No. 35

Online ISSN: 1882-2274

COMPUTER SCIENCE MONOGRAPHS

A Publication
of
The Institute of Statistical Mathematics

HIERARCHICAL SPACE-TIME POINT-PROCESS MODELS (HIST-PPM): SOFTWARE DOCUMENTATION

by

Y. Ogata, K. Katsura, M. Tanemura, D. Harte and J. Zhuang

February 2021

COMPUTER SCIENCE MONOGRAPHS

The series of *Computer Science Monographs* publishes the statistical software and the articles on computer applications originating at The Institute of Statistical Mathematics. All communications relating to this publication should be addressed to the Editorial Office, The Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan.

The electronic version of this issue is downloadable at
https://www.ism.ac.jp/

Hierarchical Space-Time Point-Process Models (HIST-PPM): Software Documentation

Y. OGATA, K. KATSURA, M. TANEMURA, D. HARTE AND J. ZHUANG

Email: `ogata@ism.ac.jp`

Revised: 21 September 2020
Beta Version: 28 March 2015

Abstract

This documentation describes some FORTRAN and R programs used for fitting and displaying the Hierarchical Space-Time ETAS (HIST-ETAS) models, 2D spatial Poisson processes, 1D space vs time Poisson processes and location-dependent *b*-value estimates. The FORTRAN programs are used for the computationally intensive work of fitting the models, including a large dataset. The R programs provide graphical summaries of characteristics of the fitted models, which can be replaced by your preferred graphical software.

The document is split into five parts. In the first part, we outline the file naming convention that we use, how to compile the source code, and execution of jobs on standard Linux systems. In the second part, documentation is given for each of the FORTRAN programs. In the third part, various R programs are described for plotting spatial images that visualize the inversion outputs of the FORTRAN programs. In the fourth part, based on the estimated HIST-ETAS models, the FORTRAN programs for forecasting future seismicity rate are explained. R programs are then described to display snapshots of the spatial distribution of forecasts. In the fifth part, programs are given for simulating spatial nonhomogeneous Poisson model, spatial magnitude simulation using location-dependent *b*-values, and space-time simulation of HIST-ETAS models. The Appendix contains mathematical background of the models and optimization procedures.

Keywords: space-time ETAS model, space-time point process, location dependent parameters, penalized log-likelihood, maximum posterior estimates, non-homogeneous spatial Poisson process, location dependent *b*-value of the Gutenberg-Richter's formula, magnitude frequency, FORTRAN, R, Short-term seismicity forecast, simulations of HIST-PPM models

Contents

The sections and subsections in the content are linked to those in the text by using PDF heading (bookmark).

Part I. File Organisation and Code Execution

The programs are written in FORTRAN and R. FORTRAN is generally used for the computationally intensive work, and R is used for graphical displays. The documentation is written for UNIX like systems, and it is assumed that a satisfactory FORTRAN compiler is installed along with the R statistical software distributed by the R Project (R Development Core Team, 2009).

Alternatively, you can use your own graphical software such as Matlab. Data is exchanged between the FORTRAN and R software as standard text files, and hence could be read by other graphic software too.

1 File Organization

1.1 Program Source Code

The original version of `HIST-PPM` is in the following program directory

`http://bemlar.ism.ac.jp/ogata/HIST-PPM/`

and its a revised version `HIST-PPM-V2` can be taken from the following program directory

`http://bemlar.ism.ac.jp/ogata/HIST-PPM-V2/`

in which the following program subdirectory

`http://bemlar.ism.ac.jp/ogata/HIST-PPM-V2/estimation/`

is equivalent to the original `HIST-PPM` package, containing the same FORTRAN source codes, but some corrected R programs from those in the original package.

The additionally provided FORTRAN and R programs in HIST-PPM-V2 are for the implementation of Short-Term Earthquake Forecasting that are taken from the program subdirectory

`http://bemlar.ism.ac.jp/ogata/HIST-PPM-V2/forecasting/`

the use of which is explained in Part IV of this manual.

Finally, simulating spatial nonhomogeneous Poisson model, spatial magnitude simulation using location-dependent b-values, and space-time simulation of HIST-ETAS models are added to those in HIST-PPM-V3.

All the programs, inputs files and outputs files in this package HIST-PPM-V3 are selected and separately located in the directories that correspond to the subsections of Sections 3 ~ 16 in this manual, so that it will be useful that you can learn the implementation of the programs by reading the manual.

1.2 File Naming Convention

Files are grouped with a common file name. This enables the user to determine the files that are associated with a particular program. It also ensures that later programs do not overwrite the output of earlier programs. The files have been named as follows. The suffix determines the nature of the file:

`FILENAME.conf:` Configuration file (i.e. input parameters to `FILENAME.f`)

`FILENAME.f:` FORTRAN source code for single processor

`FILENAME:` Compiled object code for single processor

`FILENAME.prt:` `write(6,*)` output to keep by

```
FILENAME |tee FILENAME.prt
```

or

```
FILENAME > FILENAME.prt &
```

`FILENAME.out:` Various output files for single processor `out1, out2, ...` number denotes I/O unit in Fortran code, where the transient output is `out6.`

`FILENAME.upda:` Various output of the updated maximum a posteriori solution for the weights that improved ABIC value in the searching by the simplex method.

`FILENAME.omap :` Various output of the **optimal maximum a posteriori** (OMAP) solution where "optimal" means MAP solution under the optimal weights (i.e., minimum ABIC solution).

`FILENAME.R:` R program (usually to plot a graph)

`FILENAME.pdf:` Graphics output from R

`FILENAME.ts:` Hypocenter dataset (earthquake events) in the format, as given in §3.2.

`FILENAME.etas:` Earthquake dataset in etas-format, as given in §3.2.

2. Compiling and Executing FORTRAN Programs

2.1 Compile FORTRAN Programs

The FORTRAN source code conforms to FORTRAN 77. Source code can be compiled in most Linux operating systems by using `gfortran`, as follows:

```
gfortran FILENAME.f -o FILENAME
```

You can use other FORTRAN packages such as Intel Fortran:

```
ifort FILENAME.f -o FILENAME
```

We have confirmed that both FORTRAN compilers above work well throughout the presented programs. It has been observed that Intel FORTRAN (`ifort`) works significantly faster than `gfortran` with some of the programs.

2.2 Memory Issues

The array dimensions in our FORTRAN programs are taken large enough for a moderately sized dataset. Usually, they are sufficiently large to accommodate a few tens of thousands of earthquakes. If the used memory is in excess of that defined, meaningless output can be produced. So, you have to be careful enough to check whether dimensions are set large enough. In Intel FORTRAN, for example, the following compilation command

```
ifort *.f –traceback –g –CB
```

allows a trace back when problems occur. However, there is no comparable command available in GNU FORTRAN, but you may find information by viewing the core-dump file in the Linux system.

Another potential problem is that the default FORTRAN settings may not allocate enough working memory in a standard Linux system compared to supercomputers. To increase such memory, the following command is available for Intel FORTRAN:

```
ifort *.f –mcmodel=large –shared-intel
```

2.3 Execution of FORTRAN Jobs

A job can be submitted interactively or in batch mode. Batch mode allows the user to log out of the system while the job continues to run in the background. The job could consist of a shell script (e.g. `job.sh`) or it may simply be a compiled FORTRAN binary file. The advantage of a shell script is that it can do other things before and after calling the compiled FORTRAN object.

An example script file (job.sh) is

./FILENAME

R CMD BATCH FILENAME.R
mail –s “Job Complete” –r user@localhost

This would execute the compiled FORTRAN binary called FILENAME, run an R script which may do plots, then email the user on completion.

Batch Mode – Submit Immediately: Use Linux command `nohup`, e.g.

```
nohup batch.sh &
```

The ampersand at the end of the line frees the terminal after executing the command. In the above usage, any diagnostic output, including that which would normally be written to unit 6, will be written to `FILENAME.prt`.
To write the output to a file with a specific name, e.g. `program.prt`, run:

```
nohup batch.sh > program.prt &
```

Batch Mode – Submit Later: Use Linux command `at`, e.g. To execute a shell script called `batch.sh` at 21:06 on 08Nov, run the following in an XTERM within the program directory containing `batch.sh`:

```
at -f batch.sh -t 11082106
```

The command `atq` lists jobs in the queue, and `atrm` removes jobs from the queue. More details about each can be found on the manual page (`man at`). Alternatively, jobs can be set up to run at regular time intervals by using chron.

Part II. PARAMETER ESTIMATION FOR EACH MODEL

The programs documented in this part are not used independently of the each other. They will generally need to be executed in a certain order, as the outputs from some of the programs are required for the execution of other programs. A flowchart in **Figure 1** gives a summary of the output from each that is required in other programs.

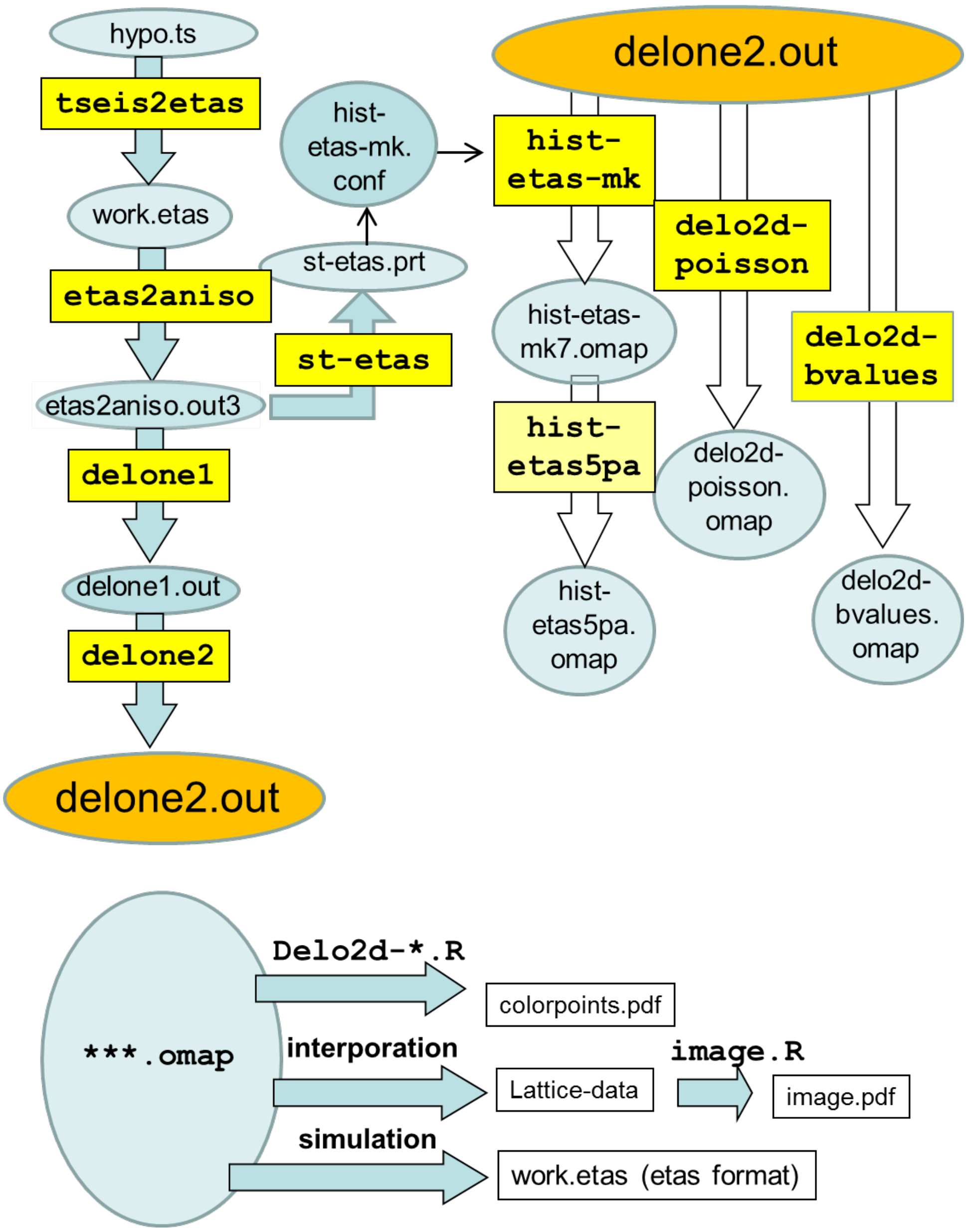


Fig. 1. The diagram shows the flow of output from each program to subsequent programs. Rectangular shapes represent programs, which are explained below. Elliptical shapes represent input/output files. The simulation components are given in Part V (§ 14 - § 16).

3 Formatting of the ETAS data from hypocenter catalog (tseis2etas)

Initially the earthquake catalog data are transformed into what we call an "etas" format. This format is more convenient for the use with the subsequent Fortran model fitting programs. Both input and output allow free format reading in our programs and several initial records at the beginning are shown. All the used files in this section are selected in the program directory of `Section3files/` in the package.

3.1 File Names

```
Program: tseis2etas.f
Object: tseis2etas
input: hypo.ts
output: work.etas
```

3.2 Program Execution

```
./tseis2etas < hypo.ts
```

The file `hypo.ts` contains the earthquake catalog, and is assumed to have the following format.

```
1973 01 01 00 00  0.00  140.8700  33.4700   56.00 -9.5
1973 01 05 05 31  5.80  140.8700  33.4700   56.00  4.5
1973 01 05 11 48  37.50  140.9100  33.1600   33.00  3.9
1973 01 06 10 21  16.30  140.8500  33.4900   33.00  4.2
1973 01 06 11 21  54.70  140.9300  33.2700   46.00  4.5
1973 01 06 14 55  52.80  140.7100  33.1500   61.00  4.7
1973 01 09 02 21  14.80  141.6900  37.8100   59.00  3.5
```

<< omitted the middle.>>

```
2011 05 30 00 05  39.30  142.6400  36.6200   32.00  4.9
2011 05 30 01 04  36.02  142.7100  36.5400    6.00  4.8
2011 05 30 19 36  42.25  140.8000  36.4200   49.00  4.9
2011 05 30 23 53  44.79  143.2300  40.3400   32.00  4.9
2011 05 31 07 50  16.83  140.8400  36.5100   42.00  4.6
2011 05 31 11 26  50.06  141.2400  37.4900   20.00  4.7
2011 05 31 12 28  36.09  141.9300  39.4000   40.00  5.6
2011 05 31 16 26  12.41  143.2000  40.2500   38.00  4.9
2011 05 31 17 14   0.38  146.5900  36.5900   14.00  4.8
2011 05 31 23 53  59.18  142.1800  38.6000   59.00  4.7
```

Columns in the order from left to right are year, month, day, hour, minute, second, longitude (deg.), latitude (deg.), depth (km) and magnitude. The first record defines the beginning of the observation period, and the very small (negative) magnitude indicates that it is a non-event. The very small magnitude ensures that it has no effect in the analyses.

If you want make an aftershock analysis, the first row above starts with the main shock hypocenter.

The present data is shown in Figure 2.

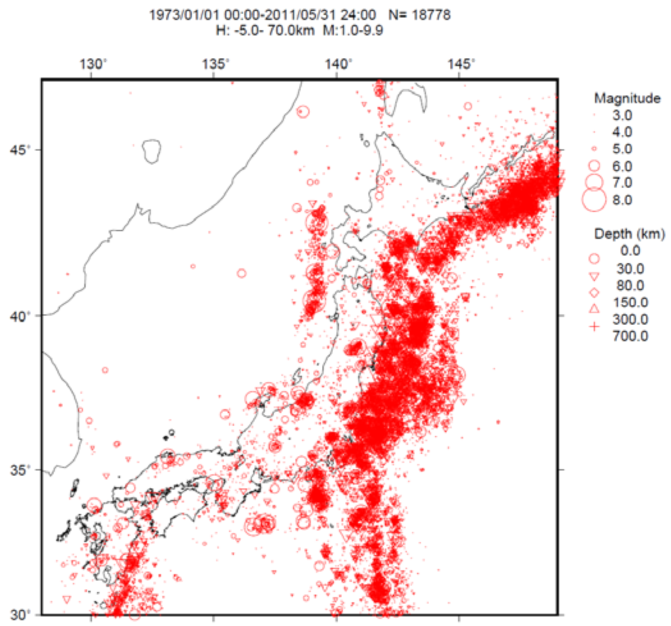


Fig. 2. All detected earthquake by the JMA catalog, drawn by TSEIS visualization program package (Tsuruoka, 1996)

We recommend using all detected earthquakes to identify anisotropic clusters using `etas2aniso` program in the next section. Then, the corresponding `work.etas` comes as follows.

```
formatted_for_etas
     1  140.87000   33.47000      -9.50    0.0000000  -56.00       1973  1  1
     2  140.87000   33.47000       4.50    4.2299282  -56.00       1973  1  5
     3  140.91000   33.16000       3.90    4.4921007  -33.00       1973  1  5
     4  140.85000   33.49000       4.20    5.4314387  -33.00       1973  1  6
     5  140.93000   33.27000       4.50    5.4735498  -46.00       1973  1  6
     6  140.71000   33.15000       4.70    5.6221389  -61.00       1973  1  6
     7  141.69000   37.81000       3.50    8.0980880  -59.00       1973  1  9
     8  137.36000   36.84000       4.40    9.5782477  -33.00       1973  1 10
     9  140.98000   33.11000       3.90   11.3227303  -20.00       1973  1 12
    10  141.06000   33.28000       3.90   12.0151331  -40.00       1973  1 13
```

<< omitted the rest >>

Columns in order from left to right, are: event numbers, longitude (deg.), latitude (deg.), magnitude, time in days from the starting observation time, depth (negative km), and calendar date in year, month and day. Note that the first record is a comment.

