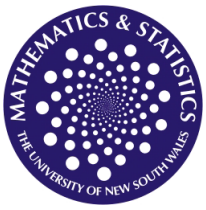

Semester 2, 2013

UNSW
THE UNIVERSITY OF NEW SOUTH WALES

A Location Transformed Method on Space-Time ETAS Model

The University of New South Wales Department of Statistics

Program:	Master of Statistics
Supervisor:	Dr. Feng, Chen
Name:	ZHOUSHU, HUANG
Student ID:	z3368430
E-MAIL:	z3368430@zmail.unsw.edu.au

Abstract

The Epidemic Type Aftershock Sequence (ETAS) model based on marked self-exciting point process (SEPP) is widely used in seismic activity analysis. Since it proposed by Ogata in 1988, it has an active area of research in seismology; also, the improvement of this model is still in an area of active research recently, such as the space-time ETAS model has been proposed. The traditional space-time ETAS model consider the rotation angle of the location ellipse distribution as a constant number (controlled by a estimated ρ in the space-time ETAS model). In this paper, after reviewing the ETAS model, we proposed a location transformed method which can solve this problem on the space-time ETAS model. Based on this method, we practiced in the real world earthquakes data, by changing the coordinate system of earthquake location and creating two new vectors called *Distance* (d) and *Length* (l) instead of the Latitude-Longitude coordinate to eliminate this ρ. Finally we analyzed the model fitted by the data set of the Sichuan earthquake in 2008; we also verified this improvement by the 2010 Chile earthquake data.

Keywords: Chile earthquake; clustering event; conditional intensity; Declustering event; ETAS model; Hawkes process; point process; self-exciting; seismic activity; seismic belt; Sichuan earthquake; spatial-temporal point;

Acknowledgments

October 31, 2013

I would like to thank all people who supported me to finish this paper, made an unforgettable experience for me.

First of all, I would like to express my deepest and sincere of gratitude to my supervisor Dr. Feng Chen for the continuous support of my project, for his patience, motivation, enthusiasm, and immense knowledge. Without his guidance, I could never have been able to accomplish this project.

Besides my supervisor, I would like to thank my parents, for their encouragement and support during my whole master program, and to make me study in UNSW possibly.

I am thankful to Dr. Donna Salopek for her guidance on the process of my project, for the absolution support of my oral presentation. Special thanks to professors William Dunsmuir and Spiridon Penev, for their insightful comments. Finally, I also acknowledge to all staff at Department of Statistics in UNSW, for the education during my master program.

Contents

1. Introduction

Multidimensional space-time point (or time series point with spatial information) analysis is widely used in many current areas of research such as finance, meteorology and seismology. The seismic activity events record (including the earthquake information like occurrence time, location and magnitude) is such a typical space-time point pattern. For this record, people have invented many mathematical tools which can simulate the record so as to understand the seismic activity. The Epidemic Type Aftershock-Sequences (ETAS) model which is based on the self-exciting point process is such a complex tool concentrating on a series of earthquake events. The idea of this model is to consider the occurrence of every occurred earthquake dividing into two factors: the clustering part and the declustering part. Based on this consideration, ETAS model utilizes the locations and time gaps in the earthquake sequence as arguments, to find the proportion of these two parts. With a large number of seismic activity events recording by GIS, this model is achievable by high performance computers. In this paper, we analyze the 2008 Sichuan earthquake ($M_{\mathrm{w}}7.9$) by ETAS model, and introduce the location transformed method which can improve the performance of this model; we also verify this improvement by analyzing the 2010 Chile earthquake data.

1.1 Self-exciting Point Process and Epidemic Type Aftershock-Sequences (ETAS) Model

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space. Let $\{t_i, i = 1, 2, ..., \}$ be a point process on the time interval $[0, T]$ $(0 \leqslant t_1 < t_2 < ... \leqslant T)$. Let $N(t) = \max\{i, t_i \leqslant t\}$ be the cumulative number of events during period $[0, t]$, and $\mathcal{H}_t$ denote the history of $N(t)$. A corresponding intensity $\lambda(t)$ of $N(t)$ is defined as an $\mathcal{H}_t-$predictable process, such that $M(t) = N(t) - \int_0^t \lambda(t)\mathrm{d}t$ is a zero mean $\mathcal{H}_t$-martingale (under P) which is locally square integrable. Based on these, the self-exciting intensity is an intensity with the form:

$$\lambda(t|\mathcal{H}_t) = \mathrm{E}(dN(t)|\mathcal{H}_t)/\mathrm{d}t = \mu + \sum_{i=1}^{N(t)} g(t - t_i) \tag{1}$$

This self-exciting form is first introduced by Hawkes in 1971 (Hawkes, 1971), where μ is a constant number refers to the baseline intensity and $g(t - t_i)$ is the self-exciting function measuring the temporal effects from the previous events. With a self-exciting property, the self-exciting process and its extensions are widely applied in many fields. Some are still in current research and seismicity is one typical topic in these fields.

By the self-exciting property of this process, the epidemic type aftershock sequence (ETAS) model has been proposed by Ogata (1988), based on the Gutenberg-Richter's law and Omori formula. In

this model, the whole seismic activity events is considered as an earthquake sequence in a certain region; each event in this sequence is followed by another earthquake, as a time series marked by their magnitude. The magnitude of an earthquake is introduced as mark in the SEPP, with such a conditional intensity

$$\lambda(t|\mathcal{H}_t) = \lambda_c(t) + \lambda_a(t) = \mu + \sum_{t_i<t} g(t - t_i; M_i) \tag{2}$$

where λ_c is the background pattern (mainshock), λ_a refers to the clustering seismicity (aftershock) and M means the magnitude (see Ogata, 1988), for any $t \leqslant T$. In this model, the self-exciting intensity is considered as the seismic rate; the corresponding λ_c refers to the independent rate (which is independent of other historical earthquakes); and the clustering part λ_a, is the total triggers rate by history events.

One important factor in the clustering part is given by Gutenberg-Richter's law. Based on this Gutenberg-Richter's law (B. Gutenberg and C. F. Richter, 1956), the log of frequency (N) of the earthquakes with magnitude not less than a baseline magnitude M_0 is proposed as

$$\log N = a - bM_0 \tag{3}$$

with constant a, b. Then, Ogata suggests for the magnitude $M \geqslant M_0$ earthquake, the self-exciting function (or trigger function) of the ETAS model with form

$$g_\theta(t; M) = A \exp(\alpha(M - M_0)) \frac{p-1}{c}(1 + \frac{t}{c})^{-p} \tag{4}$$

in (2), where $\theta = (\mu, A, \alpha, c, p)$ is the parameter. In this model, g could be regarded as the measure of a clustering event triggered by the another previous event; specifically, $A \exp(\alpha(M - M_0))$ refers to the predicted number of triggered earthquakes from an event of magnitude M; $\frac{p-1}{c}(1 + \frac{t}{c})^{-p}$ is the time density function (from the modified Omori formula) of the triggered earthquakes.

1.2 The 2008 Sichuan Earthquake

The 2008 Sichuan earthquake (the Great Sichuan Earthquake) occurred at 06:28:01 UTC on May 12, 2008 at 30.002°N, 103.322°E. The record shows the earthquake with magnitude 7.9 M_{w} centered at depth 19 km on the Sichuan seismic belt. Jiang and Zhuang (2010) point out there are 41 earthquakes with magnitude larger then 7.0 during 1500 years in this region. Smyth and Mori (2010) proposed a prediction of large earthquakes following the mainshock (7.9 M_{w}) of this earthquake by ETAS model; Jiang and Wu (2011) checked Pattern Informatics forecast by ETAS model on the Sichuan-Yunnan region; Jiang and Zhuang (2010) apply space-time ETAS model to evaluate background seismicity and potential source zones of strong earthquakes in this region.

1.3 The 2010 Chile Earthquake

United States Geological Survey reports the February 27, 2010 M_{w} 8.8 Chile earthquake occurred off the coast of central Chile (38.952°S, 72.931°W), at 03:34 local time (06:34 UTC). United Nations Environment Programme (UNEP) year book 2011 estimates the losses to economy of Chile from this earthquake about 15 to 30 billion U.S. Dollar. This seismic activity region is between the Nazca plate and South American tectonic plate. USGS points the earthquake caused by subduction of the Nazca plate beneath the South American Tectonic Plate. Since 1900, there are six earthquakes of magnitude greater than 8.0 in this region, M_{w} 8.2 1906, M_{w} 8.5 1922, M_{w} 8.2 1943, M_{w} 8.2 1960, M_{w} 9.5 1960 and this M_{w} 8.8 2010 respectively.

2. Data Processing for the 2008 Sichuan Earthquake

2.1 Data Preparation

In this section we deal with the 2008 Sichuan earthquake data. We search data from U.S. Geological Survey (USGS) between latitude (30 to 33° N) and longitude (103 ~ 106° E), covering a part of Sichuan seismic belt from 12/05/2008 to 12/05/2012 (four-year period after the 2008 Sichuan earthquake, shown in the Figure 1).

There are 1042 events in the data set. The first earthquake in the sequence is the mainshock, $M_{\mathrm{w}} = 7.9$. Figure 2 shows offspring events mainly occurred in a short period after this mainshock, not only the occurrence, but also the magnitude of these offspring shocks affected by this mainshock. Hence it is reasonable to consider the mainshock as a start point, and not including forward events before the mainshock in this case. We will fit ETAS model by this earthquake sequence.

In practice, the value of baseline magnitude M_0 choice should be careful. In the ETAS model (4), M_0 is from the Gutenberg-Richter law (3). By this law, the relationship of frequency of seismic activity with magnitude is proposed as (3). Utsu (1965) provides a maximum likelihood estimate of $b - value$:

$$b = \frac{1}{(\bar{M} - M_0 + 0.05)\log 10} \tag{5}$$

$\bar{M}$ is the mean of magnitude which magnitude in an earthquake is greater than or equivalent to M_0 in the whole data set. In this case we choose three values of M_0(4.5, 4.3 and 4.0 respectively) by the Gutenberg-Richter law (left graphs of Figure 3, 4, 5).

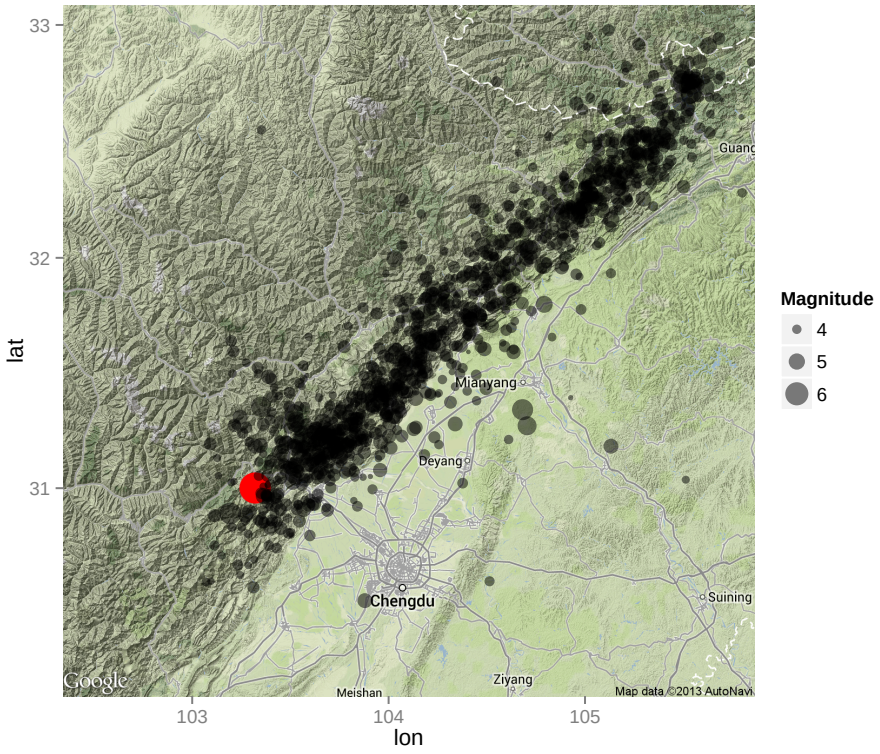

Figure 1: This figure shows four-year records of seismic activities in this region (1042 events), where the red point refers to the mainshock on 12/05/2008 (7.9 M_{w}). Obviously these events are occurred on a seismic belt. The coordinate system is original location on the earth ('lon' means longitude, positive for eastern; lat is latitude, positive for northern). The map information is from Google Earth.

