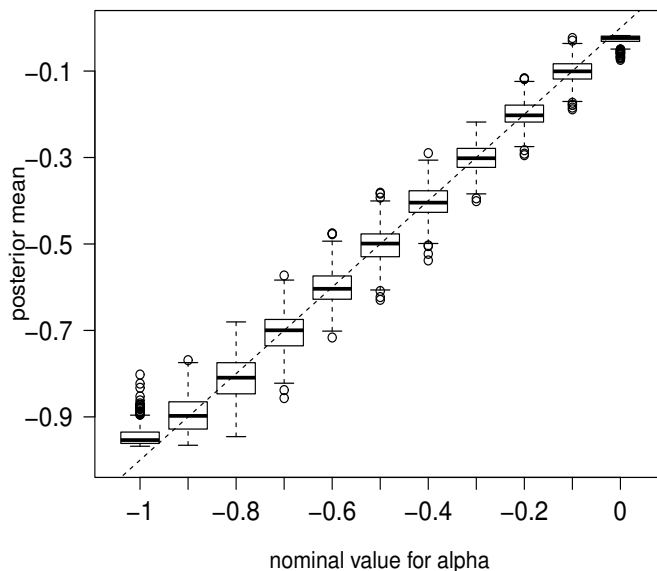
Figure 4. Sampling distribution of the posterior mean for α.
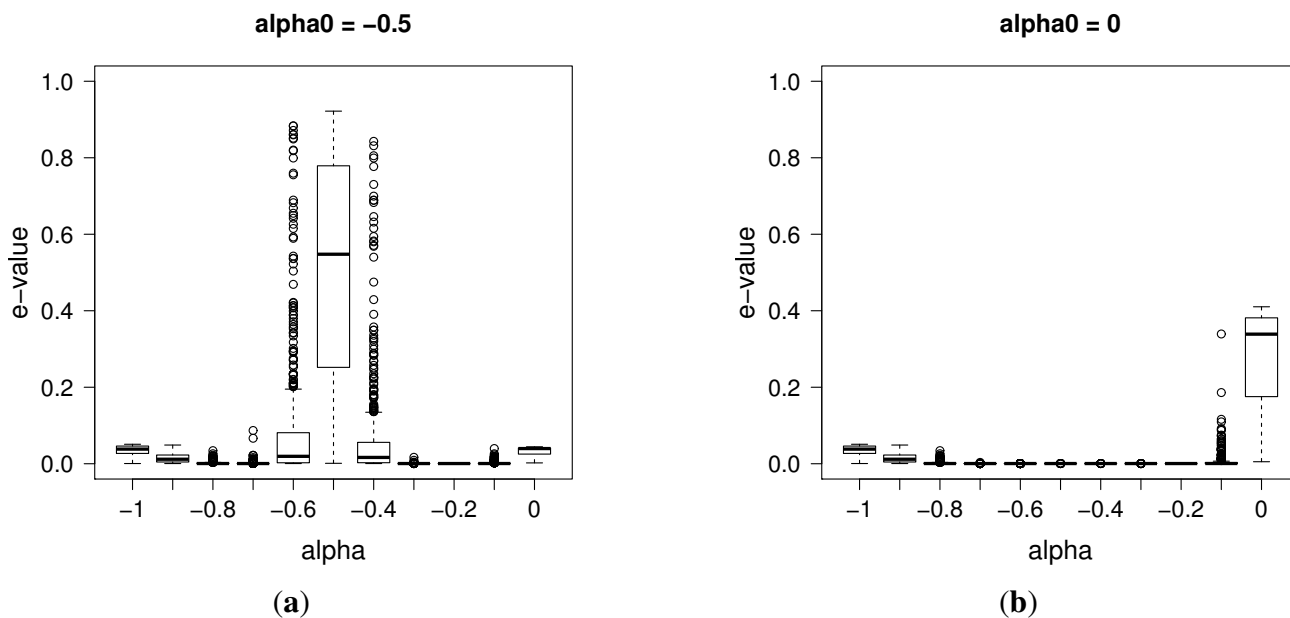

Figure 5. Sampling distribution of the e-value for: (**a**) $\alpha_0 = -0.5$; (**b**) $\alpha_0 = 0$.

In the case $\alpha_0 = 0$, shown in Figure 5b, even when the e-value is not as high for $\alpha = 0$ as in the previous case, it does allow for a good discrimination in favor of the null hypothesis.

In the general case, when μ and τ are unknown, we applied Algorithm 2, reporting the following results for the nominal values $\mu = 10$ and $\tau = 1$, as illustrated in Figure 6. The numerical results were very similar for other values in terms of precision and computational time.