4 Identify Anisotropic Clusters of Events (etas2aniso)

Before fitting the space-time models, we compile a dataset with a similar solution (but restricted on the 2-dimensional space) as the so-called centroid Moment tensor solution (Dziewonski *et al*.1981) using early aftershocks activity. This program first selects the large earthquakes and then selects their immediate aftershocks during a certain time span. This is achieved using a fixed space window centered at each large earthquake.

For each such aftershock sequence, the normalized ellipsoidal coefficients (the variances and correlations of a fitted ellipse) are calculated as shown Figure 3. A new catalogue is printed containing the original earthquake origin values together with the two variances and rotation angle, written with each identified main shock. These additional data are used to fit the anisotropic space-time ETAS model (see §5). All the used files in this section are selected in the program directory of `Section4files/` in the program package.

For more details, see §A.1.

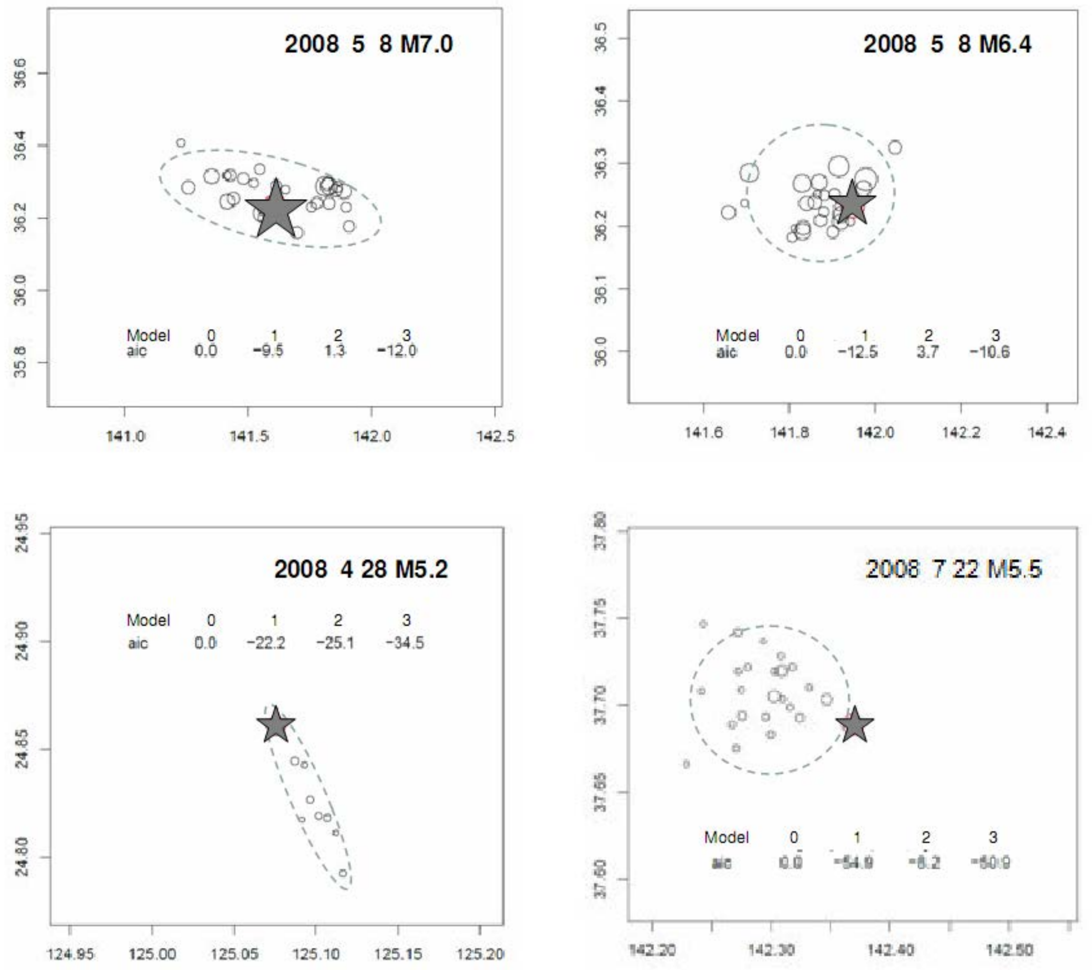


Fig. 3. Examples of identified non-anisotropies.

4.1 File Names

`Program: etas2aniso.f`

`Object: etas2aniso`

`Configuration: etas2aniso.conf` (including `work.etas` as an input) writes:

`etas2aniso.out2:`

contains lists of immediate aftershocks that are triggered by large earthquakes, specified magnitude threshold and a time span described in `etas2aniso.conf`, together with results of best selected case of anisotropy analysis by the smallest AIC value.

`etas2aniso.out3:`

contains the centroid locations and normalized ellipsoidal coefficients for all event with magnitude not less than the cutoff magnitude.

`etas2aniso.out4:`

summarises changed data with either centroid coordinates or anisotropy matrix.

`etas2aniso.out8:`

summarises changed data of the identified earthquakes.

`etas2aniso.out9:`

contains the centroid locations of immediate aftershocks of large events with their normalized ellipsoidal coefficients.

The input data are included in a file whose name is specified in the configuration file (see below). Note that the first event in the input data file is a "no event". Its time, usually zero, indicates to the program the start of the analysis interval. A negative magnitude will ensure that it has no effect.

4.2 Configuration File Format

The configuration (or initialisation) file is called `etas2aniso.conf` and has a format as in the following example.

```
./work.etas     !input data
6.5  6.0        !clms cutm
1.0              !xxx(day)= time span for analyzing centroid and anisotropy
```

The first line is the name of the data file, here `work.etas`. Here it is recommended to use all detected earthquakes without any magnitude cutoff. In the second line, the number "6.5" is the smallest magnitude (`clsm` in the FORTRAN program) of earthquake to analyse its cluster of triggering earthquakes that were followed within a certain time span and certain range of neighborhood (may be called as aftershocks). And "6.0" is used to set the cutoff magnitude (`cutm` in the FORTRAN program) of the output (`etas2aniso.out3`) for a homogeneous data. It is read in using free format.

The third line, "1.0" determines the time window in days for each cluster, here we set one day or less time span in the case where we have a larger earthquake within the considered space window. The time window can be longer in the low detected region or during old period. On the other hand, from a real time forecasting perspective, one may set "1/24 = 0.04167 day = one hour "to quickly determine the centroid location and orientation characteristics of the impending aftershock sequence after a main shock event. For the recent catalog, events within one-hour interval after the main shock will be sufficient to give a reasonably good estimate of the centroid and orientation characteristics of the evolving aftershock sequence.

If you want to use the original epicenters and isotropic clustering for all earthquakes in the original catalog, you can take either a very large magnitude `clsm=9.9` or a very small time span `xxx = 0.00001` in `etas2aniso.conf.`

4.3 Executing the Program

The current program directory must contain the configuration file `etas2aniso.conf` and the data file, whose name is specified on the first line of `etas2aniso.conf`. Other values in the configuration file must be specified by the user.

The program code can then be run by executing the following shell script, after editing the program directory location of the compiled object file called `etas2aniso.`

```
./etas2aniso | tee etas2aniso.prt
```

After execution, the current program directory will contain the following additional files: `etas2aniso.out2`, `etas2aniso.out3`, `etas2aniso.out4`, `etas2aniso.out8`, and `etas2aniso.out9`. Some of these are required by programs documented in the following sections.

Example of output of `etas2aniso.prt` is omitted here.

Example of output of `etas2aniso.out3`

```
310  0.128E+03  0.149E+03  0.206E+02  0.300E+02  0.470E+02  0.170E+02
   176  146.06919   42.94649  7.70    167.16323   1.00000   1.00000   0.00000
   205  146.04000   42.71000  6.00    167.85969   1.00000   1.00000   0.00000
   263  146.65053   43.15368  7.10    174.11349   1.00000   1.00000   0.00000
   297  146.56000   43.17000  6.60    176.93889   1.00000   1.00000   0.00000
   370  146.43000   43.45000  6.00    220.44753   1.00000   1.00000   0.00000
                              << omitted the middle >>
 13914  141.71029   36.17382  6.90  12910.69813   0.18635   0.13324  -0.48998
 14039  140.88000   39.03000  6.90  12947.98872   0.07126   0.11156   0.77508
 14210  142.50500   37.48250  7.00  12983.11075   1.00000   1.00000   0.00000
 14232  142.05000   37.19000  6.00  12985.47951   1.00000   1.00000   0.00000
 14331  144.05375   41.75250  6.80  13037.01448   0.15116   0.07155  -0.68744
                               << omitted the rest >>
```

The first row record represents number of M≥6 earthquakes, `minlong, maxlong, maxlong-minlong, minlat, maxlat, maxlat-minlat.` The following records represent earthquake number, longitudes, latitudes, magnitudes, occurrence times in days; the last three columns represent the estimate of σ_1 , σ_2 and ρ (correlation coefficients) for modified epicenters of the centroid type. Relevantly, some of the epicenters are also modified from the routine epicenters as shown in `etas2aniso.out4.`

The above output data are partially illustrated following in Figure 4.

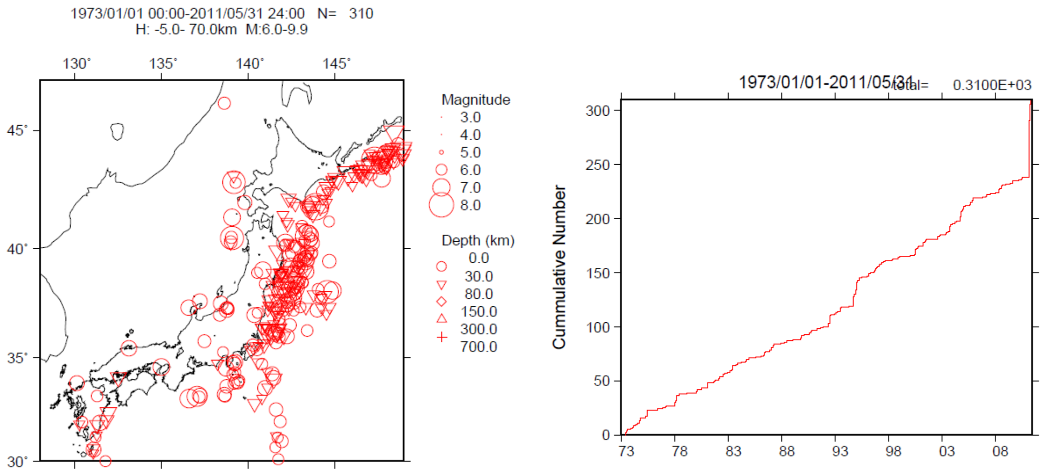


Fig. 4. M ≥ 6.0 earthquake by the JMA catalog

5 Spatial ETAS with All Parameters Constant (st-etas)

This program fits various versions of the space-time ETAS model. It contains two main classes of model. The first class is where the function in (§A.5) that determines the spatial triggering component of the intensity function is assumed to be isotropic. The second class is where it is assumed to be anisotropic. The program does not estimate the anisotropy parameters, but uses those values calculated by the program described in §4. All the used files in this section are selected in the program directory of `Section5files/` in the program package.

Within each class, there are 4 possible models. In the program, the isotropic versions of these models are referred to as models 5–8, and their anisotropic counterpart as 15–18, respectively. The intensity functions of these models are defined by Ogata and Zhuang (2006), Equations 5–7, and 10, respectively. The matrix S_j in those equations is a 2×2 positive definite matrix. In the isotropic case, it will simply be the identity matrix. In the anisotropic case, its elements will contain those values estimated by the program in §4. Further mathematical details can be found in §A.5.1.

5.1 File Names

For the estimation phase, done in FORTRAN:

`Program: st-etas.f`

`Object:  st-etas`

`Configuration: st-etas.conf` (see §5.2)

`Write outputs:  st-etas.prt`

5.2 Configuration File Format

The configuration file is called `st-etas.conf` and has a format as in the following example. Note the symbol "→" below indicates that the record has been split in this document, and the symbol is not part of the configuration file.config.

```
etas2aniso.out3              !hypodata
7                            !nfunct
21.0  17.0  14012.0  310     !tx,ty,tz,nn
128.0  30.0  6.0  0.0  0.0 730.0  2.0 !xmin,ymin,xmg0,zmin,xmg1,tstar,bi2
7                            !n=# of parameters
0.95428E-06   0.40231E-02   0.12529E-02   0.11723E+01   0.94533E+00 →
0.11215E-04   0.13821E+01    ! μ0,K0,c, α,p,d,q
7                            !ipr
```

The numbers are read in as free format and have the following interpretation.

Line 1: Name of data file.

Line 2: Indicates the required space-time model. Valid values are: 5, 6, 7, 8, 15, 16, 17, or 18. *Warning: The software has only been tested for cases 7 and 17, and others may be unstable.*

Line 3: Longitude region width (`tx` degrees), latitude region width (`ty` degrees), upper time boundary (`tz` days), and number (`nn`) of data points.

Line 4: Minimum longitude (`xmin` degrees), minimum latitude (`ymin` degrees), threshold magnitude (`xmg0`), minimum time (`zmin`), another magnitude (`xmg1`, currently not used), and starting time (`tstar` day). Parameter `bi2` is a multiplier used with the "*Utsu Spatial Distance* (*USD*)" defined explicitly in Appendix A5 (§A.5). The `bi2` is infinity (very large) in exact log-likelihood calculation, and this enables an approximation to shorten the computation time to have good initial ETAS parameter values. The *USD* is the width of a square, centred on the main shock, within which it is assumed that most of the aftershocks associated with the given main shock will occur. This assumption considerably lessens required calculations because the intensity at the location of subsequent events will only be affected by historical events if the given event is contained within the Utsu squares associated with the historical events.

Line 5: Number of initial model parameters listed on line 6.

Line 6: Initial parameter estimates.

Line 7: If ipr = 7, additional output is printed for the linear search procedure, and not printed if ipr=0

5.3 Executing the Program

The current program directory must contain `st-etas.conf` and the data file. The required data file is `etas2aniso.out3` which is one of the outputs from `etas2aniso`. See Appendix A.1 for some detail.

Appropriate initial parameter values must be edited into the configuration file by the user.

The job is executed by running the following execution command.

```
./st-etas | tee st-etas.prt
```

Note that the number of events stated in `st-etas.conf` is the number of events in `etas2aniso.out3`.