Based on these M_0, we fit three ETAS models by maximum likelihood method. As a function of parameter θ, the likelihood function $L(\theta)$ for an intensity $\lambda(t;\theta)$ of a counting process $N(t)$ on the whole time interval $[0,T]$ is given by

$$L(\theta) = \prod_{t\in[0,T]} \left\{ \lambda(t;\theta)^{\mathrm{d}N(t)} (1-\lambda(t;\theta)\mathrm{d}t)^{1-\mathrm{d}N(t)} \right\} = \prod_{i=1}^{N(T)} \lambda(t_i;\theta) \exp\left\{ -\int_0^T \lambda(t;\theta)\mathrm{d}t \right\} \tag{6}$$

where t_i is the jump time. By logarithm transformation on this likelihood function (6), we can maximize

$$\log\, L(\theta) = \sum_{i=1}^{N(t)} \log\, \lambda(t_i;\theta|\mathcal{H}_{t_i}) - \int_0^t \lambda(u;\theta|\mathcal{H}_u)\mathrm{d}u \tag{7}$$

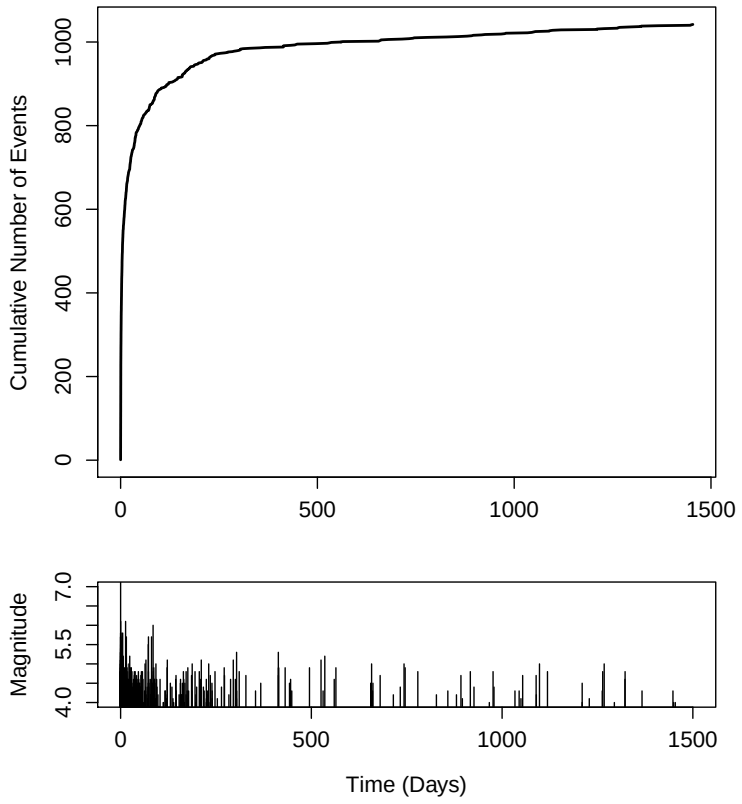

Figure 2: The cumulative number of events and magnitude against ordinary time (days) in the data set. The first event is the mainshock with 7.9 magnitude.

to obtain an optimal parameter θ, where $\theta = (\mu, A, \alpha, c, p)$. The maximum algorithm could use general unconstrained nonlinear optimization such as BFGS algorithm (Nielsen, 2000).

2.2 Model Verifying

Ogata (1988) uses the Akaike information criterion (AIC) as a comparison between different models fitted by the same data set. In this case we have three different values of M_0, hence the data scales are changed in three different M_0; as a result, the model comparison by checking the AIC is not suitable. To present a goodness-of-fit test, we can use the integral of the conditional intensity to compare the total number of events at time t_i, for instance:

$$\mathrm{E}(N(t_i)|\mathcal{H}_{t_i}) = \Lambda(t_i|\mathcal{H}_{t_i}) = \int_0^{t_i} \lambda(t|\mathcal{H}_t)\mathrm{d}t \tag{8}$$

Papangelou (1972) points out $\Lambda(t_i)$ distributed as a stationary Poisson process with intensity 1. This is from the definition of the counting process intensity: λ is an $\{\mathcal{F}_t\}-$intensity for N if satisfies

$$N(t) - \int_0^t \lambda(u)\mathrm{d}u \tag{9}$$

is a $\{\mathcal{F}_t\}-$martingale ($t \leqslant T$); then we take expectation on both sides of (9) to obtain (8), where $\Lambda(t|\mathcal{H}_t)$ is called residual point process (see Meyer 1971). Hence we can carry out the Kolmogorov-Smirnov test for the uniform distribution of $\int_0^{t_i} \hat{\lambda}(t|\mathcal{H}_t)\mathrm{d}t$ against $\mathrm{E}(N(t_i)|\mathcal{H}_{t_i})$ to check models; more precisely, we can check the p-value of the $\hat{\Lambda}(t_i)$ uniformly distributed on $[N(0), N(T)]$.

M_0	Total Number	μ	A	c	α	p	$p-value$
Initial value		0.0271	0.0509	0.1358	1.8680	1.8680	
4.0	846	0.0131	0.0428	0.2867	2.4969	1.2654	0.434
4.3	561	0.0112	0.0436	0.1577	2.6225	1.2035	0.670
4.5	366	0.0115	0.0692	0.0770	2.5226	1.1602	0.567

Table 1: The maximum likelihood estimations of ETAS model parameter for three different M_0.

Among these three ETAS models, the best choice of M_0 is 4.3 (based on the $p-value$). It is widely known that the more points to fit, the worse performance of the ETAS model; hence the least $p-value$ is in the model with $M_0 = 4.0$, which includes 846 events. However, the highest $p-value$ should have been from the case of $M_0 = 4.5$ (366 events); in fact the $p-value$ of this case is 0.567, less than the the $p-value$ in the model of $M_0 = 4.3$ (561 events, $p-value = 0.67$). The reason could be from the number of earthquake with magnitude greater than 6.2 in the whole data set (Figure 4, 5). In this data set, the number of events which magnitude is greater than 6.2 is only one: the 7.9 M_w mainshock. This would influence the result of the Gutenberg-Richter's law (3); Figure 4 ($M_0 = 4.3$) shows the zero point of $\log N$ is at magnitude 6.6, while Figure 5 ($M_0 = 4.5$) is 6.4. By this difference in the Gutenberg-Richter's law, we can verify this effect by the estimated parameters in these ETAS models. In the ETAS model, the Gutenberg-Richter's law is mainly carried out by the term $Aexp(\alpha(M - M_0))$, and the estimated A in these two cases are quite different ($A = 0.0436$ for $M_0 = 4.3$, $A = 0.0692$ for $M_0 = 4.5$, Table 1). By this we can conclude that the large magnitude earthquakes could have significant influence in the choice of the value of M_0, even though in this case the large earthquakes take a little proportion in the whole data.

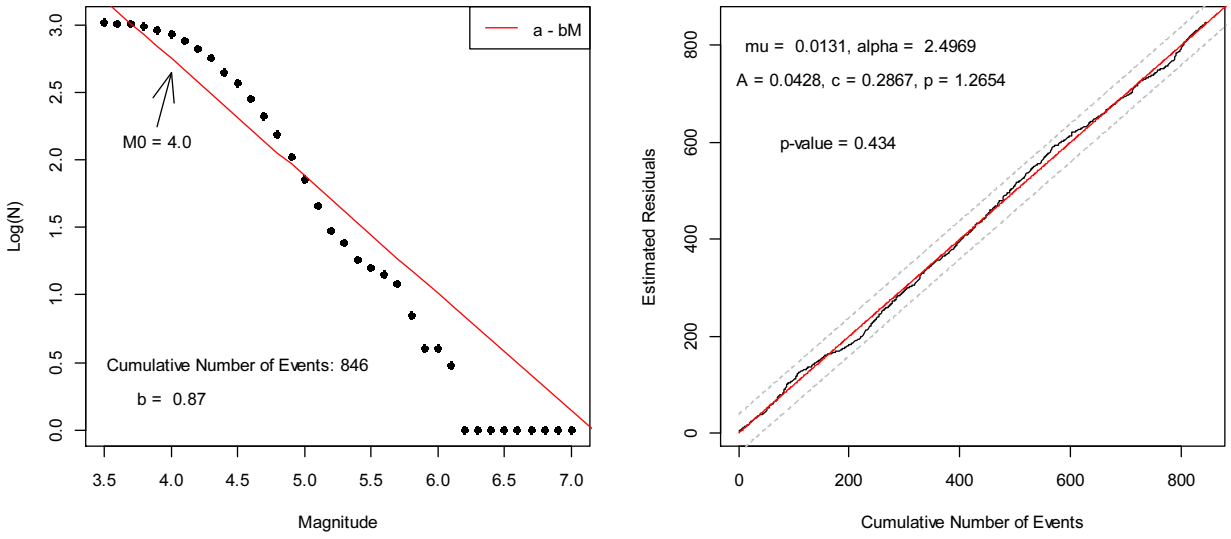

Figure 3: Gutenberg-Richter's law (3) plot (left) and goodness of fit plot (right) for ETAS model (2) with $M_0 = 4.0$, there are 846 events with magnitude not less than 4.0; the p-value is 0.434. The dash line on the right plot is 95% confidence interval of the Kolmogorov-Smirnov statistic.

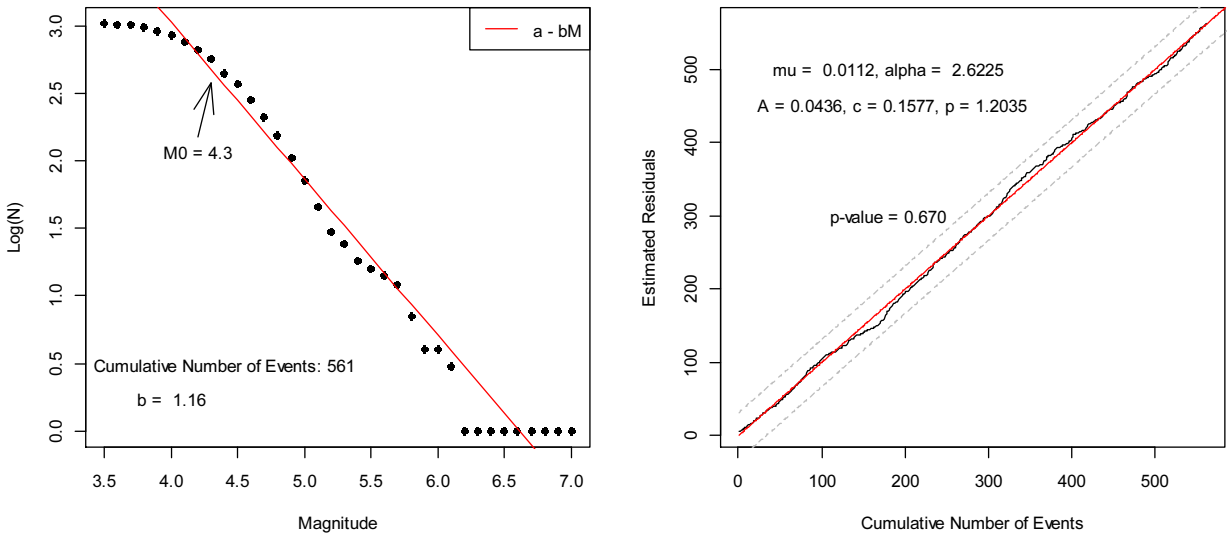

Figure 4: Like figure (3), $M_0 = 4.3$, there are 561 events with magnitude not less than 4.3; the p-value is 0.670. This is the best model among 3 different M_0.

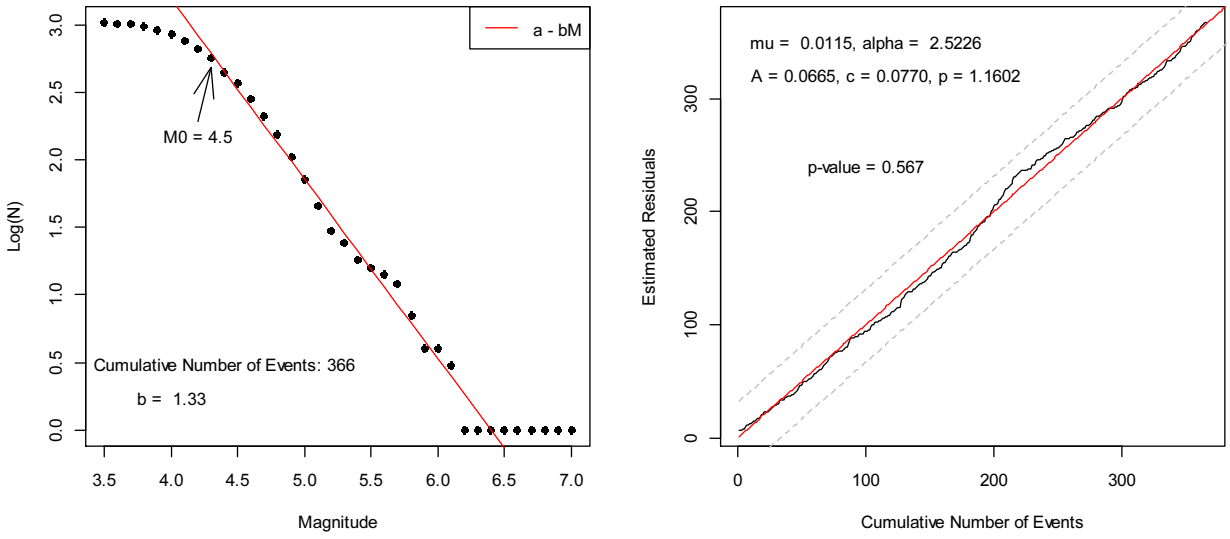

Figure 5: $M_0 = 4.5$, there are 366 events with magnitude not less than 4.0; the p-value is 0.537.

