Rodolfo Console
Centro di Geomorfologia Integrata
per l'Area del Mediterraneo
Potenza, Italy

Real-time earthquake hazard estimate for low probability events

9th Scientific Meeting of the Classification and Data Analysis Group of the Italian Statistical Society
Modena, September 18-20, 2013

***Short-term time dependent model* (*epidemic model*) I**

Every event is potentially triggered by all the previous events and every event can trigger subsequent events according to their relative time-space distance

A definition of the words ***foreshock***, ***mainshock*** *and* ***aftershock*** is not necessary

Time dependent model (epidemic model) II

$$\lambda(x,y,t,m) = f_r \cdot \lambda_0(x,y,m) + \sum_{i=1}^{N} H(t-t_i) \cdot \lambda_i(x,y,t,m)$$

where

$$\lambda_i(x,y,t,m) = K(x,y,x_i,y_i)\, g(t-t_i)\, h(m)$$

Time dependent model (epidemic model) III

$$\lambda(x,y,t,m) = f_r \cdot \lambda_0 (x,y,m) + \sum_{t_i<t} K\, f(x,y,x_i,y_i)\, g(t-t_i)\, h(m)$$

Time independent distribution of the epicenters and magnitude for the spontaneous seismicity

$$\lambda_0 (x, y, m) = \mu_0 (x, y)\, \beta\, e^{-\beta(m-m_0)}$$

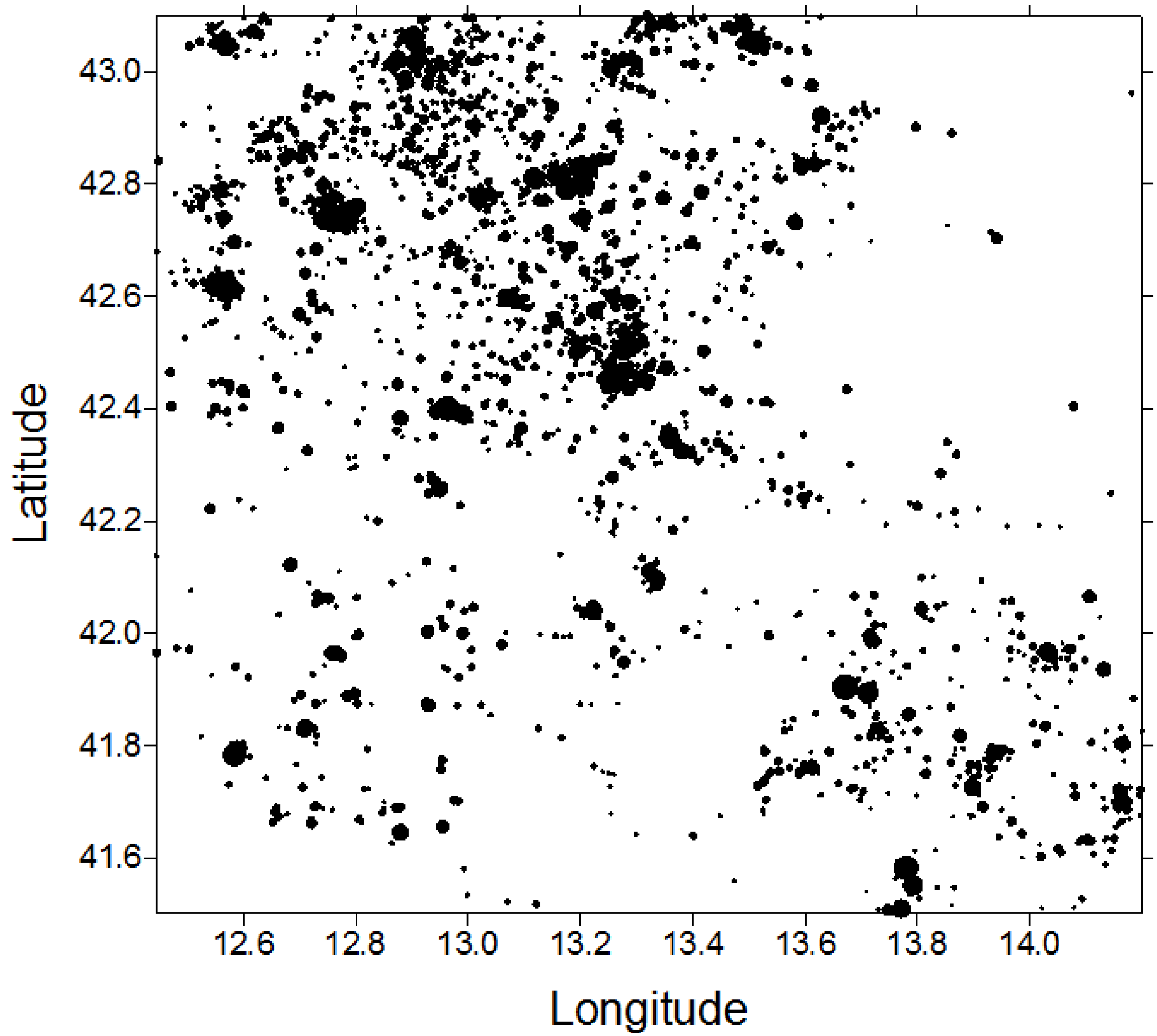
where m_0 *is the completeness magnitude of the catalog*

A smooth geographical distribution is computed at each node *k* of a regular grid through the method introduced by Frankel (1995):

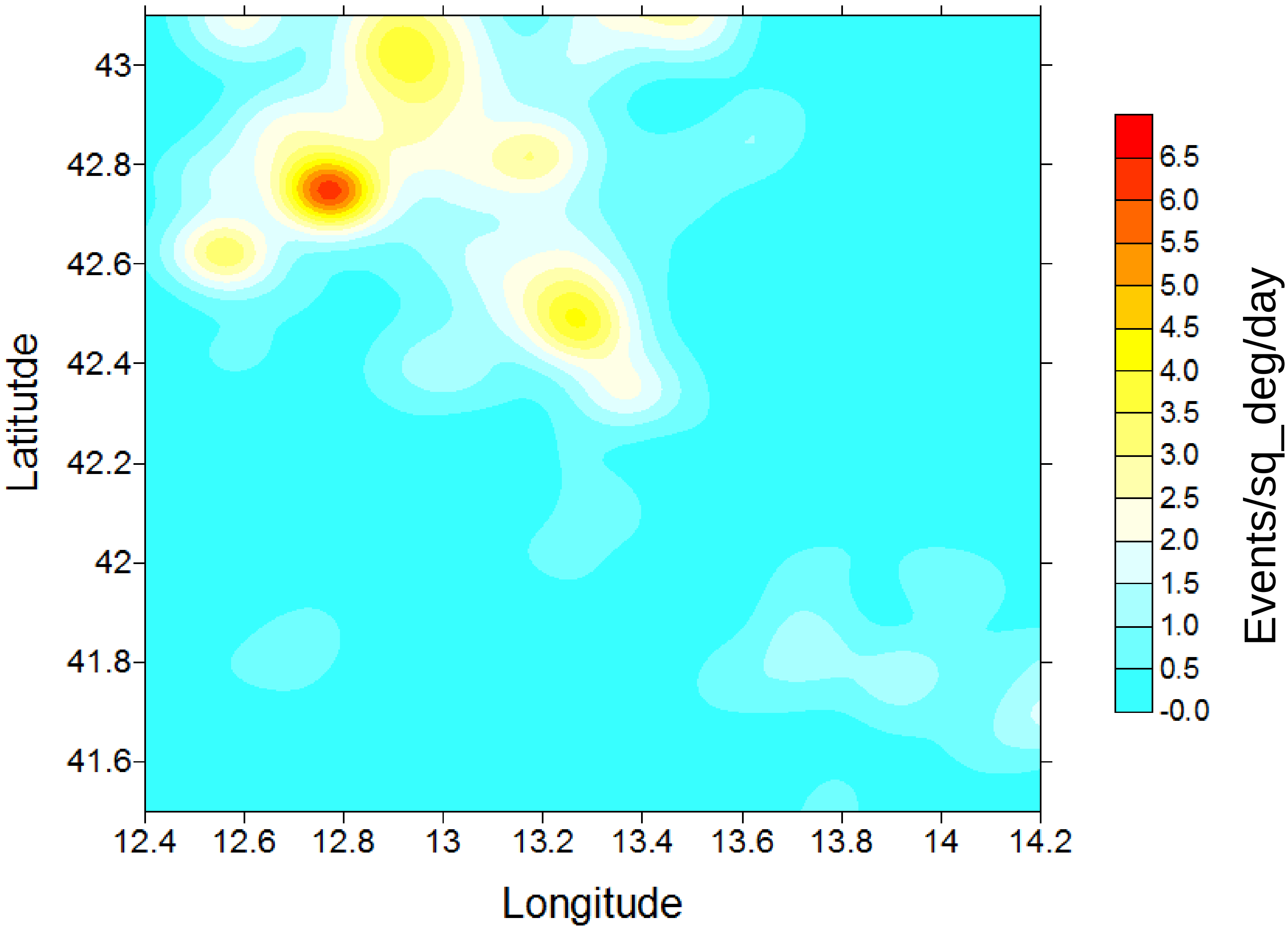
$$\tilde{N}_k = \frac{\Sigma_l N_l \exp(-\Delta_{kl}{}^2 / d^2)}{\Sigma_l \exp(-\Delta_{kl}{}^2 / d^2)}$$