It is worth noting the symmetry of the sampling distribution of the posterior mean for μ, as well as the asymmetry of the sampling distribution of the posterior mean for τ, as it was supposed to be by Equation (3). Furthermore, the estimates are consistent around the nominal value.
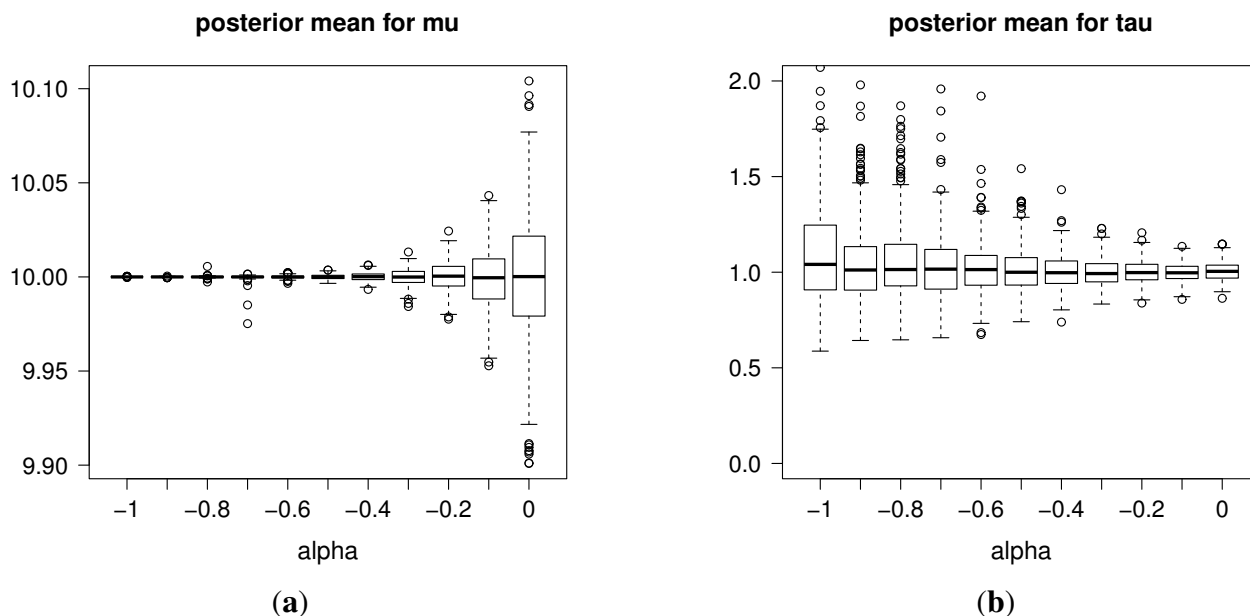

**Figure 6.** Sampling distribution of the posterior mean for: (**a**) μ; (**b**) τ.

The results suggest that, as $\alpha \to -1$, the data become more informative for the location parameter. Inversely, for the precision parameter τ, the estimates seem to be less accurate as α goes to -1. Both features are related to the behavior of the trajectories: as α decreases to -1, the increments stabilize faster around zero than for α closer to zero, and consequently, we obtain more precise estimates for μ. This very behavior explains the posterior for τ: the fast stabilization of the increments leads to underestimating the variance, that is to overestimating the precision. The same results were observed for other nominal values of μ and τ.

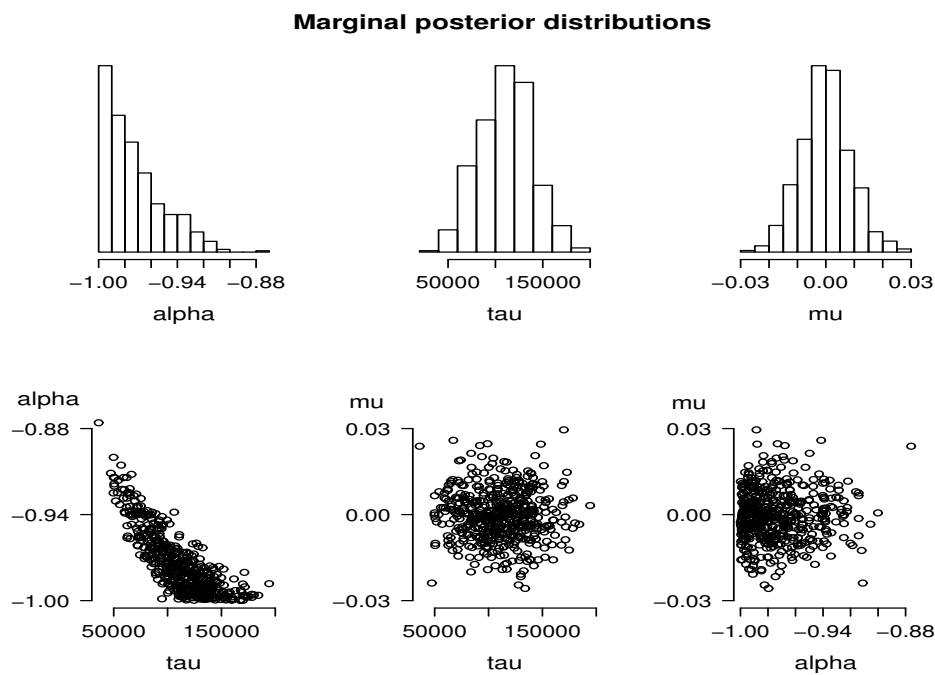### 4.2. Earthquake Acceleration Data

Our proposal is illustrated extending the ideas in [4]. We analyze sequential data obtained from an accelerometer recording the big earthquake in Southern Chile (27 February 2010), with the epicenter in Cobquecura, approximately 335 km southwest of Santiago, which reached 8.8 on the Richter scale. The dataset was obtained from the Hydrographic and Oceanographic Service of the Chilean Navy. The time series was recorded at a rate of 50 observations per second, with $n = 1653$, as shown in Figure 7a.