3. Space-Time ETAS model for 2008 Sichuan Earthquake

3.1 Space-Time ETAS model

Commonly, the records of seismic activity also include the earthquakes location information, and the location of historical earthquakes would affect the next events occurrence in an earthquake sequence. So, the development of the ETAS model from a earthquake time series to a space-time points pattern analysis is naturally proposed. Musmeci and Vere-Jones (1992) put forward a space-time clustering model for earthquake based on multi-dimensional self-exciting point process. By these previous works, Ogata (1998) establishes space-time ETAS models. These models are concentrating on the time information t and location pattern (x, y), based on the muli-dimensional self-exciting intensity $\lambda(x, y, t|\mathcal{H}_t)$ with the form

$$\lambda(x, y, t|\mathcal{H}_t) = \lim_{\Delta x, \Delta y, \Delta t \to 0} \frac{P(N(\Delta x, \Delta y, \Delta t) > 0|\mathcal{H}_t)}{\Delta x \Delta y \Delta t} \tag{10}$$

$N(x, y, t)$ is a seismic activity counting process, counts the total number of events occurring in the

space $[x_0, x] \times [y_0, y] \times [t_0, t]$, where (x,y) means the location of an earthquake in a seismic region, described by the degree of latitude and longitude on the map; $\mathcal{H}_t$ is the history of past events before time t, including time, location and magnitude information. With a stationarity assumption, Ogata posed the model

$$\lambda(x, y, t|\mathcal{H}_t) = \mu(x, y) + \sum_{t_i<t} g(t - t_i, x - x_i, y - y_i; M_i) \tag{11}$$

where

$$g(t, x, y|M) = A \exp\{\alpha(M - M_0)\} \frac{p-1}{c}(1 + \frac{t}{c})^{-p} \left[\frac{1}{\pi\sigma} f(x, y|M)\right] \tag{12}$$

$\mu(x, y)$ is a unknown function, which measures the declustering seismic rate of an earthquake, only dependent on its own location; and $\frac{1}{\pi\sigma} f(x, y|M)$ is the space distribution in the triggered part. Musmeci and Vere-Jones (1992) suggest a normal diffusion function

$$\frac{1}{\pi\sigma} f(x, y|M) = \frac{1}{2\pi\sigma_x\sigma_y u} \exp\left\{-\frac{1}{2u}(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2})\right\} \tag{13}$$

Ogata (1998) compares three different location density models

$$\frac{1}{\pi\sigma} f(x, y|M) = \frac{1}{2\pi\sigma(M)} \exp\left\{-\frac{(x^2 + y^2)}{2\sigma(M)}\right\} \tag{14}$$

$$\frac{1}{\pi\sigma} f(x, y|M) = \frac{q-1}{\pi\sigma(M)}(1 + \frac{x^2 + y^2}{\sigma(M)})^{-q} \tag{15}$$

and

$$\frac{1}{\pi\sigma} f(x, y|M) = \frac{q-1}{\pi\sigma(M)}(1 + \frac{(\frac{\sigma_y}{\sigma_x}x^2 - 2\rho_{xy}xy + \frac{\sigma_x}{\sigma_y}y^2)}{\sigma(M)\sqrt{1 - \rho_{xy}^2}})^{-q} \tag{16}$$

with parameter $\theta = (\mu, A, c, p, \alpha, \gamma, D, q)$ by the Japan Meteorological Agency (JMA) data in the time period 1926-1995, where $\sigma(M) = D \exp\{\gamma(M - M_0)\}$. The constant number σ_x, σ_y and ρ_{xy} are maximum likelihood estimates of the standard deviation and correlation for the location pattern (x, y).

Model (16) assumes that the earthquake sequence is mainly distributed on a seismic belt; hence the estimated σ_x, σ_y and ρ_{xy} are introduced in the location density function in order to obtain an isometric ellipse (rotated by the angle β, where β satisfies $\tan 2\beta = -2\rho_{xy}/(\frac{\sigma_y}{\sigma_x} - \frac{\sigma_x}{\sigma_y})$) with the same area as a circle (Ogata 1998, Ogata and Zhuang, 2006). Consequently, this rotation angle could reflect an estimated direction of the seismic belt, by a semi-isotropic ellipse. Only the third model (16) obeys on this assumption among these three models. Ogata in his paper (Ogata, 1998) suggests the third model (16) could be the best one among these three different spatial distributions. In this paper we concentrate on the third one.

In practice, the rotating angle of an isometric ellipse is controlled by an constant $\hat{\rho}_{xy}$. However in the real world, the seismic activity location could be complex; hence the real ρ_{xy} should be varying. When the seismic region is lying in a bending belt, the real ρ_{xy} or the rotating angle changes frequently, even from positive to negative. In the current research, how to deal with this varying ρ_{xy} is still unsolved. In the following section we propose a transformed location coordinate system to cancel ρ_{xy} (or let rotating angle of isometric ellipses be zero) in the model (16) as a potential solution.

3.2 Location Transformed Method for the 2008 Sichuan Earthquake

Sichuan seismic belt region is a straight narrow belt on the map (see Figure 1). In this case the latitude of seismic activity is mainly over 30 degree. Ogata (1998) suggests a flat longitude transformation when the two-dimension space pattern is on the high degree of latitude surface of the earth (Eucildean coordinate), since the unit length of longitude is shorter when locates at a higher degree of latitude. Using the equator length of one degree of longitude (about 111.1km, almost the same as latitude, hence we can obtain a coordinate with the same length in a unit of Latitude and Longitude), we can transform the raw longitude at the high degree of latitude to the equator longitude (Figure 6, $\text{longitude}_i = \text{longitude}_i * \cos(\overline{\text{latitude}})$, $\overline{\text{latitude}}$ is the mean latitude).

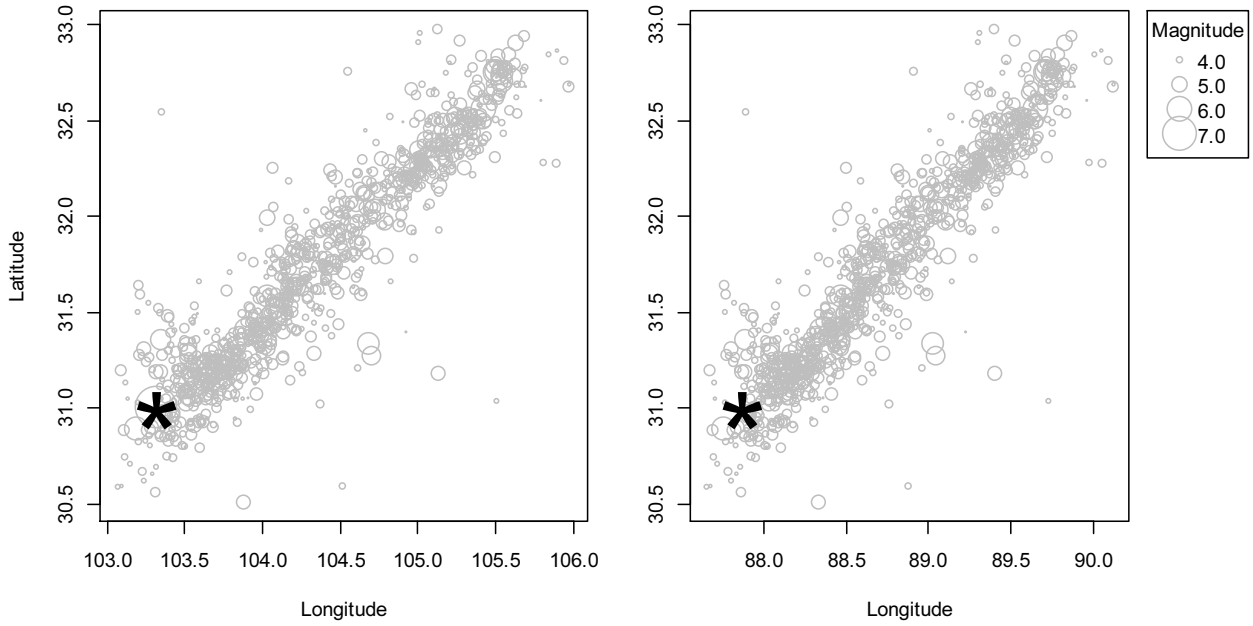

Figure 6: Two location coordinate system scatter plots for Sichuan earthquake (1042 events). The left is original longitude (without transformed); right plot is flat transformed longitude coordinate, the length of unit latitude and longitude equivalent. The big star points on both sides refer to the main shock (magnitude 7.9).

Considering the assumption of model (16), we first observe the shape of seismic belt in this case. From Figure 6, we can use some nonparametric way to estimate seismic belt, such as locally weighted regression (LOWESS, see Cleveland and William, 1979). By LOWESS, we minimize:

$$\sum_i w_i(x_j)(y_i - X_i^{\mathrm{T}}\beta_j)^2 \tag{17}$$

for the 2008 Sichuan earthquake data, where j refers to the number of subset of the whole data set;

$$w_i(x_j) = W(\frac{|x_i - x_j|}{h_j}) \tag{18}$$

$$W(x) = \begin{cases} (1-|x|^3)^3 & \text{for } |x| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

where $h_j = \max|x_j - x_i|1_{\{x_i \text{ in the } j\text{th data span}\}}$. When we apply LOWESS, we use 2/3 as the data span in order to obtain a smooth enough curve, which is close to the real seismic belt in the

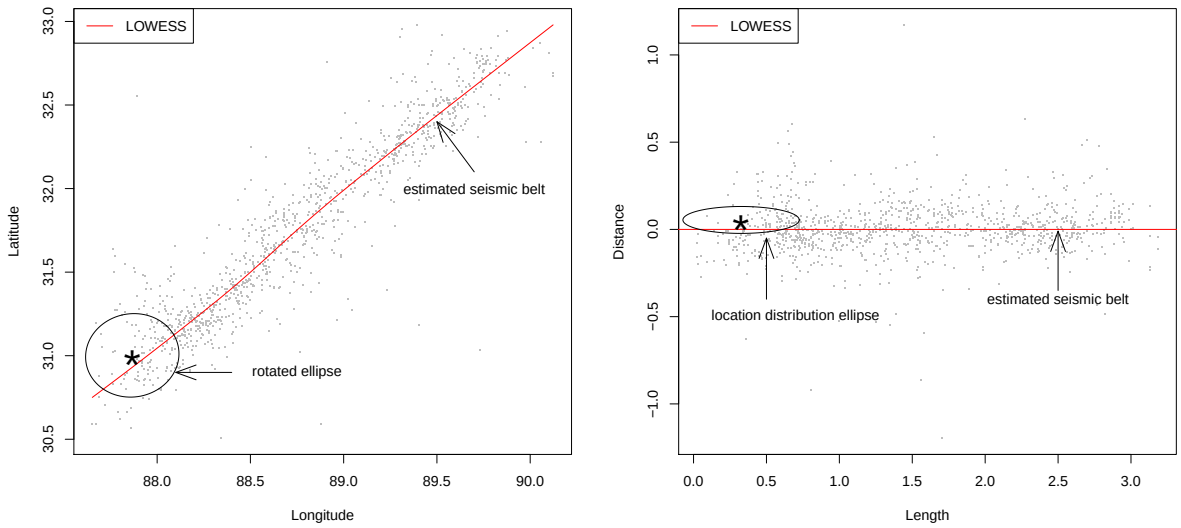

Figure 7: Estimate seismic belt by LOWESS (left) and transformed coordinate system scatter plot (right plot, *Length* against *Distance*), the big star point is the $M_{\mathrm{w}}7.9$ mainshock and gray points for other events. The ellipse is the location distribution area $\frac{\sigma_y}{\sigma_x}x^2 - 2\rho_{xy}xy + \frac{\sigma_x}{\sigma_y}y^2$ from (16) with a mainshock centre point, where the rotation angle β on the left satisfying $\tan 2\beta = -2\rho_{xy}/(\frac{\sigma_y}{\sigma_x} - \frac{\sigma_x}{\sigma_y})$, and zero on the right.

physical world; and x refers to longitude and y to latitude in (17). This underfitted curve is very smooth hence we are easy to imagine the projective point on this curve for every event.