An example of the `st-etas.prt` is as follows:

```
./etas2aniso.out3
           17
        21.          17.      14012.          310
      128.0         30.0          6.0          0.0          0.0       730.0          2.0
 data set          310   0.128E+03   0.149E+03   0.206E+02   0.300
 input device           10
 nn=             310
 nfunct=               17
0tx,ty,tz,xmin,ymin,xmg0,xmg1,zmin,tsta
     16.435       17.000 14012.000      128.000       30.000        6.000        0.000        0.000 →
0.000
 nn =    310 nnc =    294
 bi2     2.0000000000000000
 jmax               67
 tstar,nstar     730.00000000000000                       16
0    input data
       n=   7    itr=
0x=     0.95428E-06     0.40231E-02     0.12529E-02     0.11723E+01     0.94533E+00
0x=     0.11215E-04     0.13821E+01
 linear ipr                7
 - log likelihood =    0.153377032136182D+04        aic =       3081.5
 lambda =   0.5000000000D+00                 e2 =    0.10000000000000000D+31
 lambda =   0.5000000000D-01                 e4 =    0.10000000000000000D+31
 lambda =   0.5000000000D-02                 e4 =    0.10000000000000000D+31
 lambda =   0.5000000000D-03                 e4 =    0.10000000000000000D+31
 lambda =   0.5000000000D-04                 e4 =    0.10665559805887825D+06
 lambda =   0.5000000000D-05                 e4 =    0.28818211711226591D+04
 lambda =   0.5000000000D-06                 e4 =    0.11629651734330857D+04
 lambda =   0.1900108183D-05                 e5 =    0.17130527321761620D+04
 lambda =   0.8710402217D-06                 e6 =    0.13234771092168321D+04
  lmbd = 0.5000000D-06    -ll = 0.132347710921683D+04    -0.24D+11    0.24D+11
 lambda =   0.5000000000D-06                 e2 =    0.11421402460408272D+04
 lambda =   0.1000000000D-05                 e3 =    0.11725055350543980D+04
 lambda =   0.4534072998D-06                 e5 =    0.11419671670520129D+04
 lambda =   0.4581397976D-06                 e6 =    0.11419644849252979D+04
  lmbd = 0.4581398D-06    -ll = 0.114196448492530D+04    -0.97D+08    0.11D+09
```

<< skipped >>

```
 lambda =   0.2089781397D+01                 e6 =    0.84830056255183285D+03
  lmbd = 0.1393188D+01    -ll = 0.848300562551833D+03    -0.81D-15    0.15D-08
 lambda =   0.1393187600D+01                 e2 =    0.84830030469372718D+03
 lambda =   0.2786375200D+01                 e3 =    0.84830056255183172D+03
 lambda =   0.1393187594D+01                 e5 =    0.84830056255183490D+03
 lambda =   0.2089781396D+01                 e6 =    0.84830056255183433D+03
  lmbd = 0.1393188D+01    -ll = 0.848300562551834D+03    -0.12D-15    0.23D-09
  - log likelihood =    0.848300562551825D+03        aic =       1710.6
 0-----    x    -----
   -0.54093D-03    0.14630D+00    0.42172D-01    0.10343D+01    0.93246D+00    0.13550D+00 →
0.12333D+01
 0***    gradient    ***
   -0.59377D-05    0.22362D-06 -0.32350D-06    0.26929D-07 -0.13557D-06 -0.18250D-06 →
0.10456D-06
```

```
mle =   0.29261E-06 0.21404E-01 0.17785E-02 0.10697E+01 0.86948E+00 →
        0.18359E-01 0.15210E+01
```

The last 7 numbers are the MLEs of μ, K_0, c, α, p, d and q of a space-time ETAS model, which will be used (copy & pasted) for the reference parameters in `hist-etas-mk.conf` in §9.2.

5.4 Additional Advice

When the background rates in space are far from homogeneous, the MLE above may not converge well. In that case, firstly, set about a half of the average earthquake occurrence rate per unit time and unit area, say, for an initial estimate of the μ parameter as the case of the above; and set its gradient for the μ parameter being always zero. Then, program `st-etas` implements the stable optimization for the other parameters than with the unfixed μ parameter. This is implemented by additionally setting 1 in the 8th line in `st-etas.conf` as follows:

```
etas2aniso.out3              !hypodata
7                            !nfunct
21.0   17.0   14012.0   310           !tx,ty,tz,nn
128.0   30.0   6.0   0.0   0.0 730.0   2.0 !xmin,ymin,xmg0,zmin,xmg1,tstar,bi2
7                            !n=# of parameters
0.95428E-06    0.40231E-02    0.12529E-02    0.11723E+01    0.94533E+00 →
0.11215E-04    0.13821E+01            ! μ0,K0,c, α,p,d,q
7  1    ! ipr, igrd  for optimization by fixing μ–parameter by 0, otherwise 1
```

and then run the program `st-etas`.

Having done that, use the above estimated μ, K_0, c, α, p, d and q for initial estimates without the 8th line, again to run `st-etas` by the unfixed 7 parameters could lead an eventually stable MLE. This is implemented by setting the value other than 0 (say, 1 or nothing) in the 8th line in `st-etas.conf.`

6 Delaunay Tessellation for Spatial Variation (delone1, delone2)

This section describes a group of programs that are used to perform a Delaunay tessellation of the two-dimensional spatial coordinates. This tessellation is used by subsequent programs to provide spatial estimates of some or all of the ETAS parameters.

The first FORTRAN program (`delone1.f`) performs a Delaunay tessellation. It initially augments the spatial locations of the points closest to the boundary with the location of their mirror image in the boundary. The second (`delone2.f`) treats the locations where the triangle lines cross the observation region boundary as a new point, and excludes the mirror image added by `delone1.f`. Together with the original observed locations and these boundary points, it repeats the Delaunay tessellation, and then outputs the determined triangles in a satisfactory format so that the R program `delone2.R` can be used to plot all of the triangles. This output is also used by programs for estimation where one or more parameters are assumed to vary in space. All the used files in this section are selected in the program directory of `Section6files/` in the program package.

Further mathematical detail can be found in §A.2.

6.1 File Names to Perform Delaunay Tessellation

```
Program:        delone1.f
Object:         delone1
Configuration:  delone1.conf
Reads:          etas2aniso.out3
Writes:          delone1.out
```

6.2 Configuration File Format

An example of a configuration file follows.

```
1.00E-15              ! for EPS
1000   7000           ! for NEF0, NRG0
128.0   30.0          ! for xmin, ymin
  21.   17.           ! for BXLX, BXLY
  310                 ! for NP (e.g., number of earthquakes)
```

Parameters are read as free format. The above parameters are fragile for successful computation; see §6.4 to check. The error bound is already very small and can be larger, which makes the computation faster in the case where the number of (earthquake) data points `NP` is very large. A rough rule of thumb is that `NEF0` should be approximately 0.8 times the number of data points `NP`, and `NRG0` should be larger, especially in the case where points are highly clustered. Note that "21." for `BXLX` is the width of the analysis region (degrees longitude), "17." for `BXLY` is the height of the analysis region (degrees latitude), "310" for `NP` is the number of points, "128" for `xmin` is the western boundary (longitude), and "30" for `ymin` is the southern boundary (latitude). In western hemisphere xmin should be positive taking between 180 and 360 degrees, and in the southern hemisphere `ymin` is negative, taking values between -90 and 0 degrees.

6.3 Executing Delaunay Tessellation Program

The required data are contained in `etas2aniso.out3`. The source code `delone1.f` requires the configuration file (i.e. `delone1.conf`).

```
./delone1 |tee delone1.prt
```

Running this job, we get the following output file (`delone1.prt`):

```
0         ***** input parameters *****
          iperio=        0    np   =        0
          dens  =  1.00000    eps=   0.10000E-14
          nef0  =     1000    nrg0  =     7000
          nclx1 =        3    ncly1 =        3
          ilist =        1    ifile =        0
          incard=        1    idpat =        1

         32   149.000000000000       44.0300000000000
        302   131.780000000000       30.0000000000000
 np        308
         *** input coordinates ***
          np=   308   idpat=    1    bxlx,bxly=      21.00000
17.00000   dens=       0.86835
0*** detailed outputs ***
```

<< skipped >>.

```
         ***** result of voronoi division *****
idpat  np   bxlx   bxly   brasq   sum of pol.ar. box area
  1   308  21.0000   17.0000  50.820   3.57000000E+02  3.57000000E+02
 number of delaunay triangle =    618
```

Note here that the number of earthquakes in `etas2aniso.conf (NP=310)` is reduced to `308` because the two earthquakes on the rectangular boundary are removed in the computation.

In particular, in the second to last line on the right-hand side are values of "`sum of pol.ar.`" and "`box area`". The values for these should be the same if the Delaunay tessellation is correct. If they are not, then the values of `NEF0` and `NRG0` in the configuration file `delone1.conf` may need adjusting.

Another output file `delone1.out` to be used for the next subsection writes as follows:

```
  308    21.00000    17.00000   128.00000    30.00000
    1 18.06954167 12.94617946    7.70000   167.16323   1.0000   1.0000   0.0000
    2 18.03999678 12.70995838    6.00000   167.85969   1.0000   1.0000   0.0000
    3 18.65051057 13.15371107    7.10000   174.11349   1.0000   1.0000   0.0000
    4 18.56019533 13.17001224    6.60000   176.93889   1.0000   1.0000   0.0000
    5 18.43048317 13.45037851    6.00000   220.44753   1.0000   1.0000   0.0000
```

<< skipped >>

```
  306 14.86989084  9.09956961    6.00000 13991.42554   1.0000   1.0000   0.0000
  307 16.06046994  8.16971916    6.10000 14003.62383   1.0000   1.0000   0.0000
  308 13.33026396  7.40966455    6.10000 14011.98325   1.0000   1.0000   0.0000
      1       1    182    207       3
      2       1      2    207       3
      3       1      2      4       3
      4       1      4    182       3
      5       2    207    208       3
      6       2    174    208       3
```

<< skipped >>

```
    614     274    280    306       3
    615     275    280    306       3
```

```
   616    275   282   306      3
   617    275   276   282      3
   618    275   276   280      3
```

The first line contains the number of earthquakes (NP), lengths of longitude (bxlx) and latitude (bxly) spans, the origin longitude and latitude of the rectangular region, in the order from the left. Then, the following first block provides the same data as in etas2aniso.out3. Here the order of earthquakes is given in the first column up to the number NP=308. Also note here that the number of earthquakes in etas2aniso.conf (NP=310) is reduced to 308 because the two earthquakes on the rectangular boundary are removed in the computation. The second block lists the Delaunay triangles, numbered from 1 to 618 in the first column, vertex points, and the id-number of each triangle.

6.4 Generation of the Map Data with Boundary Points

The files associated with generating map data are as follows.

```
Program: delone2.f
Object:  delone2
Reads:   delone1.out
Writes:  delone2.out
```

The above FORTRAN code can be executed by running the following shell script within the current program directory.

```
./delone2 |tee delone2.prt
```

An example of the delone2.prt is as follows:

```
ss=    356.999999999999       tx*ty=    357.000000000000
          10
          12
```

We can confirm the accuracy of the tessellation program by equality of the two calculated areas in the first line; where ss represents the sum of the Delaunay triangle areas and tx*ty represents the whole rectangular area. The second line is the largest number of following connected earthquakes by the Delaunay tessellation. The last line indicates the largest number of preceding and following connected earthquakes by the Delaunay tessellation. The Incomplete Cholesky Conjugate Gradient (ICCG) method, used later, requires that the maximum number of connected edge points of the Delaunay triangulation kkmax is 12, which is given in the last line of delone2.out, and the last line of delone2.prt in the above.

An example of the delone2.out is as follows:

```
  308      342      648    21.00000    17.00000
 1 18.06954167 12.94617946    7.70   167.1632300   1.0000   1.0000   0.0000
 2 18.03999678 12.70995838    6.00   167.8596900   1.0000   1.0000   0.0000
 3 18.65051057 13.15371107    7.10   174.1134900   1.0000   1.0000   0.0000
 4 18.56019533 13.17001224    6.60   176.9388900   1.0000   1.0000   0.0000
```

```
  5 18.43048317 13.45037851    6.00   220.4475300   1.0000   1.0000   0.0000
```

<< skipped >>

```
221  8.70176461  7.25834122    6.70 12501.0291400   0.1139   0.0798   0.6972
222 10.55744694  7.48449001    6.60 12614.0509500   0.1070   0.0790   0.6728
223 14.02977645  8.50037863    6.10 12776.5865100   1.0000   1.0000   0.0000
224 13.53967847  6.18031260    6.20 12910.6680900   1.0000   1.0000   0.0000
225 13.75989075  6.15983542    6.10 12910.6782000   1.0000   1.0000   0.0000
226 13.71057900  6.17401699    6.90 12910.6981300   0.1863   0.1332  -0.4900
227 12.88007256  9.02951694    6.90 12947.9887200   0.0713   0.1116   0.7751
228 14.50513802  7.48206907    7.00 12983.1107500   1.0000   1.0000   0.0000
229 14.05029992  7.18993243    6.00 12985.4795100   1.0000   1.0000   0.0000
230 16.05360444 11.75244392    6.80 13037.0144800   0.1512   0.0716  -0.6874
```

<< skipped >>

```
305 12.30031094  5.60951590    6.20 13989.5673500   1.0000   1.0000   0.0000
306 14.86989084  9.09956961    6.00 13991.4255400   1.0000   1.0000   0.0000
307 16.06046994  8.16971916    6.10 14003.6238300   1.0000   1.0000   0.0000
308 13.33026396  7.40966455    6.10 14011.9832500   1.0000   1.0000   0.0000
309  0.00000000  0.00000000    0.00     0.0000000   0.0000   0.0000   0.0000
310 21.00000000  0.00000000    0.00     0.0000000   0.0000   0.0000   0.0000
311 21.00000000 17.00000000    0.00     0.0000000   0.0000   0.0000   0.0000
```

<< skipped >>

```
340 21.00000000  0.99002710    0.00    0.0000000   0.0000   0.0000   0.0000
341 21.00000000  7.05039114    0.00    0.0000000   0.0000   0.0000   0.0000
342 21.00000000  7.35957019    0.00    0.0000000   0.0000   0.0000   0.0000
      1      1    182    207  0.796487629875D-01
      2      1      2    207  0.113149385584D+00
      3      1      2      4  0.546448113038D-01
      4      1      4    182  0.346708339754D-01
```

<< skipped >>

```
    645    275    280    306  0.132964341353D-01
    646    275    282    306  0.158110057792D-01
    647    275    276    282  0.302560585723D-01
    648    275    276    280  0.252603344595D-01
     1     4      2     4   182   207
     2     9      3     4    97   100   165   174   190   207   208
     3     4      4    13    40   165
     4     4     13    75   159   182
     5     7     48    75   138   140   159   182   207
     6     5     77   184   295   298   299
     7     8     52    99   188   189   191   206   220   232
```

<< skipped >>

```
   335     1     339
   336     0
   337     1     338
   338     0
   339     0
   340     1     341
   341     1     342
   342     0
      12
```

The first record gives the number of earthquakes (`NP`), number of points Delaunay tessellation including those on boundaries, number of Delaunay triangles, and lengths of longitude (`BXUP`) and latitude (`BYUP`) spans, in the order from the left. Then, the first block provides the same data as `st-etas.out`. Here the order of earthquakes are given in the first column up to the number `NP=308`. The second block of the index numbers from `NP=309` to `342` includes the Delaunay vertex points on the boundary of the rectangular region. The third block lists the Delaunay triangles numbered from 1 to 648 in the first column, vertex points id-number of each triangle, and area of the triangle in the last column. The forth block indicates neighboring points connected by the sides of the triangle; the first record specifies the id-numbers of points, the second indicates the number of the connected points by the side of the triangles, and the rest of the columns show the id-numbers of the nearest points. The bottom raw number shows the largest numbers of the nearest points.

These provide necessary information to the Bayesian smoothing procedure, especially for the Hessian matrix and incomplete Cholesky conjugate gradient (ICCG) method: see Appendix B.2.

6.5 Plotting Delaunay Tessellations

The files associated with plotting the Delaunay tessellations using the R statistical language are as follows.

```
Program: delone-plot.R
Reads:   delone2.out
Writes:  delone-plot.pdf
```

The above R program can be executed by running R within the current program directory (`Section6files`), and executing the R function `source` to run the contents of the file interactively as:

```
R
> source('delone-plot.R')
```

The plot will be written into the file `delone-plot.pdf`.

6.6 Example Output

An example of the Delaunay tessellation plot example data is shown in the following figure (Fig.5).

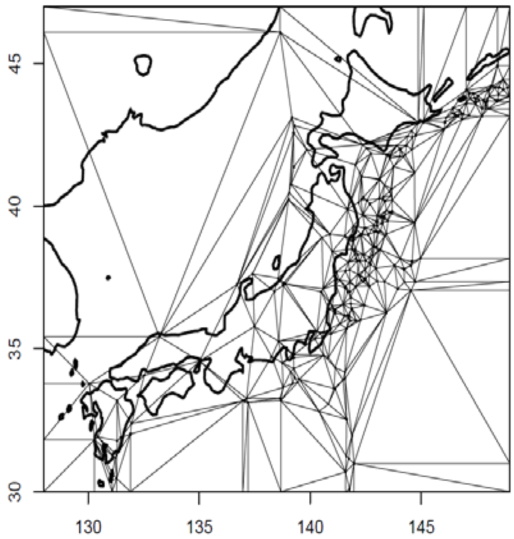


Fig. 5: Delaunay tessellation plot, `delone-plot.pdf`

7 Spatially Varying *b*-Value of Magnitude Frequency (`b-values`)

These programs calculate and plot estimates of the *b*-value over a spatial region. The program calculates *b*-values at the nodes of the Delaunay tessellations (§6). Estimates at other spatial points can be made using the interpolation program. All the used files in this section are selected in the program directory of `Section7files/` in the program package. Further mathematical detail can be found in §A.4.

7.1 File Names

For the estimation phase, done in FORTRAN:

```
Program:       delo2d-bvalues.f
Object:        delo2d-bvalues
Configuration: delo2d-bvalues.conf
Reads:         delone2.out
Writes:        delo2d-bvalues.omap
```

For the spatial plot, done in R:

```
Program: delo2d-bvalues.R
Reads:   delone2.out, delo2d-bvalues.omap
Writes:  delo2d-bvalues.pdf.
```

7.2 Configuration File Format

The configuration file `delo2d-bvalues.conf` includes the following three lines:

```
128. 30. 5.95   !xmin, ymin, threshmag = magnitude threshold
6.0d0           !w1 = initial weight of the penalty to be optimized.
7               ! ipr
```

containing the following records; the first line includes the origin of the considered region in longitude and latitude, and then magnitude threshold. The second line is an initial weight value for the penalty function. In the third line, if ipr = 7, more detailed output about the linear search procedure is given, and is not given if ipr = 1. Parameters are read as free format.