A brief exploratory analysis of the data, summarized in Figure 7b–d, led us to believe that the proposed model can be well fitted according to the properties described in Section 2.

The marginals for the posterior distribution are shown in Figure 8, with posterior means for the parameters equal to $\hat{\alpha} = -0.975(\pm 0.021)$, $\hat{\mu} = -0.0007(\pm 0.007)$, $\hat{\tau} = 106305(\pm 3800)$. Observe, in the first scatterplot, the sharp and negative posterior correlation between α and τ, as expected by Equation (3). Furthermore, the e-values for the hypotheses $\alpha_0 = -0.5$ and $\alpha_0 = 0$ are clearly zero, as confirmed by both graphs.
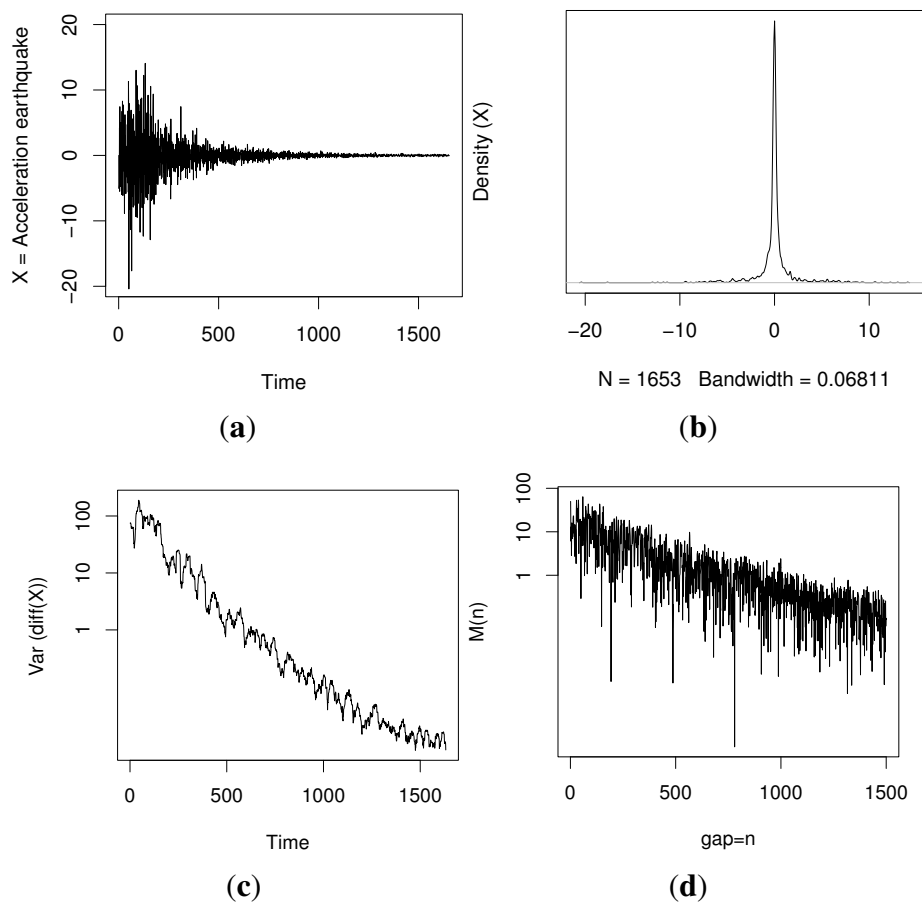
Figure 7. (**a**) Acceleration earthquake series, X_t ($n = 1653$ at 50 observations per second) and exploratory graphics: (**b**) density of the raw data; (**c**) sample variance of the increment process, $Var(\Delta(X_t))$, in log-scale; (**d**) sample auto-covariance, $M(n)$, in log-scale.

Figure 8. Marginal posterior distributions for the parameters in the acceleration earthquake series.

5. Final Remarks

The model proposed here seems to be suitable for the phenomena under study. In particular, interesting properties, including long memory and trajectory behavior, are discussed, on which our inference methodology is grounded.

For the ABC algorithm, we considered at first a minimum entropy criterion for selecting the approximate posterior sample, because that choice had a good performance in estimating the long memory parameter for stationary non-Gaussian processes as, for instance, binary and the Rosenblatt processes [7]. However, given the trajectory behavior of the G-M process, the most informative statistics we found is the m-block variance statistic T^*.