By the estimated seismic belt line, we can apply the location transformed method. First we define two vectors: *Distance* (d_i) and *Length* (l_i); *Distance* refers to the least distance from a event to the seismic belt line (if the event is locating below this line, this *Distance* is negative); *Length* means the length of one event along the seismic belt to the start of this seismic belt. If we divide the LOWESS line into $(m-1)$ intervals (m is not less than the number of total events, such as $m \geqq 1042$ in this case), each interval has the equivalent length; then we have m points uniformly distributed on this line. We write these points as (Lo_{x_j}, Lo_{y_j}) on Longitude-Latitude coordinate system; then the absolute value of *Distance* (d_i) could be calculated by

$$|d_i| = \min_j \left\{(x_i - Lo_{x_j})^2 + (y_i - Lo_{y_j})^2\right\}^{\frac{1}{2}} \quad j = 1, 2, 3 \cdots \tag{20}$$

After this, we can obtain a series of points $(Lo_x^{(k_i)}, Lo_y^{(k_i)})$ on this LOWESS line, uniquely against every event location (x_i, y_i) with the least distance *Distance* (d_i). Then the *Length* (l_i) can be calculated by

$$l_i = \sum_{j<k_i} \left\{(Lo_x^{(j+1)} - Lo_x^{j})^2 + (Lo_y^{(j+1)} - Lo_y^{(j)})^2\right\}^{\frac{1}{2}} \tag{21}$$

Finally, we obtain a new location coordinate (l_i, d_i) uniquely for every event. Then, by stretching the estimated seismic belt to a straight line, we build a new coordinate system (Figure 7). Based on this new coordinate, the rotating angle in the model (16) is naturally equivalent to zero (the earthquake events become distributed on the x-axis, which is the estimated seismic belt).

So far we have a data set with new space pattern. Consequently, the space pattern of space-time ETAS model should be changed for this new data. We call the data using *latitude* and *longitude* to describe location pattern as '**original data**'; and call '**transformed data**' if the data set uses the new vectors *Distance* (d_i) and *Length* (l_i) as location pattern. Then we apply the space-time ETAS model (16) on these two data set. Specifically, the estimated ρ_{xy} of the '**original data**' is from the maximum likelihood estimate, while $\rho_{xy} = 0$ in the '**transformed data**'. Hence the location distribution of 'transformed data' is

$$\frac{1}{\pi\sigma} f(x, y|M) = \frac{q-1}{\pi\sigma(M)} (1 + \frac{(\frac{\sigma_y}{\sigma_x} x^2 + \frac{\sigma_x}{\sigma_y} y^2)}{\sigma(M)})^{-q} \tag{22}$$

3.3 Maximum Likelihood Method and Clustering Probability

For a muti-dimensional counting process $N(t, \chi)$ on the whole space $[0, T] \times \mathrm{A}$, the likelihood function $L(\theta)$ for its intensity $\lambda(t, \chi; \theta)$ of this process is

$$\begin{aligned} L(\theta) &= \prod_{t \in [0,T],\, \chi \in A} \left\{ \lambda(t, \chi; \theta)^{\mathrm{d}N(t,\chi)} (1 - \lambda(t, \chi; \theta) \mathrm{d}\chi \mathrm{d}t)^{1 - \mathrm{d}N(t,\chi)} \right\} \\ &= \prod_{i}^{N(T,\mathrm{A})} \lambda(t_i, \chi_i; \theta) \exp\left\{ - \int_0^T \int_{\mathrm{A}} \lambda(t, \chi) \ell_{\mathrm{A}}(\mathrm{d}\chi) \mathrm{d}t \right\} \end{aligned} \tag{23}$$

where t_i is the jump time of process $N(t, \chi)$ and ℓ_{A} is the reference measure on set A. Now with the parameter $\theta = (\mu, A, c, \alpha, p, D, q, \gamma)$, we can maximize log-likelihood function

$$\log L(\theta) = \sum_{i=1}^{N} \log \lambda_\theta(t_i, x_i, y_i; M_i | \mathcal{H}_{t_i}) - \int_0^T \int \int_{\mathrm{A}} \lambda_\theta(t, x, y; M | \mathcal{H}_t) \mathrm{d}x \mathrm{d}y \mathrm{d}t \tag{24}$$

to obtain an optimal $\hat{\theta}$. Ogata (1998) estimates $\int \int_{\mathrm{A}} \lambda_\theta(t, x, y; M|\mathcal{H}_t) \mathrm{d}x \mathrm{d}y$ by dividing the set A into many radial segments and Zhuang et.al. (2002) suggest estimate $\mu(x, y)$ by thinning process (see Zhuang et.al. 2002 and Delay & Vere-Jones 1972). Based on thinning process, Zhuang and Ogata (2004) define two probabilities

$$\varphi_i = \frac{\mu(x_i, y_i)}{\lambda(t_i, x_i, y_i | \mathcal{H}_{t_i})} \tag{25}$$

and $1-\varphi_i$, where φ_i refers to the probability of the ith event to be a declustering (or background) event; correspondingly, $1-\varphi_i$ is the clustering event probability of the ith event. Moreover, note that

$$1-\varphi_i = \frac{\sum_{j<i} g(t_i-t_j, x_i-x_j, y_i-y_j; M_j)}{\lambda(t_i, x_i, y_i)} \tag{26}$$

where function g is from (11). We write$\psi_{ij} = \frac{g(t_i-t_j, x_i-x_j, y_i-y_j; M_j)}{\lambda(t_i, x_i, y_i)} 1_{j<i}$, then (26) could be write as

$$\psi_i = 1-\varphi_i = \sum_{j=1} \psi_{ij} \tag{27}$$

By a certain i, j, ψ_{ij} refers the probability of an event i triggered by the jth event (see Zhuang and Ogata, 2004; Zhuang et.al., 2005). Hence the clustering probability of ith event is measured by ψ_i. In application, the value of φ_i and ψ_i could be the evidence to classify the background area or the clustering area in a seismic region.

3.4 The 2008 Sichuan Earthquake Space-time Analysis

By previous formula, we apply models (16) and (22) by original data and location transformed data (location coordinate is *Distance* against *Length*). The results are shown in the Table 2, and the difference between the parameters of these two models is not obvious. This is because the estimated seismic belt (Figure 7, LOWESS) is nearly a straight line; hence the location transformation could be as a roughly radial rotation. The result also displays a little difference between two model, mainly from the parameters about location pattern (D, q, γ). The time parameters c and p are almost the same between two models, this because the occurrence time of events is never changed. Finally, the difference between parameters A and α might be from the optimization process; while A decreases from original data model to transformed data, the parameter α is increasing as a compensator. Consequently, the Gutenberg-Richter's law pattern $A\exp(\alpha(M-M_0))$ is almost not changed.

Data	μ_0	A	c	α	p	D	q	γ	AIC
Initial value	0.7106	0.0970	0.1094	1.5716	1.3367	0.0090	2.0339	0.2567	
Original	0.0578	0.5150	0.0358	1.8365	1.2161	0.0144	1.8433	0.6241	-935.061
Transformed	0.0571	0.5090	0.0358	1.8680	1.2163	0.0152	1.8646	0.6422	-950.052

$\text{AIC} = -2\max_{\theta} \log(L_\theta) + 2(\text{number of estimated parameters})$

Table 2: Parameters from estimated space-time ETAS model on 2008 Sichuan earthquake data, constant $M_0 = 4.3$, 561 shocks. The original data set is fitted by model (16), $\rho_{xy} \approx 0.9$, while transformed data $\rho_{xy} = 0$ (model (22)).

The goodness-of-fit test we used in previous ETAS model (only including time and magnitude information) is not appropriate for space-time ETAS model. Because the space set A is a compact

set; therefore $\exists\chi_* \in \mathrm{A}$, can let $\mu(x,y)$ be zero in (11), then the condition of residual process $\Lambda(t) = \int_0^t \lambda(u)\mathrm{d}u$ for each point t_i is uniformly distributed on $N(t_i)$ could not be satisfied, such as the condition $\int_0^\infty \lambda(t)\mathrm{d}t = \infty$ (see Schoenberg 2003, Meyer 1971). This because the residual process of the space-time ETAS model is

$$\Lambda(t, \mathrm{A}_i) = \int_0^t \int\int_{\mathrm{A}_i} \lambda(u, x, y|\mathcal{H}_u)\mathrm{d}x\mathrm{d}y\mathrm{d}u \tag{28}$$

if we check the infinite integral of the conditional intensity for this model on the subset $\mathrm{A}_i \subset \mathrm{A}$, then

$$\begin{aligned}\int_0^\infty \int\int_{\mathrm{A}_i} \lambda(t, x, y)|\mathcal{H}_t)\mathrm{d}x\mathrm{d}y\mathrm{d}t = &\int_0^\infty \int\int_{\mathrm{A}_i} \mu(x, y)\mathrm{d}x\mathrm{d}y\mathrm{d}t \\ &+ \int_0^\infty \int\int_{\mathrm{A}_i} \sum_{t_i<t} g(t - t_i, x - x_i, y - y_i; M_i)\mathrm{d}x\mathrm{d}y\mathrm{d}t\end{aligned} \tag{29}$$

Since the set A is a compact set, if A_i is little enough, the value of $\mu(x,y)$ could be zero, then the first term of right-hand-side of (29) could be zero; and the second term can be write as

$$\int\int_{\mathrm{A}_i} \sum_{t_i<t} A\exp\left\{\alpha(M_i - M_0)\right\}\left[\frac{1}{\pi\sigma} f(x - x_i, y - y_i|M_i)\right]\mathrm{d}x\mathrm{d}y < \infty$$

Overall, the goodness-of-fit test based on residual process may be neglect in application (Schoenberg 2003). To test a goodness-of-fit, Zhuang and Ogata (2004) provide a location distribution based on the standardized distance. They define the squared standardized distance of a triggered earthquake j to it ancestor i is:

$$r_{ij}^2 = \frac{\frac{\sigma_y}{\sigma_x}(x_j - x_i)^2 - 2\rho_{xy}(x_j - x_i)(y_j - y_i) + \frac{\sigma_x}{\sigma_y}(y_j - y_i)^2}{\sigma(M)\sqrt{1 - \rho_{xy}^2}} \tag{30}$$

From (16), the density function of r_{ij} could be write as

$$f_R(r) = 2r(q-1)(1 + r^2)^{-q} \tag{31}$$

Zhuang and Ogata (2004) reconstruct (31) into

$$\hat{f}_R(r) = \frac{\sum_{i,j} \psi_{i,j} 1_{|r_{i,j} - r| < \frac{\Delta r}{2}}}{\Delta r \sum_{i,j} \psi_{i,j}} \tag{32}$$

where $\psi_{i,j}$ is from (27). Figure 8 shows the distribution of standardized distance from two different models, the original data and the transformed data. In this case we choose $\Delta r = 0.03$ to obtain many enough details as comparison (when Δr is large enough these two estimated density functions

are very close to the theoretical line, almost indistinguishable). From this we can see that the reconstructed density function from the transformed data set seems to be the same performance as the original data. This is difficult to say which is better.

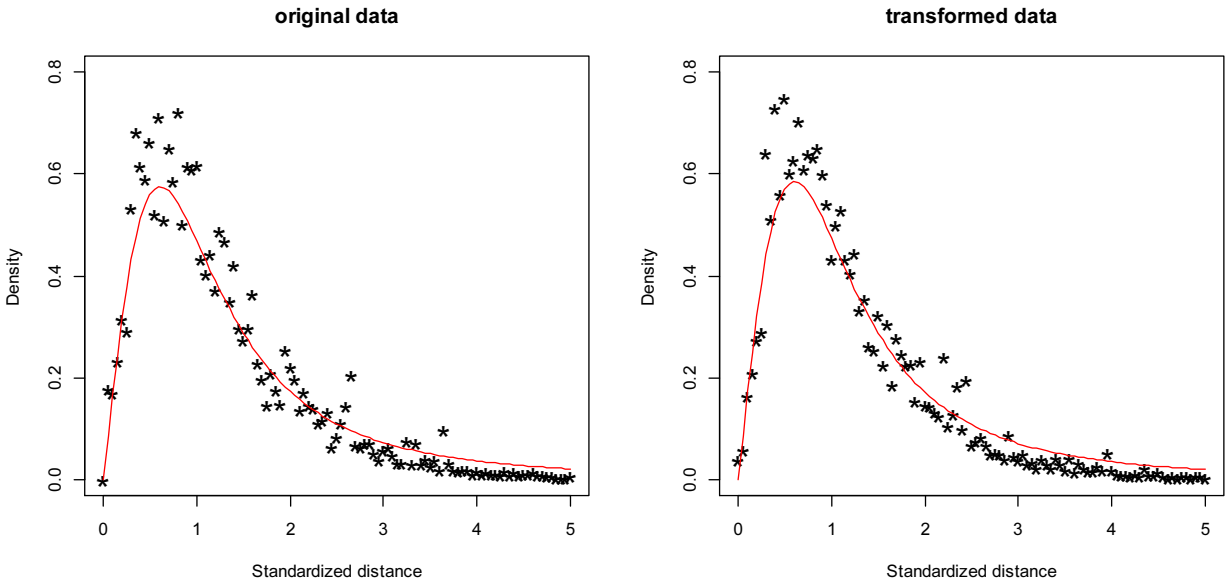

Figure 8: The standardized distance density plot. The red line in both sides is the theoretical line, and the star points refer to the estimated value of (32). The left plot is calculated by the original data, while the transformed data plot on the right ($\Delta r = 0.03$).

