Journal of Statistical Software

January 2019, Volume 88, Code Snippet 1. *doi: 10.18637/jss.v088.c01*

ETAS: An R Package for Fitting the Space-Time ETAS Model to Earthquake Data

Abdollah Jalilian
Razi University

Abstract

The epidemic-type aftershock sequence (ETAS) model is the most widely used statistical model to describe earthquake catalogs. **ETAS** is an R package for fitting the space-time ETAS model to an earthquake catalog using the stochastic declustering approach introduced by Zhuang, Ogata, and Vere-Jones (2002). The package provides two classes and several functions to facilitate data preparation, model fitting and some simple diagnostic checks. The present paper is a description of the package and illustrates modeling earthquake data using the space-time ETAS model.

Keywords: earthquake catalog, space-time ETAS model, stochastic declustering, R.

1. Introduction

Since the establishment of seismograph networks for instrumental monitoring of earthquakes in the past decades, earthquake catalogs in computer-readable formats are regularly accessible from national and international seismological agencies or institutions (Di Giacomo, Harris, Villaseñor, Storchak, Engdahl, and Lee 2015). An earthquake catalog is a chronologically ordered list of historic and/or instrumental earthquake records in a particular geographical region S and during a specific time period. It often consists of time, coordinates of epicenter, magnitude and focal depth of all recorded earthquakes with magnitudes greater than or equal to a certain threshold m_0 that occurred inside or in the vicinity of S during the specified time period. For example, Table 1 shows parts of an earthquake catalog covering the study period from 1973-01-01 to 2016-01-01 and the rectangular geographical region 22°–42°N and 40°–65°E (Iran and surrounding areas) with the magnitude threshold $m_0 = 4.0$. This catalog is described in more detail in Section 6 and is used for illustration purposes in this paper.

From a statistical point of view, an earthquake catalog is the available data on earthquakes in a space-time window (Ogata 1999). Statistical analysis can then be used to find a suitable

Event	Date	Time	Longitude	Latitude	Magnitude
1	1973-01-06	15:39:31.00	46.43	38.00	4.20
2	1973-01-06	20:01:50.90	48.26	33.10	4.80
3	1973-01-10	17:02:56.00	51.28	31.19	4.40
4	1973-01-13	14:14:41.10	63.90	25.58	5.00
5	1973-01-18	22:35:18.80	48.01	32.75	4.40
6	1973-02-07	05:27:20.00	49.43	32.05	5.20
7	1973-02-10	17:07:47.50	56.12	31.89	4.50
8	1973-02-24	00:02:40.10	52.62	28.58	5.20
10	1973-03-03	02:46:25.30	51.15	29.52	4.60
11	1973-03-12	13:21:49.10	49.34	32.05	4.90
⋮	⋮	⋮	⋮	⋮	⋮
5967	2015-11-25	21:17:18.59	49.57	31.97	5.40
5968	2015-12-02	23:27:09.25	40.26	39.22	5.50
5969	2015-12-04	19:23:17.92	51.95	28.98	5.00
5970	2015-12-24	22:39:20.17	52.09	40.30	4.60

Table 1: A catalog of earthquakes for Iran and surrounding areas from the beginning of 1973 to the end of 2015.

model for the underlying earthquake process. Such a model provides a theoretical description of the seismic activities in the study region and makes it possible to estimate the probabilities of future events (Kagan 1991; Vere-Jones 2003). Therefore, fitting an appropriate statistical model to a given earthquake catalog is of great importance in the probabilistic assessment of seismic hazard.

Among different proposed models, the epidemic-type aftershock sequence (ETAS) model, introduced by Ogata (1988) and Ogata (1998), is the most widely used statistical model to describe earthquake catalogs (see e.g., Helmstetter and Sornette 2003; Ogata 2006a; Console, Jackson, and Kagan 2010; Lombardi and Marzocchi 2010; Zhuang 2011; Harte 2013; Zhuang 2012). The space-time version of the ETAS model is a semi-parametric model that describes the background and triggering seismic activities in a geographical region and can be used for earthquake declustering (Zhuang *et al.* 2002). However, estimation of the ETAS model parameters is computationally challenging (Veen and Schoenberg 2008; Schoenberg 2013; Adelfio and Chiodi 2015a).

In this article we introduce the R (R Core Team 2018b) package **ETAS** (Jalilian and Zhuang 2019) which is an R implementation (through C and C++ ports) of the original Fortran implementation `etas8p` (Zhuang, Ogata *et al.* 2017). The package is available from the Comprehensive R Archive Network (CRAN) at `https://CRAN.R-project.org/package=ETAS` and GitHub at `https://github.com/jalilian/ETAS` and can be distributed under the GPL-2 license (GNU General Public License).

The material in the remainder of the paper is organized as follows. The ETAS model is described in Section 2, and Section 3 introduces the simultaneous estimation of model parameters and the background seismicity rate using a stochastic declustering approach. Residual analysis and model assessment are briefly discussed in Section 4. In Section 5, the main functions in the **ETAS** package are discussed and Section 6 provides a worked example of fitting the ETAS model to earthquake data based on a data set included in the package. The **ETAS**

package is compared with other available software packages in Section 7. Further possible extensions of the package are discussed in Section 8.

2. The ETAS model

2.1. Temporal marked point processes

Given an earthquake catalog consisting of N events, let t_i, x_i, y_i and m_i, $i = 1, \ldots, N$, denote, respectively, the time, longitude of epicenter, latitude of epicenter and magnitude of the ith event in the catalog. Then the catalog data (as for example exemplified in Table 2) can be regarded as a point pattern $\{(t_i, x_i, y_i, m_i) : i = 1, \ldots, N\}$ on $\mathbb{R}_+ \times \mathbb{R}^2 \times [m_0, \infty)$ which is generated by a temporal marked point process $\mathbf{X}$ that governs the occurrence of earthquakes in time and space (Vere-Jones 1970; Ogata 1998).

The distribution of the temporal marked point process $\mathbf{X}$ is uniquely determined by its conditional intensity function $\lambda(t, x, y, m|\mathcal{H}_t)$ (see Daley and Vere-Jones 2003, Section 7.3). Roughly speaking, for each $t > 0$, $(x, y) \in S$, $m \geq m_0$ and infinitesimal $\mathrm{d}t$, $\mathrm{d}x$, $\mathrm{d}y$ and $\mathrm{d}m$, the conditional intensity function $\lambda(t, x, y, m|\mathcal{H}_t)$ satisfies

$$\lambda(t, x, y, m|\mathcal{H}_t)\mathrm{d}t\mathrm{d}x\mathrm{d}y\mathrm{d}m \approx \mathsf{P}\Big\{\mathbf{X} \cap [t, t+\mathrm{d}t) \times [x, x+\mathrm{d}x) \times [y, y+\mathrm{d}y) \times [m, m+\mathrm{d}m) \neq \emptyset \Big| \mathcal{H}_t\Big\},$$

where

$$\mathcal{H}_t = \{(t_i, x_i, y_i, m_i) \in \mathbf{X} : t_i < t\}$$

is the occurrence history of the earthquakes up to time t (Ogata 1998; Zhuang *et al.* 2002).

We assume that the time t is measured in decimal days with $t = 0$ corresponding to a specific time origin, say 1945-01-01, 00:00:00, and a study period of length T days started at time $t_{\text{start}} \geq 0$ is of interest. Only events inside the study region S and the study period $[t_{\text{start}}, t_{\text{start}} + T]$ are considered as *target* events. Other events with $t < t_{\text{start}}$ or $(x, y) \notin S$, if any, are assumed to be *complementary* events (Zhuang, Ogata, and Vere-Jones 2006). Complementary events are included in the history $\mathcal{H}_t$ in order to take into account their effect on target events and to some extent resolve what is known as the boundary (edge) effect (Ogata 1998). Note that N is the total number of target and complementary events in the catalog.

2.2. The space-time ETAS model

The ETAS model is a multiple type branching process and a special case of the marked Hawkes process (see Daley and Vere-Jones 2003, pp. 202–205). The conditional intensity function of

Event	Time	Longitude	Latitude	Magnitude
1	t_1	x_1	y_1	m_1
2	t_2	x_2	y_2	m_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
N	t_N	x_N	y_N	m_N

Table 2: Example of catalog data.

the space-time ETAS model is given by

$$\lambda_{\beta,\theta}(t,x,y,m|\mathcal{H}_t) = \nu_\beta(m)\lambda_\theta(t,x,y|\mathcal{H}_t), \tag{1}$$

where

$$\nu_\beta(m) = \beta \exp\big[-\beta(m-m_0)\big], \quad \beta > 0,$$

is the probability density function (PDF) of the magnitude of an event and (Ogata 1998)

$$\lambda_\theta(t,x,y|\mathcal{H}_t) = \tilde{u}(x,y) + \sum_{i:t_i<t} \kappa_{A,\alpha}(m_i) g_{c,p}(t-t_i) f_{D,\gamma,q}(x-x_i, y-y_i; m_i), \tag{2}$$

with the following components:

$\tilde{u}(x,y)$ is the *background seismicity rate* and is assumed to be independent of time. The semi-parametric form $\tilde{u}(x,y) = \mu u(x,y)$, where $\mu > 0$ and $u(x,y)$ is a smooth function on S, is usually considered for the background seismicity rate.

$\kappa_{A,\alpha}(m_i)$ is the expected number of triggered events (aftershocks) generated from an event of magnitude m_i and it can be expressed as

$$\kappa_{A,\alpha}(m) = A \exp\big[\alpha(m-m_0)\big],$$

where $A > 0$ and $\alpha > 0$ are unknown parameters.

$g_{c,p}(t-t_i)$ is the PDF of the occurrence time of a triggered event generated from an event of magnitude m_i occurring at time t_i. It is assumed that the probability distribution of the time until the occurrence of a triggered event is a function of the time lag $t-t_i$ from its direct main shock and is independent of m_i. Based on the modified Omori's law, the following inverse power density is considered

$$g_{c,p}(t-t_i) = \begin{cases} \frac{p-1}{c}\left(1+\frac{t-t_i}{c}\right)^{-p} & t-t_i > 0 \\ 0 & t-t_i \le 0 \end{cases},$$

where $c > 0$ and $p > 1$ are unknown parameters.

$f_{D,\gamma,q}(x-x_i, y-y_i; m_i)$ is the PDF of the occurrence location of a triggered event generated from an event of magnitude m_i occurring at the location (x_i, y_i). It is assumed that the probability distribution of the location of a triggered event depends on the magnitude and the location of its direct main shock. The radially symmetric density function

$$f_{D,\gamma,q}(x-x_i, y-y_i; m_i) = \frac{q-1}{\pi D \exp\big[\gamma(m_i-m_0)\big]}\left(1+\frac{(x-x_i)^2+(y-y_i)^2}{D\exp\big[\gamma(m_i-m_0)\big]}\right)^{-q},$$

is usually considered for this distribution, where $D > 0$, $\gamma > 0$ and $q > 1$ are unknown parameters.

The ETAS model assumes that each event is either a *background* (spontaneous) event or *triggered* by a previous event (Ogata 1998). The background events are generated by a Poisson process with intensity $\tilde{u}(x,y)$, which is stationary in time. The relaxing coefficient coefficient μ in $\tilde{u}(x,y) = \mu u(x,y)$ is introduced in order to fasten the convergence of the model

fitting algorithm (see Algorithm 2 below). Previous events, whether they are background or triggered events, generate further events according to a non-stationary Poisson process with intensity function (Zhuang 2011)

$$\sum_{i:t_i<t} \kappa_{A,\alpha}(m_i) g_{c,p}(t-t_i) f_{D,\gamma,q}(x-x_i, y-y_i; m_i).$$

The expected number of triggered events generated from a typical event, regardless of its magnitude, is

$$\int_{m_0}^{\infty} \kappa_{A,\alpha}(m)\nu_\beta(m)\mathrm{d}m = \frac{A\beta}{\beta-\alpha}.$$

If $A\beta/(\beta-\alpha) < 1$, then the stability (existence of a stationary version) of the model is guaranteed (Daley and Vere-Jones 2003, p. 204).

Under the stability condition, the expected number of events in an interval of length T is proportional to T. Thus, the expected number of events tend to infinity as $T \to \infty$. Theoretically, the total spatial intensity function is defined as

$$\begin{aligned}\Lambda(x,y) &= \lim_{T\to\infty} \frac{1}{T}\int_0^T \lambda_\theta(t,x,y|\mathcal{H}_t)\mathrm{d}t \\ &= \mu u(x,y) + \lim_{T\to\infty}\left(\frac{1}{T}\sum_{i:t_i<T} \kappa_{A,\alpha}(m_i) f_{D,\gamma,q}(x-x_i, y-y_i; m_i)\int_{t_i}^T g_{c,p}(t-t_i)\mathrm{d}t\right).\end{aligned}$$

In practice, the total spatial intensity function is approximated by

$$\Lambda(x,y) \approx \mu u(x,y) + \frac{1}{T}\sum_{i:t_i<T} \kappa_{A,\alpha}(m_i) f_{D,\gamma,q}(x-x_i, y-y_i; m_i)$$

and used to obtain the clustering (triggering) coefficient

$$\omega(x,y) = 1 - \frac{u(x,y)}{\Lambda(x,y)},$$

which quantifies the clustering effect relative to the total spatial intensity at any location $(x,y) \in S$ (Zhuang *et al.* 2006).

2.3. Stochastic declustering

Zhuang *et al.* (2002) proposed a probabilistic approach for declustering an earthquake catalog to obtain the background seismicity rate. They suggested to use

$$p_{ij} = \begin{cases} \frac{\kappa_{A,\alpha}(m_i) g_{c,p}(t_j-t_i) f_{D,\gamma,q}(x_j-x_i, y_j-y_i; m_i)}{\lambda_\theta(t_j,x_j,y_j|\mathcal{H}_{t_j})}, & t_j > t_i \\ 0, & t_j \le t_i \end{cases} \tag{3}$$

as the probability that event j is triggered by event i. Consequently,

$$p_j = \sum_{i:t_i<t_j} p_{ij}, \quad j = 1,\ldots,N, \tag{4}$$

is the probability that event j is triggered by a previous event and

$$1 - p_j = \frac{\mu u(x_j, y_j)}{\lambda_\theta(t_j, x_j, y_j | \mathcal{H}_{t_j})}$$

is the probability that event j is a background event.