The rate density $\lambda_0(x,y)$ at any point is obtained by linear interpolation among the nearest grid nodes. The free parameter *d* is determined maximizing the likelihood of the seismicity contained in half catalog under the model obtained from the other half.

Learning period: 17/04/2005-15/03/2009
1431 days, 2588 events M ≥1.6



Learning period: 17/04/2005-15/03/2009
Initial spatial distribution (function $\mu_0(x, y)$)
(smoothing distance d = 8 km)



Time dependent model (epidemic model) IV

$$\lambda(x,y,t,m) = f_r \cdot \lambda_0(x,y,m) + \sum_{t_i<t} K\, f(x,y,x_i,y_i)\, g(t-t_i)\, h(m)$$

Spatial distribution (kernel) of triggered events

$$f(x,y,x_i,y_i) = \left(\frac{d_i^{\,2}}{r^2 + d_i^{\,2}} \right)^q$$

where $$d_i = d_0 10^{\alpha(m_i - m_0)}$$

and $$r = \sqrt{(x-x_i)^2 + (y-y_i)^2}$$

Time dependent model
(*epidemic model*) *V*

$$\lambda(x,y,t,m) = f_r \cdot \lambda_0(x,y,m) + \sum_{t_i<t} K\, f(x,y,x_i,y_i)\, g(t-t_i)\, h(m)$$

Temporal distribution of triggered events
(*Omori –Utsu law*)

$$g(t-t_i) = (p-1)\, c^{(p-1)}\, (t-t_i+c)^{-p} \qquad (p \neq 1)$$

Time dependent model (epidemic model) VI

$$\lambda(x,y,t,m) = f_r \cdot \lambda_0(x,y,m) + \sum_{t_i<t} K\, f(x,y,x_i,y_i)\, g(t-t_i)\, h(m)$$

Magnitude distribution of triggered events

$$\boldsymbol{h(m) = \beta\, e^{-\beta(m-m_0)}}$$

Iterative adjustment of the background component spatial distribution

We find the maximum likelihood set of free parameters using the initial distribution of the smoothed seismicity $\lambda_0(x,y,m)$.

We define the probability of independence p_i as the ratio between the independent component $f_r \cdot \lambda_0(x_i,y_i,m_i)$ and the composite rate density $\lambda(x_i,y_i,m_i,t_i)$ for any event i.

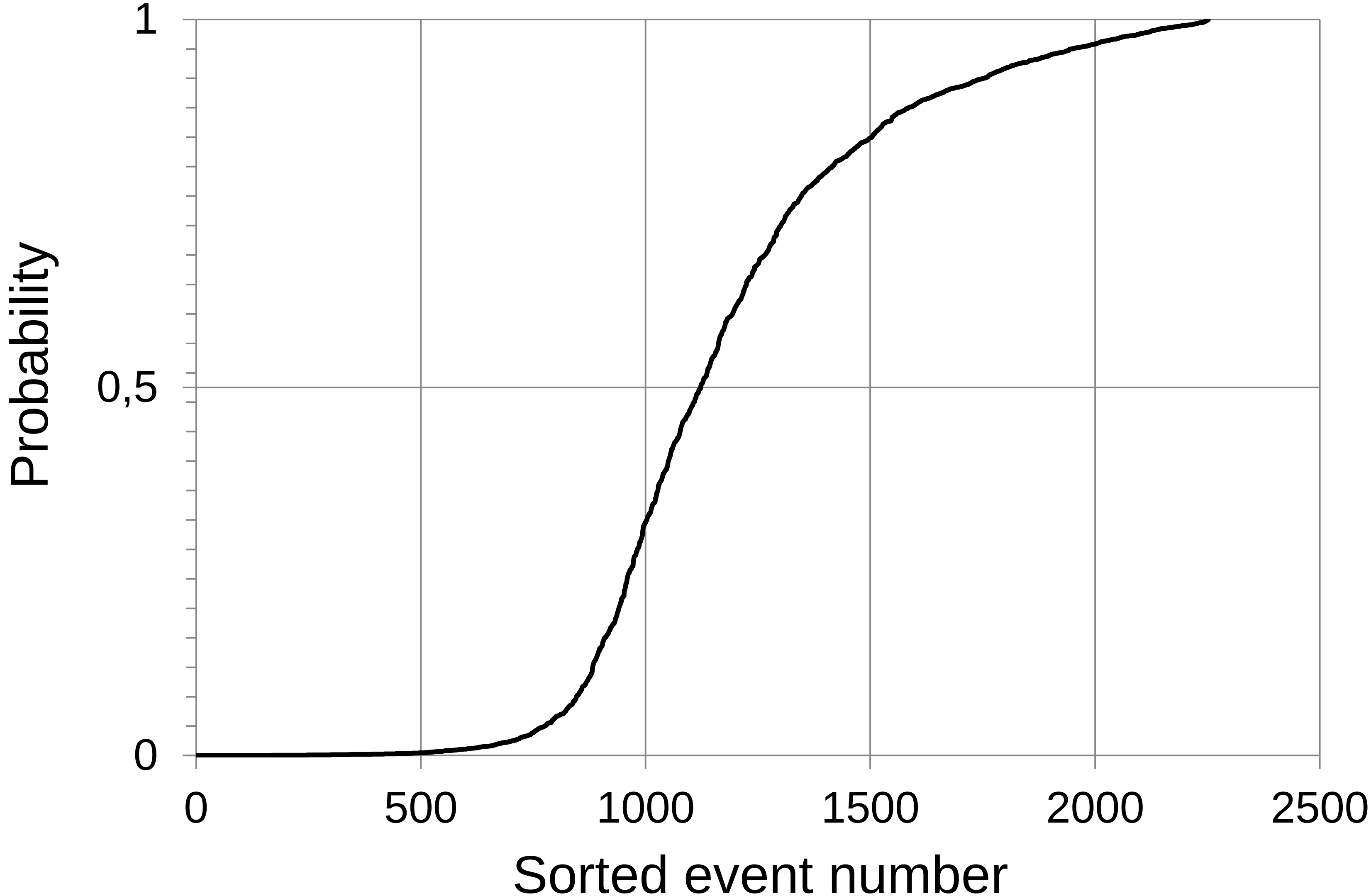
Then we compute a new distribution of $\lambda_0(x,y,m)$ introducing the weights p_i to count the number of events for each cell in the Frankel (1994) algorithm. The new distribution is normalized to the observed total number of events