3.5 Background Events Region and Clustering Events Region

In seismology, the classification of the declustering earthquake area and the clustering earthquake area is very important. These two kinds of earthquake distribution in an seismic region could be as the evidence to describe the hazard region. Jiang and Zhuang (2010) state the high probability of clustering events zones could be as the evidence of potential strong earthquakes zones, and the declustering events zones is the "normal" seismic activity zones. In this section we compare these two types of seismic zones by previous two space-time ETAS models in the 2008 Sichuan earthquake region.

The probability of declustering events and clustering events are defined as φ_i and ψ_i from (25) and (27). Figure 9 and Figure 10 show the estimated two kings of events distribution on the real world Longitude-Latitude coordinate system. From these we can see that the application of location transformed method on the 2008 Sichuan earthquake data causes little difference, or nearly no difference. The reason could be from that the previous estimated seismic belt in this region is almost

a straight line; as a consequence, the relative distances of any two points between original data and transformed are not so much difference, and the number of clustering events (or declustering) have the same value in these two model. Moreover, from these two Figures (9 & 10), we can conclude that the seismic activity region in this case could be a high clustering earthquakes region. Comparing between Figure 9 and Figure 10, the clustering events scale is much greater than declustering events. This also suggest the 2008 Sichuan earthquake region is a highly potential region of strong earthquakes; this result is the same as Jiang and Zhuang (2010).

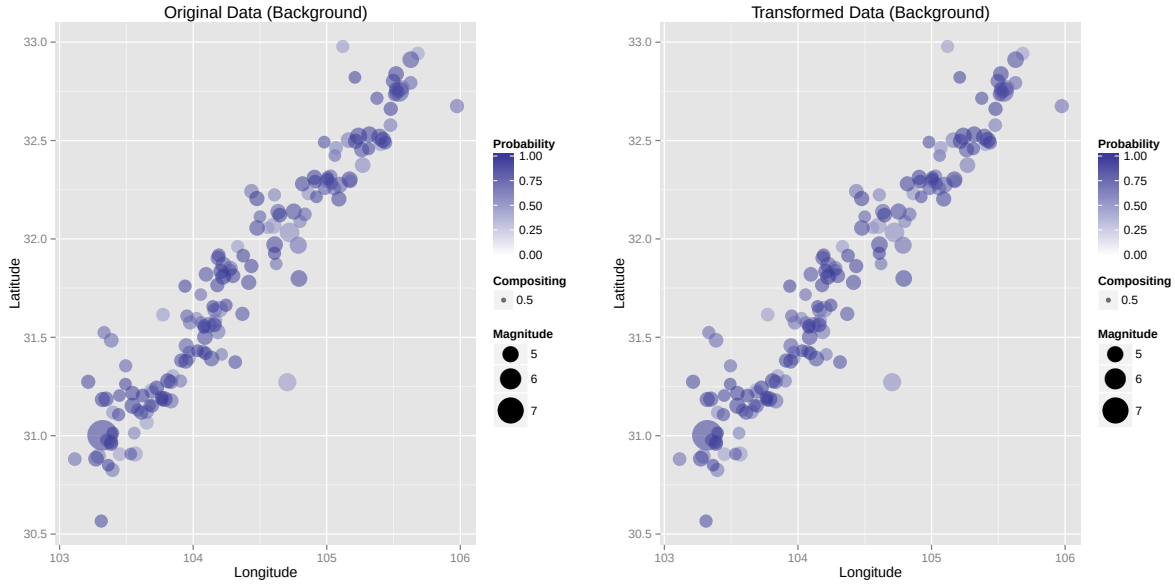

Figure 9: The background events distribution on the real world latitude and longitude scale for the 2008 Sichuan earthquake data. The left one is from original data, while the right is transformed data. Blue colour refers to the probability of being a background event (φ_i from (25)).

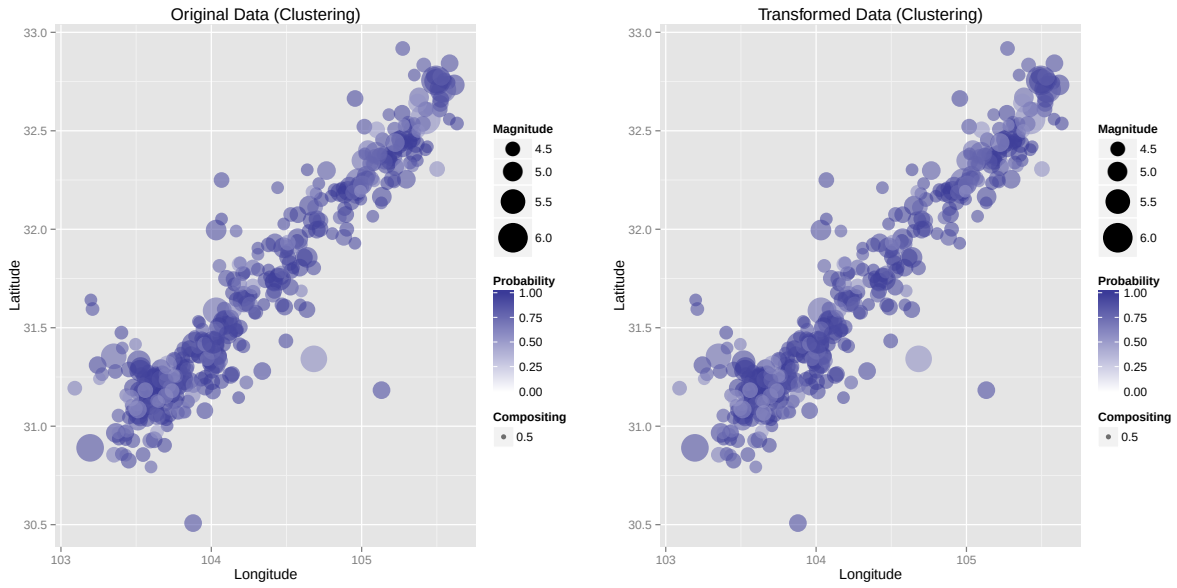

Figure 10: The clustering events (the 2008 Sichuan earthquake) distribute on plot. The left one is from original data, and the right is transformed data. Blue colour refers to the probability of being a clustering event (ψ_i from (27)).

4. The Location Transformed Method on the 2010 Chile Earthquake

This section we use the transformed method on the 2010 Chile earthquake data, as a purpose to show the improvement of this method on the space-time ETAS model. In order to show this improvement, we choose a seismic region located on a bending seismic belt. By this purpose, the Chile seismic region is an appropriate choice; as a result, we search data on the U.S. Geological Survey (USGS) website. The latitude is from -39 to -32, longitude from -74.3 to -71, and the time interval is from 27/02/2010 to 15/08/2013. From Figure 11 we can see that either frequency or magnitude, the 2010 Chile earthquake events are much stronger than the 2008 Sichuan earthquake.

For convenience, we ignore the different value of M_0 in this case. By Gutenberg-Richter's law (3) we choose $M_0 = 4.5$ (Figure 12). The total number of $M_{\mathrm{w}} \geqslant 4.5$ events is 1834, starting at the magnitude 8.8 mainshock. We fit this data first by ETAS model (which only includes time and magnitude effects), with the result displayed on Figure 13. By the Kolmogorov-Smirnov test, this model is rejected at 0.05 significant level. However this result still could be a significant reference for this seismic activity region. Compared to

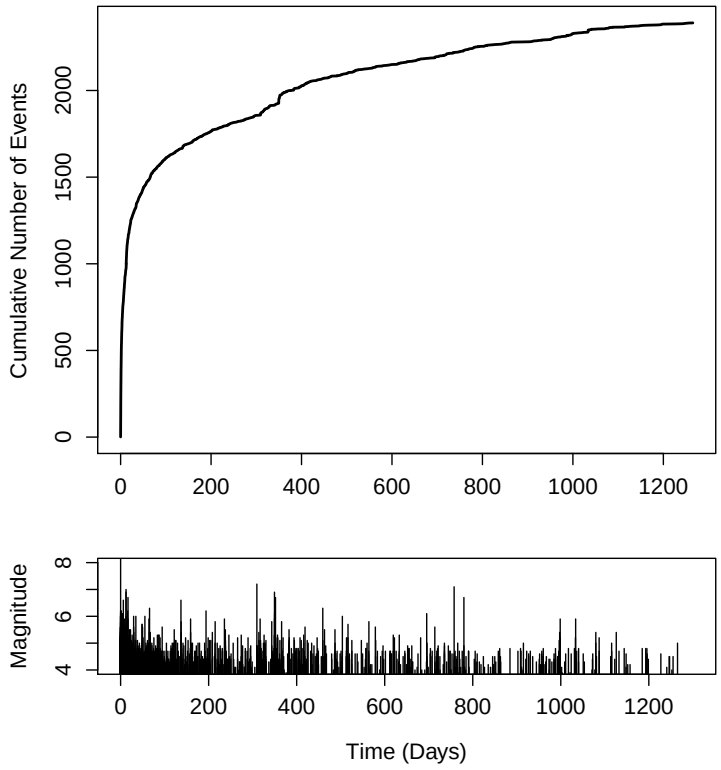

Figure 11: The cumulative number of events and magnitude against ordinary time (days) in the 2010 Chile earthquake data. The first event is the mainshock with 8.8 magnitude.

the 2008 Sichuan earthquake, the large value of μ (or background intensity) in this model suggests the higher proportion of declustering earthquakes occurred during this period.

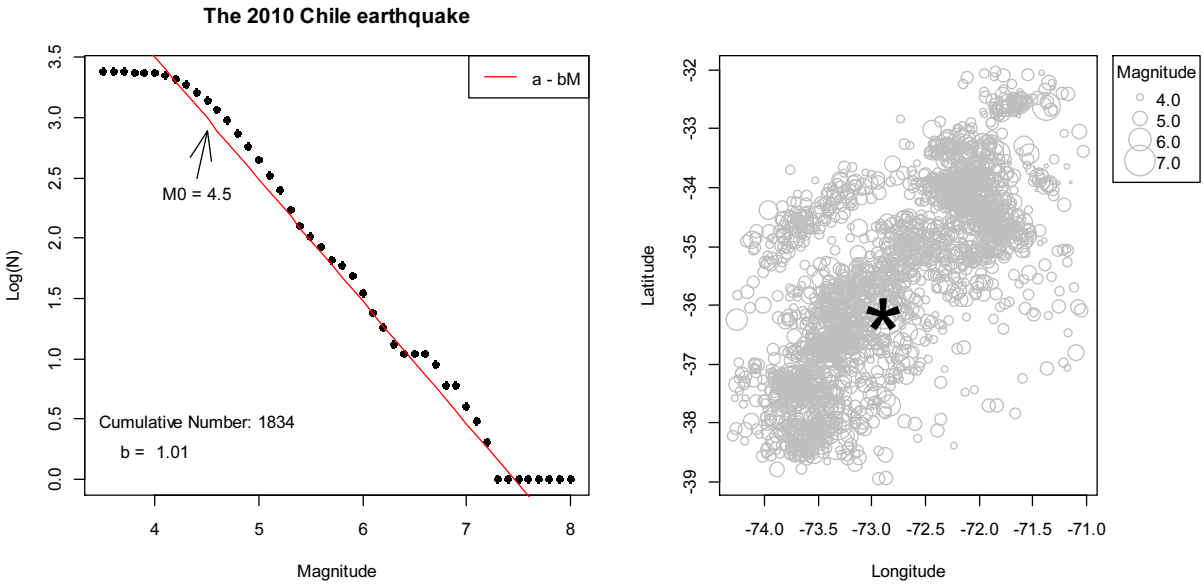

Figure 12: Gutenberg-Richter's law (3) plot (left) for $M_0 = 4.5$ and scatter plot (right), there are 1834 events with magnitude not less than 4.5. The big star refers to the main shock ($M_{\mathrm{W}} = 8.8$).