3. Parameter estimation and model fitting

The ETAS model, defined through (1) and (2), is a semi-parametric model with Euclidean parameters β and $\theta = (\mu, A, \alpha, c, p, D, \gamma, q)$ and an infinite dimensional parameter $u(x, y)$, $(x, y) \in S$ (Zhuang *et al.* 2006). The log-likelihood function of the model is given by (Ogata 1998)

$$l(\beta, \theta | \mathcal{H}_T) = \sum_{i=1}^{N} \delta_i \log\left(\lambda_{\beta,\theta}(t_i, x_i, y_i, m_i | \mathcal{H}_{t_i})\right) - \int_{m_0}^{\infty} \int_{t_{\text{start}}}^{t_{\text{start}}+T} \iint_S \lambda_{\beta,\theta}(t, x, y, m | \mathcal{H}_t) \mathrm{d}x\mathrm{d}y\mathrm{d}t\mathrm{d}m,$$

where $\delta_i = 1$ if event i is a target event and $\delta_i = 0$ otherwise. Due to separability of β and θ in (1) and the fact that $\int_{m_0}^{\infty} \nu_\beta(m)\mathrm{d}m = 1$, the log-likelihood function can be written as

$$l(\beta, \theta | \mathcal{H}_T) = l_1(\beta | \mathcal{H}_T) + l_2(\theta | \mathcal{H}_T),$$

where

$$l_1(\beta | \mathcal{H}_T) = \sum_{i=1}^{N} \delta_i \log\left(\nu_\beta(m_i)\right) = \log(\beta) \sum_{i=1}^{N} \delta_i - \beta \sum_{i=1}^{N} \delta_i (m_i - m_0),$$

and

$$l_2(\theta | \mathcal{H}_T) = \sum_{i=1}^{N} \delta_i \log\left(\lambda_\theta(t_i, x_i, y_i | \mathcal{H}_{t_i})\right) - \int_{t_{\text{start}}}^{t_{\text{start}}+T} \iint_S \lambda_\theta(t, x, y | \mathcal{H}_t) \mathrm{d}x\mathrm{d}y\mathrm{d}t, \tag{5}$$

where

$$\int_{t_{\text{start}}}^{t_{\text{start}}+T} \iint_S \lambda_\theta(t, x, y | \mathcal{H}_t) \mathrm{d}x\mathrm{d}y\mathrm{d}t = T\mu \iint_S u(x, y) \mathrm{d}x\mathrm{d}y + \sum_{i=1}^{N} \kappa_{A,\alpha}(m_i) \int_{t_{\text{start}}}^{t_{\text{start}}+T} g_{c,p}(t - t_i) \mathrm{d}t \iint_S f_{D,\gamma,q}(x - x_i, y - y_i; m_i) \mathrm{d}x\mathrm{d}y \tag{6}$$

and

$$\int_{t_{\text{start}}}^{t_{\text{start}}+T} g_{c,p}(t - t_i) \mathrm{d}t = \begin{cases} \left[1 + (t_{\text{start}} - t_i)/c\right]^{1-p} - \left[1 + (t_{\text{start}} + T - t_i)/c\right]^{1-p} & t_i \le t_{\text{start}} \\ 1 - \left[1 + (t_{\text{start}} + T - t_i)/c\right]^{1-p} & t_i > t_{\text{start}} \end{cases}.$$

Thus, by maximizing $l_1(\beta | \mathcal{H}_T)$, the maximum likelihood estimate (MLE) of β is

$$\widehat{\beta} = \frac{N'}{\sum_{i=1}^{N} \delta_i (m_i - m_0)},$$

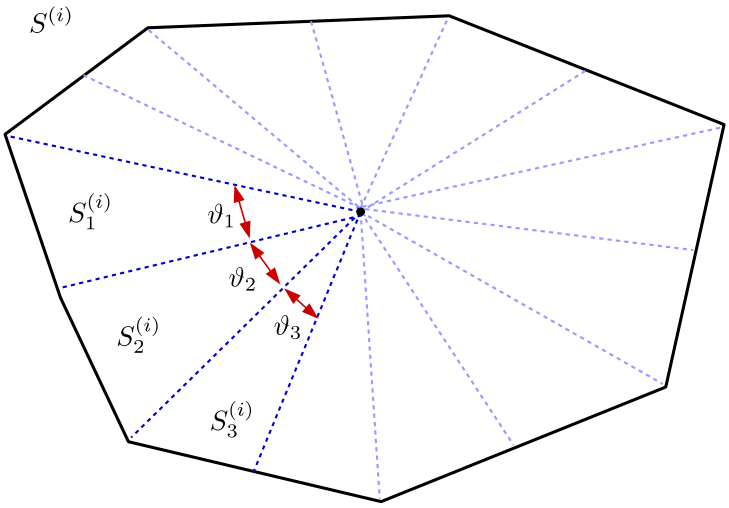

Figure 1: Radial partitioning of the geographical region S for integral approximation.

where $N' = \sum_{i=1}^{N} \delta_i$, while to obtain the MLE of θ, an estimate of $u(x, y)$, numerical approximations of the integral terms over S in (6) and an appropriate numerical maximization method for $l_2(\theta|\mathcal{H}_T)$ are required.

Ogata (1998) suggested a radial partitioning of the geographical region S in order to efficiently approximate the integral term

$$\iint_S f_{D,\gamma,q}(x - x_i, y - y_i; m_i)\mathrm{d}x\mathrm{d}y = \iint_{S^{(i)}} f_{D,\gamma,q}(x, y; m_i)\mathrm{d}x\mathrm{d}y,$$

where

$$S^{(i)} = \{(x - x_i, y - y_i) : (x, y) \in S\}, \quad i = 1, \ldots, N.$$

In this method, for each event i, $S^{(i)}$ is partitioned into n_v subregions $S_1^{(i)}, \ldots, S_{n_v}^{(i)}$ by dividing the boundary of $S^{(i)}$ into n_v knots and using radial segments to connect the origin, $(0, 0)$, and each of these knots. Then the integral on each subregion is transformed from the Cartesian coordinates (x, y) to the polar coordinates (r, ϑ) and approximated using the length of the radial segments and the angles between them (see Figure 1). As follows in (10), the integral of $u(x, y)$ over S is approximated similarly.

Zhuang *et al.* (2002) introduced an iterative approach to simultaneous estimation of the smoothed function $u(x, y)$, $(x, y) \in S$, and the parameter vector θ under the stochastic declustering framework. Given an initial estimate for $u(u, v)$ and using the above mentioned integral approximation, the MLE of θ is obtained by minimizing $\xi(\theta) = -l_2(\theta|\mathcal{H}_T)$ using the Davidon-Fletcher-Powell method, which is a gradient-based nonlinear optimization algorithm (Ogata 1998). Let $\nabla\xi(\theta)$ denote the gradient of $\xi(\theta)$. Then Algorithm 1 summarizes the Davidon-Fletcher-Powell (DFP) procedure. As $k \to \infty$ in Algorithm 1, $\widehat{\theta}_k \to \theta^*$, where θ^* is a local minimum of $\xi(\theta)$, and the matrix H_k converges to the inverse of the Hessian matrix of $\xi(\theta)$ at $\theta = \theta^*$; i.e.,

$$(H_k^{-1})_{ij} \to \frac{\mathrm{d}^2}{\mathrm{d}\theta_i \mathrm{d}\theta_j} \xi(\theta)\Big|_{\theta=\theta^*}$$

Algorithm 1 Davidon-Fletcher-Powell algorithm.

Inputs:
  $\widehat{\theta}_0$ ▷ initial guess for the model parameters
  $\widehat{u}_0(x_j, y_j), \quad j = 1, \ldots, N$ ▷ initial background rate
  H_0 ▷ initial guess for the inverse of the Hessian matrix
$k \leftarrow 0$
repeat
  $d_k \leftarrow -H_k \times \nabla\xi(\widehat{\theta}_k)$ ▷ ascent direction
  $\zeta_k \leftarrow \arg\min_\zeta \xi(\widehat{\theta}_k + \zeta d_k)$ ▷ line search along d_k
  $\widehat{\theta}_{k+1} \leftarrow \widehat{\theta}_k + \zeta_k d_k$ ▷ update the estimate
  $\eta_k \leftarrow \nabla\xi(\widehat{\theta}_{k+1}) - \nabla\xi(\widehat{\theta}_k)$
  $c_1 \leftarrow \zeta_k / (d_k^\top \times \eta_k)$
  $c_2 \leftarrow 1/(\eta_k^\top \times H_k \times \eta_k)$
  $H_{k+1} \leftarrow H_k + c_1 h_k \times d_k^\top - c_2 H_k \times \eta_k \times \eta_k^\top \times H_k$ ▷ update the inverse Hessian
  $k \leftarrow k + 1$
until $\|\widehat{\theta}_k - \widehat{\theta}_{k-1}\| < \epsilon$ **or** $\|\nabla\xi(\widehat{\theta}_k)\| < \epsilon$ ▷ convergence criteria
return $\widehat{\theta}_k$ and H_k

(Fletcher and Powell 1963). Thus, when the algorithm converges, H_k is an estimate (approximation) for the inverse of the Fisher information matrix $\mathcal{I}(\widehat{\theta}_k)$, where

$$\mathcal{I}_{ij}(\theta) = \mathsf{E}\left[-\frac{\mathrm{d}^2}{\mathrm{d}\theta_i \mathrm{d}\theta_j} l_2(\theta|\mathcal{H}_T)\right], \quad i, j = 1, \ldots, 8.$$

Therefore, by asymptotic properties ($T \to \infty$) of maximum likelihood estimators (Schoenberg 2005; Rathbun 1996), $\widehat{\mathcal{I}}(\widehat{\theta}_k) = H_k$ is an estimate of the asymptotic covariance matrix of $\widehat{\theta}$.

Once the estimates of $u(x, y)$ and θ are available, the declustering probabilities p_j can be estimated from (3) and (4) and the variable bandwidth kernel estimator

$$\widehat{u}(x, y) = \frac{1}{T}\sum_{j=1}^{N}(1 - \widehat{p}_j)\varphi(x - x_j, y - y_j; h_j) \tag{7}$$

provides a better estimate for $u(x, y)$, where

$$\varphi(x, y; h) = \frac{1}{2\pi h^2}\exp\left(-\frac{x^2 + y^2}{2h^2}\right) \tag{8}$$

is the bivariate isotropic Gaussian kernel function with bandwidth h. The bandwidth h_j is considered to be

$$h_j = \max\{h_{\min}, r(j, n_p)\}, \tag{9}$$

where $h_{\min}$ is a minimum threshold bandwidth value and $r(j, n_p)$ denotes the distance between the location of event j and its n_pth nearest neighbor. The n_pth nearest neighbor distance $r(j, n_p)$ can be very small or even zero because events may overlap at the same location. Thus, $h_{\min}$ is used to prevent very small or even zero h_i values (Zhuang *et al.* 2002). Since spatial point patterns of locations of earthquake epicenters are clustered, using a variable bandwidth in the kernel estimator (7) is more preferable than a fixed bandwidth. A fixed bandwidth

Algorithm 2 Fitting the ETAS model using stochastic declustering method.

Inputs:
  $\widehat{\theta}_0$ ▷ initial guess for the model parameters
  $\widehat{u}_0(x_i, y_i), \quad i = 1, \ldots, N$ ▷ initial background rate
$H_0 \leftarrow I$ ▷ initial guess for the inverse of the Hessian matrix
$k \leftarrow 0$
repeat
  use Algorithm 1 with inputs $\widehat{\theta}_k$, $\widehat{u}_k(x_i, y_i)$ and H_k and get $\widehat{\theta}_{k+1}$ and H_{k+1}
  $\widehat{p}_j \leftarrow 1 - \widehat{\mu}_{k+1}\widehat{u}_k(x_j, y_j)/\lambda_{\widehat{\theta}_{k+1}}(t_j, x_j, y_j|\mathcal{H}_{t_j})$ ▷ declustering probabilities
  $\widehat{u}_{k+1}(x, y) \leftarrow \sum_{j=1}^{N}(1 - \widehat{p}_j)\varphi(x - x_j, y - y_j; h_j)/T$ ▷ update the estimate of $u(x, y)$
  $k \leftarrow k + 1$
  $e_1 \leftarrow \max_l |\widehat{\theta}_{l,k}/\widehat{\theta}_{l,k-1} - 1|$ ▷ maximum relative change in parameters
  $e_2 \leftarrow \max_j |\widehat{u}_k(x_j, y_j)/\widehat{u}_{k-1}(x_j, y_j) - 1|$ ▷ maximum relative change in $\widehat{u}(x_j, y_j)$
  $e_3 \leftarrow |\xi(\widehat{\theta}_k)/\xi(\widehat{\theta}_{k-1}) - 1|$ ▷ relative change in the log-likelihood function
until $e_1 < \varepsilon$ **and** $e_2 < \varepsilon$ **and** $e_3 < \varepsilon$ ▷ convergence criteria
return $\widehat{\theta}_k$, $\widehat{u}_k(x, y)$ and H_k

produces under-smoothing in areas with sparse observations and over-smoothing in areas with dense observations. The bandwidth (9) is adaptive in the sense that $r(j, n_p)$, and hence h_j, is small in areas with dense observations and large in areas with sparse observations. As suggested by Zhuang (2011), $n_p = 3, \ldots, 6$ seems reasonable and $h_{\min}$ should be equivalent to the order of earthquake location error (due to location inaccuracy or rounded coordinates at a certain precision), say $h_{\min} = 0.05°$ equivalent to 5.56 kilometers on the earth surface.

Then,

$$\iint_S \widehat{u}(x, y)\mathrm{d}x\mathrm{d}y = \frac{1}{T}\sum_{j=1}^{N}(1 - \widehat{p}_j)\iint_{S^{(j)}} \varphi(x, y; h_j)\mathrm{d}x\mathrm{d}y \tag{10}$$

and since (8) is a radially symmetric function, integral terms in the right hand side of (10) can be approximated by transforming to the polar coordinate system and using the radial partitioning of the geographical regions $S^{(j)}$ as described above. Replacing the initial estimate of $u(x, y)$ by the new estimate from (7), estimating θ and $u(x, y)$ is repeated until the estimates converge (Zhuang *et al.* 2006). This iterative approach is summarized in Algorithm 2.

Finally, similar to (7), the total spatial intensity function can be estimated by

$$\widehat{\Lambda}(x, y) = \frac{1}{T}\sum_{j=1}^{N}\varphi(x - x_j, y - y_j; h_j), \tag{11}$$

which along with $\widehat{u}(x, y)$ yields the estimate $\widehat{\omega}(x, y) = 1 - \widehat{u}(x, y)/\widehat{\Lambda}(x, y)$ for the clustering coefficient.

4. Residuals analysis and diagnostics plots

The first-order spatio-temporal residuals of a fitted ETAS model with respect to $h(t, x, y) \geq 0$ are defined as (Baddeley, Turner, Møller, and Hazelton 2005; Zhuang 2006):

$$R(I \times B; h) = \sum_{i=1}^{N} \delta_i \mathbb{1}\left[t_i \in I, (x_i, y_i) \in B\right] h(t_i, x_i, y_i) \lambda_{\widehat{\theta}}(t_i, x_i, y_i | \mathcal{H}_{t_i}) - \int_I \iint_B h(t, x, y) \lambda_{\widehat{\theta}}(t, x, y | \mathcal{H}_t) \mathrm{d}x\mathrm{d}y\mathrm{d}t,$$

where $\mathbb{1}[\,\cdot\,]$ denotes the indicator function, $I \subset [t_{\text{start}}, t_{\text{start}} + T]$ and $B \subset S$. Similarly, the first-order temporal and spatial residuals of a fitted ETAS model can be defined, respectively, by

$$R^{\text{temp}}(I; h) = \sum_{i=1}^{N} \delta_i \mathbb{1}[t_i \in I] h(t_i) \lambda^{\text{temp}}_{\widehat{\theta}}(t_i | \mathcal{H}_{t_i}) - \int_I h(t) \lambda^{\text{temp}}_{\widehat{\theta}}(t | \mathcal{H}_t) \mathrm{d}t$$

and

$$R^{\text{spat}}(B; h) = \sum_{i=1}^{N} \delta_i \mathbb{1}[(x_i, y_i) \in B] h(x_i, y_i) \lambda^{\text{spat}}_{\widehat{\theta}}(x_i, y_i) - \iint_B h(x, y) \lambda^{\text{spat}}_{\widehat{\theta}}(x, y) \mathrm{d}x\mathrm{d}y,$$

where

$$\lambda^{\text{temp}}_{\widehat{\theta}}(t | \mathcal{H}_t) = \iint_S \lambda_{\widehat{\theta}}(t, x, y | \mathcal{H}_t) \mathrm{d}x\mathrm{d}y \quad \text{and} \quad \lambda^{\text{spat}}_{\widehat{\theta}}(x, y) = \int_{t_{\text{start}}}^{t_{\text{start}}+T} \lambda_{\widehat{\theta}}(t, x, y | \mathcal{H}_t) \mathrm{d}t.$$

If the fitted model is able to adequately describe the spatial and temporal variations in the data, then we expect to have $R(I \times B; h) \approx 0$, $R^{\text{temp}}(I; h) \approx 0$ and $R^{\text{spat}}(B; h) \approx 0$ for any $I \subset [t_{\text{start}}, t_{\text{start}} + T]$ and $B \subset S$. Thus, any systematic departure of residuals from zero reflects model inadequacy.