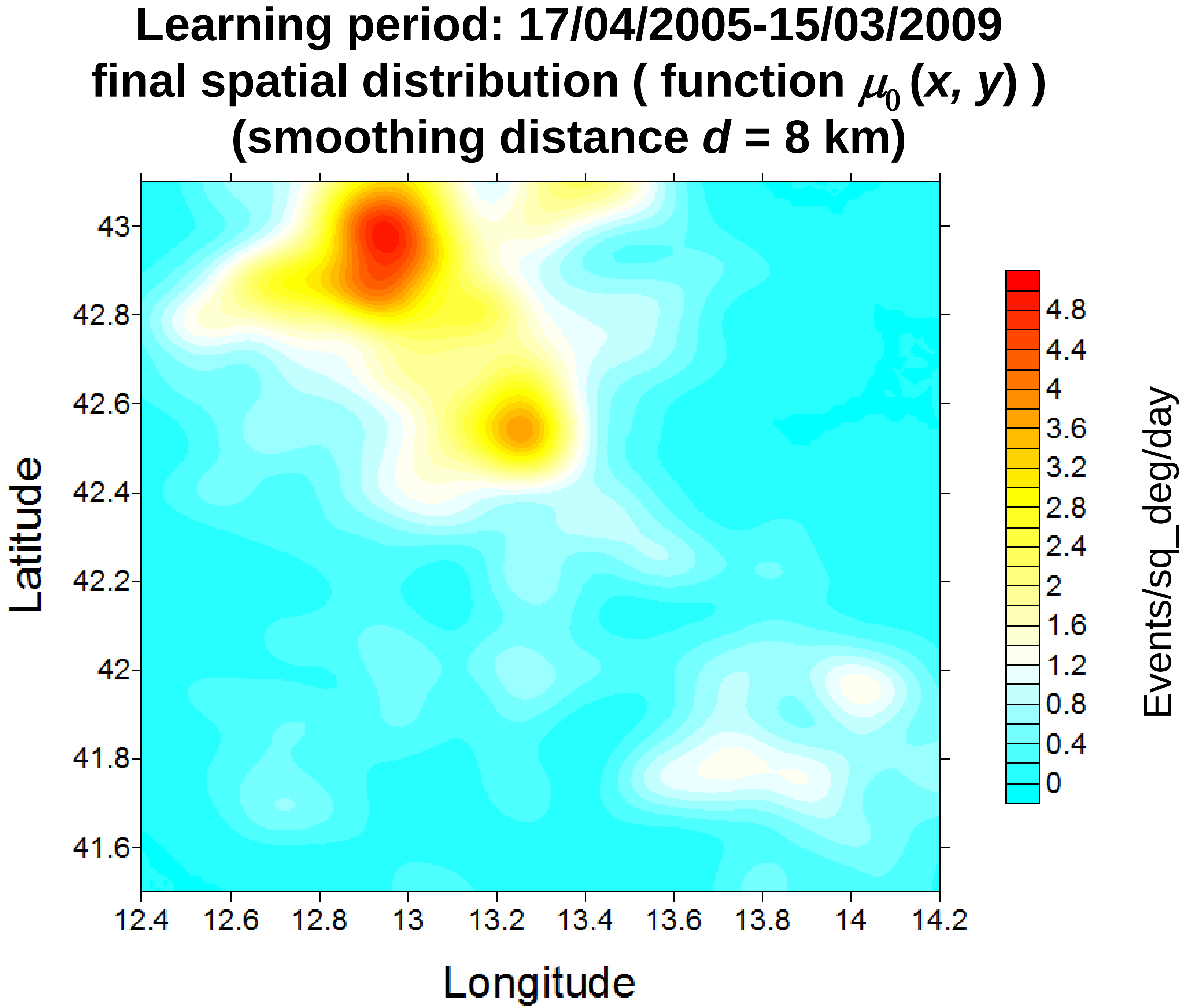
The new smoothed distribution is used in a new maximum likelihood best fit of the free parameters, and so on...

Values obtained for the parameters of the epidemic model in the iterative adjustment of the background seismicity

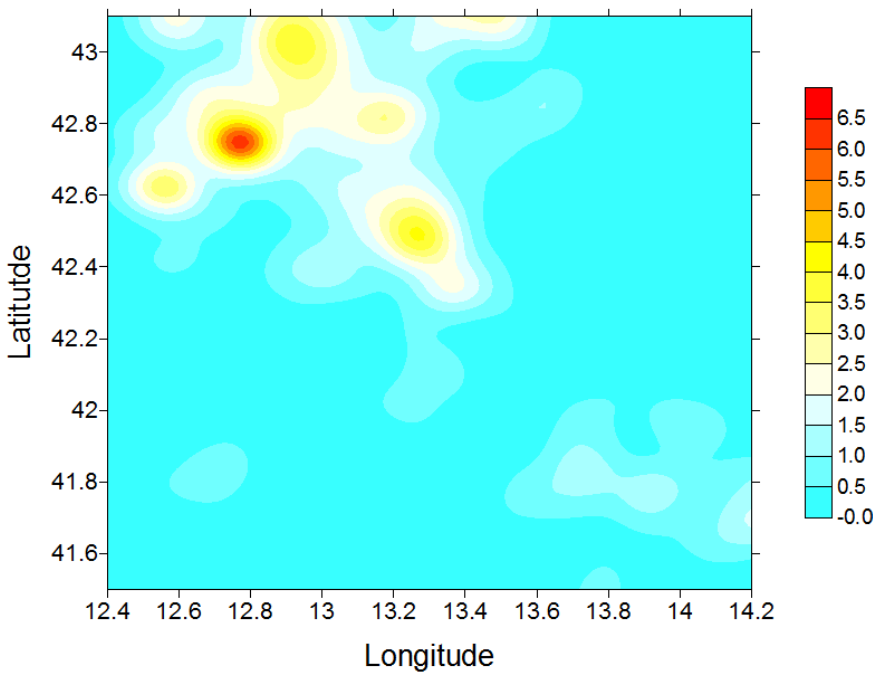
	Initial	Step 2	Step 3	Step 4
K	0.2383	0.2383	0.2384	0.2382
d_0	0.5215	0.5215	0.5215	0.5215
q	2.009	2.009	2.008	2.009
c	0.02004	0.01998	0.02005	0.02001
p	1.1047	1.1047	1.1047	1.1047
α	1.000	1.000	1.000	1.000
f_r	0.4796	0.4803	0.4796	0.4806
$\log L$	26954.2	27075.0	27080.0	27079.1

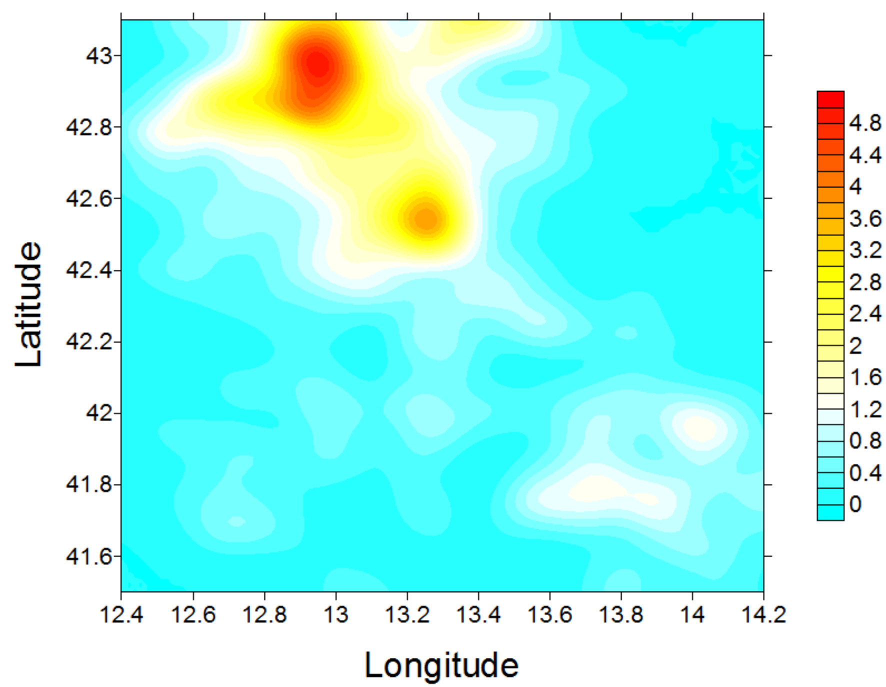
Distribution of probability for an event to be spontaneous