So far we consider space-time ETAS model in this data set. Applying the location transformed method on this data (the method described at section 3.2), we establish a new coordinate system. Not like the 2008 Sichuan earthquake, we choose $\frac{1}{3}$ data set as span units in this case; moreover, in this LOWESS estimate, we use longitude as response and Latitude refers to the predictor (note the vector *distance* d_i is negative when the ith event is on the left of the LOWESS line). By doing this, we obtain a curve with several turnings, which is close to the real seismic belt in the physical world. Therefore, ρ_{xy} (or the rotation angle of ellipse space) in the location distribution (16) should be varying in the original data set. By applying this location transformed method, the varying ρ_{xy} could reduce to zero (Figure 14).

Data	μ_0	A	c	α	p	D	q	γ	AIC
Initial value	0.7106	0.0970	0.1094	1.5716	1.3367	0.0090	2.0339	0.2567	
Original	0.1482	0.5905	0.0410	1.1854	1.2791	0.0124	1.3734	0.1410	1646.70
Transformed	0.1448	0.5514	0.0433	1.1890	1.2777	0.0134	1.4814	0.3398	1708.03

$\mathrm{AIC} = -2\max_{\theta} \log(L_\theta) + 2(\text{number of estimated parameters})$

Table 3: Parameters from estimated space-time ETAS model on 2010 Chile earthquake data, constant $M_0 = 4.5$, 1834 shocks. The original data set is fitted by model (16), $\rho_{xy} \approx 0.7$, while transformed data $\rho_{xy} = 0$ (model (22)).

The results of space-time ETAS models fitted by the 2010 Chile earthquake data (original and

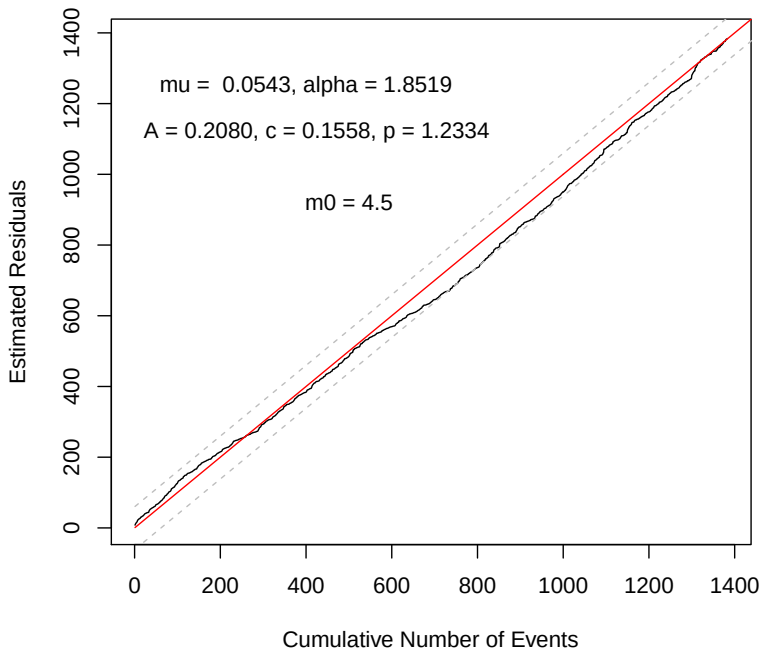

Figure 13: Cumulative number of events estimated by ETAS model on the 2010 Chile earthquake data, against the observations, the gray dash lines refer to the 99% Kolmogorov-Smirnov bounds.

transformed) are shown in the table 3. Unlike the results of the 2008 Sichuan earthquake, the estimated parameter is obviously different between the original data and transformed data in this case, especially the values of parameters about location pattern (D, q, γ). This is mainly from the application of the location transformed method. Compared to the case of Sichuan seismic region, the curved seismic belt on the Chile seismic region makes the transformed data much different with the original data. Similarly, the models comparison cannot carry out by the value of AIC, hence we test goodness-of-fit on this two models by the same method described at section 3.4 (Figure 15). Obviously, the right model (with the location transformed method) is much better than the original data model. This is easy to conclude: in the transformed data, we avoid to estimate ρ_{xy} which value should been varying.

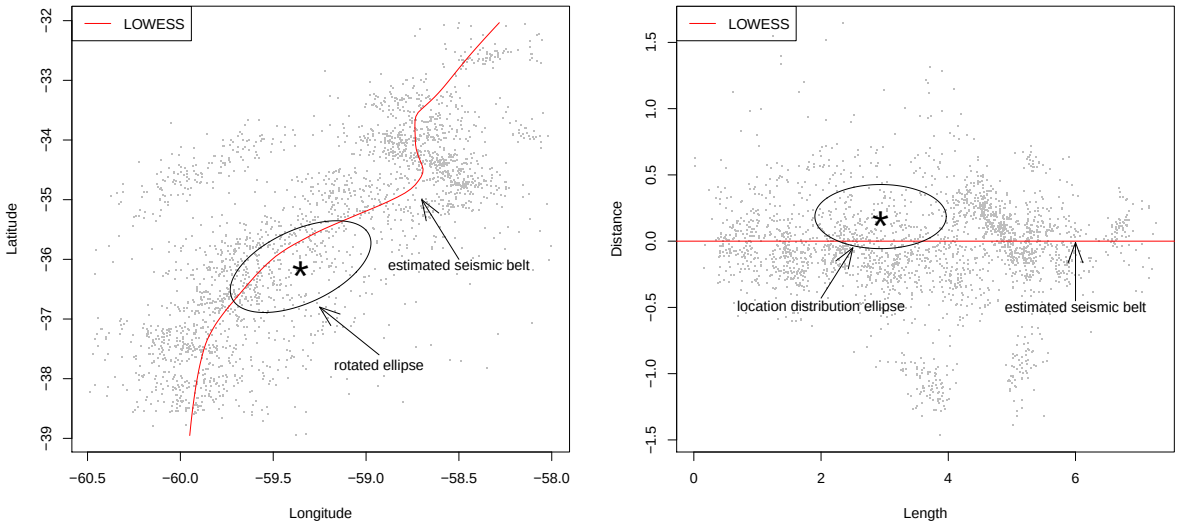

Figure 14: LOWESS (left, data span = 1/3, longitude as response) and transformed coordinate system scatter plot (right plot, if an event is on the left of LOWESS line, the location pattern *distance* is negative, see the mainshock point) for the 2010 Chile earthquake data, the big star point is the $M_{\rm w}8.8$ mainshock and gray points for other events. The ellipse is the location distribution area $\frac{\sigma_y}{\sigma_x}x^2 - 2\rho_{xy}xy + \frac{\sigma_x}{\sigma_y}y^2$ from (16) with a mainshock centre point, where the rotation angle β on the left satisfying $\tan 2\beta = -2\rho_{xy}/(\frac{\sigma_y}{\sigma_x} - \frac{\sigma_x}{\sigma_y})$, and zero on the right.

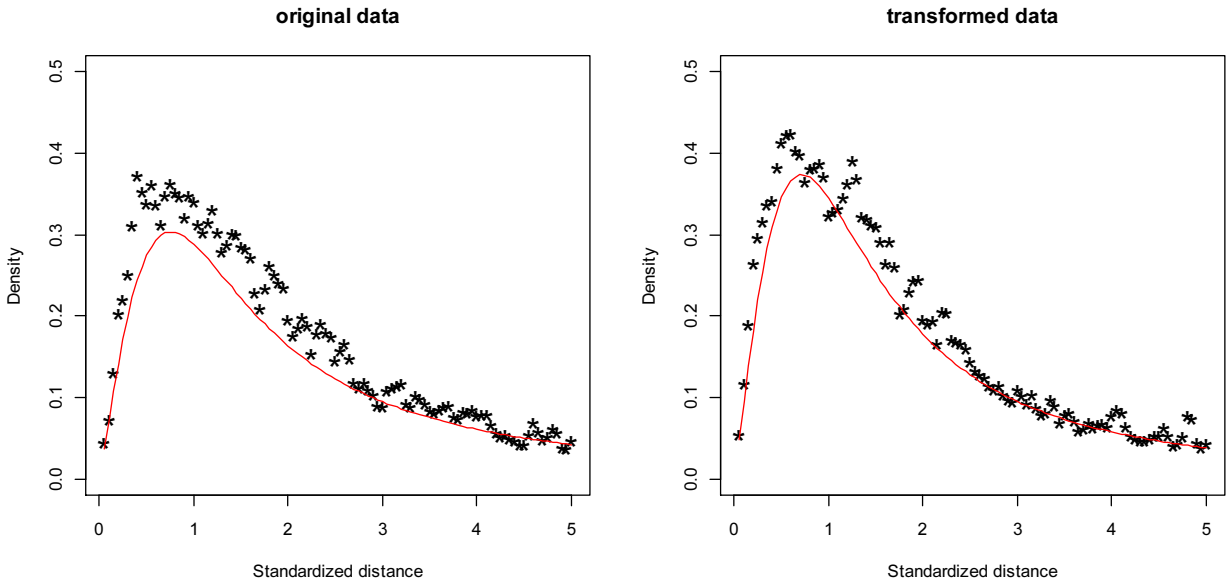

Figure 15: The standardized distance density plot. The red line in both sides is the theoretical line, and the star points refer to the estimated value of (32). The left plot is calculated by the original data, while the transformed data plot on the right ($\Delta r = 0.1$).

Finally we check the declustering events and clustering events zones in the 2010 Chile earthquake region. This seismic region is between the Nazca and South American tectonic plates, where is widely known as a highly potential zones of strong earthquakes. Figure 16 and Figure 17 display the declustering and clustering events distribution on this region by the real world coordinate. The space-time ETAS models results show that the number of declustering events ($\varphi_i \geqslant 0.5$) for original data is 497, while for the transformed is 504 (Figure 18 shows these points information). From these figures we can verify the fact of this earthquake hazardous region, not only for the clustering events, but the background earthquakes area also in high probabilities of strong earthquakes.

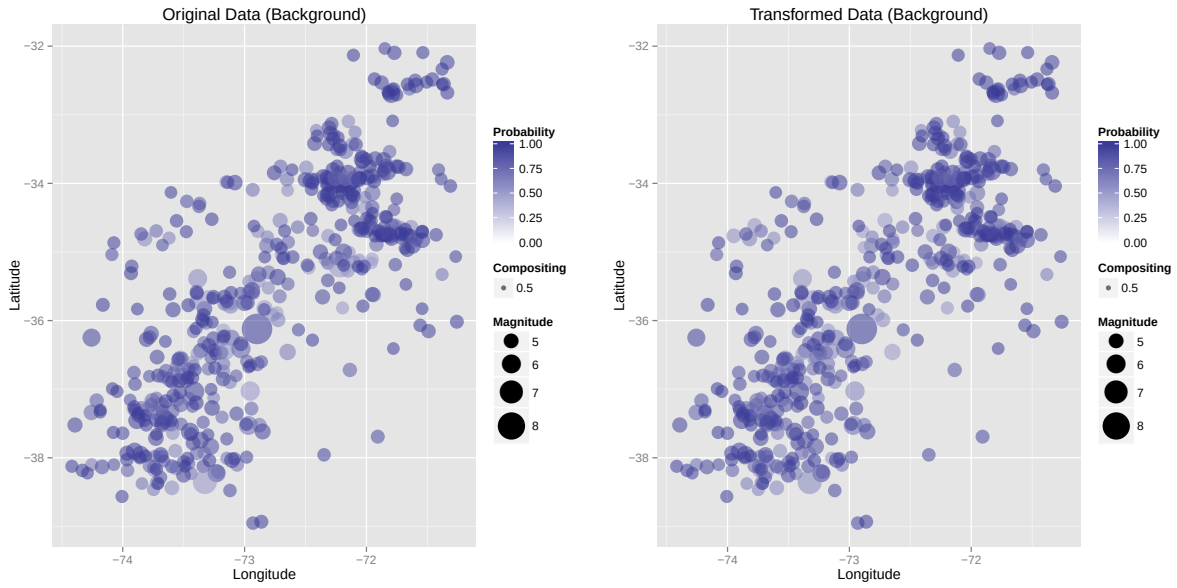

Figure 16: The background events distribution (the 2010 Chile earthquake data) on the real world latitude and longitude scale. The left one is from original data, while the right is transformed data. Blue colour refers to the probability of being a background event (φ_i from (25)).