To assess the fitted ETAS model, and for the ease of computation and visualization, we only compute the temporal residuals $R^{\text{temp}}(I_j; h)$ and spatial residuals $R^{\text{spat}}(B_i; h)$, where $I_1, \ldots, I_{n_{\text{temp}}}$ and $B_1, \ldots, B_{n_{\text{spat}}}$ form finite partitions of the study period $[t_{\text{start}}, t_{\text{start}} + T]$ and the geographical region S, respectively. For example, the temporal residuals can be computed for $I_j = \left(t_{\text{start}} + (j-1)T/n_{\text{temp}}, t_{\text{start}} + jT/n_{\text{temp}}\right]$, $j = 1, \ldots, n_{\text{temp}}$, for some n_{temp}. Plotting and investigating temporal residuals $R^{\text{temp}}(I_j; h)$ may reveal possible temporal insufficiency of the fitted model.

Similarly, a suitable quadrature scheme can be used to compute spatial residuals. For example, the Berman-Turner device adds additional dummy points to the set of locations of target events $\{(x_i, y_i) : i = 1, \ldots, N, \delta_i = 1\}$ in order to create a quadrature scheme with quadrature points $(\tilde{x}_i, \tilde{y}_i)$ with corresponding quadrature cells B_i and quadrature weights $\tilde{w}_i = |B_i|$, $i = 1, \ldots, n_{\text{spat}}$ (see Baddeley and Turner 2000). Using the Berman-Turner quadrature scheme the spatial residuals become

$$\begin{aligned} R^{\text{spat}}(B_i; h) &= \tilde{\delta}_i h(\tilde{x}_i, \tilde{y}_i) \lambda^{\text{spat}}_{\widehat{\theta}}(\tilde{x}_i, \tilde{y}_i) - \int_{B_i} h(x, y) \lambda^{\text{spat}}_{\widehat{\theta}}(x, y) \mathrm{d}x\mathrm{d}y \\ &\approx h(\tilde{x}_i, \tilde{y}_i) \lambda^{\text{spat}}_{\widehat{\theta}}(\tilde{x}_i, \tilde{y}_i)(\tilde{\delta}_i - \tilde{w}_i), \quad i = 1, \ldots, n_{\text{spat}}, \end{aligned}$$

where $\tilde{\delta}_i = 0$ if $(\tilde{x}_i, \tilde{y}_i)$ is a dummy point and $\tilde{\delta}_i = 1$ otherwise. The spatial residuals $R^{\text{spat}}(B_i; h)$, $i = 1, \ldots, n_{\text{spat}}$, then can be smoothed by, e.g., the Gaussian kernel (8) to obtain an image of smoothed spatial residuals (cf. Baddeley *et al.* 2005).

Three types of raw, reciprocal and Pearson temporal residuals are obtained, respectively, with $h(t) = 1$, $h(t) = 1/\lambda_{\widehat{\theta}}^{\text{temp}}(t|\mathcal{H}_t)$ and $h(t) = 1/\sqrt{\lambda_{\widehat{\theta}}^{\text{temp}}(t|\mathcal{H}_t)}$ and setting $h(x, y) = 1$, $h(x, y) = 1/\lambda_{\widehat{\theta}}^{\text{spat}}(x, y)$ and $h(x, y) = 1/\sqrt{\lambda_{\widehat{\theta}}^{\text{spat}}(x, y)}$ yields the same types of spatial residuals (Baddeley *et al.* 2005; Zhuang 2006).

The so-called transformed times

$$\tau_j = \int_{t_{\text{start}}}^{t_j} \lambda_{\widehat{\theta}}^{\text{temp}}(t|\mathcal{H}_t)\mathrm{d}t = \int_{t_{\text{start}}}^{t_j} \iint_S \lambda_{\widehat{\theta}}(t, x, y|\mathcal{H}_t)\mathrm{d}x\mathrm{d}y\mathrm{d}t$$

for target events $j \in J = \{i : \delta_i = 1\}$ are closely related to raw temporal residuals $R^{\text{temp}}([t_{\text{start}}, t_j]; 1)$ (Zhuang 2006). If the fitted model is close enough to the true underlying earthquake process, then $\{\tau_j : j \in J\}$ follows a standard (unit rate) Poisson process and, consequently, $U_j = 1 - \exp\big[-(\tau_j - \tau_{j-1})\big]$, $j \in J$, are independent and identically distributed uniform random variables on $(0, 1)$ (Ogata 1988; Rathbun 1996). Thus, plotting τ_j against j and a Q-Q plot of U_j yield simple diagnostic checks for the temporal adequacy of the fitted model: If the model is good enough then points of the first plot approximately lie on the line $y = x$ and the second plot shows agreement between the empirical quantiles of U_j and the theoretical quantiles of a $U(0, 1)$ distribution. Furthermore, the Kolmogorov-Smirnov test for goodness-of-fit can be employed to determine whether random variables $U_2, \ldots, U_N$ follow a $U(0, 1)$ distribution.

5. The ETAS package

Assume that the catalog data are imported into R as a '`data.frame`' object (see the R Data Import/Export manual for information about importing and working with different data formats in R; R Core Team 2018a). This '`data.frame`' object is required to have at least 5 columns with (exact or partial match of) names `date`, `time`, `lat`, `long` and `mag` containing, respectively, the date, time, latitude, longitude and magnitude of each event in the catalog. The `date` and `time` columns should be in appropriate formats such that `paste(date, time)` can be converted to date-time (calendar dates plus time to the nearest second) by the R function `as.POSIXlt`. For information on calendar dates and times in R, see the help on `"DateTimeClasses"` obtained with:

```
R> help("DateTimeClasses")
```

5.1. Data preparation with function `catalog`

As the first step, the '`data.frame`' object that contains the earthquake data should be passed to the user accessible `catalog` function in order to specify the data as an earthquake catalog. The function `catalog` creates an object of class '`catalog`' to represent an earthquake catalog. An object of this class contains the earthquake data as well as the intended study period, geographical region and the magnitude threshold. The magnitude threshold can be specified by the argument `mag.threshold`. If not specified, the minimum magnitude value in the catalog will be used as the magnitude threshold. The arguments `time.begin`, `study.start`, `study.end` and `study.length` may be used to determine the study period (see Figure 2):

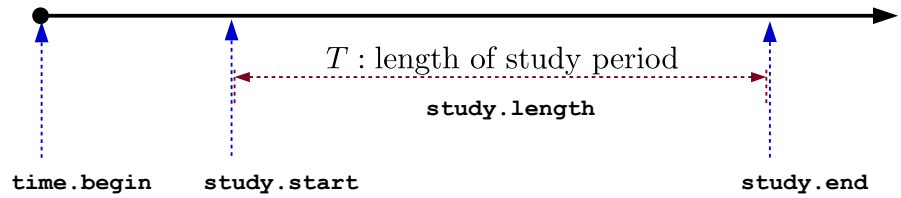

Figure 2: Specifying the study period using `time.begin`, `study.start`, `study.end` and `study.length` arguments in the `catalog` function.

- `time.begin` is the beginning of time span of the catalog. By default, it is set to the time of the first event in the catalog.
- `study.start` is the start of the study period. If not specified, then `time.begin` will be used instead.
- `study.end` is the end of the study period. By default, it is set to the time of the last even in the catalog.
- `study.length` is the length of the study period (T) in decimal days.

Arguments `time.begin`, `study.start` and `study.end` must be character strings or objects that can be converted to date-time by the `as.POSIXlt` function. Either `study.end` or `study.length` can be specified, but not both. The occurrence times are converted to elapsed decimal days since `time.begin`. If the events in the catalog are not chronologically sorted, then the `catalog` function will produce a warning and sort the events in ascending order with respect to time of occurrence.

In **ETAS**, a geographical region is defined as a set of point coordinates in a rectangular or irregular polygonal two-dimensional study region (see Figure 3). For a rectangular geographical region, the latitude and longitude ranges should be specified by `lat.range` and `long.range` arguments. A polygonal study region can be specified by the `region.poly` argument which contains coordinates of the vertices of the polygon. The `region.poly` argument must be either a list with components `lat` and `long` of equal length or a '`data.frame`' with columns `lat` and `long`. The vertices must be listed in anticlockwise order and no vertex should be repeated (i.e., the first vertex is not repeated, see Figure 3).

The function `inside.owin` in the R package **spatstat** (Baddeley and Turner 2005; Baddeley, Rubak, and Turner 2015) is used to indicate whether events lie inside the study region. As mentioned in Section 2.1, only events inside the study region and the study period (`study.start`, `study.end`) are considered as target events and other events are complementary events.

By default (`flatmap = TRUE`) the spherical coordinates `(long, lat)` on the earth surface are transformed to flat map (planar) coordinates `(x, y)` in order to approximate the great-circle distance on the sphere by the corresponding Euclidean distance on the flat map. The default case `dist.unit = "degree"` uses the Equirectangular projection (see Snyder 1997, pp. 5–8)

$$\texttt{x} = \texttt{cos}(\texttt{cnt.lat}/180 \times \texttt{pi}) \times (\texttt{long} - \texttt{cnt.long}) \quad \text{and} \quad \texttt{y} = \texttt{lat} - \texttt{cnt.lat},$$

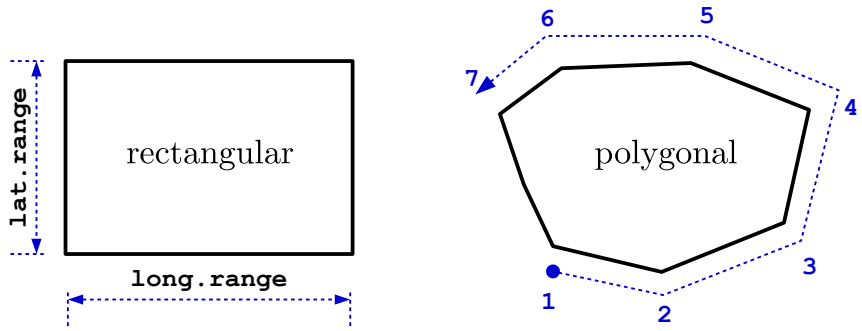

Figure 3: Specifying the rectangular or polygonal two-dimensional study region in the `catalog` function using `lat.range` and `long.range` or `region.poly` arguments.

`dist.unit`	$u(x,y)$	μ	A	α, γ	c	p	D	q
`degree`	$\frac{\text{shock}}{\text{day}\times\text{degree}^2}$	—	shock	magnitude^{-1}	day	—	degree^2	—
`km`	$\frac{\text{shock}}{\text{day}\times\text{km}^2}$	—	shock	magnitude^{-1}	day	—	km^2	—

Table 3: Units of the ETAS model parameters. Dimensionless parameters are denoted by —.

to obtain the flat map coordinates `(x, y)` in degrees, where `cnt.lat` and `cnt.long` are, respectively, the latitude and longitude of the centroid of the geographical region. Setting `dist.unit = "km"` performs the projection

$$\texttt{x} = 111.32 \times \texttt{cos}(\texttt{lat}/180 \times \texttt{pi}) \times \texttt{long} \quad \text{and} \quad \texttt{y} = 110.547 \times \texttt{lat},$$

where here `x` and `y` are in (approximate) kilometers. Note that changing `dist.unit` affects the unit of $u(x, y)$ and D parameters (see Table 3).

Specific `print` and `plot` methods associated with '`catalog`' objects are designed to allow users to see a summary of the specified earthquake catalog.

5.2. Fitting the model with function `etas`

The function `etas` fits the ETAS model to a catalog of earthquakes. The function takes an object of class '`catalog`' and calls the internal functions `decluster` and `etasfit` for simultaneous estimation of the background seismicity rate $u(x, y)$ and the parameter vector θ using the iterative approach discussed in Section 3. An initial guess for the parameter vector $\theta = (\mu, A, \alpha, c, p, D, \gamma, q)$ can be provided by the `param0` argument. Thus, `param0` needs to be a numeric vector of length 8. If `param0` is not specified, as suggested by Ogata (1998), the default values $\mu = N/(4T|S|)$, $A = 0.01$, $c = 0.01$, $\alpha = 1$, $p = 1.3$, $D = 0.01$, $q = 2$ and $\gamma = 1$ are used as components of `param0`. Note that this is a crude initial estimate and does not guarantee the convergence of Algorithm 2. The algorithm is relatively sensitive to the choice of the initial estimate of θ and, if possible, providing more informative initial estimate is highly recommended.

The bandwidth for each event in the kernel estimators (7) and (11) of the background seismicity rate and total spatial intensity can be specified by the `bwd` argument, which needs to be a numeric vector of length N. If `bwd` is not specified, the `nnp` and `bwm` arguments determine,

respectively, the number of nearest neighbors n_p and the minimum threshold bandwidth $h_{\min}$ and (9) is used to obtain the bandwidth for each event. The values `nnp = 5` and `bwm = 0.05` are used as default values for `nnp` and `bwm` (see Zhuang 2011). The argument `ndiv`, with default value `ndiv = 1000`, sets the number of knots on each side of the geographical region S in radial partitioning of S for integral approximation as described in Section 3.

The `no.itr` argument specifies the maximum number of iterations in the iterative estimation method (Algorithm 2). The estimates often converge in less than ten iterations. If `no.itr` is not specified, `no.itr = 11` will be used. Moreover, the relative iteration convergence tolerance ε in Algorithm 2 and the optimization convergence tolerance ϵ in Algorithm 1 are, respectively, determined by the `rel.tol` and `eps` arguments. The progress of the computations can be traced by setting the the `verbose` and `plot.it` arguments to `TRUE`. For example, the following output snippet shows the value of the target function $\xi(\theta) = -l_2(\theta|\mathcal{H}_T)$ and its gradient $\nabla\xi(\widehat{\theta})$ for the current estimate of θ and the value of the optimal coefficient ζ of the line search procedure in one iteration of Algorithm 1.

```
estimating:
        start Davidon-Fletcher-Powell procedure ...
Function Value = 3895.6343
Gradient[1] =  -120.51        theta[1] = 0.707107
Gradient[2] =   111.44        theta[2] = 0.447214
Gradient[3] =   -57.87        theta[3] = 0.223607
Gradient[4] =    25.07        theta[4] = 1.643168
Gradient[5] =   261.80        theta[5] = 1.095445
Gradient[6] =   156.03        theta[6] = 0.141421
Gradient[7] =   -31.10        theta[7] = 1.516575
Gradient[8] =     1.32        theta[8] = 0.173205

line search along the specified direction ... zeta = 0.000085
```

The `etas` function returns an object of class '`etas`'. Specific `print` and `plot` methods associated with '`etas`' objects are designed to allow users to see a summary of the fitted model. An object of class '`etas`' can be passed to the `rates` function to obtain and plot estimates of the total spatial intensity, background seismicity rate and clustering coefficient from the fitted model over a grid on S. The `probs` and `resid.etas` functions take an object of class '`etas`' and return estimates of declustering probabilities p_j and model residuals and diagnostic plots, respectively.

Lack of completeness and time-non-stationarity of earthquake catalogs produce some problems in the statistical analysis of the catalog data. This version of the ETAS model assumes that the earthquake catalog is complete and the data are stationary in time (Zhuang 2011). If the catalog is incomplete or there is non-stationarity (e.g., increasing or seasonal trend) in the time of events, then the results of this function are not reliable.