Magnitude rounding issue: if magnitude data are rounded to 0.1 units, the threshold magnitude here should be modified to `5.95` (= *Mc* -0.05) to avoid the *b*-value MLE bias. This is because a rounded value of 6.0 may have been as small as 5.95 or large as 6.05. This applies to the traditional catalogs such as the JMA, NEIC-PDE, and ISC catalog. Otherwise, namely, less than 0.01 magnitude unit, we can keep `threshmag = 6.0`.

7.3 Program Execution

FORTRAN execution command:

```
./delo2d-bvalues |tee delo2d-bvalues.prt
```

The contents of `delo2d-bvalues.prt` includes the calculation processes as follows:

```
xmin,ymin,threshmag=    128.000000000000          30.0000000000000
   5.95000000000000
 weight=    6.00000000000000
 linear ipr              7
          308           342           648    21.0000000000000
   17.0000000000000
 an =    1.00000000000000
 npex          342
 w1,w2,w3    6.00000000000000          0.000000000000000E+000
  0.000000000000000E+000
 ptdet = 0.1286030538956D+04
 #1: w1 =   0.60000000D+01
 penalized-log-likelihood =   0.463684636109451D+02
 lambd2 =   0.5000000000D+00             e2 =   0.30828181796202939D+04
 lambd4 =   0.5000000000D-01             e4 =   0.62135776017985577D+02
 lambd4 =   0.5000000000D-02             e4 =   0.45023517036283550D+02
 lambd5 =   0.1285750683D-01             e5 =   0.44224067911271810D+02
 lambd6 =   0.1284875160D-01             e6 =   0.44224065397758963D+02
 1        1 lambda =   0.1284875D-01     pell =   0.442240653977590D+02   0.33D+03
 lambd2 =   0.1284875160D-01             e2 =   0.50878298842753544D+02
 lambd4 =   0.1284875160D-02             e4 =   0.43818395118639856D+02
 lambd5 =   0.2832311965D-02             e5 =   0.43645759409649770D+02
 lambd6 =   0.2832374825D-02             e6 =   0.43645759408993605D+02
 1        2 lambda =   0.2832375D-02     pell =   0.436457594089936D+02   0.41D+03
 lambd2 =   0.2832374825D-02             e2 =   0.43616306798161546D+02
 lambd3 =   0.5664749651D-02             e3 =   0.43586937733633853D+02
 lambd3 =   0.1132949930D-01             e3 =   0.43528450272154494D+02
 lambd3 =   0.2265899860D-01             e3 =   0.43412478220408836D+02
 lambd3 =   0.4531799721D-01             e3 =   0.43184547213016458D+02
```

```
lambd3 =   0.9063599442D-01           e3 =   0.42744750536460081D+02
lambd3 =   0.1812719888D+00           e3 =   0.41929523194600279D+02
lambd3 =   0.3625439777D+00           e3 =   0.40557386459155495D+02
lambd3 =   0.7250879553D+00           e3 =   0.38853488561817301D+02
lambd3 =   0.1450175911D+01           e3 =   0.39668601393578882D+02
lambd5 =   0.9826636726D+00           e5 =   0.38493436628796665D+02
lambd6 =   0.9834376482D+00           e6 =   0.38493427204469839D+02
1        3 lambda =   0.9834376D+00    pell =   0.384934272044698D+02   0.49D+01
lambd2 =   0.9834376482D+00           e2 =   0.38489324810747796D+02
lambd3 =   0.1966875296D+01           e3 =   0.38493117194215735D+02
lambd5 =   0.1002746347D+01           e5 =   0.38489323394303419D+02
lambd6 =   0.1002091162D+01           e6 =   0.38489323392487314D+02
1        4 lambda =   0.1002091D+01    pell =   0.384893233924873D+02   0.67D-02
lambd2 =   0.1002091162D+01           e2 =   0.38489323390150034D+02
lambd3 =   0.2004182325D+01           e3 =   0.38489323392506961D+02
lambd5 =   0.9999924475D+00           e5 =   0.38489323390150027D+02
lambd6 =   0.9992774566D+00           e6 =   0.38489323390150084D+02
1        5 lambda =   0.9999924D+00    pell =   0.384893233901500D+02   0.18D-08
penalized log likelihood =   0.384893233901500D+02
#e: w1 = 0.60000000D+01
  abic = 0.8410057591D+02   -l =-0.2930057833D+03   pn = 0.1298666099D+04
-------- xd -------     1.00000000000000      6.00000000000000     84.1005759128927
```

<< skipped >>

```
w1,w2,w3     2.00974825755177            0.000000000000000E+000 →
  0.000000000000000E+000
ptdet = 0.9130617889564D+03
#1: w1 =   0.20097483D+01
penalized-log-likelihood =   0.287800079161731D+02
lambd2 =   0.5000000000D+00           e2 =   0.29546166935190861D+02
lambd4 =   0.5000000000D-01           e4 =   0.28777011048782104D+02
lambd5 =   0.3346968607D-01           e5 =   0.28776042319021148D+02
lambd6 =   0.3346159299D-01           e6 =   0.28776042318756339D+02
1        1 lambda =   0.3346159D-01    pell =   0.287760423187563D+02   0.24D+00
lambd2 =   0.3346159299D-01           e2 =   0.28782452945428989D+02
lambd4 =   0.3346159299D-02           e4 =   0.28775473430799885D+02
lambd5 =   0.8752834473D-02           e5 =   0.28775122438782077D+02
lambd6 =   0.8752857664D-02           e6 =   0.28775122438782141D+02
1        2 lambda =   0.8752834D-02    pell =   0.287751224387821D+02   0.21D+00
lambd2 =   0.8752834473D-02           e2 =   0.28775048199836448D+02
lambd3 =   0.1750566895D-01           e3 =   0.28774974613552200D+02
lambd3 =   0.3501133789D-01           e3 =   0.28774829398984622D+02
lambd3 =   0.7002267579D-01           e3 =   0.28774546801964874D+02
lambd3 =   0.1400453516D+00           e3 =   0.28774012937290273D+02
lambd3 =   0.2800907031D+00           e3 =   0.28773070532623031D+02
lambd3 =   0.5601814063D+00           e3 =   0.28771687079862836D+02
lambd3 =   0.1120362813D+01           e3 =   0.28770926064465890D+02
lambd3 =   0.2240725625D+01           e3 =   0.28777431321669166D+02
lambd5 =   0.9995936669D+00           e5 =   0.28770863947998595D+02
lambd6 =   0.9996405857D+00           e6 =   0.28770863947986673D+02
1        3 lambda =   0.9996406D+00    pell =   0.287708639479867D+02   0.39D-02
lambd2 =   0.9996405857D+00           e2 =   0.28770863945859837D+02
lambd3 =   0.1999281171D+01           e3 =   0.28770863947983585D+02
lambd5 =   0.1000003763D+01           e5 =   0.28770863945859844D+02
lambd6 =   0.9954337900D+00           e6 =   0.28770863945859880D+02
1        4 lambda =   0.9996406D+00    pell =   0.287708639458598D+02   0.38D-08
penalized log likelihood =   0.287708639458598D+02
#e: w1 = 0.20097483D+01
  abic = 0.8153752949D+02   -l =-0.1162398678D+03   pn = 0.9425712218D+03

-------- xd -------     6.00000000000000          2.00974825755177 →
   81.5375294926255
####   iteration, f, epsilon =       6   0.81537529D+02   0.35904191D-03
x =     0.13960189D+01
    0.20097E+01      0.81538E+02      342
```

The records including `lambd#` ('`#`' for a number) show linear search for the minimum of the negative penalized log likelihood (`pell`), and the rows including `pell` show the minimized value and the sum of squares of the gradient vector components of the `pell` function with respect to the minimizing parameters. Furthermore, the `abic` value is minimized with respect to a weight `w1,` assuming isotropic smoothing constraint. The rows with “-------- xd -------“ shows every step where the minimum was updated by the simplex algorithm. The third to last rows from the bottom starting at #### show that the iterated simplex algorithm updated the ABIC for 6 times with the minimum abic = 0.8153752949D+02 and the difference with the previous smallest ABIC is `0.82721787D-04`. This is attained by `w1` = 0.20097483D+01 (5th row from the bottom), and the bottom row shows its logarithm. See Appendix A for the definitions and Appendix B for the numerical procedures.

The file `delo2d-bvalues.prt` includes a large volume of output. It may be useful to use UNIX command `egrep (grep)` to restrict output to records of interest. For example,

```
egrep xd delo2d-bvalues.prt
```

and

```
egrep xd |abic delo2d-bvalues.prt
```

shows you a series of only the updated smallest ABIC values and of all searched ABIC values in the simplex minimization procedure, respectively.

```
./delo2d-bvalues > delo2d-bvalues.omap
```

```
R
> source('delo2d-bvalues.R')
```

The output shows Fig. 6.

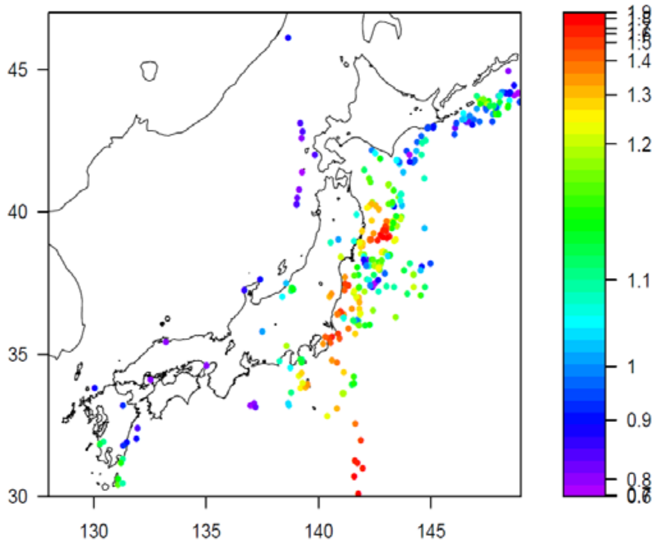


Fig. 6. `bvalues.pdf;` colors are ordered in frequency-linearized scale.

8 Spatial Occurrence Rate (delo2d-poisson)

This program fits a nonhomogeneous spatial Poisson model with no time component to the location of earthquakes. This is done be estimating the Poisson rates at the nodes of the Delaunay tessellations (§6). All the used files in this section are selected in the program directory of `Section8files/` in the program package. Further mathematical detail can be found in §A.5.2.

8.1 File Names

For the estimation phase, done in FORTRAN:

```
Program: delo2d-poisson.f
Object:  delo2d-poisson
Configuration: delo2d-poisson.conf
Reads:   delone2.out
Writes:  delo2d-poisson.omap
```

For the spatial plot, done in R:

```
Program: delo2d-poisson.R
Reads:   delone2.out, delo2d-poisson.omap
Writes:  delo2d-poisson.pdf
```

8.2 Configuration File Format

The configuration file `delo2d-poisson.conf` includes the following three lines:

```
128. 30. 5.95  ! xmin, ymin, threshmag  = magnitude threshold
6.0d0          !w1= initial weight of the penalty to be optimized.
  7            ! ipr
```

containing the following records; the first line includes the origin of the considered region in longitude and latitude, and then magnitude threshold. The second line is an initial weight value for the penalty function. In the third line, if ipr = 7, more detailed output about the linear search procedure is given, and is not given if ipr = 0. Parameters are read as free format.

8.3 Program Execution

FORTRAN execution command :

```
./delo2d-poisson |tee delo2d-poisson.prt
```

The example of `delo2d-poisson.prt` includes the calculation processes as follows:

```
     308          342          648     16.434778690338135      17.000000000000000 →
0.95798319327731085
 an =     1.0000000000000000
 tx,ty       16.435       17.000       nn,np,npex,nd =       308       308       342       648
 ptdet = 0.1218006057961D+04
 #1: w1,w2,w3,w4 =    0.60000000D+01 0.60000000D+01 0.00000000D+00 0.10000000D+01
 wx,wy= 0.60000D+01wxx,wyy= 0.00000D+00     pell =    0.293388497473962D+03
 lambda =    0.5000000000D+00               e2 =    0.27484230578836838D+03
 lambda =    0.1000000000D+01               e3 =    0.26290412013045670D+03
 lambda =    0.2000000000D+01               e3 =    0.25722684413266461D+03
 lambda =    0.4000000000D+01               e3 =    0.31070799318271708D+03
 lambda =    0.1762692124D+01               e5 =    0.25648282432909440D+03
 lambda =    0.1745673777D+01               e6 =    0.25647816623906681D+03
 lambda =    0.1745674D+01      pell =    0.256478166239067D+03    -0.44D+02      0.27D+04
 cgres_0              31    2.51817502773355539E-009    5.18160438220919444E-013
 #iteration=                1
 cgres_0              35    8.16996669971836904E-010    4.73376310296145925E-013
 lambda =    0.1745673777D+01               e2 =    0.12575332143434689D+04
 lambda =    0.1745673777D+00               e4 =    0.18238027528344378D+03
 lambda =    0.4214208914D+00               e5 =    0.10619977032668238D+03
 lambda =    0.5037843862D+00               e6 =    0.89583483392868573D+02
 lambda =    0.5037844D+00      pell =    0.895834833928686D+02    -0.47D+03      0.17D+04
 cgres_0              31    1.96048463382195580E-010    4.13813896662551007E-013
 #iteration=                2
 cgres_0              31    3.02592639283617553E-010    7.26217704144932776E-013
 lambda =    0.5037843862D+00               e2 =    0.69744102731447668D+02
 lambda =    0.1007568772D+01               e3 =    0.63752199016021812D+02
 lambda =    0.2015137545D+01               e3 =    0.96660591021549394D+02
 lambda =    0.9574016612D+00               e5 =    0.63709613227398854D+02
 lambda =    0.9673758963D+00               e6 =    0.63706622422671018D+02
 lambda =    0.9673759D+00      pell =    0.637066224226710D+02    -0.53D+02      0.42D+03
 cgres_0              29    9.76411103241434597E-011    2.51836405349217590E-013
 #iteration=                3
 cgres_0              33    8.39914763820138480E-013    3.42533485175028649E-013
 lambda =    0.9673758963D+00               e2 =    0.63386328627224401D+02
 lambda =    0.1934751793D+01               e3 =    0.63625716518613871D+02
 lambda =    0.1037296374D+01               e5 =    0.63385217407971851D+02
 lambda =    0.1029415522D+01               e6 =    0.63385196860218286D+02
 lambda =    0.1029416D+01      pell =    0.633851968602183D+02    -0.63D+00      0.25D+01
 cgres_0              33    1.23078004902940045E-012    5.04589483633406912E-013
 #iteration=                4
 cgres_0              36    1.23479568721858537E-015    6.82639669277830308E-013
 lambda =    0.1029415522D+01               e2 =    0.63385147363602783D+02
 lambda =    0.2058831044D+01               e3 =    0.63385202811382449D+02
 lambda =    0.1000227585D+01               e5 =    0.63385147321246720D+02
 lambda =    0.1000170025D+01               e6 =    0.63385147321246102D+02
 lambda =    0.1000170D+01      pell =    0.633851473212461D+02    -0.99D-04      0.18D-02
 cgres_0              36    1.38678649094504771E-015    7.66666188308590379E-013
 #iteration=                5
 cgres_0              41    3.70006474981530149E-023    7.93982954620509432E-013
 penalized log likelihood =    0.633851473212461D+02          rss1 = 0.00000D+00
 #2: w1,w2,w3,w4 = 0.60000000D+01 0.60000000D+01 0.00000000D+00 0.10000000D+01
  abic = 0.1423320648D+03    -l =-0.2322598418D+03    pn = 0.1233567828D+04

 -------- xd -------      1.0000000000000000              142.33206476053772
```