Our results show, firstly, a clear and easy way of implementing the ABC-MCMC algorithms in a standard software. For the real data ($n = 1653$), the computational cost was moderate, obtaining a posterior sample of a size of 500 from the ABC-MCMC proposal in one hour.

In our simulations, we obtained a very high precision in the estimates given by this procedure. In addition, the estimates for the precision parameter, τ, are affected reasonably by scale changes. For instance, for the rescaled data, $\mathbf{z} = c \times \mathbf{y}$, we estimated τ as $\tau^* = \hat{\tau}/c^2$, approximately.

The chosen parameterization allows for $\alpha \in (0, 1]$. However, we did not study such a case, since that process diverges as time increases, making the inference procedure harder for α. We recommend then to treat this problem as a separate case.

Finally, we believe that this model has wider applications, mainly for its parsimony and the straightforward interpretation of the parameters. For diagnostic purposes, for example, the predictive series could be used to perceive the increasing fore-shocks before the arrival of a new earthquake.

Acknowledgments

Author Contributions

All of the authors developed the model and the methodology presented in this paper. Plinio Andrade performed the simulations and data analysis and wrote the paper. All of the authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

Appendix: Discussion of Other Models

Let us consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbf{P}; \{\mathcal{F}_t\})$ that supports a Brownian motion, $\{B_t\}$, a positive random variable, W, and a Gamma process $\{\gamma_t\}$ with parameters $\alpha = 1/2$, $\beta = 2$. Suppose that all of the processes are adapted to the filtration $\{\mathcal{F}_t\}$ and are mutually independent. Furthermore, assume that $\{\mathcal{F}_t\}$ is the right continuous filtration associated with (B, W, γ).

Consider the following three processes defined by:

1. $\xi_t = B_t$;
2. $T_t = B_t/\sqrt{W}$, where W is a positive random variable and;
3. $\mathcal{T}_t = B_{t+1}\, \gamma_{t+1}^{-1/2}$.

The increment of the process $\{\xi_t\}$ defines a random i.i.d. noise with a Gaussian distribution, representing the most common noise used in the literature for regression models. The increments of the second noise describe a perturbation of a Gaussian i.i.d. error in the distributional sense. This process will be called t-homoscedastic, given that the increments remain independent and with the same distribution, whose variance is related to the degrees of freedom of the independent variable W. Specifically, we are interested in the case where W is a χ^2 random variable with ν degrees of freedom.

Finally, the third noise is the Gamma-modulated process, as defined in [4].

A. Some Properties of the Models

A.1. Heteroscedastic t-Student Process

Proposition 4. *The one-dimensional densities of $\mathcal{T}_t$, $\Delta\mathcal{T}_t = \mathcal{T}_t - \mathcal{T}_s$ and $(\Delta\mathcal{T}_t, \mathcal{T}_s)$ for fixed $1 < s < t$ are given by:*

$$
f_{\mathcal{T}_t}(x) = \int_0^\infty \phi\left(x; 0, \frac{t}{\gamma}\right) d\Gamma_t(\gamma),
$$

$$
f_{\Delta\mathcal{T}_t}(x) = \int \phi\left(x; 0, g(\gamma_1, \gamma_2)\right) d\Gamma_{t-s}(\gamma_1) d\Gamma_s(\gamma_2),
$$

$$
f_{(\Delta\mathcal{T}_t, \mathcal{T}_t)}(\bar{x}) = \int \phi\left(\bar{x}; 0, \Sigma\right) d\Gamma_{t-s}(\gamma_1) d\Gamma_s(\gamma_2),
$$

where Γ_t is the cdf of γ_t, $\phi(x; \mu, \sigma^2)$ denote the univariate density at x of a normal distribution with mean μ and variance σ^2,

$$
g(\gamma_1, \gamma_2) = s\left(\frac{1}{\sqrt{\gamma_2}} - \frac{1}{\sqrt{\gamma_1 + \gamma_2}}\right)^2 + \frac{t - s}{\gamma_1 + \gamma_2},
$$

and

$$
\Sigma = \begin{pmatrix} \frac{t-s}{\sqrt{\gamma_1}} + s\left(\frac{1}{\sqrt{\gamma_1} - \sqrt{\gamma_2}}\right)^2 & \frac{s}{\sqrt{\gamma_2}}\left(\frac{1}{\sqrt{\gamma_1} - \sqrt{\gamma_2}}\right) \\ \frac{s}{\sqrt{\gamma_2}}\left(\frac{1}{\sqrt{\gamma_1} - \sqrt{\gamma_2}}\right) & \frac{s}{\sqrt{\gamma_2}} \end{pmatrix}.
$$

Proof. The densities are consequences of standard computation. □

A.2. Homoscedastic t-Student Process

Proposition 5. *The homoscedastic t-Student process is conditionally independent and has stationary increments. Furthermore,*

$$
\begin{aligned}
E(T_t) &= 0\,, \\
E(T_t^2) &= tE(W^{-1}), \quad \text{if } E(W^{-1}) < \infty\,, \\
E(T_t^k) &= t^{k/2}E(Z^k)E(W^{-k/2}), \quad \text{if } E(W^{-1/2}) < \infty\,,
\end{aligned}
$$

where Z is a standard Gaussian random variable with $E(Z^k) = 0$, if k is odd, and $E(Z^k) = 2^{k/2}\Gamma\left(\frac{k+1}{2}\right)/\sqrt{\pi}$, if k is even.