The `etas` function is based on a C port of the original Fortran program by Jiancang Zhuang, Yosihiko Ogata and their colleagues (Zhuang *et al.* 2017). Computations of the conditional intensity function, the log-likelihood function, declustering probabilities and the Davidon-Fletcher-Powell algorithm for optimization are all written in the low-level C code. As of version 0.3, a new C++ code is implemented using the **Rcpp** package (Eddelbuettel and

François 2011; Eddelbuettel 2013) which allows multi-thread parallel computing on multi-core processors with OpenMP (OpenMP Architecture Review Board 2013). It is known that the computation time of the log-likelihood function (5), and consequently the whole model fitting procedure, is approximately proportional to the square of the number of events (Zhuang 2011). Thus, parallel computing of the computations of the log-likelihood function (5) reduces the computation time for large earthquake catalogs.

By default (`cxxcode = TRUE`) the C++ code is used and the argument `nthreads` determines the number of threads in the parallel region of the code. If `nthreads = 1` (the default case), then a serial version of the C++ code carries out the computations. As pointed out in the Writing R Extensions manual (R Core Team 2018c):

> The performance of OpenMP varies substantially between platforms. The Windows implementation has substantial overheads, so is only beneficial if quite substantial tasks are run in parallel.

Thus, in particular on Windows, the parallel computing (`nthreads > 1`) should only be considered for catalogs with a large number of events (say more than 2000). Also, the number of threads `nthreads` should be carefully specified. The `detectCores` function in the **parallel** package (R Core Team 2018b) can be consulted to find out the overall number of available threads on a given machine. For example, on a laptop with an Intel Core i7 processor we get

```
R> parallel::detectCores()

[1] 8
```

However, resource usage and limitations should be considered when setting `nthreads`.

6. Data example

The following example shows how to fit the space-time ETAS model to an earthquake catalog using the R package **ETAS**. Three sample earthquake catalog data sets `iran.quakes`, `italy.quakes` and `japan.quakes` are included within the package. They are objects of class '`data.frame`' and can be directly accessed by loading the package.

```
R> library("ETAS")
```

The `iran.quakes` data set contains data of an earthquake catalog of Iran and surrounding areas (the rectangular geographical region 22°–42°N and 40°–65°E) from 1973-01-01 to 2016-01-01 with the magnitude threshold $m_0 = 4.0$ in body-wave magnitude (mb) scale. The data are extracted from the ANSS global comprehensive catalog (available at `http://earthquake.usgs.gov/earthquakes/search/`).

There are five columns in `iran.quakes` that represent occurrence date, time, longitude, latitude and magnitude of 5970 shallow earthquakes (depth < 100 km).

```
R> summary(iran.quakes)
```

```
        date               time           long            lat
2014-08-18:  44   00:04:50.70:   2   Min.   :40.01   Min.   :22.13
2011-10-23:  42   04:46:33.00:   2   1st Qu.:48.34   1st Qu.:28.32
2013-04-09:  31   07:02:27.00:   2   Median :51.88   Median :32.13
2013-05-11:  27   00:00:19.42:   1   Mean   :52.20   Mean   :32.80
2011-10-24:  19   00:00:41.44:   1   3rd Qu.:56.39   3rd Qu.:37.52
1990-06-21:  18   00:00:46.52:   1   Max.   :65.00   Max.   :41.98
(Other)   :5789   (Other)    :5961
     mag
Min.   :4.000
1st Qu.:4.200
Median :4.400
Mean   :4.468
3rd Qu.:4.700
Max.   :6.200
```

For the purpose of illustration, we consider the study period from 1991-01-01 to 2011-01-01 ($t_{\text{start}} = 6574$ and $T = 7305$ days), the magnitude threshold $m_0 = 4.5$ and a polygonal geographical region with the vertices at (26, 52), (25, 59), (29, 58), (38, 45) and (35, 43). Thus, all events in `iran.quakes` with magnitude greater than or equal to 4.5 that occurred between 1991-01-01 and 2011-01-01, and inside the specified geographical region, are target events. Events with magnitude greater than or equal to 4.5 that occurred before 1991-01-01 or outside the specified geographical region act as complementary events. Note that the start date 1991-01-01 of the study period is $t_{\text{start}} = 6574$ days after the beginning date 1973-01-01 of the time span of the catalog and that the length of the study period is $T = 7305$ days.

```
R> diff(as.Date(c("1973-01-01", "1991-01-01")))

Time difference of 6574 days

R> diff(as.Date(c("1991-01-01", "2011-01-01")))

Time difference of 7305 days
```

To create a catalog with the specified study period, geographical region and magnitude threshold, we use the following commands.

```
R> gregion <- list(lat = c(26, 25, 29, 38, 35),
+    long = c(52, 59, 58, 45, 43))
R> iran.cat <- catalog(iran.quakes, study.start = "1991/01/01",
+    study.end = "2011/01/01", region.poly = gregion, mag.threshold = 4.5)
```

The `print` method for the '`catalog`' class prints a summary of the created catalog. The catalog can be printed by simply typing its name.

```
R> iran.cat
```

```
earthquake catalog:
  time begin 1973-01-06 15:39:31
  study period: 1991-01-01  to  2011-01-01 (T = 7305 days)
geographical region:
    polygonal with vertices:
     lat long
[1,]  26   52
[2,]  25   59
[3,]  29   58
[4,]  38   45
[5,]  35   43
threshold magnitude: 4.5
number of events:
  total events 2500 : 695 target events,  1805 complementary events
  ( 438 events outside geographical region, 1367 events outside study period)
```

The catalog can also be plotted using the generic function `plot`.

```
R> plot(iran.cat)
```

It produces a plot consisting of six subplots: a plot to show the spatial distribution of events and the selected region, three plots to see how latitude, longitude and magnitude of events change over time and two plots to visually inspect the completeness and time stationarity of the catalog. Let N_m denote the number of events in the catalog with magnitude greater than or equal to m. The Gutenberg-Richter law states that $\log_{10}(N_m) = a - bm$ for some a and b, which is equivalent to assuming that m follows an exponential distribution (Ogata 1998). Departure from linearity in the plot of $\log_{10}(N_m)$ versus m is an indication of the catalog incompleteness. In addition, the minimum m value for which this linear relation holds can be used as the magnitude threshold. Time stationarity of the catalog, on the other hand, can be assessed by inspection of the plot of N_t versus t, where N_t is the number of events in the catalog that occurred up to t. If the earthquake process is stationary in time, then we expect $N_t = \lambda_0 t$ for some $\lambda_0 > 0$.

The resulting plot is shown in Figure 4. The plot of $\log_{10}(N_m)$ versus m shows the linear relation expected from the Gutenberg-Richter law. The plot of N_t versus t is linear during the considered study period, but a nonlinear trend is evident over the time span 1973-01-01 to 2011-01-01 (13879 days). Nevertheless, we assume that the specified catalog is complete and stationary in time.

To fit the ETAS model to the created catalog `iran.cat`, we first provide an initial guess for the model parameters.

```
R> param01 <- c(0.5, 0.2, 0.05, 2.7, 1.2, 0.02, 2.3, 0.03)
```

Now, the model can be fitted by the following command.

```
R> iran.fit <- etas(iran.cat, param0 = param01)
```

It takes about 18.4 minutes on a Linux laptop (2.4 GHz Intel Core i7-4700MQ CPU) with the default serial computing (`nthreads = 1`). The elapsed computation time reduces to 8.9 and 6.4 minutes by using parallel computing with `nthreads = 4` and `nthreads = 8`, respectively.

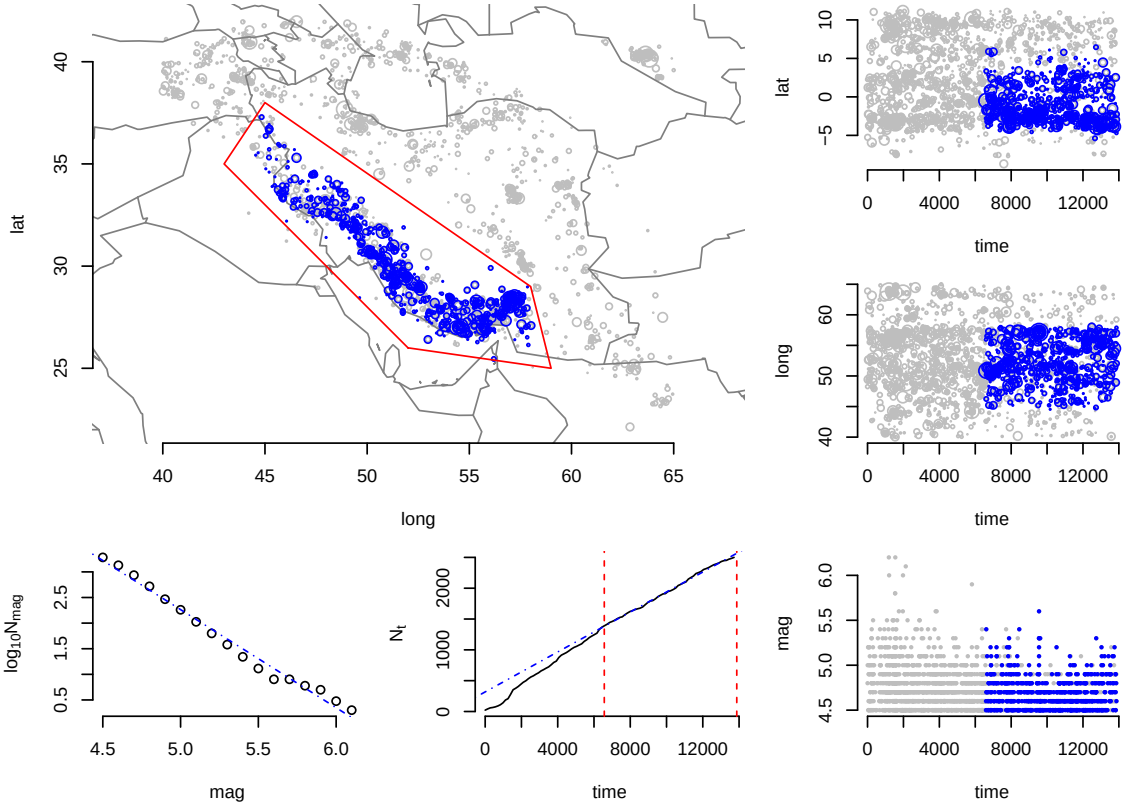

Figure 4: Location of epicenters (top left panel), logarithm of frequency by magnitude (bottom left panel), cumulative frequency over time (bottom middle panel) and latitude, longitude and magnitude against time (right panels) of 5970 earthquakes with magnitude greater than or equal to 4.5 occurred between 1973-01-01 and 2016-01-01 in Iran and its vicinity (22°–42°N and 40°–65°E), extracted from the ANSS global comprehensive catalog (available at http://earthquake.usgs.gov/earthquakes/search/). The plots are created by calling `plot` on an object of class '`catalog`'; an earthquake catalog created by the `catalog` function.

The output `iran.fit` is an object of class '`etas`'. The maximum likelihood estimates of the model parameters, their estimated asymptotic standard errors, the value of the log-likelihood function at its maximum point and the Akaike information criterion (AIC) of the model can be seen by printing `iran.fit`.

```
R> iran.fit

ETAS model: fitted using iterative stochastic declustering method
converged after 4 iterations: elapsed execution time 18.37 minutes

ML estimates of model parameters:
           beta     mu      A      c  alpha      p      D      q  gamma
Estimate 5.6094 0.5484 0.1862 0.0471 2.7071 1.1548 0.0160 2.3234 0.0238
StdErr   0.0453 0.0133 0.0519 0.1093 0.0334 0.0106 0.1016 0.0361 5.7553
```

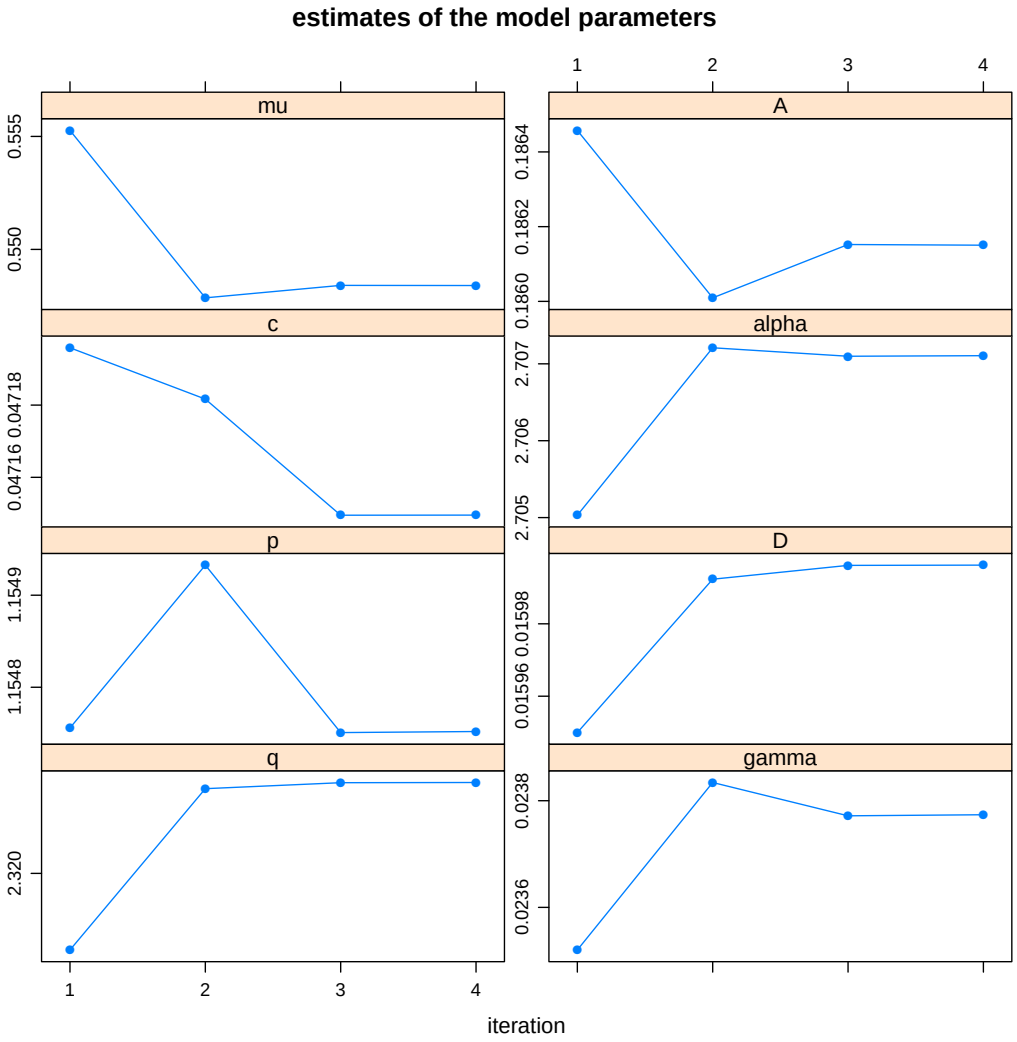

Figure 5: Estimated parameters of the fitted ETAS model to the Iranian catalog in each iteration. The plots are created by calling `plot` on an object of class '`etas`'.

```
Declustering probabilities:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.0000  0.1094  0.9032  0.6365  0.9770  1.0000

log-likelihood:  -3888.709         AIC:  7793.418
```

The estimated asymptotic standard errors of $\widehat{c}$ and specially $\widehat{D}$ and $\widehat{q}$ are much larger than their corresponding estimates. This may be caused by insufficient data (small catalog) or model inadequacy, as discussed in what follows.

Plot of `iran.fit` displays the estimated parameters in each iteration.

```
R> plot(iran.fit)
```

Figure 5 shows the plot. It can be seen that the estimates converge after four iterations.