<< skipped >>

```
cgres_0              11    1.54020947332860286E-015    2.29884192443332537E-013
 lambda =    0.1043609807D+01               e2 = -0.22403882242063287D+03
 lambda =    0.2087219613D+01               e3 = -0.22403690586438574D+03
 lambda =    0.9992959948D+00               e5 = -0.22403882557855775D+03
```

```
 lambda =   0.9994767724D+00                e6 = -0.22403882557861556D+03
 lambda =   0.9994768D+00       pell = -0.224038825578616D+03   -0.32D-02     0.67D-02
 cgres_0              11   1.30812015597880167E-015   1.95274516420379816E-013
 #iteration=               2
 cgres_0              15   3.46131678207414862E-021   5.61917090100347095E-013
 penalized log likelihood = -0.224038825578616D+03        rss1 = 0.00000D+00
 #2: w1,w2,w3,w4 = 0.24251537D+00 0.24251537D+00 0.00000000D+00 0.10000000D+01
   abic =-0.2578557275D+03    -l = 0.2735685405D+02    pn = 0.3141466440D+03

 -------- xd -------     7.0000000000000000            -257.85572753556761
 ####    iteration, f, epsilon =       10 -0.25785573D+03    0.40149378D-02
                                         <<   skipped   >>
ptdet = 0.1266242306393D+03
 #1: w1,w2,w3,w4 =    0.24444285D+00 0.24444285D+00 0.00000000D+00 0.10000000D+01
 wx,wy= 0.24444D+00wxx,wyy= 0.00000D+00      pell = -0.223521871356295D+03
 lambda =   0.5000000000D+00                e2 = -0.22352266609096029D+03
 lambda =   0.1000000000D+01                e3 = -0.22352295814051013D+03
 lambda =   0.2000000000D+01                e3 = -0.22352203375141028D+03
 lambda =   0.1040406199D+01                e5 = -0.22352295978543751D+03
 lambda =   0.1040441881D+01                e6 = -0.22352295978543987D+03
 lambda =   0.1040442D+01       pell = -0.223522959785440D+03   -0.21D-02     0.79D-02
 cgres_0              11   1.61880861331118845E-015   1.50704974760771075E-013
 #iteration=               1
 cgres_0              11   3.85521606748695143E-016   2.26906058674522855E-013
 lambda =   0.1040441881D+01                e2 = -0.22352336816547259D+03
 lambda =   0.2080883761D+01                e3 = -0.22352289158507637D+03
 lambda =   0.1000350532D+01                e5 = -0.22352336882549560D+03
 lambda =   0.1000261750D+01                e6 = -0.22352336882549969D+03
 lambda =   0.1000262D+01       pell = -0.223523368825500D+03   -0.82D-03     0.17D-02
 cgres_0              11   3.70289506047125967E-016   2.17923635050981849E-013
 #iteration=               2
 cgres_0              16   3.06587908896622114E-023   7.84642264502102871E-014
 penalized log likelihood = -0.223523368825500D+03        rss1 = 0.00000D+00
 #2: w1,w2,w3,w4 = 0.24444285D+00 0.24444285D+00 0.00000000D+00 0.10000000D+01
   abic =-0.2578532621D+03    -l = 0.2652255568D+02    pn = 0.3158177062D+03

 ####    iteration, f, epsilon =       11 -0.25785573D+03    0.87167224D-03
 x =     0.49245849D+00
```

The rows including "`lambda =`" show values of the negative penalized log likelihood (`pell`) in the linear searching procedure. The rows including `lambda` without a number attached show the minimized value and sum of squares of the gradient vector components of the `pell` function with respect to the minimizing parameters. Furthermore, the `abic` value is minimized with respect to a weight `w1`. The rows with "-------- xd -------" shows every step where the minimum is updated by the simplex algorithm. The second to last rows show that the iterated simplex algorithm updated the ABIC for 11 times with the minimum `abic = -0.2578532621D+03`. This is attained by `w1 = w2 = 0.24444285D+00` (4th row from the bottom in the last second block, and the bottom row shows its logarithm). See Appendix A for the definitions and Appendix B for the numerical procedures.

The file `delo2d-poisson.prt` also includes a large volume of outputs. It may be useful to use UNIX command `egrep (grep)` to select specific items of interests. For example,

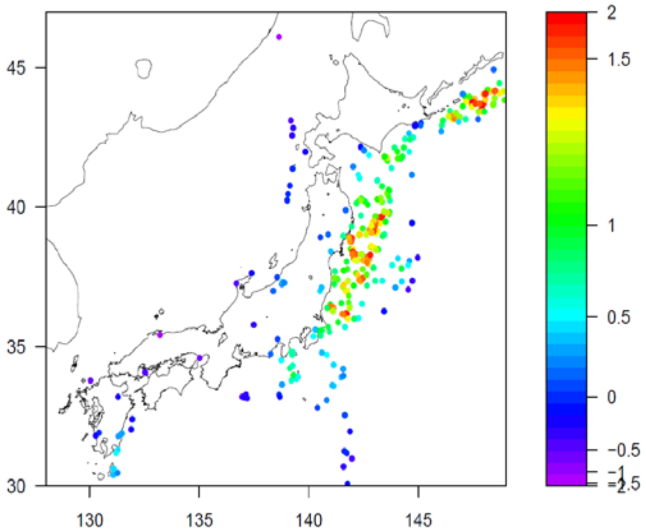
```
egrep xd delo2dpoisson.prt


egrep 'xd|abic' delo2d-poisson.prt
```

shows you just updated and all history of ABIC values, respectively.

For the spatial plot, done in R:

```
Program: delo2d-poisson.R
Reads:   delone2.out, delo2d-poisson.omap
Writes:  delo2d-poisson.pdf
```

which gives the following plot.



`delo2d-poisson.pdf`:

Fig. 7. Rainbow colors are in frequency-linearlised associated with logarithmic scale values

9 ETAS: Spatially Varying μ and K_0 parameters (`hist-etas-mk`)

This model in §A5.2 is almost the same as the space-time ETAS model as described in §5 except that the background rate μ and aftershock productivity K are location dependent. The parameters μ and K use a piecewise linear function defined on the Delaunay tessellations (§6). On the other hand, the location-independent parameters α, c, p, d and q are compensated from those obtained as the MLEs calculated by the above space-time ETAS program `st-etas`, depending on the estimation of location-dependent μ and K. All the used files in this section are selected in the program directory of `Section9files/` in the program package.

The program can take a considerable amount to time to converge, as the data size get large. An approximation of the model for a faster likelihood calculation is adjusted by `bi2`; that restrict the range of spatial distance of interaction between earthquakes; see "Line 4" of the configuration file in §5.2.

9.1 File Names

For the estimation phase, done in FORTRAN:

```
Program: hist-etas-mk.f
Object:  hist-etas-mk
Configuration: hist-etas-mk.conf
Reads:   delone2.out
Writes:  hist-etas-mk.prt
         simplex.root
         hist-etas-mk.upda
         hist-etas-mk.omap
```

9.2 Configuration File Format

An example of the configuration file is as follows. Parameters are read as free format. Note that "→" indicates that the record continues onto the following line, i.e. it is not split in the configuration file. It is *not* part of the input data configuration.

```
./delone2.out                  !maindata
21.0  17.0  14012.0  308       !tx,ty,tz,nn=#erthquakes
128.0 30.0 6.0 0.0 730.0 2.0 !xmin,ymin,xmg0,zmin,tstart,bi2
0                             !init
0                             !inits
1                             !initf
./hist-etas-mk.upda           !to be used in case init=1 to succeed calculations
0.0 1.d0 1.d0                 !w01,w1,w2
7                             !n=#of parameters
0. 29261E-06 0.21404E-01 0.17785E-02 0.10697E+01 0.86948E+00 0.18359E-01 →
0.15210E+01                   ! μ0,K0,c, α,p,d,q
0                             !if ipr = 7, printing the linear search
0 1.d0 1.d-0                  !nhesapp, dist,eps ( in subroutine simplex)
```

The data are interpreted as follows.

Line 1: Name of the data file, preceded by `./`.

Line 2: Width of region (`tx` degrees longitude), height of region (`ty` degrees latitude), end of observation period (`tz` days), number of events (`nn`) in dataset.

Line 3: Minimum longitude (`xmin` degrees), minimum latitude (`ymin` degrees), reference magnitude (`xmg0`) that can be usually a threshold magnitude of completely detected (`cutm` in §4.2), starting time of all data including precursory period for the history (`zmin` = 0 day), starting time of target period of estimation (`tstart = 730` days, in the current case), and an adjustment parameter called `bi2` (`=2.0`, in the current case), which restrict the range of spatial distance of interaction between earthquakes. For an explanation of `bi2`, see "Line 4" in §5.2.

Line 4: Value of `init`. If `init` is 0, then estimation starts at the beginning using the data file specified in Line 1 and initial parameter estimates given in Line 10. If `init` is 1, estimation continues from where a previous run was terminated.

The results of the previous run are placed in the file specified in Line 8 (`hist-etas-mk.upda`).

Line5: Value of `inits`. If `1`, the file containing the simplex optimization history from the previous run is used, `0` if it is not to be read. This information is contained in the file with `simplex.root`. There is a possibility that this will not work, in which case `inits` should be set to `0`.

Line 6: Value of `initf`. This is related to the grid search of the weights `w3,w5`, and `w7` the "`hist-etas5pa`" model (see §5 and §11.5 and §A5).

Line 7: File name containing estimation information from a previously incomplete run. It is the file `hist-etas-mk.upda`. This information can be used as a good starting point for the new run. This file includes the updated estimates of baseline parameters and the numbers in lines 10~11 are ignored in the case where `init=1.`

Line 8: Weights for the flatness constraints (`w1,w2`) of Delaunay piecewise linear function. The first weight `w01` represents the dumping penalty that is imposed only on the vertices on the boundary of the region. See A.6.2 for definition and details.

Line 9: Number of initial model parameters listed on lines 10~11.

Line 10~11: Initial estimates of baseline parameters. We recommend using the estimates is computed by the program `st-etas` given in `st-etas.prt` (see §5). These inputs estimates are ignored in the case where `init=1`

Line 12: Index `ipr` for printing the linear search results in `hist-etas-mk.prt.` If `ipr` = `0`, no printing, otherwise printing the linear search result.

Line 13: Adoption of the approximated Hessian matrix (`nhesapp=1`); initial distance for simplex search; and error bounds for the criteria of the simplex convergence (penalized log-likelihood) used in subroutine `simplex.` The other parameters that may require adjusting within the FORTRAN code are `dist` and `eps`. The first adjusts the search criterion (size of the simplex), and the second sets the convergence criterion.

9.3 Executing the Program

When `init=0`, parameter estimation does not start from where a previous run of `hist-etas-mk` terminated. Hence the file specified in Line 6 of the configuration file is not used. The model fitting starts from the initial parameter values specified in the configuration file. Execute as

```
./hist-etas-mk |tee hist-etas-mk.prt
```

When `init=1` information from a previous run is used, namely that contained in the file `hist-etas-mk.upda`. For more accurate estimation, we set a larger value of `bi2` such as 4 or 8, and we can use the previously obtained estimates. Since the new job will also write a file with the same name, we recommend copying and keeping the original `hist-etas-mk.upda.` Hence this name is specified on Line 9 of the configuration file.

9.4 Output of Calculation Process

An example of the program output (`hist-etas-mk.prt`) is as follows.

```
delone2.out
       21.0         17.0     14012.0          308
      730.0
     128.00        30.00         6.00         0.00       730.00         2.00
 input device             10
 nfunct=                17
 nn,ntstar,nnc               308                16               292
 tx,ty,tz,xmin,ymin,xmg0,zmin,tsta
     16.435       17.000 14012.000     128.000      30.000       6.000       0.000     730.000
 nn =    308 nnc =    292
 jmax               61
 n=             7
 para-init=    0.29261E-06     0.21404E-01     0.17785E-02     0.10697E+01     0.86948E+00 →
0.18359E-01     0.15210E+01
 linear ipr              0
 nhesapp,simplex(dist,eps)                 0     1.0000000000000000             1.0000000000000000
 n=             7
 non-pos diag.              339   -8.7469312074247600            586.78388345246196
 non-pos diag.              339   -8.7469312074247600            586.78388345246196
    588.42278020120909             588.42278020120909
 ptdet = 0.1176845560402D+04
 #s: w1,w2 =    0.100D+01    0.100D+01
 Initial Penalized log likelihood =      10089.469022236997
 lambda =    0.4228466D+00       pell =    0.478286492088103D+04    -0.51D+05      0.49D+06
 lambda =    0.3465761D+00       pell =    0.303540668222858D+04    -0.36D+05      0.75D+05
 lambda =    0.1063103D+01       pell =    0.194855632224969D+04    -0.18D+04      0.20D+05
 lambda =    0.1064288D+01       pell =    0.185052756355960D+04    -0.16D+03      0.61D+03
 lambda =    0.7674186D+00       pell =    0.183141972939415D+04    -0.52D+02      0.18D+03
 lambda =    0.8041680D+00       pell =    0.182171531754999D+04    -0.23D+02      0.49D+02
 lambda =    0.7680201D+00       pell =    0.181722348272042D+04    -0.12D+02      0.31D+02
 lambda =    0.7644997D+00       pell =    0.181447949500491D+04    -0.69D+01      0.98D+01
 lambda =    0.7265978D+00       pell =    0.181309094850632D+04    -0.39D+01      0.68D+01
 lambda =    0.7025772D+00       pell =    0.181225563224721D+04    -0.23D+01      0.26D+01
 lambda =    0.6575894D+00       pell =    0.181181504512178D+04    -0.14D+01      0.18D+01
 lambda =    0.6419689D+00       pell =    0.181154264428305D+04    -0.82D+00      0.81D+00
 lambda =    0.6188474D+00       pell =    0.181139001000143D+04    -0.50D+00      0.62D+00
 lambda =    0.6180905D+00       pell =    0.181129162444832D+04    -0.31D+00      0.29D+00
 lambda =    0.6092636D+00       pell =    0.181123423083046D+04    -0.19D+00      0.23D+00
 lambda =    0.6101669D+00       pell =    0.181119729576200D+04    -0.12D+00      0.11D+00
 lambda =    0.6068614D+00       pell =    0.181117528517862D+04    -0.73D-01      0.87D-01
 lambda =    0.6069101D+00       pell =    0.181116129443409D+04    -0.46D-01      0.40D-01
 lambda =    0.6059907D+00       pell =    0.181115284215355D+04    -0.28D-01      0.33D-01
 lambda =    0.6053561D+00       pell =    0.181114752142969D+04    -0.17D-01      0.15D-01
 lambda =    0.6055761D+00       pell =    0.181114427741159D+04    -0.11D-01      0.13D-01
 lambda =    0.6045714D+00       pell =    0.181114224762131D+04    -0.67D-02      0.59D-02
 lambda =    0.6053656D+00       pell =    0.181114100240355D+04    -0.41D-02      0.48D-02
 lambda =    0.6041784D+00       pell =    0.181114022595194D+04    -0.26D-02      0.22D-02
 lambda =    0.6052692D+00       pell =    0.181113974761236D+04    -0.16D-02      0.18D-02
 lambda =    0.6039944D+00       pell =    0.181113944988561D+04    -0.98D-03      0.85D-03
 lambda =    0.6052458D+00       pell =    0.181113926593057D+04    -0.61D-03      0.70D-03
 lambda =    0.6039214D+00       pell =    0.181113915152839D+04    -0.38D-03      0.33D-03
 lambda =    0.6052650D+00       pell =    0.181113908069490D+04    -0.23D-03      0.27D-03
 lambda =    0.6039132D+00       pell =    0.181113903665419D+04    -0.15D-03      0.12D-03
 lambda =    0.6053139D+00       pell =    0.181113900934350D+04    -0.90D-04      0.10D-03
 lambda =    0.6039305D+00       pell =    0.181113899236161D+04    -0.56D-04      0.48D-04
 lambda =    0.6054994D+00       pell =    0.181113898181821D+04    -0.35D-04      0.40D-04
 lambda =    0.6035771D+00       pell =    0.181113897526049D+04    -0.22D-04      0.18D-04
 lambda =    0.6059521D+00       pell =    0.181113897118520D+04    -0.13D-04      0.15D-04
 lambda =    0.6035625D+00       pell =    0.181113896864956D+04    -0.84D-05      0.70D-05
 lambda =    0.6059704D+00       pell =    0.181113896707257D+04    -0.52D-05      0.59D-05
 lambda =    0.6035741D+00       pell =    0.181113896609097D+04    -0.33D-05      0.27D-05
 lambda =    0.6060786D+00       pell =    0.181113896548008D+04    -0.20D-05      0.23D-05
 lambda =    0.6035508D+00       pell =    0.181113896509967D+04    -0.13D-05      0.10D-05
```

```
lambda =   0.6062551D+00     pell =   0.181113896486279D+04   -0.78D-06    0.87D-06
lambda =   0.6048568D+00     pell =   0.181113896471523D+04   -0.49D-06    0.40D-06
lambda =   0.6036937D+00     pell =   0.181113896462330D+04   -0.30D-06    0.34D-06
lambda =   0.6061852D+00     pell =   0.181113896456600D+04   -0.19D-06    0.16D-06
lambda =   0.6038110D+00     pell =   0.181113896453029D+04   -0.12D-06    0.13D-06
lambda =   0.6061742D+00     pell =   0.181113896450802D+04   -0.73D-07    0.60D-07
lambda =   0.6061742D+00     pell =   0.181113896449415D+04   -0.46D-07    0.50D-07
lambda =   0.6016804D+00     pell =   0.181113896448548D+04   -0.29D-07    0.23D-07
lambda =   0.6085610D+00     pell =   0.181113896448009D+04   -0.18D-07    0.20D-07
lambda =   0.5957944D+00     pell =   0.181113896447671D+04   -0.11D-07    0.90D-08
lambda =   0.6163351D+00     pell =   0.181113896447461D+04   -0.68D-08    0.75D-08
lambda =   0.5983139D+00     pell =   0.181113896447330D+04   -0.44D-08    0.35D-08
lambda =   0.6122871D+00     pell =   0.181113896447248D+04   -0.27D-08    0.29D-08
lambda =   0.6122871D+00     pell =   0.181113896447197D+04   -0.17D-08    0.13D-08
lambda =   0.5848831D+00     pell =   0.181113896447165D+04   -0.11D-08    0.11D-08
lambda =   0.5848831D+00     pell =   0.181113896447145D+04   -0.64D-09    0.52D-09
lambda =   0.6719067D+00     pell =   0.181113896447133D+04   -0.37D-09    0.43D-09
lambda =   0.5502216D+00     pell =   0.181113896447125D+04   -0.28D-09    0.20D-09
lambda =   0.5983607D+00     pell =   0.181113896447120D+04   -0.14D-09    0.17D-09
penalized log likelihood =   0.181113896447120D+04
#e: w1,w2 =   0.100D+01   0.100D+01
  abic = 0.3655784386D+04 -l = 0.1850351202D+04 pn = 0.1210352017D+04
repeated davidn =             1
Surface Sliding: Old a1, a2=   2.92610000000000024E-007   2.14039999999999994E-002
Surface Sliding: sss1, sss2=    1678.3115727728357          -1349.0295398812370
Surface Sliding: New a1, a2=   3.95841650338503000E-005   4.14387812724830850E-004
-------- xd -------        1   0.365578438575E+04         0.000
a1-7 0.396E-04 0.414E-03 0.178E-02 0.107E+01 0.869E+00 0.184E-01 0.152E+01

w1-2 0.100E+01 0.100E+01

                                  << skipped >>
lambda =   0.1591954D+00     pell =   0.167273796149464D+04   -0.64D-09    0.11D-09
lambda =   0.3881579D+00     pell =   0.167273796149461D+04   -0.16D-09    0.98D-10
lambda =   0.1665214D+00     pell =   0.167273796149459D+04   -0.27D-09    0.76D-10
lambda =   0.4590164D+00     pell =   0.167273796149457D+04   -0.95D-10    0.61D-10
lambda =   0.1529797D+00     pell =   0.167273796149455D+04   -0.22D-09    0.45D-10
penalized log likelihood =   0.167273796149455D+04
#e: w1,w2 =   0.823D-01   0.809D+00
  abic = 0.3526282973D+04 -l = 0.2172458180D+04 pn = 0.4366366499D+03
repeated davidn =             1
Surface Sliding: Old a1, a2=   1.13588618541079640E-004   6.72600220400718243E-005
Surface Sliding: sss1, sss2=    2.5486238908657164          -238.50222149897596
Surface Sliding: New a1, a2=   1.14438256042319838E-004   3.34881323324280578E-005
-------- xd -------        9   0.352628297294E+04        -5.110
a1-7 0.114E-03 0.335E-04 0.647E-02 0.146E+01 0.104E+01 0.210E-01 0.233E+01
w1-2 0.823E-01 0.809E+00
                                  << skipped >>
lambda =   0.2309417D+00     pell =   0.167026661988171D+04   -0.11D-09    0.47D-10
lambda =   0.3459302D+00     pell =   0.167026661988169D+04   -0.84D-10    0.42D-10
lambda =   0.1827652D+00     pell =   0.167026661988168D+04   -0.13D-09    0.38D-10
lambda =   0.4000000D+00     pell =   0.167026661988167D+04   -0.58D-10    0.33D-10
lambda =   0.1842672D+00     pell =   0.167026661988166D+04   -0.11D-09    0.30D-10
penalized log likelihood =   0.167026661988166D+04
#e: w1,w2 =   0.916D-01   0.443D+00
  abic = 0.3533595745D+04 -l = 0.2254182084D+04 pn = 0.2805016139D+03
repeated davidn =             1
####   iteration, f, epsilon =      43   0.35262830D+04   0.96234391D+00
x =   -0.24971914D+01 -0.21167911D+00 -0.50405642D+01   0.37993810D+00 →
x =     0.37616517D-01 -0.38631250D+01   0.84729410D+00
```