Additionally, for $s < t$, the covariance is given by:

$$
E(T_t T_s) = sE(W^{-1}).
$$

Finally, for given t, the density of T_t is:

$$
f_{T_t}(x) = \int_0^\infty N(x; 0, t/W)dF_W(w).
$$

Proof. The proof is a consequence of algebraic calculations. □

B. Diffusion Models

In this section, we will consider the Euler discretization of the linear stochastic differential equation:

$$
dY_t = \mu dt + \sigma dE_t\,, \tag{5}
$$

where the noise E_t has one of the following characteristics:

1. standard Brownian motion or Gaussian noise;
2. homoscedastic t-Student noise or;
3. Gamma-modulated process with heteroscedastic t-Student noise.

Then, the discrete version of Equation (5) is given by $y_i = \mu\Delta(t) + \sigma\Delta E_t$. In each case, the discretization of the errors can be represented as a combination of the scale and location parameters in a Gaussian model.

B.1. Linear SDE with Gaussian Independent Errors

In this section, we will consider the following discretized model of Equation (5), when the noise is given by the increment of a Brownian motion. In this case, y_i, $1 \leq i \leq n$, is a discrete version of Equation (5), with ($\Delta = 1$) given by:

$$
y_i = \mu + \sigma(B_{i+1} - B_i),
$$

for $i = 1, \ldots, n$. We will assume that (μ, σ^2) is independent of $\{B_t\}_{t \geq 0}$. Under this framework, the conditional setting is given by:

$$y \mid \mu, \sigma^2, \Sigma \sim N_n(\mu X, \sigma^2 \Sigma),$$

where $\Sigma = I_n$, $X = 1_n$ is the n-dimensional vector of ones, and $\mu = \beta$.

Considering the following prior specification:

$$\begin{aligned} \mu \mid \sigma^2 &\sim N(b_0, \tau_0^2) \\ \sigma^2 &\sim IG(a_0, d_0), \end{aligned}$$

where $B_0 = \tau_0^2$ and b_0, a_0, d_0 and τ_0^2 are known, then it is straightforward to obtain the complete conditional distributions. To approximate the posterior distributions, we need to sample from:

$$\begin{aligned} \mu \mid \sigma^2, y &\sim N(m, \tau^2) \\ \sigma^2 \mid \mu, y &\sim IG(a, d), \end{aligned}$$

where $m = (\sigma^2 b_0 + \tau_0^2 \bar{y}_n)/(\sigma^2 + n\tau_0^2)$, $\tau^2 = \sigma^2 \tau_0^2/(\sigma^2 + n\tau_0^2)$, $a = \frac{n}{2} + a_0$ and $d = \frac{1}{2}\sum_{i=1}^{n}(y_i - \mu)^2 + d_0$.

B.2. Linear SDE with t-Student Homoscedastic Errors

In this model, we will consider fixed *t*-Student distributed errors, which means that we will contaminate the noise with a fixed χ^2 random variable with ν degrees of freedom. In this model, the increments have a *t*-Student distribution with the same degrees of freedom, that is the increments are stationary. Then, the discretized version of the *t*-Student homoscedastic model for the SDE Equation (5), with $\Delta = 1$, is given by:

$$y_i = \mu + W^{1/2}\sigma(B_{i+1} - B_i),$$

for $i = 1, \ldots, n$, $W \sim IG(\nu/2, \nu/2)$, that is if $G = W^{-1}$, then $f_G(g) = \frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)} g^{\nu/2-1} e^{-g/2}$ for $w \geq 0$; and W is independent of $\{B_t\}_{t \geq 0}$. We will also assume that the vector (μ, σ^2) is independent of $W, \{B_t\}_{t \geq 0}$. As before, we adopt the conditional setting:

$$\begin{aligned} y \mid \mu, \sigma^2, \omega &\sim N_n(\mu, \omega\sigma^2 I_n), \\ \omega &\sim IG(\nu/2, \nu/2), \end{aligned}$$

and

$$(\mu, \sigma^2) \perp \omega.$$

It is straightforward to prove that if we integrate out W, we obtain:

$$y \mid \mu, \sigma^2 \sim t_n(\mu 1_n, \sigma^2 I_n, \nu), \tag{6}$$

where $t_n(\mu, \Sigma, \nu)$ denote the n-dimensional t distribution.