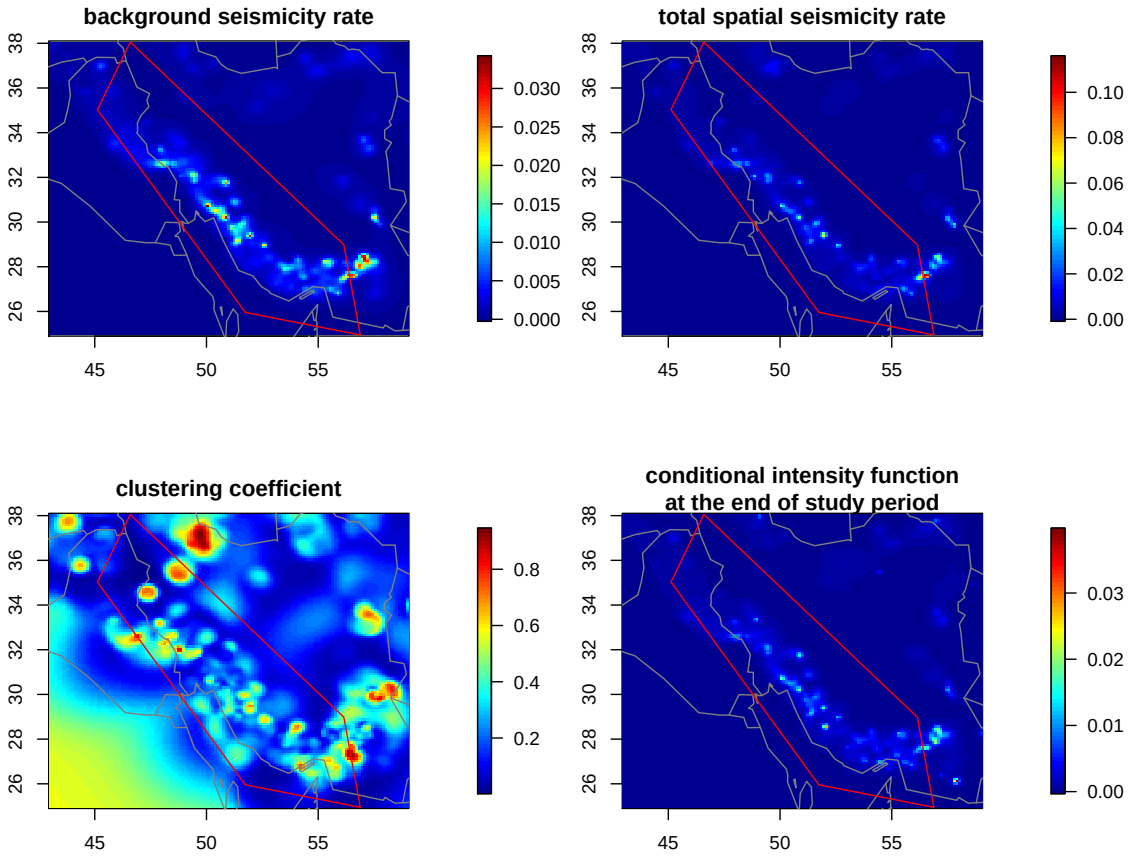

Figure 6: Plots of estimates of the background seismicity rate, total spatial intensity, clustering (triggering) coefficient and conditional intensity at $t = t_{\text{start}} + T$ for the fitted ETAS model to the Iranian catalog. The plots are created by calling the `rates` function on an object of class '`etas`'.

Estimates of the background seismicity rate $u(x, y)$, total spatial intensity $\Lambda(x, y)$, clustering coefficient $\omega(x, y)$ and conditional intensity function $\lambda_\theta(x, y, t|\mathcal{H}_t)$ at $t = t_{\text{start}} + T$ can be obtain by passing `iran.fit` to the `rates` function.

```
R> rates(iran.fit)
```

This function computes and plots the estimates on a rectangular grid over the geographical region (see Figure 6). The grid span and dimension can be changed by arguments `lat.range`, `long.range` and `dimyx`.

In addition, the function `probs` returns estimates of the declustering probabilities p_j.

```
R> pr <- probs(iran.fit)
R> summary(pr$prob)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.00000 0.02299 0.09680 0.36346 0.89058 1.00000
```

These probabilities may be used to indicate the most likely background and triggered target events (see Figure 7).

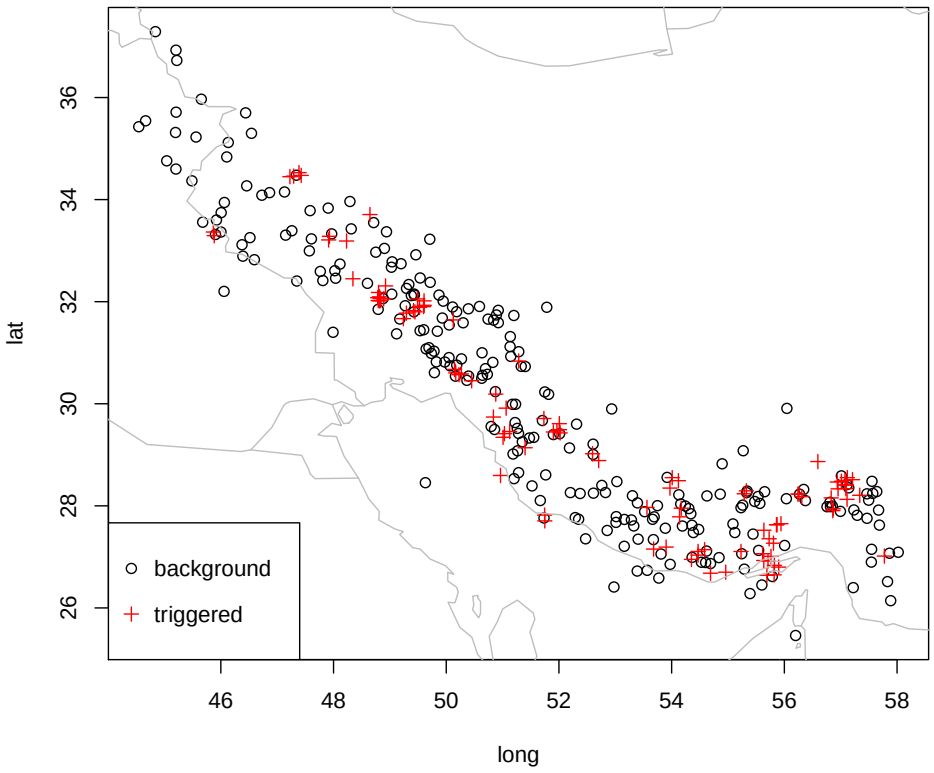

Figure 7: Most likely (probability > 0.95) background and triggered events detected by estimated declustering probabilities p_j.

```
R> plot(iran.cat$longlat.coord[pr$target & (1 - pr$prob > 0.95), 1:2])
R> points(iran.cat$longlat.coord[pr$target & (pr$prob > 0.95), 1:2],
+    pch = 3, col = 2)
R> map("world", add = TRUE, col = "grey")
R> legend("bottomleft", c("background", "triggered"), pch = c(1, 3),
+    col = 1:2)
```

The first-order residual analysis for a fitted ETAS model can be performed by using the `resid.etas` function. The `resid.etas` function computes the temporal residuals $R^{\text{temp}}(I_j; h)$, where $I_j = (t_{\text{start}} + (j-1)T/n_{\text{temp}}, t_{\text{start}} + jT/n_{\text{temp}}]$, $j = 1, \ldots, n_{\text{temp}}$. The argument `n.temp` with default value `n.temp = 1000` specifies the number of partition points n_{temp}. For spatial residuals, the **spatstat** function `quadscheme` is used to create a Berman-Turner quadrature over S and compute $R^{\text{spat}}(B_i; h)$, as described in Section 4. The `resid.etas` function by default (`type = "raw"`) returns raw temporal residuals, raw spatial residuals, transformed times τ_j and their related quantities U_j. It also produces plots of the temporal residuals, smoothed spatial residuals, transformed times τ_j against j and the Q-Q plot of U_j.

```
R> iran.res <- resid.etas(iran.fit)
R> summary(iran.res$tres)
```

```
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-5.71186 -0.55493 -0.53452  0.01888  0.40576 11.95278
```

```
R> summary(na.omit(iran.res$sres$z))
```

```
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.0376927 -0.0000682  0.0606504  0.1067355  0.1971959  0.4459727
```

Figure 8 shows the resulting plots for the fitted ETAS model to the Iranian catalog. It can be seen that points on the plots of transformed times τ_j and the Q-Q plot of U_j lie approximately on the $y = x$ line. Likewise, the Kolmogorov-Smirnov test for goodness-of-fit indicates that U_j follows a $U(0, 1)$ distribution.

```
R> ks.test(iran.res$U, punif)
```

```
        One-sample Kolmogorov-Smirnov test

data:  iran.res$U
D = 0.031565, p-value = 0.4938
alternative hypothesis: two-sided
```

Therefore, the fitted model is able to adequately describe the temporal variation in the data. However, a systematic northwest-southeast trend is evident in the raw spatial residuals. The large (inflated) estimated asymptotic standard errors of $\widehat{D}$ and $\widehat{q}$ perhaps are another sign that the fitted model is not capable of describing the complex spatial clustering of earthquakes in this region. In fact, the geographical region considered here contains the Zagros fold-and-thrust belt, which due to the ongoing collision between the Arabian Plate and the Eurasian Plate, is the most seismically active intra-continental belts on Earth (Talebian and Jackson 2004). Hence, the ETAS model with fixed parameters may not be suitable to model the occurrence of earthquakes in such an active tectonic region. Although this earthquake catalog provides as an example for the **ETAS** package, finding an appropriate model for it is well beyond the scope of this paper. It is possible that improved versions of the ETAS model with spatially varying parameters (Ogata 2011) may be more suitable for this region.

7. Other software packages

The temporal ETAS model can be fitted to the occurrence time and magnitude of earthquakes using the Fortran programs `SASeis2006` (Ogata 2006b) and `etas_solve` (Kasahara, Yagi, and Enescu 2016) and the R packages **PtProcess** (Harte 2010) and **SAPP** (The Institute of Statistical Mathematics 2016). To fit the space-time ETAS model to the occurrence time, location of epicenter and magnitude of earthquakes, the R package **etasFLP** (Chiodi and Adelfio 2017) and the original Fortran code by Jiancang Zhuang (Zhuang *et al.* 2017) are publicly available.

7.1. Comparison with the etasFLP package

The R package **etasFLP** is an alternative to **ETAS** and fits both the temporal and space-time ETAS models. The **etasFLP** package uses the forward likelihood predictive (FLP) method (Chiodi and Adelfio 2011) to estimate the bandwidth of the kernel density estimator (7) for

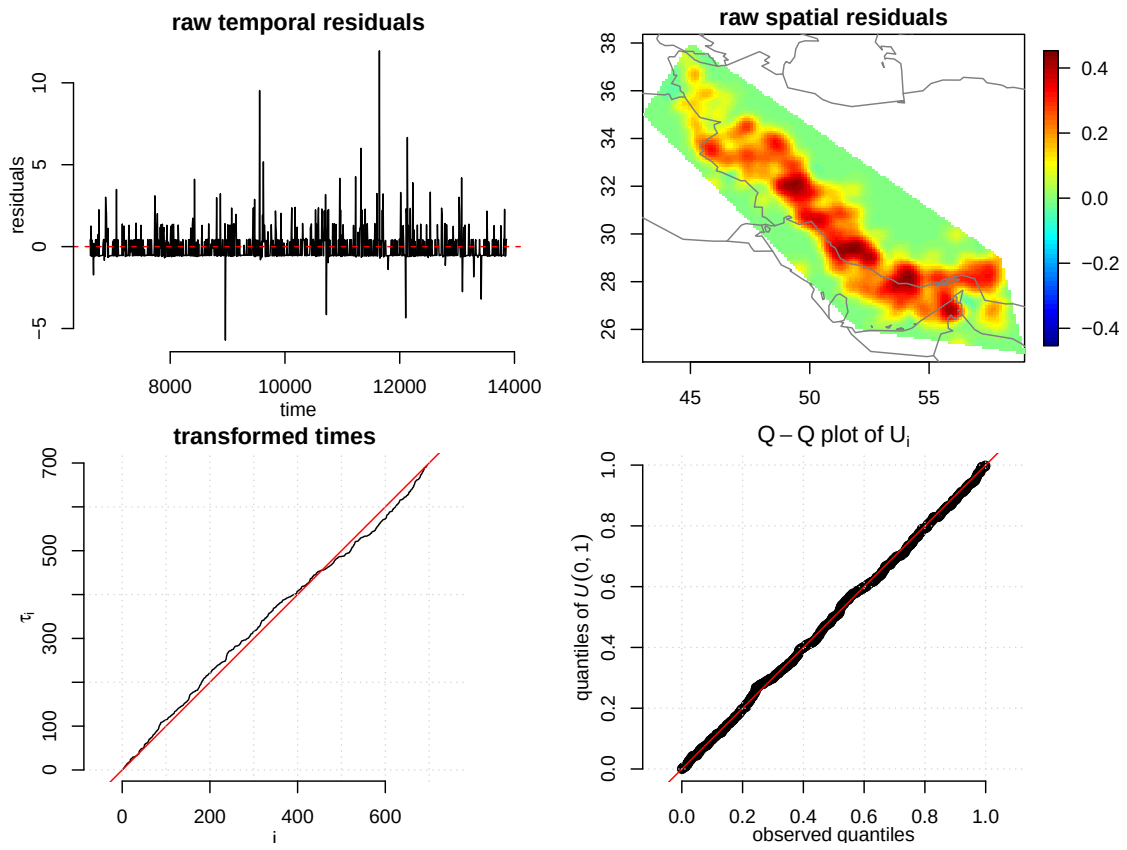

Figure 8: Diagnostic plots for the fitted ETAS model to the Iranian catalog: the temporal residuals (top left), smoothed spatial residuals (top right), transformed times τ_i against i (bottom left) and the Q-Q plot of U_i (bottom right). The plots are created by calling the `resid.etas` function on an object of class '`etas`'.

the background seismicity rate (Adelfio and Chiodi 2013, 2015a). Also, instead of (2), the conditional intensity function in the **etasFLP** package is parameterized as

$$\lambda_\theta(t, x, y|\mathcal{H}_t) = \mu u(x, y) + \sum_{i:t_i<t} \frac{\kappa_0 \exp\left[(\alpha - \gamma)(m_i - m_0)\right]}{(c + t - t_i)^p} \left(D + \frac{(x - x_i)^2 + (y - y_i)^2}{\exp\left[\gamma(m_i - m_0)\right]}\right)^{-q},$$

where

$$\kappa_0 = \frac{1}{\pi} A(p - 1)c^{p-1}(q - 1)D^{q-1}.$$

The package comes with a sample catalog of 2158 Italian earthquakes of magnitude greater than or equal to 3.0 that occurred between 2005 and 2013, extracted from the Italian Seismological Instrumental and Parametric Data Base (ISIDE) available at `http://iside.rm.ingv.it/`.

```
R> library("etasFLP")
R> data("italycatalog", package = "etasFLP")
```

The space-time ETAS model can be fitted to a given catalog by function `etasclass`.

```
R> italy.fit.flp <- etasclass(italycatalog, magn.threshold = 3,
+    magn.threshold.back = 3, mu = 1, k0 = 0.005, c = 0.005, p = 1.01,
+    a = 1.05, gamma = 0.6, q = 1.52, d = 1.1, ndeclust = 11)
```

By default the package uses the transformation `x = radius * long * pi/180` and `y = radius * lat * pi/180`, where `radius = 6371.3` is the mean earth radius, to change the distance unit from degrees to kilometers. The `magn.threshold.back` argument allows to set a different magnitude threshold for the events to be used in the kernel estimator of the background seismicity rate. The maximum number of iterations for the Algorithm 2 can be specified by the argument `ndeclust`.

```
R> summary(italy.fit.flp)
```

```
etasclass(cat.orig = italycatalog, magn.threshold = 3,
    magn.threshold.back = 3,  mu = 1, k0 = 0.005, c = 0.005, p = 1.01,
    a = 1.05, gamma = 0.6, d = 1.1, q = 1.52, ndeclust = 11)

Execution started:                   2016-12-23 11:10:47
Elapsed time of execution (hours)    0.2350042
Number of observations              2158
Magnitude threshold                 3
declustering                         TRUE
Number of declustering iterations    6
Kind of declustering                 weighting
flp                                  TRUE
sequence of AIC values for each iteration
49465.75 48449.99 48417.31 48415.07 48414.87 48414.86

-----------------------------------------------------
ETAS Parameters:
            Estimates       std.err.
mu           0.355891       0.011295
k0           0.008375       0.002055
c            0.009405       0.001795
p            1.121677       0.016273
a            1.509596       0.064070
gamma        0.858127       0.084662
d            1.915362       0.306437
q            1.836685       0.067106
-----------------------------------------------------
```

The `etasclass` function by default (`usenlm = TRUE`) calls the R function `nlm` without supplying any analytical gradient or Hessian in order to perform numerical optimization of the log-likelihood function of the ETAS model. The `nlm` function carries out an unconstrained Newton-type algorithm and uses numerical derivatives to approximate the gradient and Hessian of the target function if the user does not provide them (R Core Team 2018b).