The above lists ABIC values, the final parameter estimates, and the penalised log-likelihoods. The numbers in last column are the sum of squares of all the gradient vector components of the coefficients. The progression to smaller values as one goes down the output indicates that the computations are converging. The third to last rows show that the iterated (7-dimensional) simplex algorithm updated the ABIC value, 9 times updates, out of 43 ABIC calculation trials to get the smallest value of ABIC = 0.352628297294E+04 (indicated by `'- xd -';` the 16th row from the output bottom) which is attained by `w1 =` 0.823E-01 and `w2 =` 0.809E+00 ('w1-2'; the 15th row from the bottom), and the third last row (indicated by ####) from the output bottom summarizes iterated numbers, the smallest ABIC value, and the difference from the second smallest ABIC is 0.96234391D+00. The 16th row from the output bottom ('a1-7') shows the baseline parameters of μ, K_0, c, α, p, d and q, and the bottom two rows show their logarithmic values. See Appendix A for the definitions and Appendix B for the numerical procedures.

The file `hist-etas-mk.prt` includes a large amount of output. It may be useful to use the UNIX command `egrep(grep)` to extract items of interest, as done earlier,

```
egrep xd hist-etas-mk.prt
egrep xd|abic hist-etas-mk.prt
```

These will show you just the updated and all history of ABIC values, respectively.

An example of the program output (`hist-etas-mk.omap`) is as follows.

```
-0.2497E+01    -0.2117E+00      3533.60      684
 0.114438261954E-03  0.334881320043E-04  0.647009682182E-02  0.146219408001E+01→
 0.103833297334E+01  0.210022642758E-01  0.233332455954E+01
 -0.768090420081E+01 -0.980188281878E+01 -0.830621370392E+01 -0.797115464193E+01
 -0.893480205512E+01 -0.695269893920E+01 -0.889522839754E+01 -0.867387791505E+01
 -0.935157675729E+01 -0.898946493741E+01 -0.679076775508E+01 -0.912639592929E+01
 -0.793894681451E+01 -0.944728426317E+01 -0.942753542139E+01 -0.848239207321E+01
```

<< skipped till the end >>

Here the first line writes ln(w1), ln(w2), abic and twice of the number of coefficients in the Delaunay functions representing μ and K. The second and third lines give the optimized baseline parameters $\mu_0, K_0, c, \alpha, p, d$ and q. The fourth and following lines down to the bottom give logarithm of deviations from the baseline parameters μ_0 and K_0.

See R display procedure and example figures of the optimal maximum a posterior (OMAP) estimate in §11.5.

For a good initial estimation with the program `hist-etas5pa` in the next section, and the forecasting in §13.1, the following is the updated output file `hist-etas-mk.updа:`

```
   0.82316E-01     0.80922E+00     3526.28     684
  0.114438256042E-03   0.334881323324E-04   0.647009682182E-02   0.146219408001E+01→
  0.103833297334E+01   0.210022642758E-01   0.233332455954E+01
  0.139457092658E+01 -0.726407678229E+00   0.769261411269E+00   0.110432047804E+01
  0.140673083157E+00   0.212277617056E+01   0.180246755834E+00   0.401597166657E+00
 -0.276101629785E+00   0.860101851354E-01   0.228470737143E+01 -0.509208017719E-01
  0.113652828785E+01 -0.371809145881E+00 -0.352060305758E+00   0.593083039677E+00
 -0.683893507324E+00 -0.816700951127E+00   0.902848610295E-01   0.557331993929E+00
 -0.138820519275E+01   0.108533576988E+01 -0.180759606983E+00 -0.107075073680E+00
```

<< skipped >>

```
 0.241695424598E+00   0.511075481092E+00 -0.283118737557E+00   0.953208296924E+00
  0.114626199422E+00   0.231870568842E+00 -0.398161180796E+00 -0.180628788391E+00
 -0.114291272430E+01 -0.378213066618E+00 -0.859583923866E-01 -0.129659654388E+00
 -0.117930858809E+00 -0.528690530145E+00 -0.605507186335E+00 -0.173845062437E+00
 -0.152130753387E+00 -0.108836609546E+01 -0.907255257880E+00 -0.737753342053E+00
 -0.460987874419E+00 -0.232974758562E+00 -0.557155856504E+00 -0.591710718465E+00
 -0.471950174999E+00 -0.289090736345E+00 -0.239500503238E+00 -0.261449294363E+00
 -0.377911671516E+00 -0.547351463463E+00 -0.225341192861E+00 -0.606005843722E+00
 -0.246721505150E+00 -0.638621609165E+00 -0.383947873387E+00 -0.540608762369E+00
 -0.244800152526E+00 -0.179074966412E+00 -0.137299967512E+00 -0.135008778881E+00
```

Here the first line has w1, w2, abic and twice of number of coefficients in the Delaunay functions representing μ and K. The second and third lines give the optimized baseline parameters $\mu_0, K_0, c, \alpha, p, d$ and q. The fourth line down to the bottom give logarithm of the OMAP estimates taking account of the baseline parameters μ_0 and K_0.

9.5 Additional Advice

The current program `hist-etas-mk` is the most time consuming because of the 7 dimensional simplex optimization procedure for the reference parameters of μ, K_0, c, α, p, d and q besides the high-dimensional quasi-Newton and Newton optimizations. Nevertheless, assuming that we can use initial reference parameters with the MLEs of (`st-etas`) in §5 that converged well (see §5.4), it can take a shorter time in converging the `hist-etas-mk` program than the default case; specifically, try a short distance in the step-size of the simplex procedure. This implementation corresponds to replacing `dist=1.0` (the default) by `dist=0.05`, for example, in the last line of `hist-etas-mk.conf` and then run it.

10 ETAS: Spatial Variation in 5 Parameters (hist-etas5pa)

This model is referred to as the five-parameter model because it allows five of the parameters to vary in space, i.e. μ, K_0, α, p and q. The parameters c and d are assumed to be constant in space, and fixed throughout the computation procedure. For further mathematical detail, see §A.5.3. This program should be undertaken after having obtained the optimal estimates by `hist-etas-mk` as described in the previous section. All the used files in this section are selected in the program directory of `Section10files/` in the program package.

10.1 File Names

For the estimation phase, done in FORTRAN:

```
Program:        hist-etas5pa.f
Object:         hist-etas5pa
```

```
Configuration:  hist-etas5pa.conf
Reads:          delone2.out, hist-etas-mk.upda
Writes:         hist-etas5pa.upda,
                hist-etas5pa.omap
```

10.2 Configuration File Format

The program can take a considerable amount of time to converge, depending on the number of earthquakes. It is possible that a job may exceed queue time and be terminated by the system before it has converged. An approximation of the model, giving a faster likelihood calculation, is provided by `bi2`; see “Line 5” in §5.2. To restart the job at roughly the same place, specifically where it last wrote solution information to the disk, the configuration file needs modification. The files that track the convergence process are `hist-etas5pa.prt` and `simplex.rootu` (simplex root information).

An example of the configuration file `hist-etas5pa.conf` is as follows. Parameters are read as free format. Note that “→” indicates that the record continues onto the following line, i.e. it is not split in the configuration file. It is *not* part of the input data. The configuration file would generally have the following format when one first runs this program (see detailed explanation of each line below). Notice that, unlike the previous programs, `init` on line 6 is generally set to one. In the present example, this has the effect of using the output in file `hist-etas-mk.omap.`

```
./delone2.out                   !main data
21.0  17.0  14012.0  308        !tx,ty,tz,nn=#earthquakes
128.0 30.0  6.0  0.0 730.0 2.0  !xmin,ymin,xmg0,zmin,tstart,bi2
0                               !init
0                               !inits
1                               !initf
./hist-etas-mk.upda             !approximate solution for initial estimate
0. 1000.                        !w00, w01,
10. 100. 1000.                  !w3, w4, w5
0                               !if ipr = 7, printing the linear search results
0.1d-3  0.1d-3                     !tau1,tau2(davidn)
0.1d-3  0.1d-3                     !eps1,eps2(davidn)
0 1.d0 0.5d-0                   !nhesapp, dist,eps ( in subroutine simplex)
```

The data are interpreted as follows:

Line 1: Name of the data file, preceded by `./`.

Line 2: Width of region (`tx` degrees longitude), height of region (`ty` degrees latitude), end of observation period (`tz` days), number of events (`nn`) in dataset.

Line 3: Minimum longitude (`xmin` degrees), minimum latitude (`ymin` degrees), threshold magnitude (`xmg0`), minimum time (`zmin`), starting time (`tstart` 730 days), and an adjustment parameter called `bi2`. For and explanation of `bi2`, see “Line 5” in §5.2.

Line 4: Value of `init`. If `init` is 0, then estimation starts at the beginning using the data file `delone2.out` as specified in Line 1 and `hist-etas-mk.upda` as given in Line 7. If `init` is 1, estimation starts by replacing `hist-etas-mk.upda` in Line 7 by `hist-etas5pa.upda`.

Line 5: Value of `inits`. If `1`, the file containing the simplex optimization history from a previous run is used, `0` if it is not to be read. This information is contained in the file with `simplex.root`. There is a possibility that this will not work, in which case `inits` should be set to `0`.

Line 6: Value of `initf`. If `initf` = 1 then the program will only utilise the weights w_3,w_4, and w_5 as given in line 8. If `initf` = 0 then the simplex program searches for optimal weights (w_1, w_2, w_3, w_4, w_5) by minimizing ABIC, which takes a substantial CPU time. For the grid search of (w_3, w_4, w_5) with the fixed (w_1, w_2) that are optimized by `hist-etas-mk .f` the former should be used.

The coefficient parameters may not always be converged in case of `initf` = 1 because the Hessian matrix does not become positive-definite, when, for example, the weights of (w_3, w_4, w_5) is too small. Usually, weights for α, p and q of the HIST-ETAS model are not necessary to seek the values in accurate, and it is not bad idea to make a grid search. To execute grid searches, set `initf` = 0. Regarding grid exploration, the ninth line provides the default weights, but if the data size is a large, they can be (1000., 1000., 1000.), for example, to converge it in one loop, so remember its ABIC value for the comparisons as follows: namely, in addition for example, (1000., 10000., 10000.), (100., 100., 10000.), (100., 1000., 10000.) (10., 10., 1000.), and so on, to find the smaller ABIC value. Ignore combinations of smaller weights that still do not converge. In our experience, if the weight of (w_3, w_4, w_5) is too small, the Hessian matrix will not be positive-definite and the coefficient parameters will not converge. Especially, the weight w5 may be large because the variable parameter q does not likely change so much.

Line 7: File name containing estimation information from a previously incomplete run. It is the file with the suffix `.upda`. This information can be used as a starting point for the new run. In the case where `init` is 0, `hist-etas-mk.upda` as given in Line 7 provides an optimal initial estimates of the baseline parameters μ_0, K_0, c, α, p, d and q, and the coefficients of Delaunay functions representing μ and K. The coefficients of the Delaunay functions representing α, p and q are all set to 0 in the program. In the case where `init` is 1, the baseline parameters are the same, but the coefficients of Delaunay functions representing μ, K, α, p and q are all going to be updated, starting from those in `hist-etas5pa.upda.`

Line 8: Weights for the flatness constraints of Delaunay piecewise linear function. The first weight w_{00} represents the dumping penalty for all parameters at all vertices of Delaunay triangles, and w_{01} represents the same dumping penalty imposed only on the vertices on the boundary of the region. See A.6.2 for definition and details.

Line 9: Initial weights for the flatness constraints (w_3,w_4,w_5) of the Delaunay piecewise linear functions. See A.6.2 for definition and details. In the case of grid search of weights w_3, w_4 and w_5 for the penalty of α, p and q, these are different by exponential orders as given in Line 8, according to our experience in finding optimal weights by minimizing ABIC value.

Line 10: Index `ipr` for printing the linear search results in `hist-etas5pa.prt.` If `ipr` = `0`, no printing, otherwise printing the linear search result.

Lines 11 and **12:** convergence criteria in subroutine `davidn`.

Line 13 Adoption of the approximated Hessian matrix (`nhesapp=1`); initial distance for simplex search; and error bounds for the criteria of the simplex convergence (penalized log-likelihood) used in subroutine `simplex.` The other parameters that may require adjusting within the FORTRAN code are `dist` and `eps`. The first adjusts the search criterion (size of the simplex), and the second sets the convergence criterion.