Thus, if we consider:

$$\begin{aligned} \mu \mid \sigma^2 &\sim N(b_0, g(\sigma^2)B_0), \\ \sigma^2 &\sim IG(a_0, d_0), \end{aligned}$$

and put $\Sigma = \omega I_n$, with $\omega \sim IG(\nu/2, \nu/2)$, $g(\sigma^2) = 1$, $B_0 = \tau_0^2$ and $X = 1_n$, we obtain:

$$
\begin{aligned}
\mu \mid \sigma^2, \omega, y &\sim N(m, \tau^2) \\
\sigma^2 \mid \mu, \omega, y &\sim IG(a, d) \\
\omega \mid \mu, \sigma^2, y &\sim IG(a_1, d_1),
\end{aligned}
$$

where $m = (\sigma^2 b_0 w + \bar{y}\tau_0^2)/(w\sigma^2 + \tau_0^2 n)$, $\tau^2 = \sigma^2 w\tau_0^2/(w\sigma^2 + \tau^2 n)$, $a = \frac{n}{2} + a_0$, $d = \sum_{i=1}^{n}(y_i - \mu)^2/(2w) + d_0$, $a_1 = (\nu + n)/2$ and $d_1 = \sum_{i=1}^{n}(y_i - \mu)^2/(\sigma^2) + \nu/2$.

B.3. Gamma Modulated Process-t-Student Heteroscedastic Model

In this case, the discrete version of the process has the following representation:

$$
y_i = \mu + \sigma \left(\frac{B_{i+1}}{\sqrt{\gamma_{i+1}}} - \frac{B_i}{\sqrt{\gamma_i}} \right), \tag{7}
$$

for $i = 1, \ldots, n$. Again, we assume (μ, σ^2) independent of $\{B_t\}_{t\geq 0}$, $\{\gamma_t\}_{t\geq 0}$ and the prior distribution of (μ, σ^2) specified as before. Hence, the model is reduced to:

$$
y \mid \mu, \sigma^2, \gamma \sim N(\mu 1_n, \sigma^2 \Sigma(\gamma)),
$$

where γ is a gamma process with parameters $\alpha = \beta = 1/2$, $\Sigma(\gamma) = A(\gamma)\Lambda A(\gamma)^t$, Λ is the covariance matrix of $B = (B_1, \ldots, B_{n+1})$ given by $\lambda_{i,j} = i \wedge j$ and $A(\gamma)$ is the $n \times (n+1)$ matrix given by:

$$
a_{ij} = \begin{cases}
-\gamma_i^{-1/2} & \text{if } i = j, \\
\gamma_{i+1}^{-1/2} & \text{if } j = i+1, \\
0 & \text{otherwise.}
\end{cases}
$$

Simple computation gives:

$$
\begin{aligned}
\Sigma_{ii} &= i \left(\frac{1}{\sqrt{\gamma_i}} - \frac{1}{\sqrt{\gamma_{i+1}}} \right)^2 + \frac{1}{\sqrt{\gamma_{i+1}}}, \\
\Sigma_{ij} &= \left(\frac{1}{\sqrt{\gamma_j}} - \frac{1}{\sqrt{\gamma_{j+1}}} \right) \left(\frac{i}{\sqrt{\gamma_i}} - \frac{i+1}{\sqrt{\gamma_{i+1}}} \right) \quad \text{for} \quad i < j.
\end{aligned}
$$

On the other hand, the distribution of γ is computed from the fact that $(\gamma_1, \gamma_2 - \gamma_1, \ldots, \gamma_{n+1} - \gamma_n)$ are independent χ_1^2. Let $T = t_{ij}$ the matrix defined by:

$$
t_{ij} = \begin{cases}
1, & \text{if } i \leq j \\
0, & \text{otherwise.}
\end{cases}
$$

Then:

$$
\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{pmatrix} = T \begin{pmatrix} \gamma_1 \\ \gamma_2 - \gamma_1 \\ \vdots \\ \gamma_n - \gamma_{n-1} \end{pmatrix}.
$$

By the independence of the increments, we obtain that the density of γ is proportional to:

$$\prod_{i=1}^{n+1}(\gamma_i - \gamma_{i-1})^{-1/2}e^{\gamma_{n+1}/2}, \quad 0 < \gamma_1 < \gamma_2 < \ldots < \gamma_{n+1},$$

where we are assuming $\gamma_0 = 0$. Hence, Model (7) can be specified as follows:

$$\begin{aligned}
y \mid \mu, \sigma^2, \gamma &\sim N(\mu 1_n, \sigma^2 \Sigma(\gamma)), \\
\mu \mid \sigma^2, \gamma &\sim N(b_0, g(\sigma^2)B_0), \\
\sigma^2 \mid &\sim IG(a_0, d_0), \\
\mu \perp \sigma^2, &\qquad (\gamma_1, \ldots, \gamma_{n+1}) \perp (\mu, \sigma^2).
\end{aligned}$$

The distribution of Σ can be obtained running the Gibbs sampling on y, μ, σ^2 and γ. Usual calculations give the posterior distribution of (μ, σ^2).