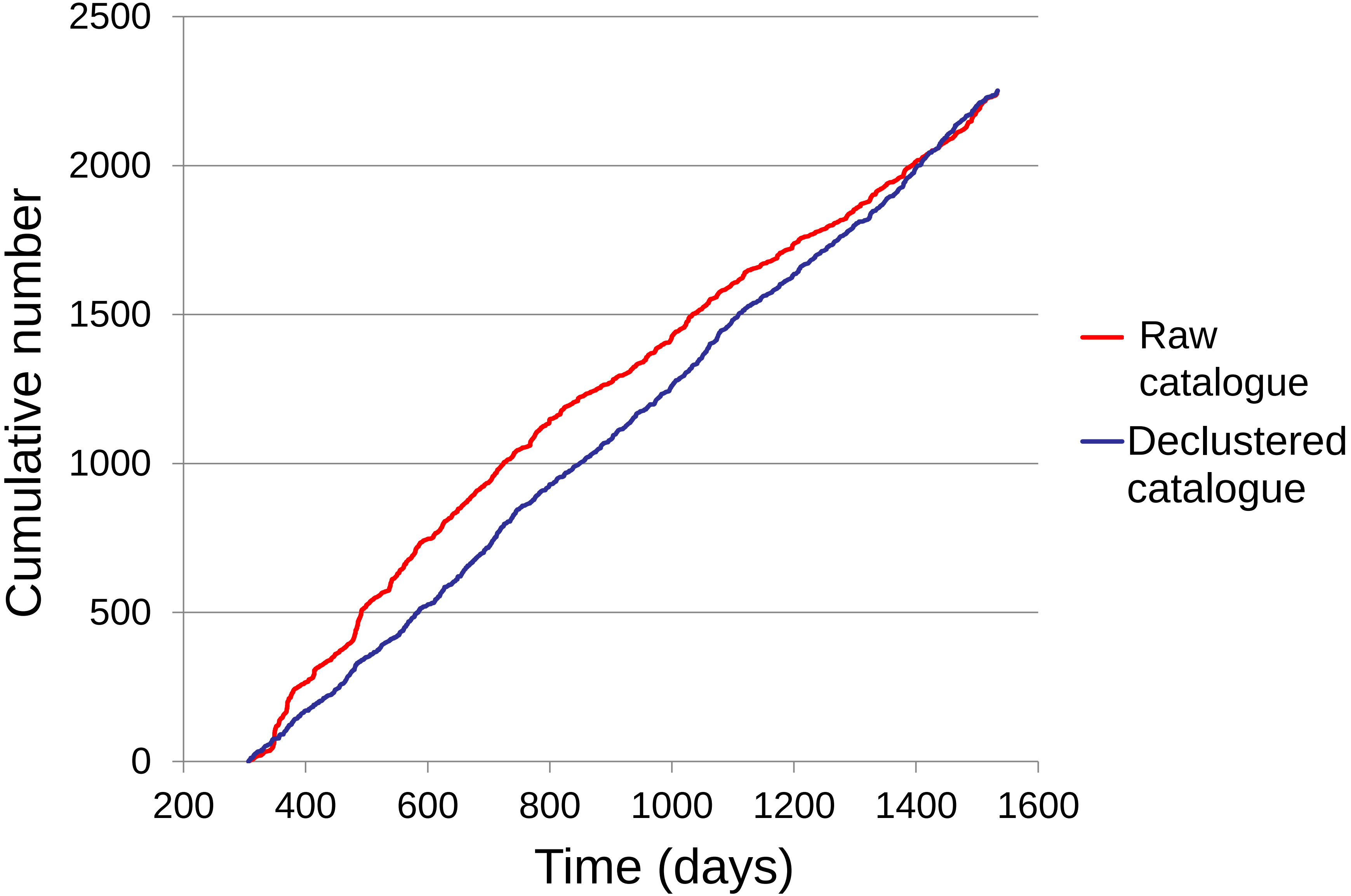

Learning period: 17/04/2005-15/03/2009
final spatial distribution (function $\mu_0(x, y)$)
(smoothing distance d = 8 km)
Latitude
43
42.8
42.6
42.4
42.2
42
41.8
41.6
Longitude
12.4
12.6
12.8
13
13.2
13.4
13.6
13.8
14
14.2
4.8
4.4
4
3.6
3.2
2.8
2.4
2
1.6
1.2
0.8
0.4
0
Events/sq_deg/day

Learning period: 17/04/2005-15/03/2009
Comparison between the inital and final spatial distribution (ev/sq_deg/day)





We draw the cumulative distribution of p_i over the time spanned by the catalog, expecting that it should be closer to a linear trend.



We define a parameter D_n to express the mismatch between the actual cumulative distribution and the expected linear trend with the same final value:

$$D_n = \max\left|\hat{F}_n(i) - F_0(i)\right|$$

where $\hat{F}_n(i)$ is the observed value of the normalized cumulative distribution at the event *i*