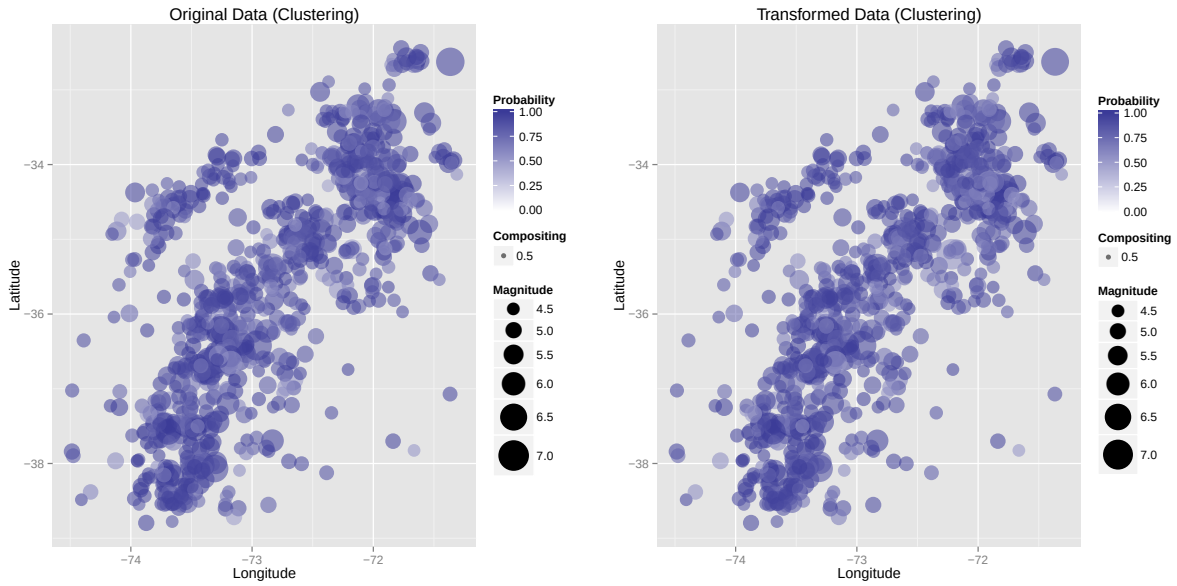

Figure 17: The clustering events distribute on plot for the 2010 Chile earthquake. The left one is from original data, and the right is transformed data. Blue colour refers to the probability of being a clustering event (ψ_i from (27)).

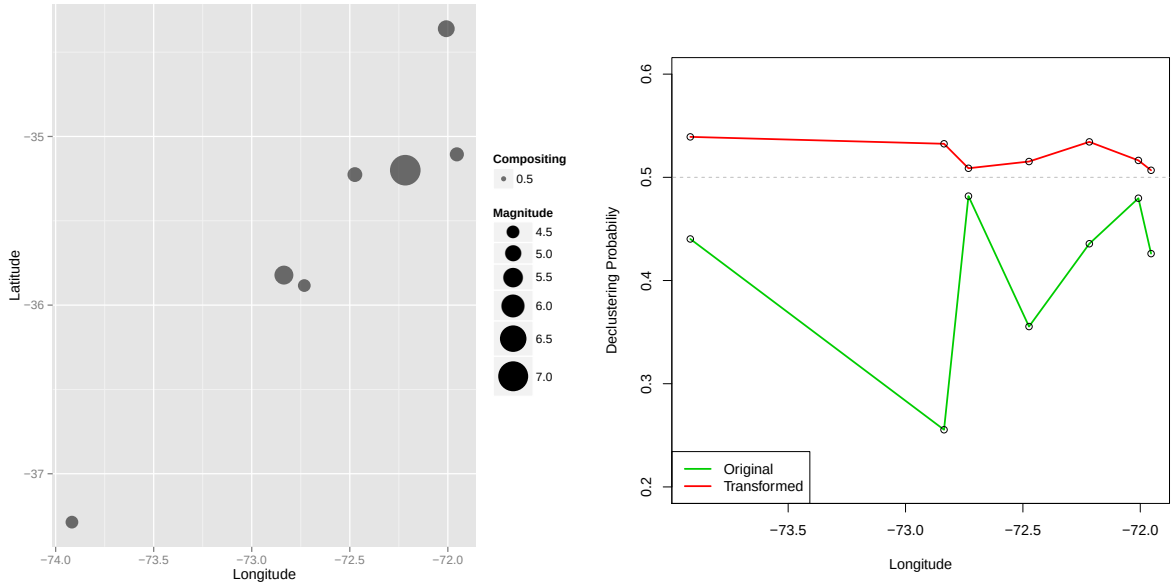

Figure 18: The seven points which regard as clustering events in the original data model; from the transformed data model we think these are background earthquakes. The dash line on the right plot is the 0.5 probability line.

5. Summary

In this paper we reviewed the Epidemic Type Aftershock-Sequence (ETAS) model and analyzed the 2008 Sichuan earthquake data by applying this model. We also put forward a location transformed method which can improve the performance of the space-time ETAS model, and verified this improvement by analyzing the 2010 Chile earthquake data. The location (Latitude and Longitude Coordinate) transformed method algorithm was also introduced. By this method, we firstly used locally weighted regression line to estimate the seismic belt in a seismic activity region. Based on this estimated seismic belt line, we established a new coordinate system to cancel the rotation angle of location distribution ellipses in the space-time ETAS model. The improvement was verified by the standard distance distribution method introduced by Zhuang and Ogata (2004). From this verification, we concluded that the improvement of the location transformed method seems to be not significant on a straight seismic belt region (such as the 2008 Sichuan earthquake region); however, when we applied this transformed method on a bending seismic belt (the 2010 Chile earthquake), the improvement was obvious.

We also compared the distribution of declustering events and clustering events in different mod-

els (the original data model and the transformed method model). The difference of classification between these two model is not obvious. In the 2010 Chile earthquake data set, there were seven points classified in different results by transformed method, while in the results of the 2008 Sichuan earthquake, the classification by two data sets were the same. From the clustering distribution in these two cases, we believe both of these two seismic activity regions are obviously potential hazard zones of strong earthquakes; the local governments of these regions should take action as possible as they can to reduce the losses in the future disaster.

The potential problem we still do not solve is the location transformed method in cross seismic belts regions, or how can we define the two vectors *Distance* (d) and *Length* (l) in a region with two or more seismic belts. Separating regions by seismic belts will lose the history information $\mathcal{H}_t$ in the self-exciting point process, and the combination of several seismic belts into one line would affect the semi-isotropy assumption on the location distribution.

A suggested development of this location transformed method is to use the physical world information about seismic belt instead of the estimation by locally weighted regression; also, the ellipse parameters σ_x and σ_y in the location distribution (16) should be constant and physically deterministic. The inference on these parameters in the space-time ETAS model is also an interesting problem.

References

[1] Alan, G. Hawkes (1971). Spectra of some self-exciting and mutually exciting point processes, Biometrik:a, 58, 1, p. 83.

[2] C. S. Jiang and Z. L. Wu (2011). PI forecast with or without de-clustering: an experiment for the Sichuan-Yunnan region. Nat. Hazards Earth Syst. Sci., 11, 697-706.

[3] Daley, D. J. and Vere-Jones, D. (1972). A summary of the theory of point processes. In Stochastic Point Processes, ed P.A.W. Lewis, Wiley-Interscience, New York, 299-383.

[4] F. Musmeci and D. Vere-Jones (1992). A space-time clustering model for historical earthquakes. Ann. Inst. Statist. Math. Vol. 44, No. 1, 1-11.

[5] F. Papangelou (1972). Integrability of expected increments of point processes and a related random change of scale. TRANS AMER MATH SOC , vol. 165, pp. 483-483.

[6] Gutenberg, B. and Richter, CF. (1956). Earthquake magnitude, intensity, energy and acceleration. Bul Seismol Soc Am 46: 105–145.

[7] Jiang, C. S. and Zhuang, J. C. (2010). Evaluation of background seismicity and potential source zones of strong earthquakes in the Sichuan-Yunan region base on the space-time ETAS model. Chinese J. Geophys. 53(2): 305-317.

[8] Meyer, P (1971). Demonstration Simplifiee d'un Theoreme de Knight. In Seminaire de Probabilites V, Universite Strasbourg, Lecture Notes in Mathematics, 191, 191–195.

[9] Nielsen, H. B. (2000) 'UCMINF - An Algorithm For Unconstrained, Nonlinear Optimization', Report IMM-REP-2000-18.

[10] Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. J. Am. Statist. Assoc. 83, 9–27.

[11] Ogata, Y. (1998). Space–time point-process models for earthquake occurrences. Ann. Inst. Stat. Math. 50, 379–402.

[12] Ogata, Y. and J. Zhuang (2006). Space-time ETAS models and an improved extension, Tectonophysics, 413, 13–23.

[13] Schoenberg, F. P. (2003). Multi-dimensional residual analysis of point process models for earthquake occurrences. J. Am. Statist. Ass., 98, 789–795.

[14] Smyth, C. and Mori, J. (2010). Model Ensembles for Prediction of Wenchuan Aftershock Activity. Bulletin of the Seismological Society of America, 100, 2532-2538.

[15] "USGS Earthquake Details". United States Geological Survey. Archived from the original on 1 March 2010. Retrieved 27 February 2010.

[16] UNEP Year Book 2011, An Overview of Our Changing Environment, United Nations Environment Programme 2011 page 2.

[17] Utsu, T. (1965). A method for determining the b-value in a formula log N = a-bm showing magnitude frequency relation for earthquakes. Hokkaido Univ Geophys Bull 13: 99–103.

[18] William S. Cleveland (1979). Robust Locally Weighted Regression and Smoothing Scatterplots, Journal of the American Statistical Association, 74:368, 829-836.

[19] Zhuang, J., Y. Ogata and D. Vere-Jones (2002). Stochastic declustering of space-time earthquake occurrences, Journal of the American Statistical Association, 97, 369-380.

[20] Zhuang, J., Y. Ogata (2004). Analyzing earthquake clustering features by using stochastic reconstruction, Journal of Geophysical Research, VOL. 109, B05301

[21] Zhuang, J., Ogata, Y. and Vere-Jones, D. (2005). Diagnostic analysis of space-time branching processes for earthquakes. Lecture Note in Statistics: Case Studies in Spatial Point Process Models, Springer-Verlag, New York, 185, 276–292.