In this way, the conditional distributions needed to implement the Gibbs sampling are obtained as usual by taking $X = 1_n$, $\beta = \mu$, $g(\sigma^2) = 1$, $B_0 = \nu_0^2$ and $\Sigma = \Sigma(\gamma)$, such that:

$$\begin{aligned}
\mu \mid \sigma^2, \gamma, y &\sim N(m, \tau^2), \\
\sigma^2 \mid \mu, \gamma, y &\sim IG(a, d),
\end{aligned}$$

where:

$$m = \left(\left(\frac{\sigma}{\nu_0}\right)^2 + \sum_{i,j}\Gamma_{i,j}\right)^{-1} \times \left(\left(\frac{\sigma}{\nu_0}\right)^2 b_0 + \sum_i \sum_j \Gamma_{i,j}y_i\right),$$

$\Gamma = \Sigma^{-1}$ is a $n \times n$ matrix,

$$\tau^2 = \sigma^2 \left(\left(\frac{\sigma}{\nu_0}\right)^2 + \sum_{i,j}\Gamma_{i,j}\right)^{-1},$$

$a = (n/2 + a_0 + 1)$, and $d = \frac{1}{2}(y - \mu)^t \Sigma^{-1}(y - \mu) + d_0$.

Finally, the posterior distribution for the random vector $\gamma = (\gamma_1, \ldots, \gamma_{n+1})$ given μ, σ^2, $y = (y_1, \ldots, y_n)$, such that $0 < \gamma_1 < \gamma_2 < \ldots < \gamma_{n+1}$, is:

$$\gamma \mid \mu, \sigma^2, y \sim \left(\prod_{i=1}^{n+1}(\gamma_i - \gamma_{i-1})^{-1/2}e^{\gamma_{n+1}/2}\right)\sigma^{-n}|\Sigma^{-n/2}|exp((\gamma - \mu)^t\sigma^{-1}\Sigma^{-1}(\gamma)(\gamma - \mu)).$$

C. Comparing the Models

In this section, we present two models in order to make a comparison between our proposal and other noises.

C.1. Mixed Brownian Model

Consider the processes defined by $T_t = B_t/\sqrt{W}$, where W is a positive random variable and B is a standard Brownian motion. It has stationary increments, and:

$$\begin{aligned}
E(T_t) &= 0, \\
E(T_t^2) &= tE(W^{-1}), \quad \text{if} \quad E(W^{-1}) < \infty, \\
E(T_t^k) &= t^{k/2}E(Z^k)E(W^{-k/2}), \quad \text{if} \quad E(W^{-1/2}) < \infty,
\end{aligned}$$

where Z is a standard Gaussian random variable with $E(Z^k) = 0$, if k is odd, and $E(Z^k) = (2^{k/2}\Gamma\left(\frac{k+1}{2}\right))/\sqrt{\pi}$, if k is even.

Additionally, for $s < t$, the covariance is given by:

$$E(T_t T_s) = sE(W^{-1}).$$

Finally, given $t > 0$, the density of T_t is:

$$f_{T_t}(x) = \int_0^\infty N(x; 0, t/W) dF_W(w).$$

If $W = 1$ almost sure, we can recover the case of the Brownian motion. Let us consider the following increments process:

$$T_t^\alpha = (T_t - T_{t-1})t^\alpha.$$

This process has independent increments, but is not stationary. The first two moments of the process are $E(T_t^\alpha) = 0$, $V(T_t^\alpha) = t^{2\alpha}$.

The covariance structure is given by:

$$COV(T_t^\alpha, T_s^\alpha) \quad = \quad E(T_t^\alpha T_s^\alpha) = (st)^\alpha E((T_t^\alpha - T_{t-1}^\alpha)(T_s^\alpha - T_{s-1}^\alpha)) = 0.$$

and finally, the memory of the process is given by:

$$E(T_1^\alpha(T_{n+1}^\alpha - T_n^\alpha)) \quad = \quad 0.$$

C.2. Fractional Brownian Motion

Fractional Brownian motion with Hurst parameter $H \in (0, 1)$ is a centered Gaussian process $(B_t^H)_{t\in[0,1]}$, whose covariance function can be written as $E(B_t^H B^H s) = \frac{1}{2}\left(s^{2H} + t^{2H} - |t - s|^{2H}\right)$. The family of processes $\{B^H; H \in (0, 1)\}$ enjoys several nice properties:

- for $H = 1/2$, one recovers the classical Brownian motion;
- for any $H \in (0, 1)$, the paths of B^H are almost sure $(H - \rho)$-Hölder continuous for any arbitrarily small $\rho > 0$. Specifically, we have:

$$|B_t^H - B_s^H| < F_0|t - s|^{H-\rho}$$

  almost sure, $t, s \in [0, T]$, where $F_0 = F_0(\omega)$ is a positive random variable, such that $E(F_0^p) < \infty$, for all $p \ge 1$;
- the covariance of the increments of B^H on intervals decays asymptotically as a negative power of the distance between the intervals;
- the fractional Brownian motion is the only finite-variance process, which is self-similar (with index H) and has stationary increments.