and $F_0(i)$ is the theoretical normalized linear trend,

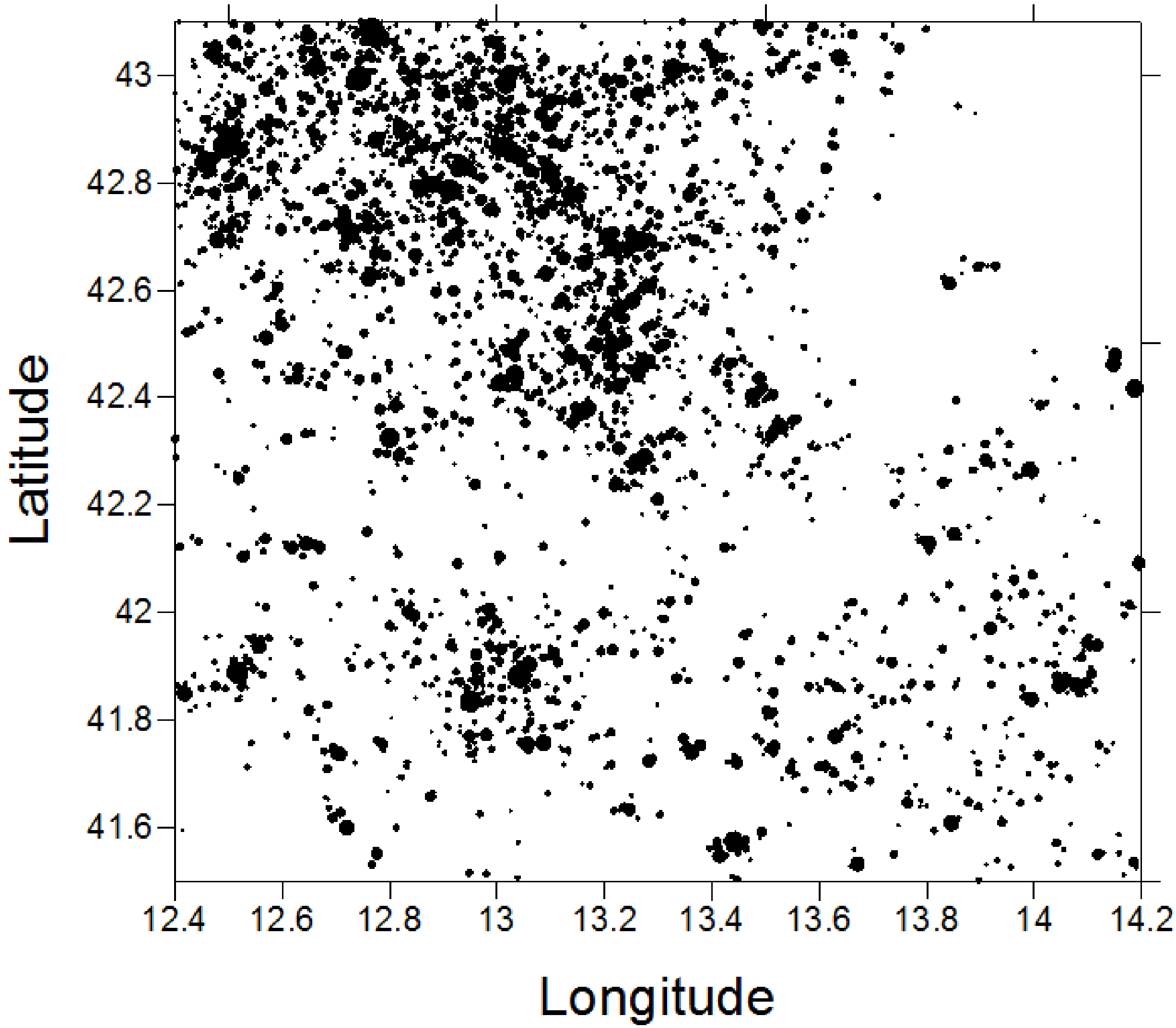
so that $\hat{F}_n(n) = F_0(n)$

KS-test

Raw catalogue	Dn = 0.108
Declustered catalogue	Dn = 0.019

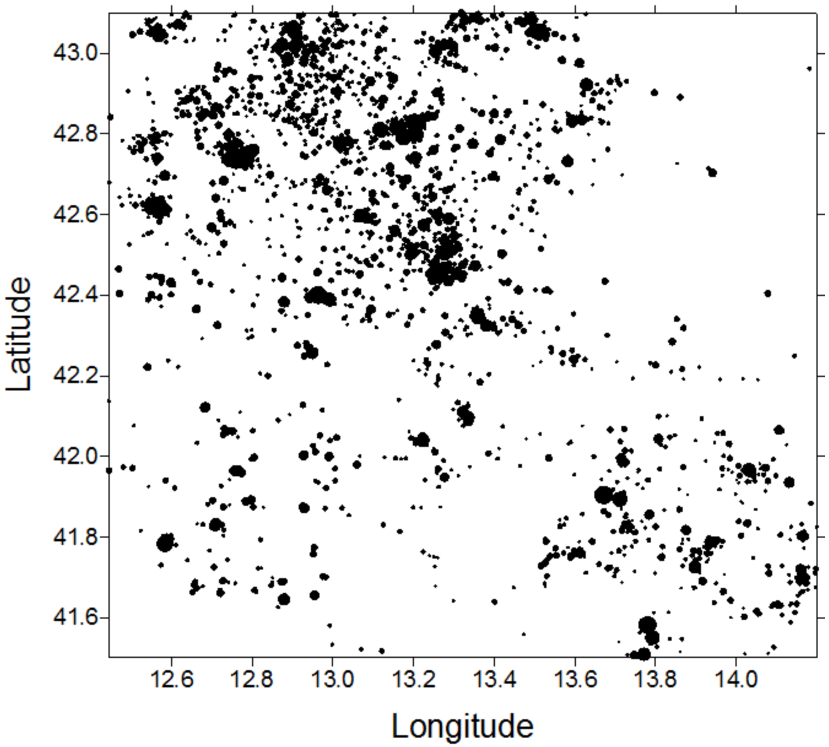
Level of significance	Critical value for Dn	Result Raw	Result Decl.
20%	0.023	Reject	Accept
10%	0.026	Reject	Accept
5%	0.029	Reject	Accept
1%	0.034	Reject	Accept

Synthetic catalog
1425 days, 3759 events M ≥1.6

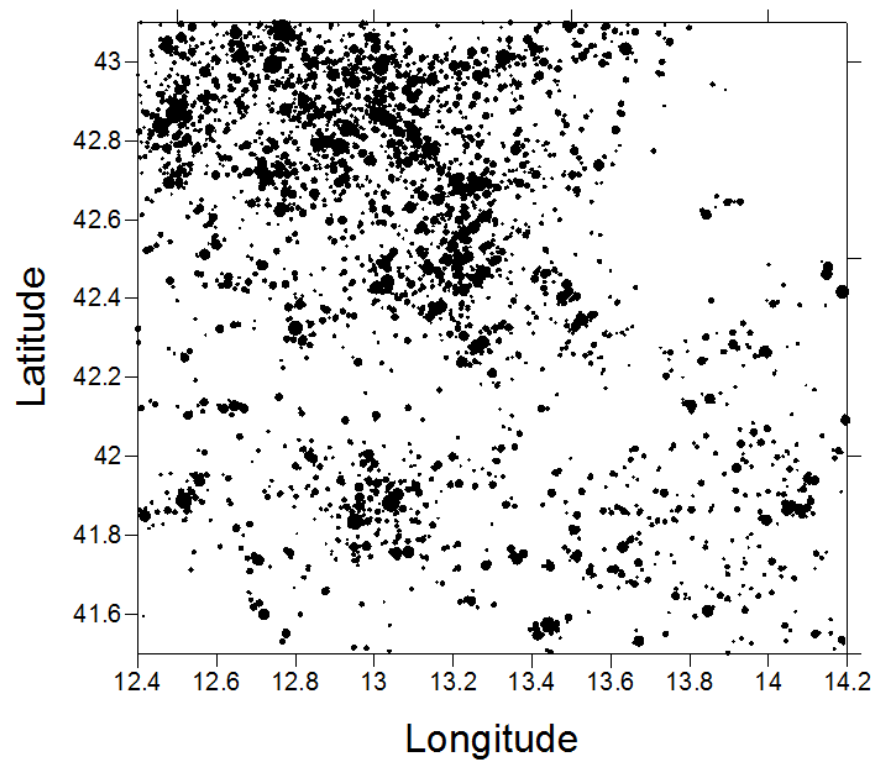


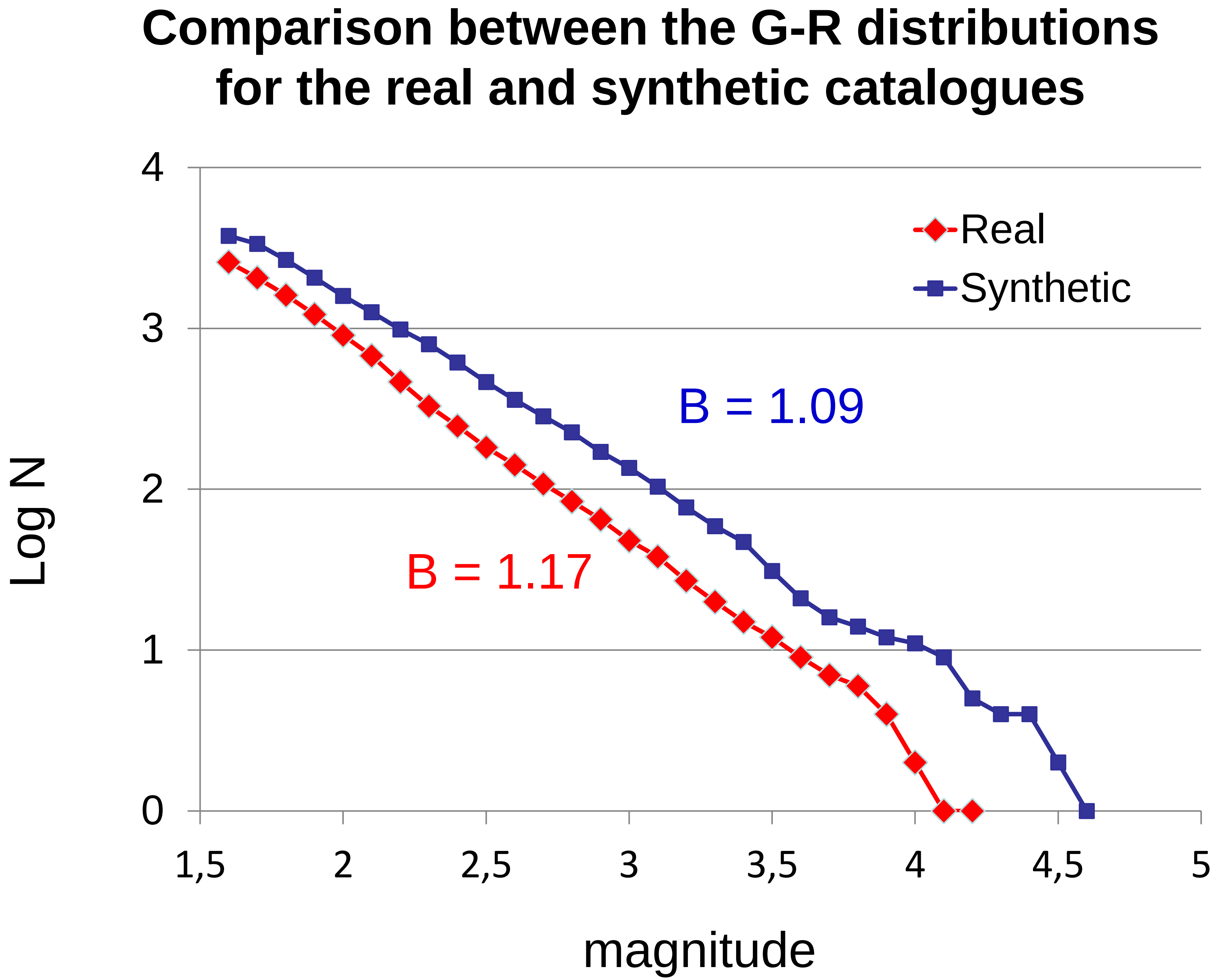
Comparison between the real and synthetic catalogues