10.3 Executing the Program

The following command executes the compiled FORTRAN code.

```
./hist-etas5pa |tee hist-etas5pa.prt
```

10.4 Output Produced by Program

An example of the program output (`hist-etas5pa.prt`) is as follows.

```
delone2.out
       21.0          17.0      14012.0            308
       128.           30.            6.              0.         730.            2.
 input device            10
0tx,ty,tz,xmin,ymin,xmg0,zmin,tsta
     16.435       17.000 14012.000      128.000       30.000        6.000         0.000      730.000
 nn =    308 nnc =    292
 jmax,bi2                  61      2.00000000000000
            308
 w00-w7    0.000000000000000E+000      1000.00000000000           8.232000000000000E-002
  0.809200000000000              10.0000000000000              100.000000000000 →
   1000.00000000000
 linear ipr                   0
 davidn tau    1.000000000000000E-004    1.000000000000000E-004
 davidn eps    1.000000000000000E-004    1.000000000000000E-004
 nhesapp,simplex(dist,eps)                  0      1.00000000000000       0.500000000000000
 n=                  7
 w1-w7    8.232000000000000E-002    0.809200000000000              10.0000000000000 →
   100.000000000000              1000.00000000000
 a1-a7    1.144382542340000E-004    3.348813210360000E-005    6.470096821820000E-003 →
  1.46219408001000    1.03833297334000    2.100226427580000E-002    2.33332455954000
 w00, w01 =    0.000000000000000E+000      1000.00000000000
 non-pos diag.                339    -106.26040147013342             -257.26626201915627
 non-pos diag.                339    -10.809089255604786              515.23453657230607
 ptdet = 0.7637790019860D+04
 repeated davidn =                1
 #s: w1,w2,w4,w5,w7 =    0.823D-01    0.809D+00    0.100D+02    0.100D+03    0.100D+04
 Initial Penalized log likelihood =      1672.7379602189433
 lambda =    0.8744497D+00       pell =    0.167044234940854D+04    -0.52D+01      0.60D+04
 lambda =    0.9592622D+00       pell =    0.166605016259747D+04    -0.93D+01      0.27D+04
 lambda =    0.3613464D+00       pell =    0.166533962998344D+04    -0.38D+01      0.22D+04
 lambda =    0.1055977D+01       pell =    0.166495986340342D+04    -0.73D+00      0.47D+02
 lambda =    0.3649882D+00       pell =    0.166475860787496D+04    -0.11D+01      0.79D+03
 lambda =    0.6678868D+00       pell =    0.166467378583788D+04    -0.25D+00      0.19D+02

                                   <<skipped>>

 lambda =    0.9000000D+00       pell =    0.166426648388209D+04    -0.55D-10      0.32D-07
 lambda =    0.1377750D+00       pell =    0.166426648388207D+04    -0.31D-09      0.25D-07
```

```
lambda =   0.1152778D+01     pell =   0.166426648388205D+04   -0.41D-10    0.24D-07
lambda =   0.1151316D+00     pell =   0.166426648388203D+04   -0.32D-09    0.18D-07
lambda =   0.2285714D+01     pell =   0.166426648388200D+04   -0.28D-10    0.17D-07
penalized log likelihood =   0.166426648388199D+04
#e: w1,...,w5 =   0.823D-01   0.809D+00   0.100D+02   0.100D+03   0.100D+04
 abic=   0.3569086244E+04   -l= -0.5841625728E+03   pn=   0.7878343296E+04
-------- xd -------              1     3569.0862435900335
a1-7 0.114E-03 0.335E-04 0.647E-02 0.146E+01 0.104E+01 0.210E-01 0.233E+01
w1-7 0.000E+00 0.100E+04 0.823E-01 0.809E+00 0.100E+02 0.100E+03 0.100E+04
```

<< skipped >>

```
lambda =   0.7812500D+00     pell =   0.167574329557346D+04   -0.60D-10    0.34D-07
lambda =   0.1547237D+00     pell =   0.167574329557345D+04   -0.23D-09    0.27D-07
lambda =   0.1063830D+01     pell =   0.167574329557342D+04   -0.44D-10    0.26D-07
lambda =   0.1424074D+00     pell =   0.167574329557340D+04   -0.27D-09    0.18D-07
lambda =   0.1904255D+01     pell =   0.167574329557337D+04   -0.28D-10    0.18D-07
penalized log likelihood =   0.167574329557337D+04
#e: w1,...,w5 =   0.125D+00   0.599D+00   0.152D+02   0.415D+02   0.152D+04
 abic=   0.3562608863E+04   -l= -0.5869629136E+03   pn=   0.7877466597E+04
-------- xd -------              4     3562.6088628260377
a1-7 0.114E-03 0.335E-04 0.647E-02 0.146E+01 0.104E+01 0.210E-01 0.233E+01
w1-7 0.000E+00 0.100E+04 0.125E+00 0.599E+00 0.152E+02 0.415E+02 0.152E+04

####   iteration, f, epsilon =       2   0.35626089D+04   0.10703394D+01
x =   -0.41543796D+01 -0.10233689D+01   0.54451702D+01   0.74503404D+01 →
x =     0.14655511D+02
```

<< skipped >>

```
lambda =   0.1500000D+00     pell =   0.168438504518700D+04   -0.78D-09    0.25D-07
lambda =   0.1542857D+01     pell =   0.168438504518697D+04   -0.39D-10    0.24D-07
lambda =   0.1542857D+00     pell =   0.168438504518694D+04   -0.30D-09    0.15D-07
lambda =   0.1266667D+01     pell =   0.168438504518693D+04   -0.24D-10    0.15D-07
lambda =   0.1449091D+00     pell =   0.168438504518692D+04   -0.16D-09    0.10D-07
penalized log likelihood =   0.168438504518692D+04
#e: w1,...,w5 =   0.159D+00   0.506D+00   0.130D+02   0.670D+02   0.135D+04
 abic=   0.3563238642E+04   -l= -0.6202776694E+03   pn=   0.7944725887E+04
####   iteration, f, epsilon =       5   0.35626089D+04   0.30092985D+00
x =   -0.41543796D+01 -0.10233689D+01   0.54451702D+01   0.74503404D+01 →
x =     0.14655511D+02
```

The numbers in last column are the sum of squares of all the gradient vector components of the coefficients. The progression to smaller values as one goes down the output indicates that the computations are converging. The second to last row shows that the iterated simplex algorithm updated the ABIC for 16 times with the smallest abic= 0.3562608863E+04. This is attained by w1,...,w5 = 0.125D+00 0.599D+00 0.152D+02 0.415D+02 0.152D+04 (in the two lines before " -------- xd ------- 4 3562.6088628260377"), and the bottom row shows their logarithms. See Appendix A for the definitions and Appendix B for the numerical procedures.

The file `hist-etas5pa.prt` includes a large amount of output. It may be useful to use the UNIX command `egrep (grep)` to extract lines of interest, for example,

```
egrep xd hist-etas5pa.prt
egrep 'xd | abic' hist-etas5pa.prt
```

show you just updated and all history of ABIC values, respectively.

An example of the program output (`hist-etas5pa.omap`) is as follows.

```
0.125287875259E+00   0.599470104176E+00   0.152196155562E+02   0.414782911682E+02 →
0.152196155562E+04   0.356323892700E+04        1710              0.100000000000E+04
```

```
0.114438254234E-03  0.334881321036E-04  0.647009682182E-02  0.146219408001E+01 →
0.103833297334E+01  0.210022642758E-01  0.233332455954E+01
0.374513210667E-03  0.675979122716E-04  0.265190771687E-03  0.354908735402E-03
0.146433535062E-03  0.916947720816E-03  0.146338679517E-03  0.177204790397E-03
0.922917100686E-04  0.114314194061E-03  0.106254935918E-02  0.112398922637E-03
0.359109874846E-03  0.917853243391E-04  0.926876136674E-04  0.210983087983E-03
0.513378081538E-04  0.610030246154E-04  0.151535542538E-03  0.237911207613E-03
```

<< skipped >>

Here the first and second lines contain w_1, w_2, w_4, w_5, w_7, *abic*, and number of all coefficients 1710 = 5 * (308+34) where 308 represents the number of earthquake and 34 represents Delaunay apex on the boundary of the region; the last column represents the fixed dumping weight w_{01} in the 8th row of the configuration file `hist-etas5pa.conf`.

The third and fourth lines give the optimized baseline parameters $\mu_0, K_0, c, \alpha, p, d$ and q. The remaining values from fifth line to the bottom give logarithm of the location-dependent deviations from the baseline parameter values μ_0, K_0, α, p and q.

See R display procedure and example figures of the optimal maximum a posterior (OMAP) estimate in §11.5.

Part III. PLOTTING SPATIAL PARAMETER ESTIMATES

11 Plot Spatial Variation of Parameters

This R program plots the Delaunay tessellation of various datasets; spatial intensity rate, location-dependent b-values of Gutenberg-Richter magnitude distribution, the spatial estimates of the ETAS parameters μ and K_0, and location-dependent 5 parameters μ, K_0, α, p and q. All of these are defined based on the Delaunay tessellations, over the observed spatial region. Incidentally, the users can use any available graphical packages for the display such as Matlab, Mathematica, GMT, etc., by making their own program scripts using the present Fortran programs. The provided R and below figures are to show the examples. All the used files in this section are selected in the program directory of `Section11files/` in the program package.

11.1 Delaunay Tessellation for Spatial Variation

```
Program:   delone-plot.R
Reads:     delone2.out
Requires:  drawmap.R, f2.R
Writes:    delone-plot.pdf;
```
see §6.5 for an example figure.

11.2 Spatial Occurrence Rate

```
Program:   delo2d-poisson.R
Reads:     delone2.out, delo2d-poisson.omap
Requires:  drawmap.R, f2.R
Writes:    delo2d-poisson.pdf;
```
see §8.3 for an example figure.

11.3 Spatially Varying *b*-Value of Magnitude Frequency

```
Program:   delo2d-bvalues.R
Reads:     delone2.out, delo2d-bvalues.omap
Requires:  drawmap.R, f2.R
Writes:    delo2d-bvalues.pdf;
```
see §7.3 for an example figure.

11.4 ETAS: Spatially Varying μ and K_0

```
Program:   hist-etas-mk.R
Reads:   delone2.out, hist-etas-mk.omap
Requires:  drawmap.R, f2.R
Writes:  hist-etas-mk.pdf;
```
see the following for an example figure (Fig. 8).

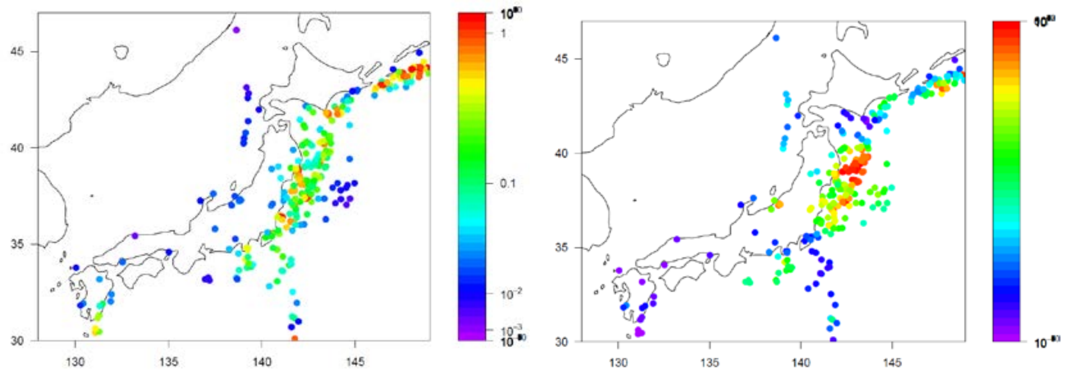


Fig. 8. `hist-etas-mk.pdf:` μ and K_0 in the order from the left to the right. The color table of K_0-values indicate that range of K_0-values change are very narrow.

11.5 ETAS: Spatial Variation in 5 Parameters

```
Program:   hist-etas5pa.R
Reads:   delone2.out, hist-etas5pa.omap
Requires:  drawmap.R, f2.R
Writes:  hist-etas5pa.pdf;
```
see the following for an example figure (Fig. 9).

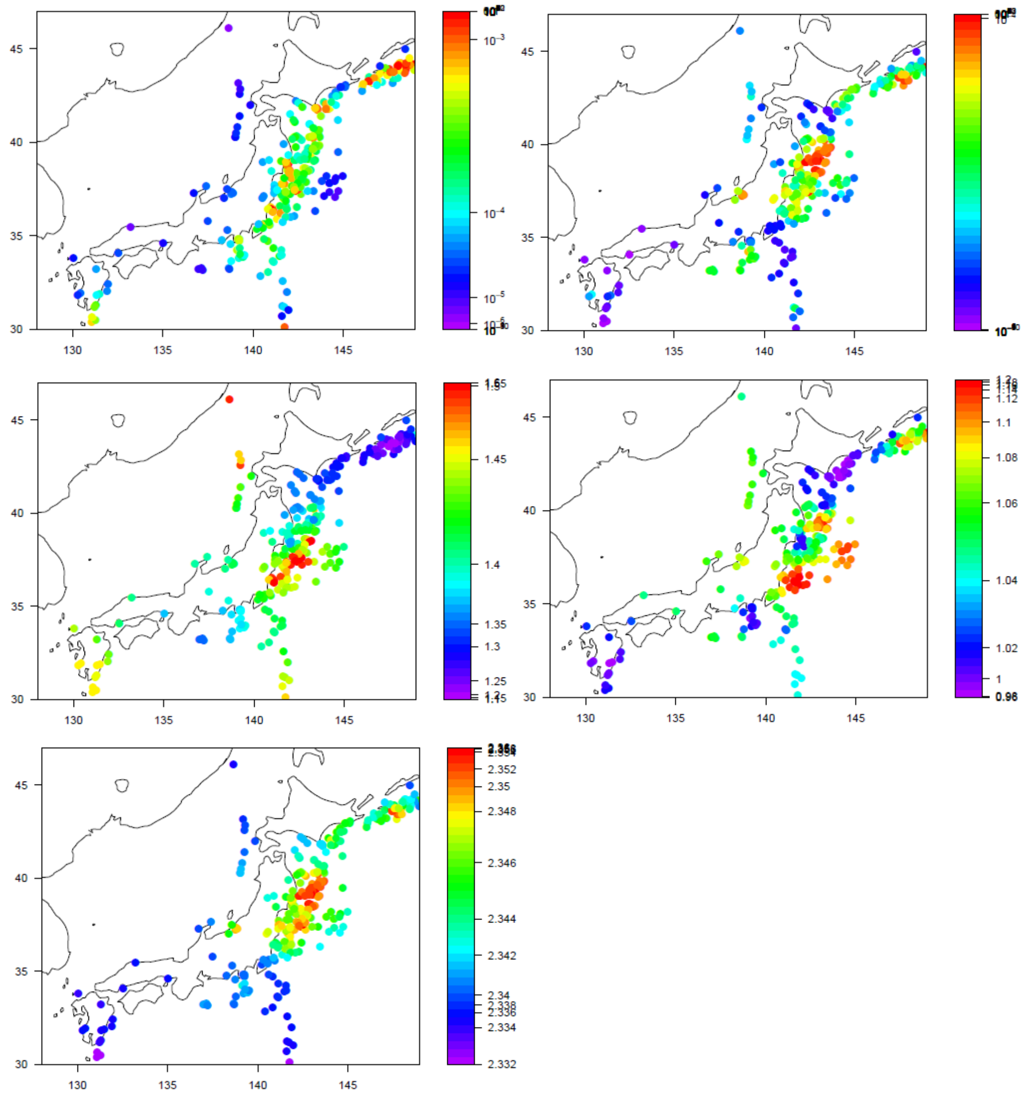


Fig. 9. `hist-etas5pa.pdf`: The estimated parameters (μ, K, α, p and q) in the order from the left to the right. The color table of K_0-values and q-values indicate that range of K_0-values and q-values change are very narrow.

12 Plot interpolated images by Delaunay triangles

The plotted color points in the last section shows the optimal maximum a posteri (MAP) estimates on earthquake event locations which are also vertices of the Delaunay triangles (see the figure in §6.6). The MAP estimates are subject to the interpolation on any lattice points by the Delaunay triangles which include the lattice point.

Note: The R-plotting procedures have been partly modified because the sub-module "`filled1.contour`" that was used in the previous version is no more available in the current R programme. Please use the followings from the program directory "`estimations`" in HIST-PPM-V2. In this version, we use `f2.r` instead, and to understand the new module, please consult "`help(filled contour)`" in `R` command. All the used files in this section are selected in the program directory of `Section12files/` in the program package.