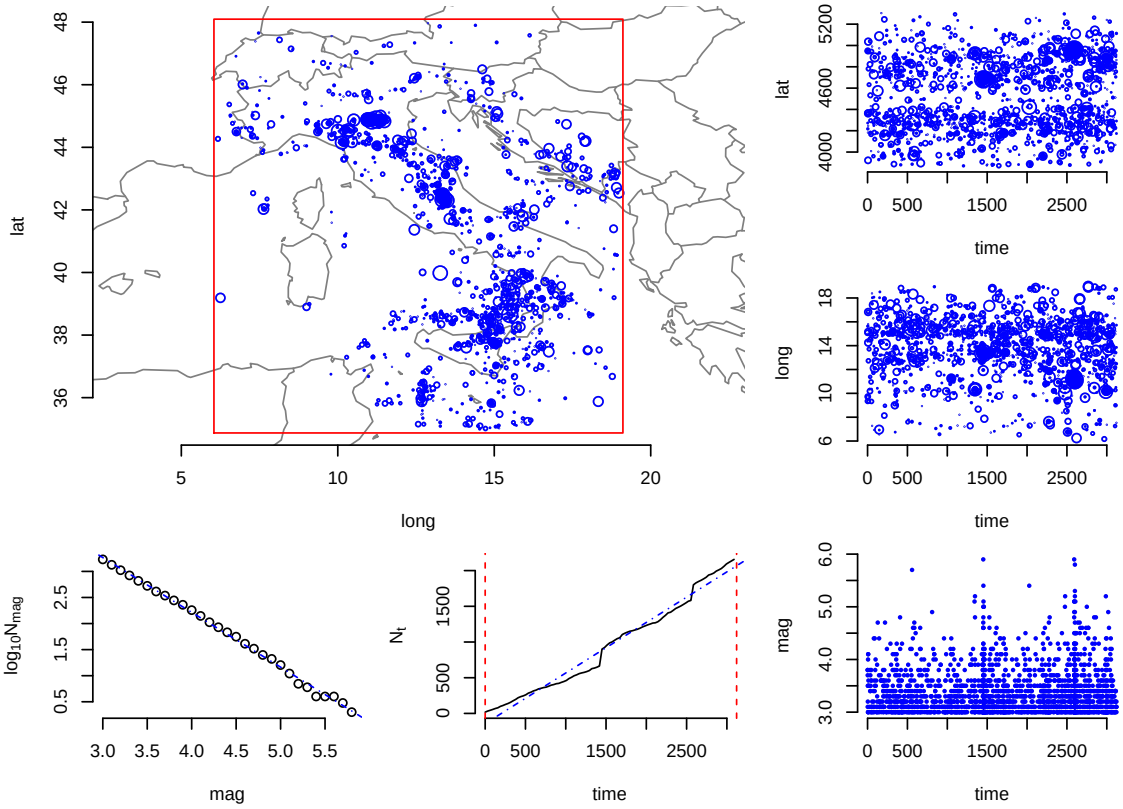

Figure 9: Location of epicenters (top left panel), logarithm of frequency by magnitude (bottom left panel), cumulative frequency over time (bottom middle panel) and latitude, longitude and magnitude against time (right panels) of 2158 earthquakes with magnitude greater than or equal to 3.0 that occurred between 2005-04-16 and 2013-11-01 in Italy and its vicinity (35°–48°N and 6.15°–19°E), extracted from the Italian Seismological Instrumental and Parametric Data Base (ISIDE) available at http://iside.rm.ingv.it/. The catalog is also available in the R package **etasFLP**.

The data set `italy.quakes` in the **ETAS** package contains data of the same Italian catalog. For compatibility with the **etasFLP** package, we use the argument `dist.unit = "km"` in the `catalog` function to create the earthquake catalog.

```
R> library("ETAS")
R> italy.cat <- catalog(italy.quakes, dist.unit = "km")
R> italy.cat

earthquake catalog:
  time begin 2005-04-16 12:27:54
  study period: 2005-04-16 12:27:54 to 2013-11-01 04:44:33 (T=3120.678 days)
geographical region:
    polygonal with vertices:
         lat     long
[1,] 34.87237  6.04186
```

```
[2,] 34.87237 19.11214
[3,] 48.09463 19.11214
[4,] 48.09463  6.04186
threshold magnitude: 3
number of events:
  total events 2158 : 2158 target events,  0 complementary events
  ( 0 events outside geographical region, 0 events outside study period)
```

```
R> plot(italy.cat)
```

Figure 9 shows the created catalog in the **ETAS** package. This catalog has no complementary events and $t_{\text{start}} = 0$ and $T = 3120.7$ days. We also use the same initial values for the model parameters as those used for the `etasclass` function in the **etasFLP** package and convert κ_0 to A.

```
R> k0 <- 0.005
R> c <- 0.005
R> p <- 1.01
R> D <- 1.1
R> q <- 1.52
R> A <- pi * k0 / ((p - 1) * c^(p - 1) * (q - 1) * D^(q - 1))
R> param0 <- c(mu = 1, A, c, alpha = 1.05, p, D, q, gamma = 0.6)
```

The catalog and initial parameters are passed to the `etas` function to fit the space-time ETAS model.

```
R> italy.fit <- etas(italy.cat, param0)
R> italy.fit
```

```
ETAS model: fitted using iterative stochastic declustering method
converged after 4 iterations: elapsed execution time 30.96 minutes

ML estimates of model parameters:
           beta     mu      A      c  alpha      p      D      q  gamma
Estimate 2.6333 1.0173 0.2115 0.0123 1.5596 1.1688 1.3185 1.8895 0.9123
StdErr   0.0032 0.0076 0.0202 0.0473 0.0102 0.0039 0.0412 0.0101 0.0227

declustering probabilities:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.0000  0.0004  0.8534  0.5350  0.9967  1.0000

log-likelihood:  -23394.52         AIC:  46805.03
```

The elapsed execution time on a Linux laptop (2.4 GHz Intel Core i7-4700MQ CPU) was 14 minutes for the **etasFLP** package and 31 minutes for the **ETAS** package. Using parallel computing in the **ETAS** package, the elapsed execution time reduces to 15 minutes with `nthreads = 4` and 10 minutes with `nthreads = 8`.

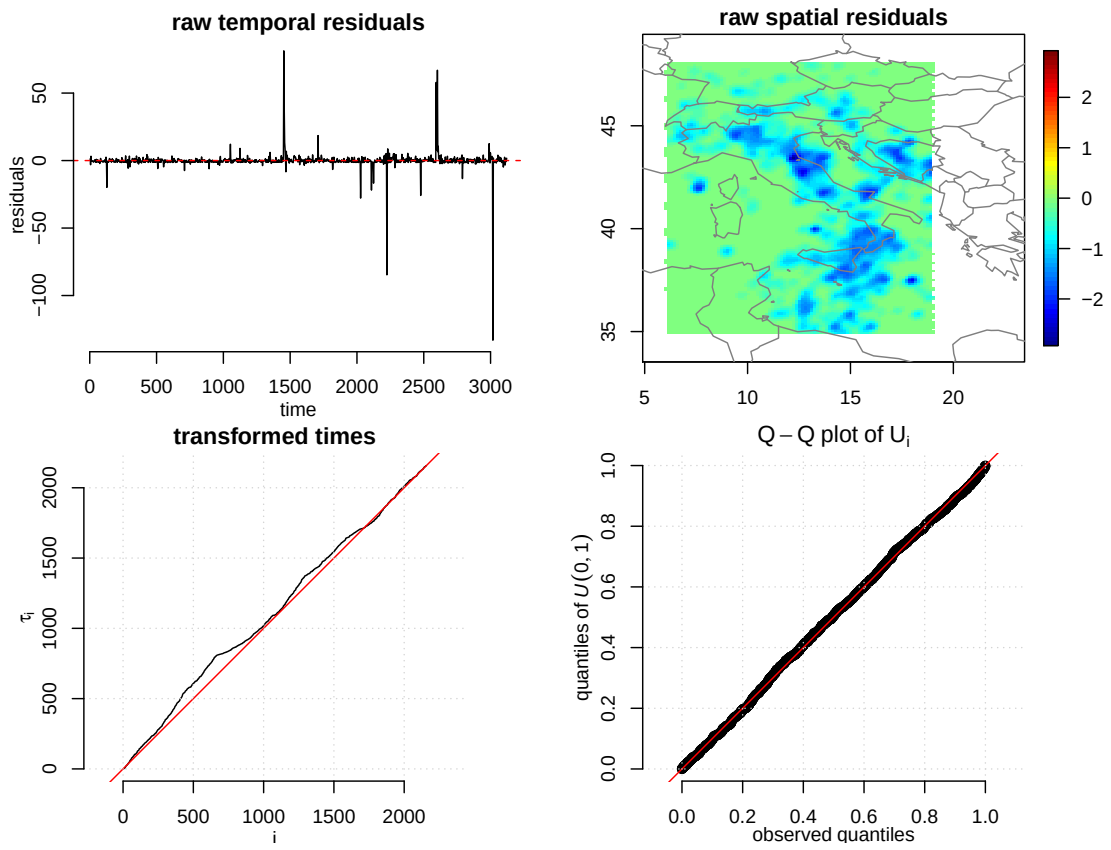

Figure 10: Diagnostic plots for the fitted ETAS model to the Italian catalog in the **etasFLP** package using the **ETAS** package: the temporal residuals (top left), smoothed spatial residuals (top right), transformed times τ_i against i (bottom left) and the Q-Q plot of U_i (bottom right).

Although the estimates obtained with both packages **etasFLP** and **ETAS** are to some extent similar (particularly the estimates of α, p and q), there are some differences between them (particularly the estimates of μ and D) due to the different bandwidth selection methods for kernel estimator of the background seismicity rate, different optimization algorithms and different transformations of `(long, lat)` to `(x, y)` coordinates. The AIC of the fitted ETAS model by the **ETAS** package is 46805.03 which is smaller than the one by **etasFLP** (48414.86). Figure 10 shows the diagnostic plots for the fitted model by the **ETAS** package.

```
R> italy.res <- resid.etas(italy.fit)
R> ks.test(italy.res$U, punif)

        One-sample Kolmogorov-Smirnov test

data:  italy.res$U
D = 0.020461, p-value = 0.3271
alternative hypothesis: two-sided

Warning message:
```

```
In ks.test(italy.res$U, punif) :
  ties should not be present for the Kolmogorov-Smirnov test
```

Except for a few extreme temporal residuals, the plot of the transformed times, the Q-Q plot of the U_j values and the Kolmogorov-Smirnov test for goodness-of-fit indicate a good temporal fit. No systematic pattern is evident in the spatial residuals, but they are not symmetric around zero and a few small clusters of extreme residuals on the image of spatial residuals may suggest room for improvements in the fitted PDF of spatial distribution of triggered events, $f_{D,\gamma,q}$, or in estimating the background seismicity rate.

The **etasFLP** package also computes and visualizes transformed times τ_i, background and triggered seismicity rates, total spatial intensity as well as different types of spatial residuals (Adelfio and Chiodi 2015b). The following commands compute U_j values and perform the Kolmogorov-Smirnov test for temporal goodness-of-fit of the fitted ETAS model to the Italian catalog using the **etasFLP** package.

```
R> italy.res.flp <- plot(italy.fit.flp)
R> tau.flp <- italy.res.flp$t.trasf
R> U.flp <- 1 - exp(-diff(tau.flp))
R> ks.test(U.flp, punif)

        One-sample Kolmogorov-Smirnov test

data:  U.flp
D = 0.018422, p-value = 0.4569
alternative hypothesis: two-sided
```

The p value indicates an adequate temporal fit to the catalog.

7.2. Comparison with the Fortran program

The original Fortran program for fitting the space-time ETAS model to earthquake data is available at http://bemlar.ism.ac.jp/zhuang/software/sd.tar.gz (accessed on November 3, 2018). The main program `etas8p` calls several subroutines for integral approximation by radial partitioning of the geographical region (`frint` and `polyint`), computing the likelihood function (`xlamb` and `xint`), declustering computations (`bkgdcalc`) and Davidon-Fletcher-Powell optimization (`linear` and `davidn`). The program estimates the model parameters and background seismicity rate using the stochastic declustering approach as in Algorithm 2, except that it does not check the convergence criteria in the repeat loop and uses a fixed number of iterations (eleven times). Moreover, in each iteration, the Davidon-Fletcher-Powell optimization described in Algorithm 1 (subroutine `davidn`) starts with the identity matrix, $H_0 = I$, as the initial guess for the inverse of Hessian matrix while in the **ETAS** package, the estimate of inverse Hessian matrix in each iteration is used as initial guess for the next iteration. These differences generally result in a faster convergence of **ETAS**. After convergence, `etas8p` calls two subroutines (`outprob` and `outrates`) to report declustering probabilities p_j and the values of the estimated background seismicity rate on a rectangular grid.

The `etas8p` program uses parallel computing, implemented by the message passing interface (MPI) standard (The MPI Forum 2015). The program's tarball `sd.tar.gz` contains a sample data set (`alljpM45.etas` file) of 13724 shallow (depth < 100 km) earthquakes that occurred between 1926-01-08 and 2007-12-29 in Japan and its vicinity (27°–45°N and 128°–145°E) with magnitudes greater than or equal to 4.5. As instructed on http://bemlar.ism.ac.jp/zhuang/software.html and assuming GNU Fortran (`gfortran`) is your compiler, the program can be compiled by

```
$ mpif77 *.f -o etas8p
```

to produce the executable file `etas8p`. The file `jap.in` contains all program inputs including the name of the data file, the beginning and length of the study period, the magnitude threshold, the coordinates of vertices of the boundary of the target geographical region, the initial values for the model parameters, the span and the dimension of the rectangular grid for estimation of the background seismicity rate and values of n_p and $h_{\min}$ for computing the bandwidth h_j for each event according to (9). The program can be run on a computer with 8 threads processor by

```
$ export GFORTRAN_UNBUFFERED_ALL='y'   # unbuffered IO with gfortran
$ mpiexec -n 8 ./etas8p < jap.in
```

The execution time takes 422.3 minutes on the same Linux laptop as before (2.4 GHz Intel Core i7-4700MQ CPU) with GNU Compiler Collection, Version 5.4.0 (https://gcc.gnu.org/), and MPICH, Version 3.2 (https://www.mpich.org/). Here is the final output of the program.

```
 - log likelihood =  0.153109518614449D+05     aic =   30637.9
 -----  x  -----
   0.74203D+00  0.40714D+00  0.17209D+00  0.12876D+01  0.10740D+01
   0.42823D-01  0.13966D+01  0.10330D+01
 ***  gradient  ***
  -0.71766D-03  0.16731D-05  0.36575D-03 -0.53808D-03  0.14690D-03
  -0.11811D-02  0.16704D-03  0.13040D-03

 mle =  0.55061E+00 0.16576E+00 0.29615E-01 0.16579E+01 0.11534E+01
        0.18338E-02 0.19505E+01 0.10670E+01
```

The program, in addition, returns estimates of the background seismicity rate, the total spatial rate and the conditional intensity function at $t = t_{\text{start}} + T$ on a 401×401 grid over 31.8°–44.3°N and 134°–145°E (`rates.dat` file) and the background seismicity rate, declustering probability and bandwidth for each target event (file `probs.dat`). Figure 11 shows the estimated background seismicity rate, imported to and plotted in R.