These characteristics have converted the fractional Brownian family into the most natural generalization of the Brownian motion among the probability community, but also for practitioners, in recent years.

In this section, we will consider the following process related to the fractional Brownian motion with Hurst parameter H,

$$B_t^{H,\alpha} = (B_t^H - B_{t-1}^H)t^\alpha.$$

This process has two components. The first one is related to the long memory process and the second one to the variance. The first two moments of the process are $E(B_t^{H,\alpha}) = 0$ and $V(B_t^{H,\alpha}) = t^{2\alpha}$.

The covariance structure is given by:

$$
\begin{aligned}
COV(B_t^{H,\alpha}, B_s^{H,\alpha}) &= E(B_t^{H,\alpha} B_s^{H,\alpha}) = (st)^\alpha E((B_t^H - B_{t-1}^H)(B_s^H - B_{s-1}^H)) \\
&= \frac{(st)^\alpha}{2}((s-t+1)^{2H} + (s-t-1)^{2H} - 2(s-t)^{2H})
\end{aligned}
$$

Finally, the memory of the process is:

$$
\begin{aligned}
E(B_1^{H\alpha}(B_{n+1}^{H,\alpha} - B_n^{H,\alpha})) &= (st)^\alpha E((B_t^H - B_{t-1}^H)(B_s^H - B_{s-1}^H)) \\
&= \frac{(n+1)^\alpha}{2}((n-1)^{2H} + (n+1)^{2H} - 2n^{2H}).
\end{aligned}
$$

By Taylor expansion, the last term is of order $n^{\alpha+2H-2}$. If $\alpha + 2H - 2 < 1/2$ the process has long memory.

References

1. Black, F.; Scholes, M. The pricing of options and corporate liabilitites. *J. Polit. Econ.* **1973**, *81*, 635–654.
2. Hurst, H. The problem of long-term storage in reservoirs. *Hydrol. Sci. J.* **1956**, *1*, 13–27.
3. Finlay, R.; Seneta, E. Stationary-increment variance-Gamma and t models: Simulation and parameter estimation. *Int. Stat. Rev.* **2003**, *76*, 167–186.
4. Iglesias, P.; San Martin, J.; Torres, S.; Viens, F. Option pricing under a Gamma-modulated diffusion process. *Ann. Financ.* **2011**, *7*, 199–219.
5. Heyde, C.; Leonenko, N. Student processes. *Adv. Appl. Probab.* **2005**, *37*, 342–365.
6. Marin, J.M.; Pudlo, P.; Robert, C.; Ryder, R. Approximate Bayesian computational methods. *Stat. Comput.* **2012**, *22*, 1167–1180.
7. Andrade, P.; Rifo, L. Long-range dependence and approximate Bayesian computation. *Commun. Stat. Simul Comput.* **2015**, doi:10.1080/03610918.2014.995816.
8. Pereira, C.; Stern, J.; Wechsler, S. Can a significance test be genuinely bayesian? *Bayesian Anal.* **2008**, *3*, 15–36.
9. Polpo, A.; Coque, M.; Pereira, C. Statistical analysis for weibull distributions in presence of right and left censoring. In Proceedings of the 8th International Conference on Reliability, Maintainability and Safety, ICRMS 2009, Chengdu, China, 20–24 July 2009; pp. 219–223.
10. Rifo, L.; Torres, S. Full Bayesian analysis for a class of jump-diffusion models. *Commun. Stat. Theory Methods* **2009**, *38*, 1262–1271.
11. Rodrigues, J. Full Bayesian Significance Test for Zero-Inflated Distributions. *Commun. Stat. Theory Methods* **2006**, *35*, 299–307.
12. Kershaw, D. Some Extensions of W. Gautschi's Inequalities for the Gamma Function. *Math. Comput.* **1983**, *41*, 607–611.

13. Pritchard, J.K.; Seielstad, M.T.; Perez-Lezaun, A.; Feldman, M.W. Population growth of human Y chromosomes: A study of Y chromosome microsatellites. *Mol. Biol. Evol.* **1999**, *16*, 1791–1798.

14. Joyce, P.; Marjoram, P. Approximately Sufficient Statistics and Bayesian Computation. *Stat. Appl. Genet. Mol. Biol.* **2008**, *7*, doi:10.2202/1544-6115.1389.

15. Blum, M. Choosing the Summary Statistics and the Acceptance Rate in Approximate Bayesian Computation. In *Proceedings of COMPSTAT'2010*; Lechevallier, Y., Saporta, G., Eds.; Physica-Verlag HD: Heidelberg, Germany, 2010; pp. 47–56.

16. Beran, J.; Feng, Y.; Ghosh, S.; Kulik, R. *Long-Memory Process—Probabilistic Properties and Statistical Methods*; Springer: Berlin/Heidelberg, Germany, 2013.

17. Grelaud, A.; Robert, C.P.; Marin, J.; Rodolphe, F.; Taly, J. ABC likelihood-free methods for model choice in Gibbs random fields. *Bayesian Anal.* **2009**, *4*, 317–335.

18. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2015.