2588 events M ≥1.6

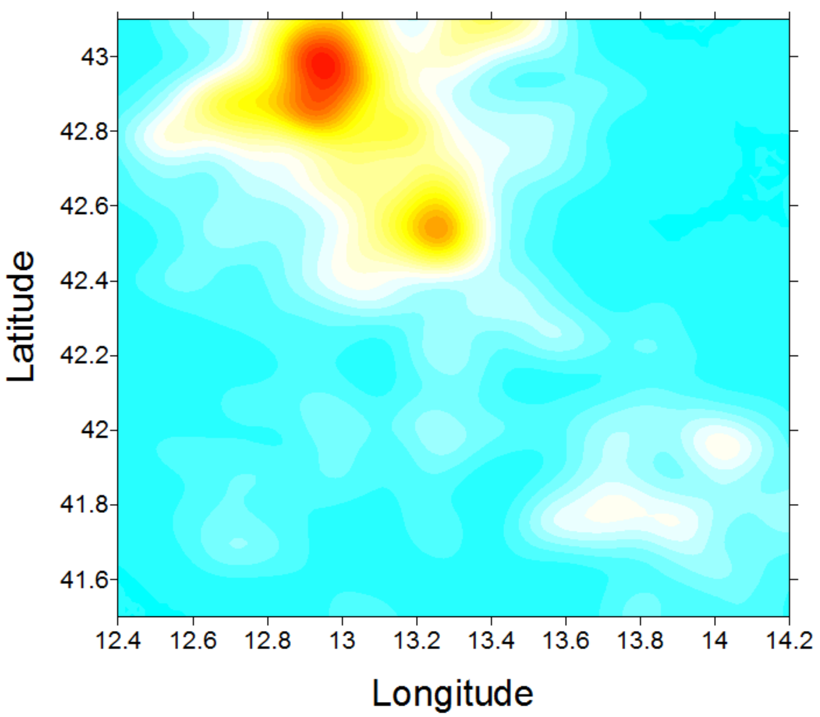


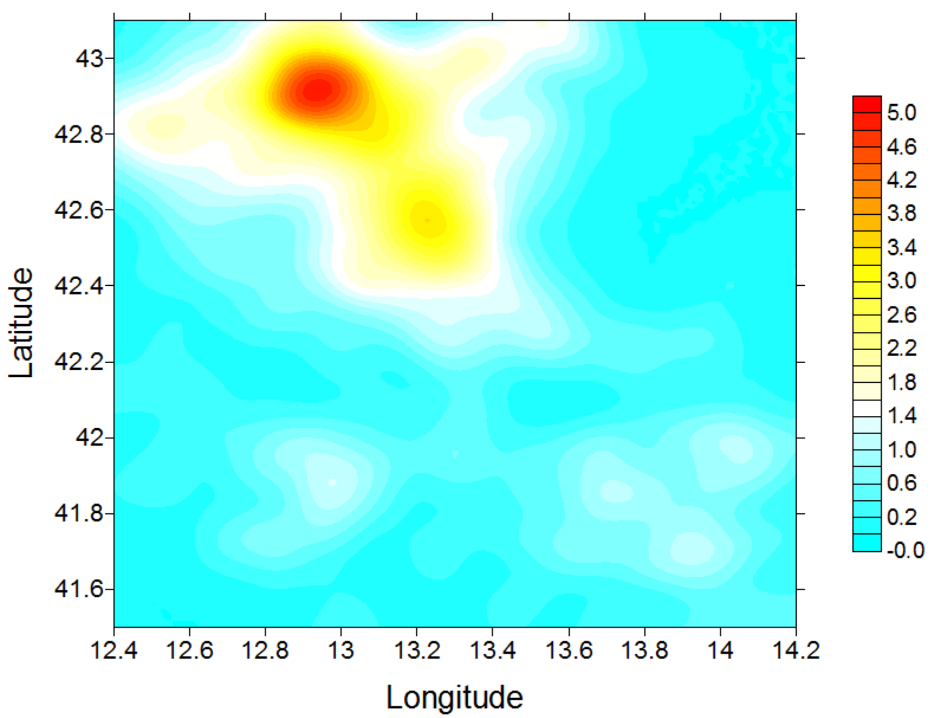
3759 events M ≥1.6




Comparison between the G-R distributions
for the real and synthetic catalogues
Log N
magnitude
Real
Synthetic
B = 1.09
B = 1.17
0
1
2
3
4
1,5
2
2,5
3
3,5
4
4,5
5

Comparison between the real and synthetic spatial distribution (ev/sq_deg/day)