For comparison, the data set of the Japanese earthquakes is included in **ETAS** as `japan.quakes`. To compare the results obtained by the `etas8p` program with the **ETAS** package, we create an earthquake catalog using `japan.quakes` and set the same polygonal geographical region and the same study period ($t_{\text{start}} = 10000$ and $T = 13376$ days) as in the `jap.in` file.

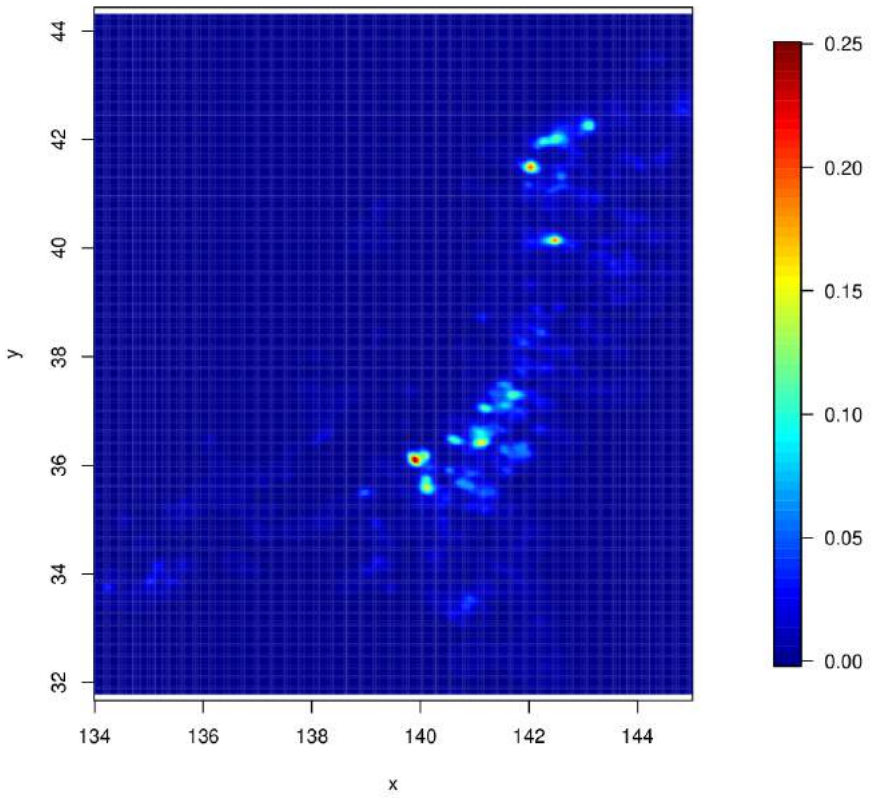

Figure 11: Estimate of the background seismicity rate for the Japanese catalog using the `etas8p` program.

```
R> jpoly <- list(long = c(134.0, 137.9, 143.1, 144.9, 147.8, 137.8, 137.4,
+    135.1, 130.6), lat = c(31.9, 33.0, 33.2, 35.2, 41.3, 44.2, 40.2, 38.0,
+    35.4))
R> japan.cat <- catalog(japan.quakes, study.start = "1953-05-26",
+    study.end = "1990-01-08", region.poly = jpoly, mag.threshold = 4.5)
R> japan.cat

earthquake catalog:
  time begin 1926-01-08
  study period: 1953-05-26  to  1990-01-08 (T = 13376 days)
geographical region:
    polygonal with vertices:
       lat  long
 [1,] 31.9 134.0
 [2,] 33.0 137.9
 [3,] 33.2 143.1
 [4,] 35.2 144.9
 [5,] 41.3 147.8
 [6,] 44.2 137.8
 [7,] 40.2 137.4
```

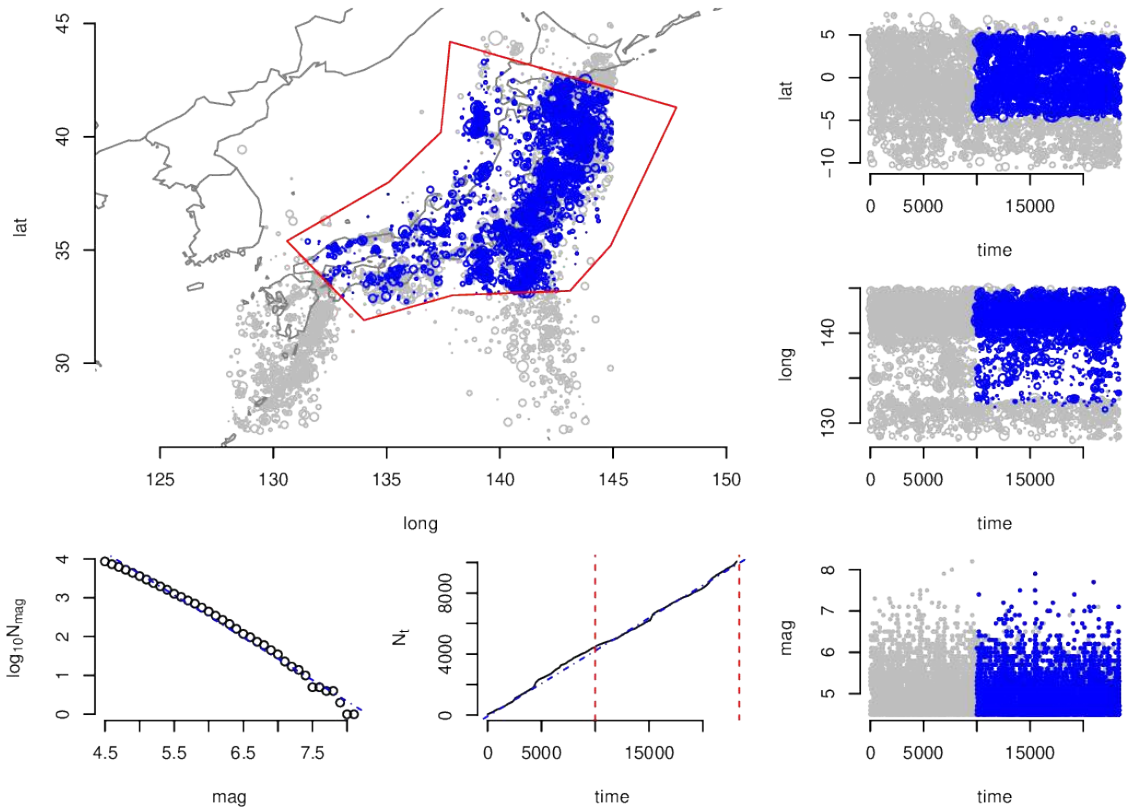

Figure 12: Location of epicenters (top left panel), logarithm of frequency by magnitude (bottom left panel), cumulative frequency over time (bottom middle panel) and latitude, longitude and magnitude against time (right panels) of 13724 earthquakes with magnitude greater than or equal to 4.5 that occurred between 1926-01-08 and 2007-12-29 in Japan and its vicinity (27°–45°N and 128°–145°E), accompanying the Fortran program by Jiancang Zhuang (Zhuang *et al.* 2017).

```
 [8,] 38.0 135.1
 [9,] 35.4 130.6
threshold magnitude: 4.5
number of events:
  total events 10072 : 4656 target events, 5416 complementary events
  (1022 events outside geographical region, 4394 events outside study period)
```

The created catalog is shown in Figure 12. We also use the same initial values for the model parameters as those specified in the `jap.in` file and call the `etas` function to fit the model.

```
R>  param0 <- c(0.592844590, 0.204288231, 0.022692883, 1.495169224,
+     1.109752319, 0.001175925, 1.860044210, 1.041549634)
R> japan.fit <- etas(japan.cat, param0)
R> japan.fit

ETAS model: fitted using iterative stochastic declustering method
converged after 4 iterations: elapsed execution time 185.26 minutes
```

```
ML estimates of model parameters:
           beta     mu      A      c  alpha      p      D      q  gamma
Estimate 1.9734 0.5505 0.1658 0.0296 1.6579 1.1534 0.0018 1.9507 1.0670
StdErr   0.0004 0.0062 0.0162 0.0316 0.0060 0.0027 0.0346 0.0089 0.0139

declustering probabilities:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.0000  0.0126  0.7565  0.5452  0.9330  1.0000

log-likelihood:  -15310.96         AIC:  30637.91
```

The execution time takes 185.3 minutes on the same Linux laptop and using parallel computing, it reduces to 104.2 and 68.1 minutes with `nthreads = 4` and `nthreads = 8`, respectively. Thus, with parallel computing and 8 threads and for the considered Japanese catalog, package **ETAS** is at least 5 times faster than the Fortran program `etas8p`.

The maximum absolute difference between the estimates of the model parameters obtained with the `etas8p` program and **ETAS** package is smaller than 10^{-3}. The absolute difference between the values of the log-likelihood function and AIC are less than 10^{-2}. The estimate of the background seismicity rate, obtained with the command (elapsed execution time 5.7 minutes)

```
R> rates(japan.fit, long.range = c(134, 145), lat.range = c(31.8, 44.3),
+    dimyx = c(401, 401))
```

and shown in the top left panel of Figure 13, is also in total agreement with Figure 11. Thus, as expected, results obtained with the `etas8p` program and **ETAS** are practically the same.

For completeness, we perform residual analysis of the fitted ETAS model.

```
R> japan.res <- resid.etas(japan.fit)
R> summary(japan.res$tres)

    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-78.7790  -1.7704  -0.3662   0.2415   1.1445  94.7057

R> summary(na.omit(japan.res$sres$z))

     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
-0.525076 -0.014207 -0.001284  0.028893  0.057943  0.593824
```

The resulting diagnostic plots are shown in Figure 14. Plot of transformed times τ_j and the Q-Q plot of U_j as well as the Kolmogorov-Smirnov test for goodness-of-fit indicate that the temporal variations in the data are sufficiently explained by the fitted model.

```
R> ks.test(japan.res$U, punif)
```

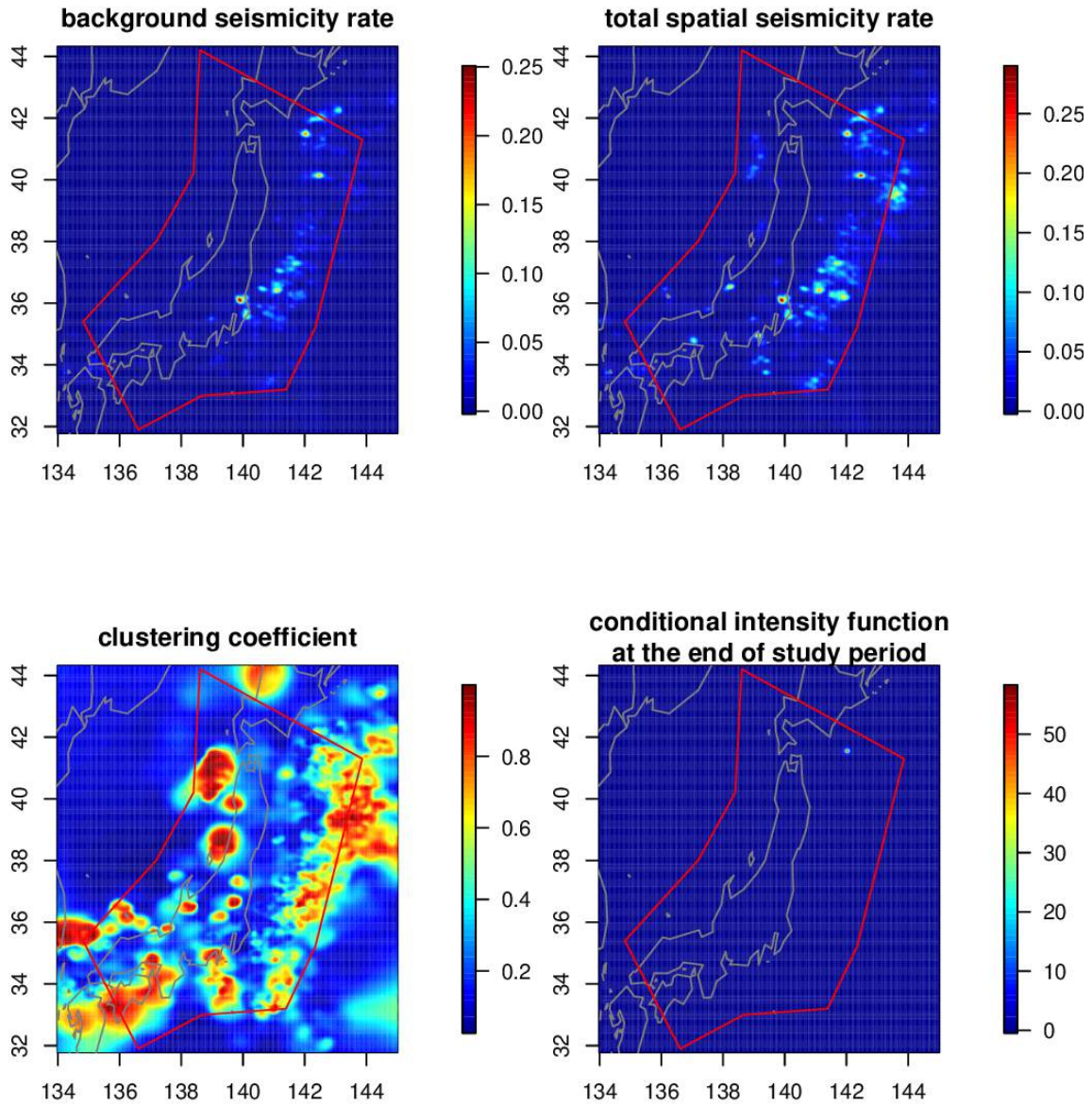

Figure 13: Plots of estimates of the background seismicity rate, total spatial intensity, clustering (triggering) coefficient and conditional intensity at $t = t_{\text{start}} + T$ for the fitted ETAS model to the Japanese catalog accompanying the Fortran program `etas8p` using the **ETAS** package.

```
        One-sample Kolmogorov-Smirnov test

data:  japan.res$U
D = 0.0093191, p-value = 0.8136
alternative hypothesis: two-sided
```

The raw spatial residuals do not show any obvious trend and are symmetric around zero, which suggests an overall good fit.