Appendix: R code for the 2008 Sichuan earthquake data

```
#This ETAS package is rebuilt to add a parameter on etas() in order to
#fit ETAS model by the rotated ellipse location distribution.
library(ETAS)

#data preparation for the 2008 Sichuan earthquake
x = read.csv("wenchuan2008.txt")
names(x)
x = x[,c(1,3,2,5)]
x[which(x$Magnitude>=6),]
x = x[which(x$Latitude>=30.5),]
n = length(x$Latitude)
x$DateTime = as.character(x$DateTime)
x$DateTime = chartr("T"," ",x$DateTime)
x$DateTime = as.POSIXlt(x$DateTime)
x = x[order(x$DateTime),]
x1 = x[2:n,]
mag = sort(unique(x$Magnitude))
main.shock = x[1,]

#Map plot for the 2008 Sichuan earthquake
library(ggmap)
map2 = get_map(location = c(103, 30.5, 105.9,33),zoom = 8)
ggmap(map2) + geom_point(aes(x = Longitude, y = Latitude), size = 7.9,
          data = main.shock, colour = "red") +
          geom_point(aes(x = Longitude, y = Latitude,
          size = Magnitude), alpha = 0.5, data = x, colour = "black")

#scatter plot
par(mar=c(5.1, 4.1, 4.1, 8.1), xpd=TRUE)
plot(x1[,2:3],cex = (x1[,4]-min(x1[,4])),
     col = "gray",main ="")
points(main.shock[2:3],pch = "*",cex = 6)
legend.names = c("4.0","5.0","6.0","7.0")
```

```
legend("topright", inset=c(-0.27,0), legend=legend.names,
        pch = 1,col="gray",pt.cex = (4:7-min(x1[,4])),title = "Magnitude")

#GutenbergCRichter law plot
mc = 4.5
cum.n = sapply(35:70/10,function(i) sum(x$Magnitude >= i))
b = 1/(mean(x$Magnitude[which(x$Magnitude>=mc)])-mc+0.05)/log(10)
bm = mag[12:25]*b
cost = sapply(5000:10000/1000,function(a) sum((a - bm-log(cum.n,10)[9:22])^2))
a = (4999+which(cost == min(cost)))/1000
windows(width = 10, height = 5, record = T)
layout(matrix(c(1, 1, 2, 2), 2), c(4, 4), c(4, 4))
plot(35:70/10,log(cum.n,10),pch = 16,ylab = "Log(N)",xlab = "Magnitude")
lines(mag,a-mag*b,col = "red")
leg.names = "a - bM"
legend("topright",leg.names,col = "red",lty = 1)
arrows(mc-0.1,a - b*mc-0.5,mc,a - b*mc-0.1,angle = 20,lwd = 1)
text(mc-0.1,a - b*mc-0.6,paste("M0 = 4.5"),cex = 1)
text(4,.25,paste("b = ",1.33),cex = 1)
text(4.4,.5,"Cumulative Number of Events: 1834")

#temporal ETAS model
tm = x[which(x[,4] >= mc),c(1,4)]

#clustering lambda
lam.t = function(x,A,c,p,alpha,m0)
{
  n = length(x[,1])
  t = x[n,1]-x[1:(n-1),1]
  m = x[1:(n-1),2] - m0
  A*((p-1)/c)*sum(exp(alpha*m)*((t/c+1)^(-p)))
}
#integral of clustering lamda
f.t = function(x,A,c,p,alpha,m0)
{
  n = length(x[,1])
  t = x[n,1]-x[1:(n-1),1]
  t2 = t - x[n,1] + x[(n-1),1] + 0.00001
```

```
  m = x[1:(n-1),2] - m0
  -A*sum(exp(alpha*m)*((t/c+1)^(-p+1)))+ A*sum(exp(alpha*m)*((t2/c+1)^(-p+1)))
}
#log likelihood function for lambda
fi = function(beta)
{
  n = length(tm[,1])
  A = beta[1]
  c = beta[2]
  p = beta[3]
  mu = beta[5]
  alpha = beta[4]
  m0 = mc
  delta.t = tm[2:n,1] - tm[1:(n-1),1]
  Integ = sapply(2:n, function(i) f.t(tm[1:i,],A,c,p,alpha,m0))
  l = numeric(n)
  l[2:n] = sapply(2:n, function(i) lam.t(tm[1:i,],A,c,p,alpha,m0))
  l = l + mu
  -sum(log(l)) + sum(Integ) + mu*tm[n,1]
}

#MLE by BFGS algorithm
library(ucminf)
est = ucminf(c(0.20798366, 0.15579496, 1.23340819, 1.85185842, 0.05426907),
      fi,control = list(stepmax = 0.0000001))
tht = est$par
n = length(tm[,1])
#Integral of lambda
lt = function(beta)
{
  n = length(tm[,1])
  A = beta[1]
  c = beta[2]
  p = beta[3]
  alpha = beta[4]
  mu = beta[5]
  m0 = mc
  delta.t = tm[2:n,1] - tm[1:(n-1),1]
```

```
  l = sapply(2:n, function(i) f.t(tm[1:i,],A,c,p,alpha,m0))
  l + mu*delta.t
}
#K-S test for uniformly distributed residual process
ltt = lt(tht)
plot(cumsum(ltt),type = "l",lwd = 1)
lines(1:2200,1:2200,col = "red",lwd = 1)
ad.test(cumsum(ltt),punif,max = n)
ks.test(cumsum(ltt),punif,max = n)
k.alpha = 1.358*sqrt(n)
lines((0:2000) - k.alpha,lwd = 1,lty = 2,col = "gray")
lines((0:2000) + k.alpha,lty = 2,lwd = 1,col = "gray")

n = length(x$Latitude)
#Transform Longitude to the equator degree
x[,1] = date2day(x[,1],start = x[1,1])
tperiod <- c(0, max(x[,1]))
xrange = tperiod
yrange = c(min(x[,2]),max(x[,2]))
zrange = c(min(x[,3]),max(x[,3]))
domain = list(xrange,yrange,zrange)
names(domain) = c("xrange","yrange","zrange")
win <- owin(yrange,zrange)
x1 = ppx(x,domain = domain,coord.type=c("t","s","s","m"))
x = long2flat(x1,win)$X$data
x = data.frame(x)

#Cumlative number & magnitude against time
par(mar = c(4.5,5,0.2,2),cex.lab = 1.2,cex.axis = 1.2)
layout(matrix(c(1, 2, 1, 2), 2), c(2, 2), c(2, 0.8))
plot(x[,1],1:length(x[,1]),type = "l",xlab = "",
     ylab = "Cumulative Number of Events",lwd = 2)
plot(0,ylim = c(4,7.0),xlim = c(0,1500),xlab = "Time (Days)",ylab = "Magnitude")
temp = sapply(1:length(x[,1]),function(i) lines(c(x[i,1],x[i,1]),c(x[i,4],0)))

#lowess plot
plot(x[,2],x[,3])
low1 = lowess(x$Latitude~x$Longitude,f = 2/3)
```

```
low.n = length(low1$x)
plot(x$Longitude,x$Latitude,xlab = "Longitude",ylab = "Latitude",
     pch = ".",cex = 2,col = "gray")
lines(low1,col = "red",cex = 3)
legend("topleft","LOWESS",lty = 1,col = "red")
points(x[1,2:3],pch = "*",cex = 3)
arrows(89.7,32.1,89.5,32.4,angle = 20,lwd = 1)
text(89.7,32.0,"estimated seismic belt",cex = 1)

#distance and length, location transformed method
min.distance = function(X,low1)
{
  dist = ((X$Latitude-low1$y)^2)+((X$Longitude-low1$x)^2)
  p = which(dist == min(dist))
  p = p[1]
  res = sqrt(min(dist))
  if( X$Latitude - low1$y[p] - (X$Longitude - low1$x[p]) < 0)
  {
     res = -res
  }
  c(res,p)
)
min.distance1 = sapply(1:n,function(i) min.distance(x[i,],low1))
min.distance1 = unlist(min.distance1)
min.distance1 = t(min.distance1)
dis.p = min.distance1[,2]
low.length = sapply(1:n,function(i)
sum(sqrt(((low1$x[dis.p[i]:2]-low1$x[(dis.p[i]-1):1])^2)+
    ((low1$y[dis.p[i]:2]-low1$y[(dis.p[i]-1):1])^2))))

#rotated ellipse
library(plotrix)
A = sd(x[,3])/sd(x[,2])
B = 1/A
ang = atan(-2*cor(x[,2],x[,3])/(A-B))/2/pi*180
plot(x[,2],x[,3],xlab = "Longtiude",ylab = "Latitude",
     pch = ".",cex = 2,col = "gray")
lines(low1, col = "red", lwd = 1)
```

```
points(x[1,2:3],pch = "*",cex = 3)
legend("topleft","LOWESS",lty = 1,col = "red")
draw.ellipse(x[1,2],x[1,3],a = 1/sqrt(A)/2, b = 1/sqrt(B)/2, angle = ang)

# initial parameters values
param01 = c(0.710586536, 0.097032749, 0.109417969, 1.571568841,
            1.336733399, 0.008964095, 2.033918358, 0.256665745)

#Orignal space-time ETAS model
z = x[which(x$Magnitude >= 4.3),]
tperiod <- c(0, max(z[,1]))
xrange = tperiod
yrange = c(min(z[,2]),max(z[,2]))
zrange = c(min(z[,3]),max(z[,3]))
domain = list(xrange,yrange,zrange)
names(domain) = c("xrange","yrange","zrange")
win = owin(yrange,zrange)
z11 = ppx(z,domain = domain,coord.type=c("t","s","s","m"))
covr1 = c(sd(x[,2]),sd(x[,3]),cor(x[,2],x[,3]))
res = etas(z11, win, tperiod, m0 = 4.3, param0=param01, no.itr=1, covr = covr1)
int = lambda(z[,1], z[,2], z[,3], res$param, z11, 4.3,covr = covr1)
int[1:560] = int[1:560] + (res$param[1])*res$bk

#Location Transformed method for Space-time ETAS model
z2 = x
z2[,2] = min.distance1[,1]
z2[,3] = low.length
covr2 = c(sd(z[,2]), sd(z[,3]), 0)
z2 = z2[which(z2$Magnitude>=4.3),]
names(z2) = c("DateTime","Distance","Length","Magnitude")
tperiod = c(0, round(max(z2[,1])+1))
xrange = tperiod
yrange = c(min(z2[,2]),max(z2[,2]))
zrange = c(min(z2[,3]),max(z2[,3]))
domain = list(xrange,yrange,zrange)
names(domain) = c("xrange","yrange","zrange")
win = owin(yrange,zrange)
z22 = ppx(z2,domain = domain,coord.type=c("t","s","s","m"))
```

```
res1 = etas(z22, win, tperiod, m0 = 4.3, param0 = param01, no.itr = 1, covr = covr2)
int1 = lambda(z2[,1], z2[,2], z2[,3], res1$param, z22, 4.3,covr = covr2)
int1[1:560] = int1[1:560] + (res1$param[1])*res1$bk

#zeta for space-time ETAS model
zeta = function(data,pa,m0,covr,lambda)
{
  n = length(data[,1])
  t = data[n,1]-data[1:(n-1),1]
  x = data[n,2] - data[1:(n-1),2]
  y = data[n,3] - data[1:(n-1),3]
  m = data[1:(n-1),4] - m0
  A = pa[2]
  c = pa[3]
  alpha = pa[4]
  p = pa[5]
  D = pa[6]
  q = pa[7]
  gamma = pa[8]
  sig = D*exp(gamma*m)
  r2 = 1/sqrt(1-covr[3]^2)*((x^2)/covr[1]*covr[2]+(y^2)/covr[2]*covr[1]-2*x*y*covr[3])
  r = sqrt(r2/sig)
  zeta = A*(exp(alpha*m)*((p-1)/c)*((t/c+1)^(-p)))*(q-1)/pi/sig/((1+r^2)^q)
  rho = zeta/lambda
  return(rbind(r,rho))
}
#location reconstruct
fr = function(data,pa,m0,covr,lambda,bk,delta,pred)
{
  q = pa[7]
  m = length(pred)
  n = length(lambda)
  if(length(data[,1]) != n)
    stop("must be the same length")
  tot = (n - sum(c(0,bk*pa[1])/lambda))*delta
  wht = unlist(sapply(2:n,function(i) zeta(data[1:i,],pa,m0,covr,lambda[i])))
  r = wht[(1:sum(1:(n-1)))*2-1]
  rho = wht[(1:sum(1:(n-1)))*2]
```

```
  Y1 = sapply(1:m, function(i) sum(rho*(abs(pred[i]-r) < delta/2))/tot)
  plot(X,Y1,cex = 2,pch = "*")
  Y2 = 2*X*(q-1)/((1+X^2)^q)
  lines(X,Y2,col = "red")
  res = list(Y1,Y2)
  names(res) = c("est","real")
  return(cbind(Y1,Y2))
}
X = (0:100)/20
dens1 = fr(z, res$param, 4.3, covr1, int, res$bk, .03, X)
dens2 = fr(z2, res1$param, 4.3, covr2, int1, res1$bk, .03, X)
#model comparison
windows(width = 10, height = 5, record = T)
layout(matrix(c(1, 1, 2, 2), 2), c(4, 4), c(4, 4))
plot(X,dens1[,1],cex = 2,pch = "*",ylim = c(0,0.8),main = "original data",
     xlab = "Standardized distance",ylab = "Density")
lines(X,dens1[,2],col = "red")
plot(X,dens2[,1],cex = 2,pch = "*",ylim = c(0,0.8),main = "transformed data",
     xlab = "Standardized distance",ylab = "Density")
lines(X,dens2[,2],col = "red")

#clustering probabilities and declustering probabilities
x1 = x[which(x$Magnitude >= 4.3),]
x1 = x1[-561,]
x1$pb = res$pb
x1.bg = x1[which(x1$pb > 0.5),]
x1.cl = x1[which(x1$pb <= 0.5),]
x1$tpb = res1$pb
x1.tbg = x1[which(x1$tpb > 0.5),]
x1.tcl = x1[which(x1$tpb <= 0.5),]

#clustering region and declustering region
d = qplot(Longitude, Latitude, data= x1.bg, colour = pb, size = Magnitude,
          alpha = 0.5, main = "Original Data (Background)")
d + guides(alpha = guide_legend(title = "Compositing")) + scale_size(range=c(5,12)) +
    scale_colour_gradient2(limits = c(0,1), name = "Probability")

d = qplot(Longitude, Latitude, data= x1.cl, colour = 1 - pb, size = Magnitude,
```

```
          alpha = 0.5, main = "Original Data (Clustering)")
d + guides(alpha = guide_legend(title = "Compositing")) + scale_size(range=c(5,12)) +
    scale_colour_gradient2(limits = c(0,1), name = "Probability")

d = qplot(Longitude, Latitude, data= x1.tbg, colour = tpb, size = Magnitude,
          alpha = 0.5, main = "Transformed Data (Background)")
d + guides(alpha = guide_legend(title = "Compositing")) + scale_size(range=c(5,12)) +
           scale_colour_gradient2(limits = c(0,1), name = "Probability")

d = qplot(Longitude, Latitude, data= x1.tcl, colour = 1 - tpb, size = Magnitude,
          alpha = 0.5, main = "Transformed Data (Clustering)")
d + guides(alpha = guide_legend(title = "Compositing")) + scale_size(range=c(5,12)) +
    scale_colour_gradient2(limits = c(0,1), name = "Probability")
```