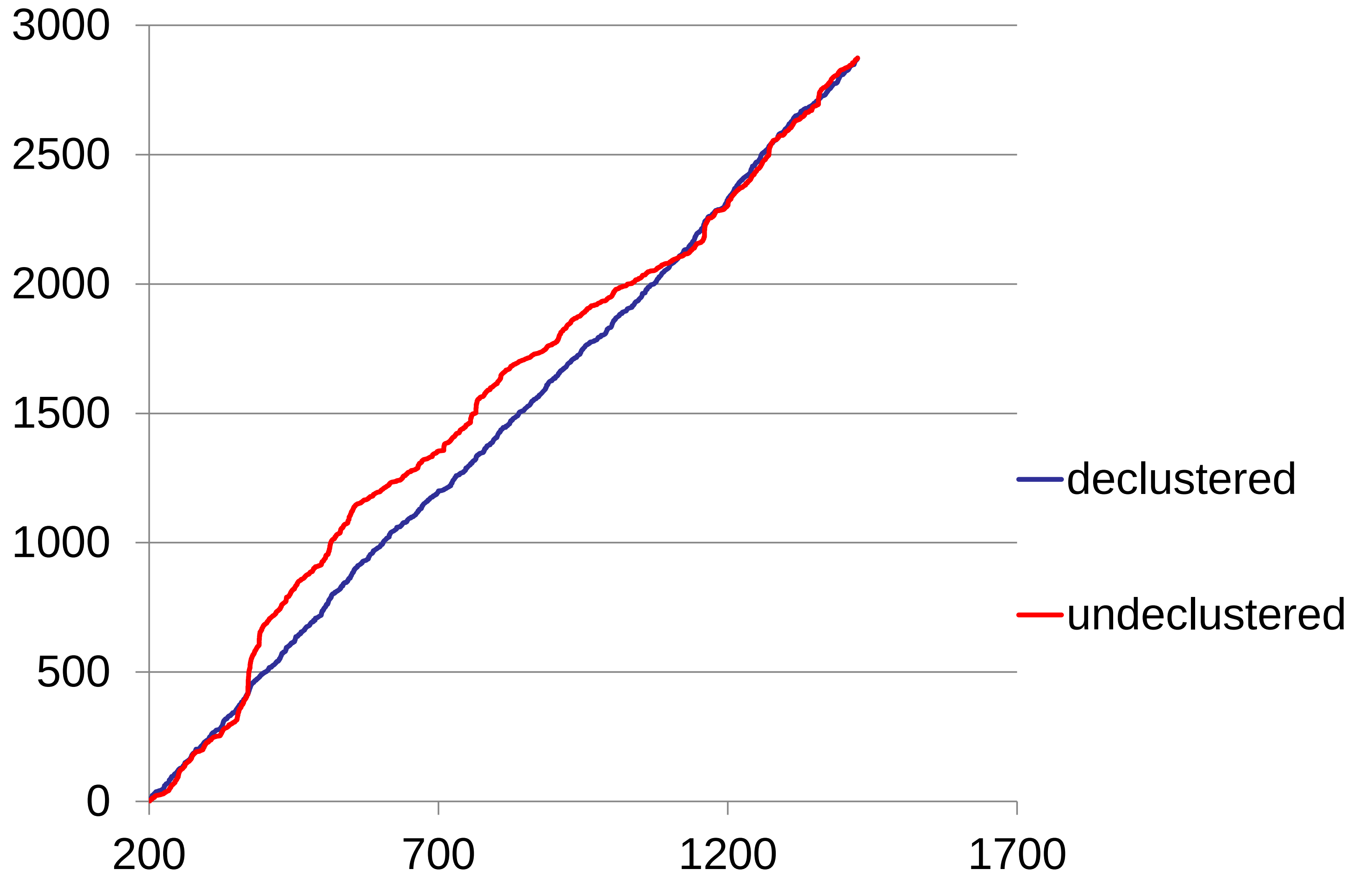




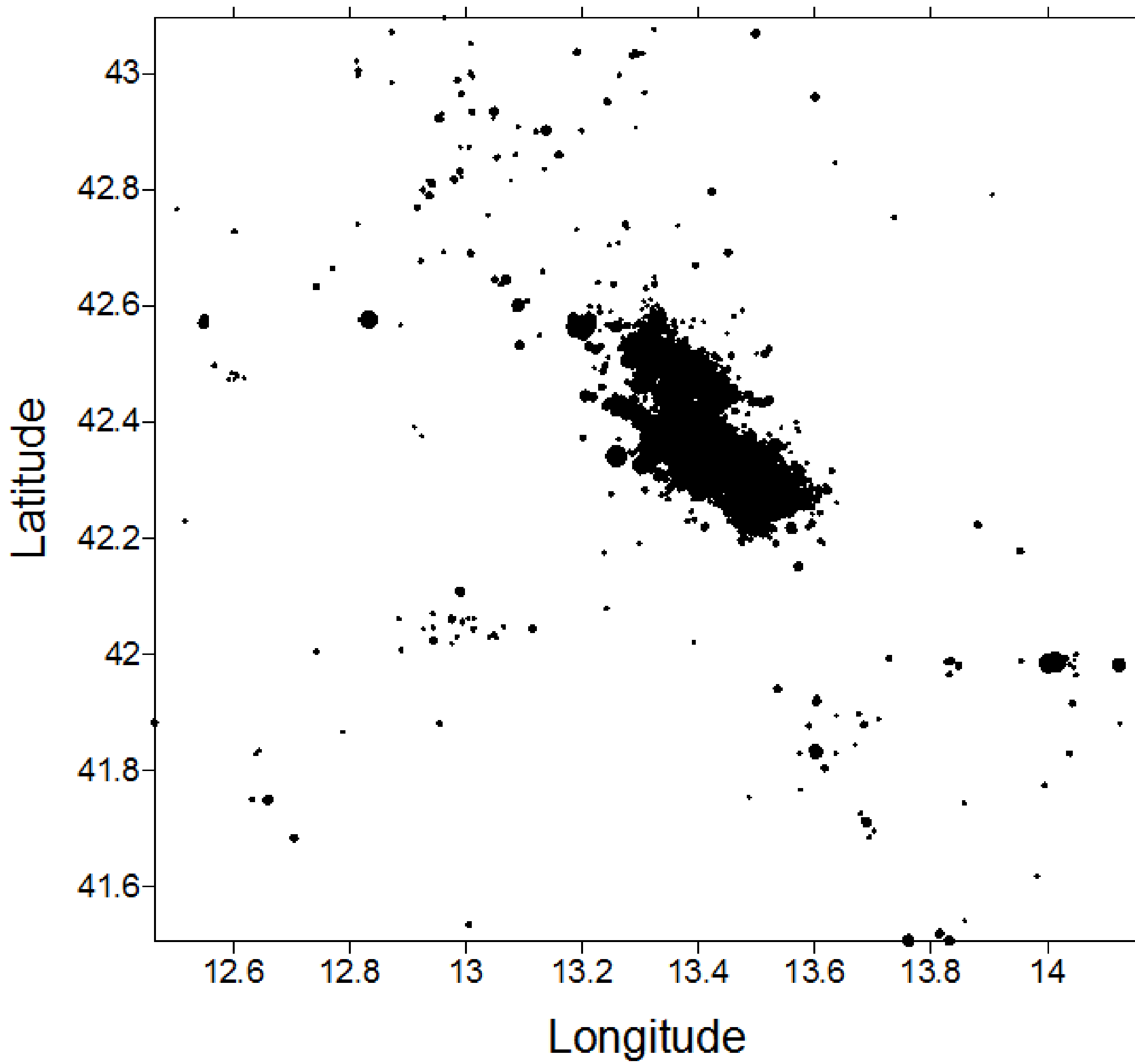
Comparison between the ML best fit parameters for the real and synthetic catalogues

	Real	Synthetic
K	0.2382	0.1697
d_0	0.5215	0.5224
q	2.009	2.008
c	0.02001	0.03495
p	1.1047	1.2177
α	1.000	1.000
f_r	0.4806	0.3286

Comparison between the raw and declustered cumulative distributions of events for the synthetic catalogue