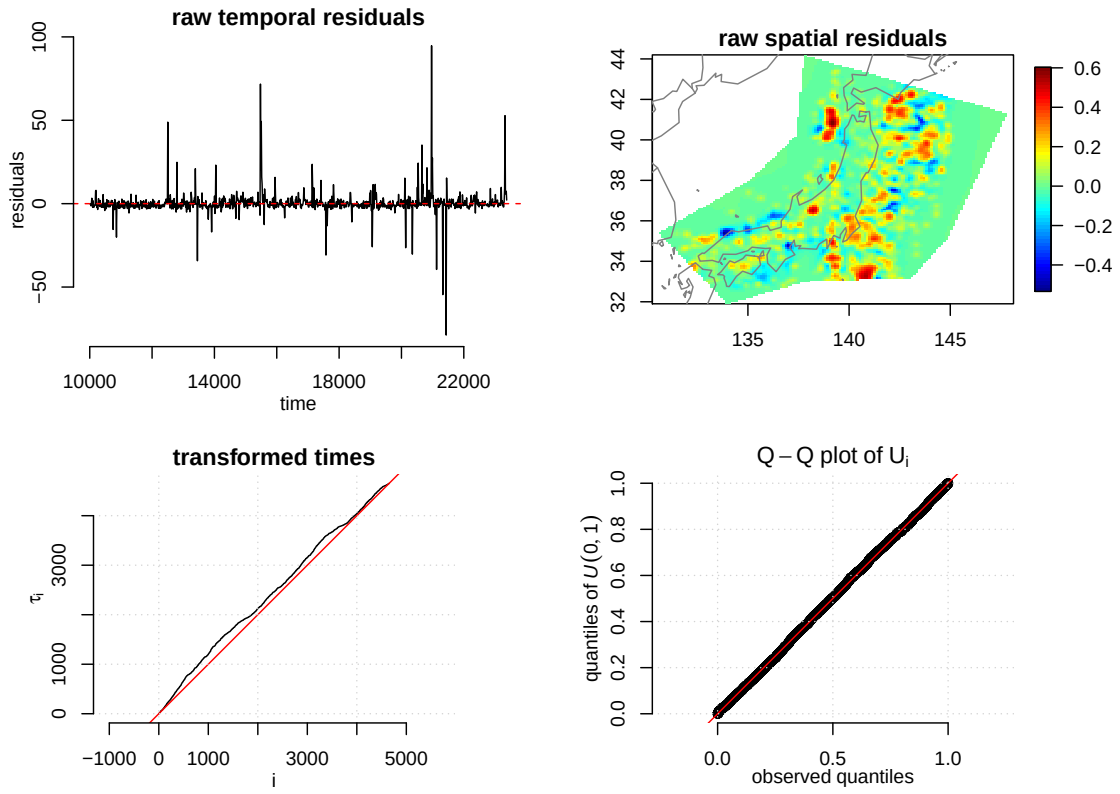

Figure 14: Diagnostic plots for the fitted ETAS model to the sample Japanese catalog accompanying the Fortran program `etas8p` using the **ETAS** package: the temporal residuals (top left), smoothed spatial residuals (top right), transformed times τ_i against i (bottom left) and the Q-Q plot of U_i (bottom right).

8. Conclusion

This paper introduces the R package **ETAS** that fits the epidemic type aftershock sequence (ETAS) model to an earthquake catalog. Given data on an earthquake catalog, the package provides two classes and several functions to facilitate data preparation and model fitting and offers some simple residual analysis and diagnostic checks for model assessment.

The **ETAS** package is based on a Fortran program by Jiancang Zhuang and colleagues. The results obtained with the **ETAS** package and the Fortran program are practically the same, but the **ETAS** package achieves a faster execution time by using the estimated inverse Hessian matrix at each iteration as an initial guess for the next iteration, and using modified convergence criteria. The R package **etasFLP** is an alternative to **ETAS**. These packages produce different, although comparable, results because they use different bandwidth selection methods for the kernel estimator of the background seismicity rate, different optimization algorithms and different flat map transformations. The **etasFLP** package is faster than **ETAS** with serial computing, but for large catalogs (more than 2000 events) the **ETAS** package with parallel computing can be as fast as **etasFLP**.

As of version 0.3, the **ETAS** package uses parallel computing with OpenMP to allow for

faster computations, particularly in case of large earthquake catalogs. The package can be extended in several ways. For example, different modifications of the space-time ETAS model have been recently proposed (Ogata 2011; Harte 2014; Nicolis, Chiodi, and Adelfio 2015). Moreover, the ETAS model has been generalized to introduce nonstationarity into the ETAS model (Kumazawa and Ogata 2014). These modifications and generalizations could be implemented and integrated into the package. Simulation of the space-time ETAS model and more comprehensive residual analyses and model diagnostics could also be included in the package.

Acknowledgments

The author is grateful to three anonymous reviewers for their helpful comments that led to a significant improvement of this paper.

References

Adelfio G, Chiodi M (2013). "Mixed Estimation Technique in Semi-Parametric Space-Time Point Processes for Earthquake Description." In *Proceedings of the 28th International Workshop on Statistical Modelling*, volume 1, pp. 65–70.

Adelfio G, Chiodi M (2015a). "Alternated Estimation in Semi-Parametric Space-Time Branching-Type Point Processes with Application to Seismic Catalogs." *Stochastic Environmental Research and Risk Assessment*, **29**(2), 443–450. doi:10.1007/s00477-014-0873-8.

Adelfio G, Chiodi M (2015b). "FLP Estimation of Semi-Parametric Models for Space-Time Point Processes and Diagnostic Tools." *Spatial Statistics*, **14**(Part B), 119–132. doi:10.1016/j.spasta.2015.06.004.

Baddeley A, Rubak E, Turner R (2015). *Spatial Point Patterns: Methodology and Applications with R*. Chapman & Hall/CRC, London.

Baddeley A, Turner R (2000). "Practical Maximum Pseudolikelihood for Spatial Point Patterns." *Australian & New Zealand Journal of Statistics*, **42**(3), 283–322. doi:10.1111/1467-842x.00128.

Baddeley A, Turner R (2005). "**spatstat**: An R Package for Analyzing Spatial Point Patterns." *Journal of Statistical Software*, **12**(6), 1–42. doi:10.18637/jss.v012.i06.

Baddeley A, Turner R, Møller J, Hazelton M (2005). "Residual Analysis for Spatial Point Processes." *Journal of the Royal Statistical Society B*, **67**(5), 617–666. doi:10.1111/j.1467-9868.2005.00519.x.

Chiodi M, Adelfio G (2011). "Forward Likelihood-Based Predictive Approach for Space-Time Point Processes." *Environmetrics*, **22**(6), 749–757. doi:10.1002/env.1121.

Chiodi M, Adelfio G (2017). "Mixed Non-Parametric and Parametric Estimation Techniques in R Package **etasFLP** for Earthquakes' Description." *Journal of Statistical Software*, **76**(3), 1–29. doi:10.18637/jss.v076.i03.

Console R, Jackson DD, Kagan YY (2010). "Using the ETAS Model for Catalog Declustering and Seismic Background Assessment." *Pure and Applied Geophysics*, **167**(6–7), 819–830. doi:10.1007/s00024-010-0065-5.

Daley DJ, Vere-Jones D (2003). *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods*. 2nd edition. Springer-Verlag, New York. doi:10.1007/b97277.

Di Giacomo D, Harris J, Villaseñor A, Storchak DA, Engdahl ER, Lee WHK (2015). "ISC-GEM: Global Instrumental Earthquake Catalogue (1900-2009), I. Data Collection from Early Instrumental Seismological Bulletins." *Physics of the Earth and Planetary Interiors*, **239**, 14–24. doi:10.1016/j.pepi.2014.06.003.

Eddelbuettel D (2013). *Seamless R and C++ Integration with **Rcpp***. Springer-Verlag, New York. doi:10.1007/978-1-4614-6868-4.

Eddelbuettel D, François R (2011). "**Rcpp**: Seamless R and C++ Integration." *Journal of Statistical Software*, **40**(8), 1–18. doi:10.18637/jss.v040.i08.

Fletcher R, Powell MJD (1963). "A Rapidly Convergent Descent Method for Minimization." *The Computer Journal*, **6**(2), 163–168.

Harte D (2010). "**PtProcess**: An R Package for Modelling Marked Point Processes Indexed by Time." *Journal of Statistical Software*, **35**(8), 1–32. doi:10.18637/jss.v035.i08.

Harte DS (2013). "Bias in Fitting the ETAS Model: a Case Study Based on New Zealand Seismicity." *Geophysical Journal International*, **192**(1), 390–412. doi:10.1093/gji/ggs026.

Harte DS (2014). "An ETAS Model with Varying Productivity Rates." *Geophysical Journal International*, **198**(1), 270–284. doi:10.1093/gji/ggu129.

Helmstetter A, Sornette D (2003). "Importance of Direct and Indirect Triggered Seismicity in the ETAS Model of Seismicity." *Geophysical Research Letters*, **30**(11), 1–4. doi:10.1029/2003gl017670.

Jalilian A, Zhuang J (2019). ***ETAS**: Modeling Earthquake Data Using ETAS Model*. R package version 0.4.6, URL https://CRAN.R-project.org/package=ETAS.

Kagan YY (1991). "Likelihood Analysis of Earthquake Catalogues." *Geophysical Journal International*, **106**(1), 135–148. doi:10.1111/j.1365-246x.1991.tb04607.x.

Kasahara A, Yagi Y, Enescu B (2016). "**etas_solve**: A Robust Program to Estimate the ETAS Model Parameters." *Seismological Research Letters*, **87**(5), 1143–1149. doi:10.1785/0220150240.

Kumazawa T, Ogata Y (2014). "Nonstationary ETAS Models for Nonstandard Earthquakes." *The Annals of Applied Statistics*, **8**(3), 1825–1852. doi:10.1214/14-aoas759.

Lombardi AM, Marzocchi W (2010). "The ETAS Model for Daily Forecasting of Italian Seismicity in the CSEP Experiment." *The Annals of Geophysics*, **53**(3), 155–164. doi:10.4401/ag-4848.

Nicolis O, Chiodi M, Adelfio G (2015). "Windowed ETAS Models with Application to the Chilean Seismic Catalogs." *Spatial Statistics*, **14**(Part B), 151–165. doi:10.1016/j.spasta.2015.05.006.

Ogata Y (1988). "Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes." *Journal of the American Statistical Association*, **83**(401), 9–27. doi:10.2307/2288914.

Ogata Y (1998). "Space-Time Point-Process Models for Earthquake Occurrences." *The Annals of the Institute of Statistical Mathematics*, **50**(2), 379–402. doi:10.1023/a:1003403601725.

Ogata Y (1999). "Seismicity Analysis through Point-Process Modeling: A Review." *Pure and Applied Geophysics*, **155**(2–4), 471–507. doi:10.1007/s000240050275.

Ogata Y (2006a). "Monitoring of Anomaly in the Aftershock Sequence of the 2005 Earthquake of M7.0 Off Coast of the Western Fukuoka, Japan, by the ETAS Model." *Geophysical Research Letters*, **33**(1), 1–4. doi:10.1029/2005gl024405.

Ogata Y (2006b). "Satistical Analysis of Seismicity – Updated Version (**SASeis2006**)." *Computer Science Monographs*, **33**, 1–28.

Ogata Y (2011). "Significant Improvements of the Space-Time ETAS Model for Forecasting of Accurate Baseline Seismicity." *Earth Planets and Space*, **63**(3), 217–229. doi:10.5047/eps.2010.09.001.

OpenMP Architecture Review Board (2013). "**OpenMP** Application Program Interface Version 4.0." URL http://www.openmp.org/wp-content/uploads/OpenMP4.0.0.pdf.

Rathbun SL (1996). "Asymptotic Properties of the Maximum Likelihood Estimator for Spatio-Temporal Point Processes." *Journal of Statistical Planning and Inference*, **51**(1), 55–74. doi:10.1016/0378-3758(95)00070-4.

R Core Team (2018a). *R Data Import/Export*. R Foundation for Statistical Computing. URL https://CRAN.R-project.org/doc/manuals/r-release/R-data.html.

R Core Team (2018b). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

R Core Team (2018c). *Writing R Extensions*. R Foundation for Statistical Computing. URL https://CRAN.R-project.org/doc/manuals/r-release/R-exts.html.

Schoenberg FP (2005). "Consistent Parametric Estimation of the Intensity of a Spatial-Temporal Point Process." *Journal of Statistical Planning and Inference*, **128**(1), 79–93. doi:10.1016/j.jspi.2003.09.027.

Schoenberg FP (2013). "Facilitated Estimation of ETAS." *Bulletin of the Seismological Society of America*, **103**(1), 601–605. doi:10.1785/0120120146.

Snyder JP (1997). *Flattening the Earth: Two Thousand Years of Map Projections*. University of Chicago Press.

Talebian M, Jackson J (2004). "A Reappraisal of Earthquake Focal Mechanisms and Active Shortening in the Zagros Mountains of Iran." *Geophysical Journal International*, **156**(3), 506–526. doi:10.1111/j.1365-246x.2004.02092.x.

The Institute of Statistical Mathematics (2016). ***SAPP**: Statistical Analysis of Point Processes*. R package version 1.0.7, URL https://CRAN.R-project.org/package=SAPP.

The MPI Forum (2015). ***MPI**: A Message-Passing Interface Standard*. URL http://mpi-forum.org/docs/mpi-3.1/mpi31-report.pdf.

Veen A, Schoenberg FP (2008). "Estimation of Space-Time Branching Process Models in Seismology Using an EM-Type Algorithm." *Journal of the American Statistical Association*, **103**(482), 614–624. doi:10.1198/016214508000000148.

Vere-Jones D (1970). "Stochastic Models for Earthquake Occurrence." *Journal of the Royal Statistical Society B*, **32**(1), 1–62.

Vere-Jones D (2003). "Probabilities and Information Gain for Earthquake Forecasting." In DK Chowdhury, JC De Bremaecker, K Lashgari, E Nyland, R Odom, M Sen, MM Vishik, VI Keilis-Borok, AL Levshin, GM Molchan, BM Naimark (eds.), *Selected Papers From Volume 30 of Vychislitel'naya Seysmologiya.*, pp. 104–114. American Geophysical Union, Washington DC. doi:10.1029/cs005p0104.

Zhuang J (2006). "Second-Order Residual Analysis of Spatiotemporal Point Processes and Applications in Model Evaluation." *Journal of the Royal Statistical Society B*, **68**(4), 635–653. doi:10.1111/j.1467-9868.2006.00559.x.

Zhuang J (2011). "Next-Day Earthquake Forecasts for the Japan Region Generated by the ETAS Model." *Earth, Planets and Space*, **63**(3), 207–216. doi:10.5047/eps.2010.12.010.

Zhuang J (2012). "Long-Term Earthquake Forecasts Based on the Epidemic-Type Aftershock Sequence (ETAS) Model for Short-Term Clustering." *Research in Geophysics*, **2**(1), 52–57. doi:10.4081/rg.2012.e8.

Zhuang J, Ogata Y, Vere-Jones D (2002). "Stochastic Declustering of Space-Time Earthquake Occurrences." *Journal of the American Statistical Association*, **97**(458), 369–380. doi:10.1198/016214502760046925.

Zhuang J, Ogata Y, Vere-Jones D (2006). "Diagnostic Analysis of Space-Time Branching Processes for Earthquakes." In A Baddeley, P Gregori, J Mateu, R Stoica, D Stoyan (eds.), *Case Studies in Spatial Point Process Modeling*, pp. 275–292. Springer-Verlag. doi:10.1007/0-387-31144-0_15.

Zhuang J, Ogata Y, *et al.* (2017). `etas8p`*: Code for the Space-Time ETAS Model and Stochastic Declustering*. Fortran implementation, URL http://bemlar.ism.ac.jp/zhuang/software.html.

Affiliation:

Abdollah Jalilian
Department of Statistics
Faculty of Sciences
Razi University
Baq-e Abrisham, Kermanshah, 6714415111, Iran
E-mail: jalilian@razi.ac.ir

Journal of Statistical Software http://www.jstatsoft.org/
published by the Foundation for Open Access Statistics http://www.foastat.org/
January 2019, Volume 88, Code Snippet 1 *Submitted:* 2016-02-07
doi:10.18637/jss.v088.c01 *Accepted:* 2017-12-21