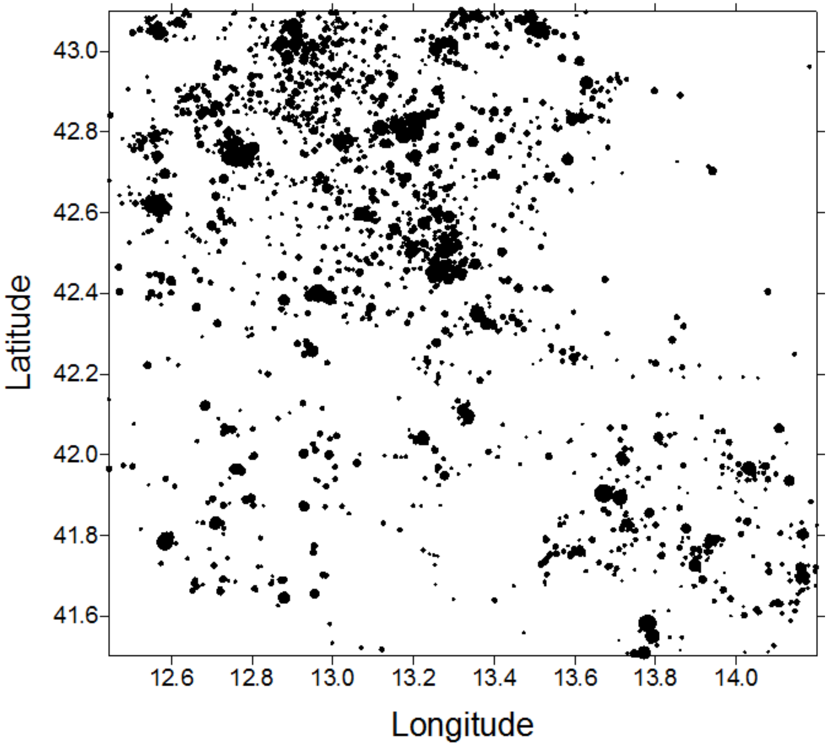


Testing period: 16/03/2009-30/06/2009
7149 events M ≥1.6

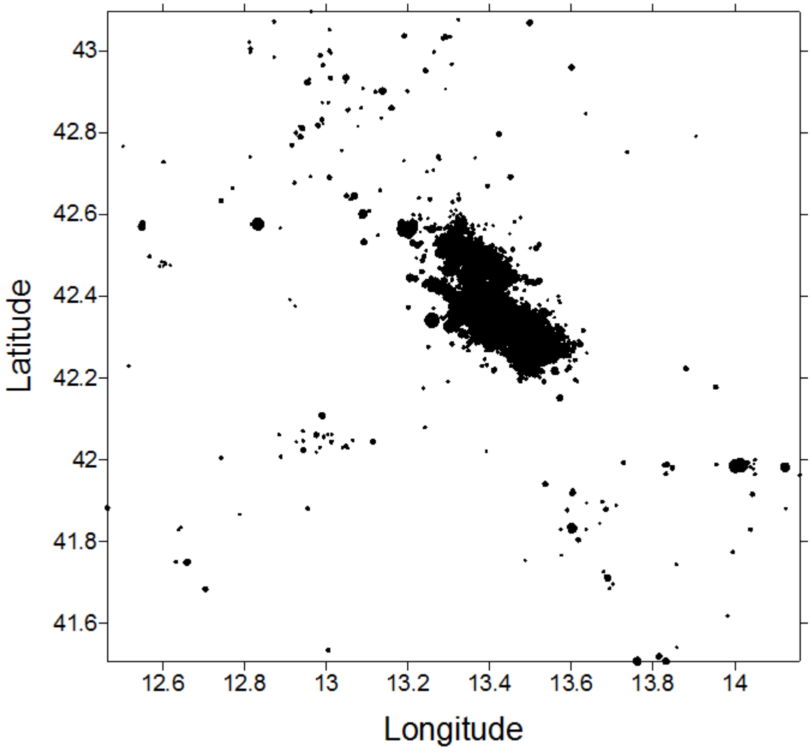


Comparison between the spatial distributions in the learning and testing periods

17/04/2005-15/03/2009



16/03/2009-30/06/2009

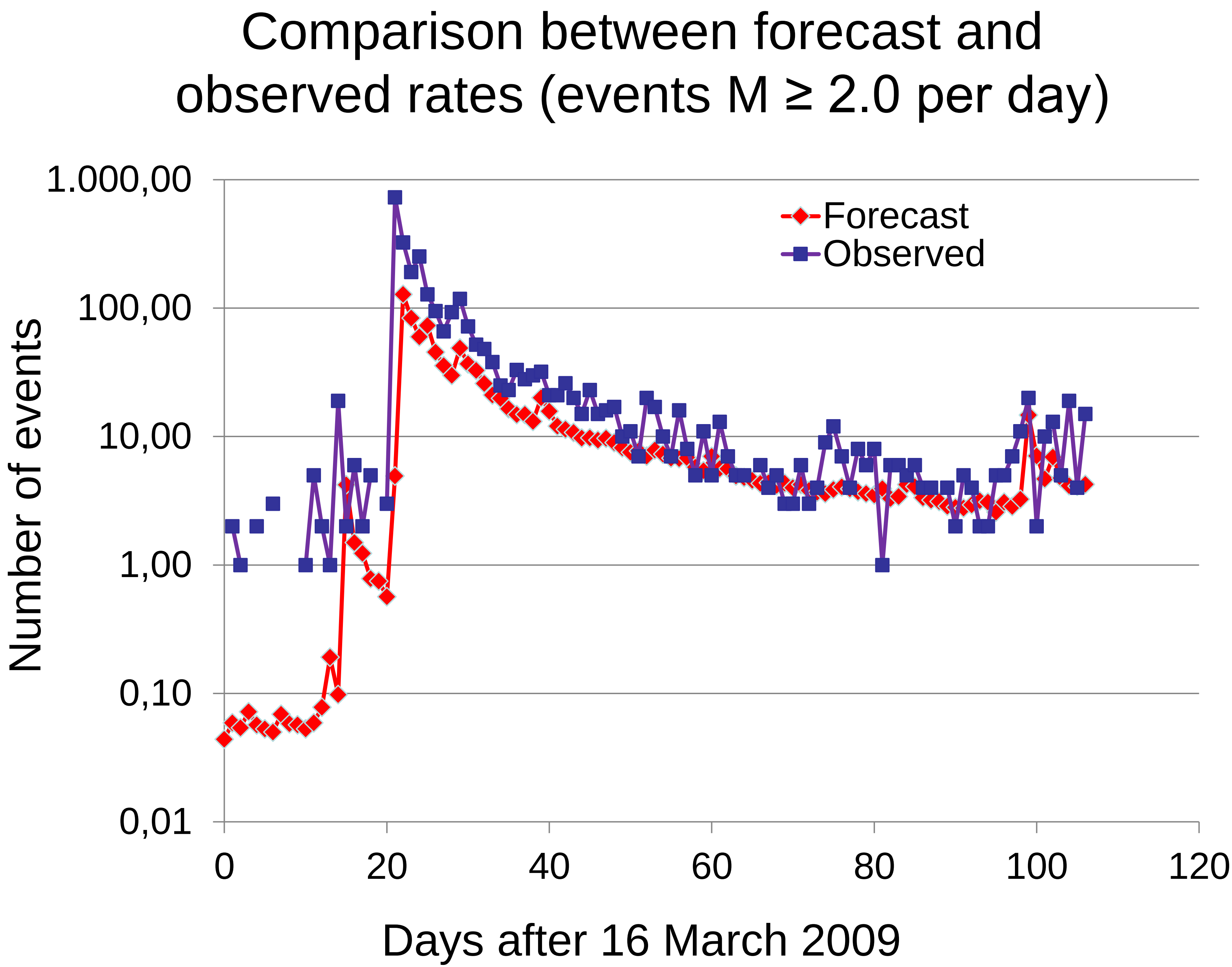


Overall results

Applying the ETAS model with the parameters obtained from the ML best fit in the learning phase, we obtain a very large performance factor with respect to the time-independent Poisson model.

The average log-performance factor per event (information gain) is equal to 50,241/7,149 = 7.03 (natural logarithms are used).

It means that for each event the average probability gain is of the order of 1,000.



Comparison between forecast and observed rates (events M ≥ 2.0 per day)
Forecast
Observed
Number of events
1.000,00
100,00
10,00
1,00
0,10
0,01
0
20
40
60
80
100
120
Days after 16 March 2009

Before the mainshock

The total conditional probability for an earthquake of magnitude $M \geq 5.0$ during the week preceding the 5 April mainshock was 0.39 %.
This probability was about 30 times larger than the background probability, due to the occurrence of some “foreshock” activity. However, this level seems still low for justifying the implementation of effective risk mitigation measures.
The expected instantaneous occurrence rate density increased by several times in the few hours before the mainshock

After the mainshock

The forecasted number of events with M≥5.0 was systematically smaller than the real one in the first month of the aftershock sequence. Afterwards, the forecasted and observed occurrence rates became more similar.

CONCLUSIONS

The ETAS model allows statistical declustering of a seismic catalogue and the simulation of catalogues with characteristics similar to the real ones.

The retrospective application of the ETAS model to the 2009 seismic sequence occurred in Central Italy has shown its capability of forecasting the behaviour of seismic activity during an aftershock sequence.

However, despite the fairly high probability gain achieved through the ETAS model, the forecast of main shocks preceded by moderate foreshocks is characterized by rather low occurrence rates for magnitudes larger than 5.0.

Thank you !