`Program: interpolated.f !` Interpolation of the optimal MAP solution to lattice image

`Reads:    interpolated.conf, delone2.out,` and

either `delo2d-bvalues.omap` or `delo2d-poisson.omap`

`Writes: interplated.pixel !` Output pixel images on lattice points

The FORTRAN program `interpolated.f` works for both *b*-value images and Poisson intensity-rate image, whose configuration file `interpolated.conf` includes the following three lines:

```
delone2.out
delo2d-bvalues.omap
128.0 30.0 141. 144. 35. 41. 100 100!lon0, lat0, minlon maxlon, minlat, maxlat, nx, ny
```

For *b*-value images, this contains the following records; the first line includes the Delone structural data, and the second line includes the optimal MAP solution of *b*-values. The first two items (`128.0, 30.0`) in the third line indicate the origin of the considered region in longitude and latitude, and the following four items are longitudes and latitudes for the restricted region, and the last two numbers indicate division of the restricted rectangular region into pixels.

Then the following are output example of `interpolated.f`, with filename `Interplated.pixel.`

```
141.01   35.03   0.299E+00
141.01   35.09   0.306E+00
141.01   35.15   0.312E+00
141.01   35.21   0.319E+00
141.01   35.27   0.325E+00
141.01   35.33   0.331E+00
```

<< skipped >>

```
143.99   40.73   0.727E-01
143.99   40.79   0.773E-01
143.99   40.85   0.727E-01
143.99   40.91   0.682E-01
143.99   40.97   0.636E-01
```

Then we can use:

`Program:  interpolated-bvalues.R`

`Reads:    interplated-bvalue.pixel, interpolated-bvalues-conf`

`Requires: drawmap.R, delone2.out`

`Writes:   interpolated-bvalues.pdf;` see the right side for an example figure (Fig. 10).

Also, we can use:

Program: enlarge.R

Reads: delone2.out, interpolated.conf

Requires: drawmap.R, f2.R

Writes: enlarged.pdf (= enlarged-bvalues.pdf); see the left-hand-side figure (Fig. 10).

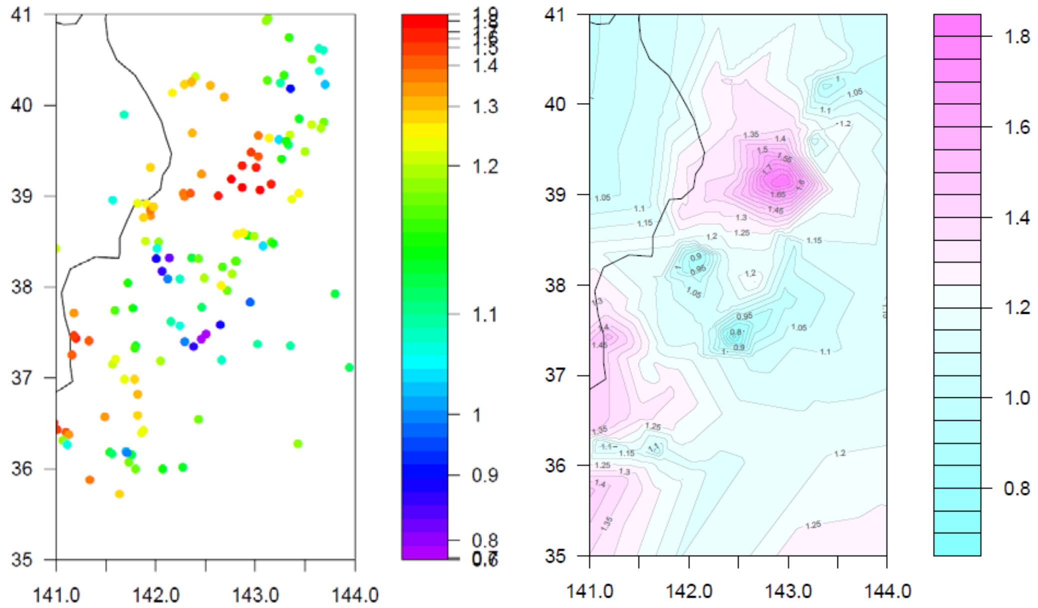


Fig. 10: enlarged-bvalues.pdf,image.pdf(=interpolated-bvalues.pdf)

For Poisson intensity rate image, the configuration file interplated-poisson.conf includes the following three lines:

```
delone2.out
delo2d-poisson.omap
128.0 30.0 141. 144. 35. 41. 100 100 !lon0, lat0, minlon maxlon, minlat, maxlat, nx, ny
```

containing the following records; the first line includes the Delone data, and the second line includes the OMAP solution of Poisson intensity rates. The first two items (128.0, 30.0) in the third line indicate the origin (longitude, latitude) of the full region, and the following four items are longitude and latitude for the enlarged region, and the last two numbers indicate division of the enlarged rectangular region into pixels.

Then the following interpolated.poisson.pixel are output example of interpolated.f.

141.01	35.03	0.129E+01
141.01	35.09	0.134E+01
141.01	35.15	0.140E+01
141.01	35.21	0.145E+01
141.01	35.27	0.150E+01

<< skipped >>

143.99	40.73	0.162E+01
143.99	40.79	0.166E+01
143.99	40.85	0.163E+01
143.99	40.91	0.160E+01
143.99	40.97	0.157E+01

Then we can use:

`Program:   interplated-poisson.R`

`Reads:     interplated-poisson.pixel, interplated-poisson.conf`

`Requires:  drawmap.R, f2.R`

`Writes: image.pdf (= interpolated-poisson.pdf);` see the right-side figure (Fig. 11).

Also, we can use:

`Program:   enlarge.R`

`Reads:   delone2.out, interpolated.conf`

`Requires:  drawmap.R, f2.R`

`Writes: enlarged.pdf(=enlarged-poisson.pdf);` see left-side figure below (Fig. 11).

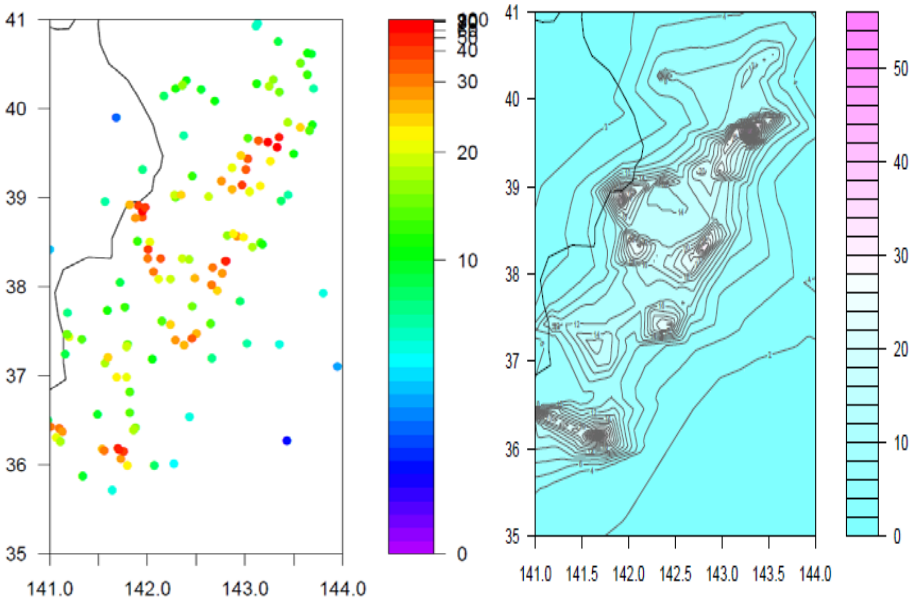


Fig. 11: `enlarged-poisson.pdf, image.pdf (=interpolated-poisson.pdf)`

Part IV. Short-Term Earthquake Forecasting

The estimated HIST-ETAS models of the previous period until a certain time instant is used to implement space-time forecasting of history-dependent seismicity rate after the previous period as moving images. Here, we assume that the model parameters do not change during the updated data until the present, and that the predictions are made on the basis of consecutively observed earthquakes.

The diagram (Fig. 12) below shows the flow chart of programs for estimations of the HIST-ETAS models and their forecasting. The hypocenter data `hypo.ts` and `hypo.dat` is connecting in time, that the first row of hypocenter data is last date of the `hypo.dat` in the same region. The flow chart details in the top block is the estimating procedure that were already explained in the above sections.

A job can be submitted interactively or in batch mode. Batch mode allows the user to log out of the system while the job continues to run in the background. The job could consist of a shell script (e.g. `job.sh`) or it may simply be a compiled FORTRAN binary file. The advantage of a shell script is that it can do other things before and after calling the compiled FORTRAN object.

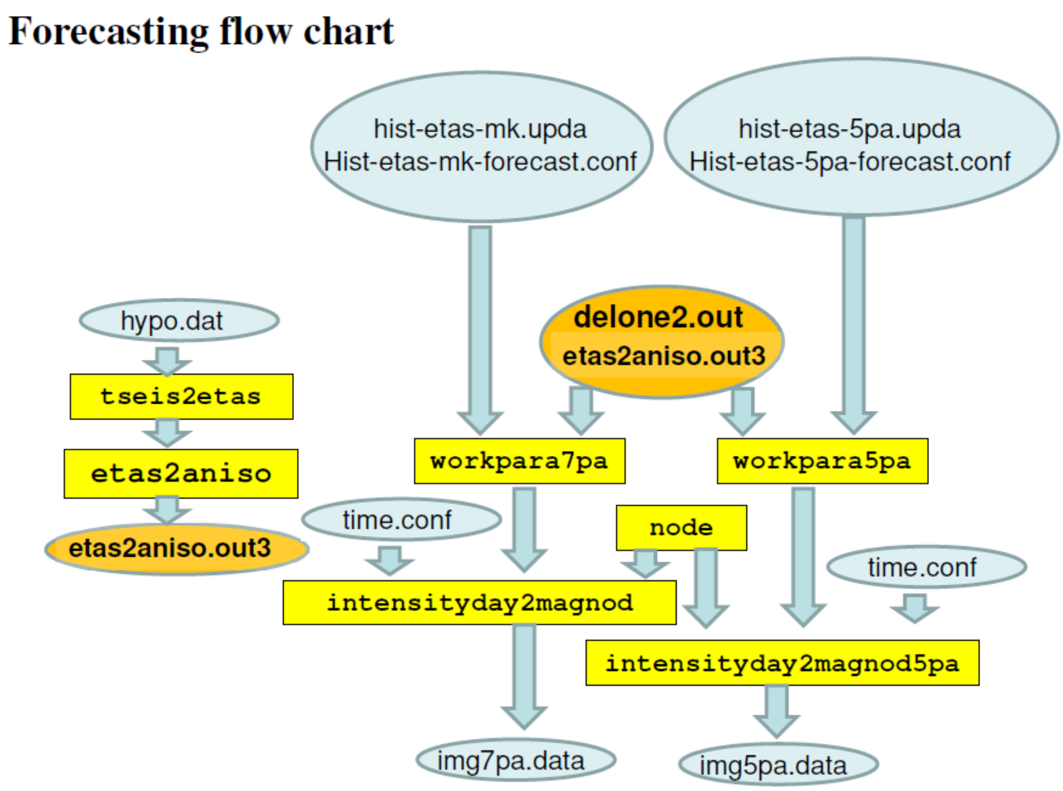


Fig. 12: The diagram shows the flow of output from each program to subsequent programs. Rectangular shapes represent programs, which are explained below. Elliptical shapes represent input/output files.

13 Seismicity forecasts by the HIST-ETAS models

All the used files in this section are selected in the program directory of `Section13files/` in the program package.

13.1 hist-etas-mk forecast

Specifically, assume the estimated `hist-etas-mk` model (`hist-etas-mk.upda`) calculated in §9 based on the data configuration in §9.2 for the whole Japan M ≥ 6 earthquake data as illustrated in §4.3. Remember that the last date of the data was the end of May 2011. After that, consider the following hypocenter data of earthquakes of M ≥ 4 in the same Japan region as §11.4. The following explains the consecutive implementation of unix (linux) shell script of [`japan.sh`]:

```
ifort tseis2etas.f -o tseis2etas
./tseis2etas < hypo.dat (output-file, work.etas)
```

```
ifort etas2aniso.f -o etas2aniso
./etas2aniso (input-files, work.etas, etas2aniso.conf; output-file,
     etas2aniso.out2, etas2aniso.out3, etas2aniso.out4,
     etas2aniso.out8, etas2aniso.out9)
```

```
ifort  workpara7pa.f -o  workpara7pa
./workpara7pa (input-files,, hist-etas-mk-forecast.conf,
hist-etas-mk.upda, delone2.out, node.dat, etas2aniso.out3; output-file,
work.para)
```

```
ifort node.f -o node
./node (input-files, node.conf, work.etas; output-file, node.dat)
```

```
ifort intensityday2magnod.f  -o intensityday2magnod
./intensityday2magnod (input-files, time.conf, hist-etas-mk.conf,
hist-etas-mk.upda, node.dat, work.para; output-file, img1.data)
```

Instead of `Intel-Fortran`, `gfortran` can be also used here.

Firstly, by using the HIST-ETAS model estimations based on [`hypo.ts`] in §3.2, we will forecast sequentially using the following updating earthquake data [`hypo.ts`]:

```
2011 06 01 01 26   7.97  143.3522  40.2497   11.65  5.1
```

```
2011 06 01 01 30  57.07  143.2218  35.2540   76.00  4.4
2011 06 01 01 41  19.63  141.7620  37.6593   43.71  4.2
2011 06 01 02 15  17.19  141.9967  38.8785   49.52  4.1
2011 06 01 06 27  33.93  143.4100  39.8478   27.23  4.5
2011 06 01 07 07  45.40  143.4283  39.8302   20.98  4.1
2011 06 01 07 28  40.90  143.8588  37.6817   36.00  4.1
2011 06 01 08 53  59.04  141.9093  38.6377   48.79  4.2
2011 06 01 12 14  11.03  143.7070  39.7765   33.00  5.1
2011 06 01 13 00   1.01  142.2350  36.7177   16.24  4.5

                            . . .
2018 09 25 14 19  23.31  148.1022  43.9925    0.00  4.4
2018 09 25 22 03  13.75  148.4427  44.0177    0.00  4.0
2018 09 26 01 22  13.29  148.2313  44.0835    0.00  4.0
2018 09 27 10 25  21.45  141.9518  34.1040   34.65  4.3
2018 09 28 04 32  25.27  141.1032  37.1130   52.21  4.0
2018 09 28 10 01   2.37  148.3172  44.0750    0.00  4.2
2018 09 29 18 25  54.33  142.0007  42.7707   35.36  4.2
2018 09 29 20 56  34.07  140.9535  35.8075   29.75  4.0
2018 09 30 17 54   4.49  141.9897  42.5498   36.86  4.9
2018 10 01 11 22   3.35  142.0100  42.7940   34.81  4.7
```

Here, to save file memory size, we restrict hypo.dat to including only M ≥ 4 earthquakes, but practically for accuracy of the analysis of the anisotropy, it is certainly preferred to take all detected earthquakes with hypocenter data.

Then, the program tseis2etas transforms this data to [work.etas] as given in the same format as given in §3.2. We use same program etas2aniso with the same configuration file, etas2aniso.conf:

```
./work.etas     !input data
6.0  6.0        !clms cutm
0.04666667      !xxx(day)= time span for analyzing centroid and anisotropy
```

Here, from a real-time forecasting perspective, we usually set ”xxx =1/24 = 0.041666667 day = one hour “to quickly determine the centroid location and orientation characteristics of the impending aftershock sequence after a main shock event. For the recent catalog, events within an hour interval after the main shock to give a reasonably good estimate of the centroid and orientation characteristics of the evolving aftershock sequence.

Then, by implementing the program etas2aniso that is actually the same program in §3 and §4, we get the output etas2aniso.out3:

```
 82  0.128E+03  0.149E+03  0.209E+02  0.300E+02  0.469E+02  0.169E+02
 31  143.83320   37.30250  6.10    2.37850  1.00000  1.00000  0.00000
125  143.58270   37.81170  6.00   13.92144  1.00000  1.00000  0.00000
155  141.82130   37.61770  6.00   17.85491  1.00000  1.00000  0.00000
187  142.59080   39.94780  6.90   22.28531  1.00000  1.00000  0.00000
289  143.29852   38.06312  7.30   39.41467  1.00000  1.00000  0.00000
405  142.09120   38.87370  6.40   52.56555  1.00000  1.00000  0.00000
414  141.62670   37.70870  6.30   54.16071  1.00000  1.00000  0.00000
457  141.22130   36.90320  6.50   60.16239  1.00000  1.00000  0.00000
```

```
  473  138.54880   34.70700  6.20     61.99874   1.00000   1.00000   0.00000
                                 . . .
 5379  144.48870   38.03600  6.30   2304.06758   1.00000   1.00000   0.00000
 5387  142.45530   40.26670  6.10   2310.22374   1.00000   1.00000   0.00000
 5396  143.94830   37.43530  6.30   2319.70802   1.00000   1.00000   0.00000
 5468  144.80580   38.00620  6.00   2357.30843   1.00000   1.00000   0.00000
 5473  140.74530   32.35200  6.00   2360.78026   1.00000   1.00000   0.00000
 5568  142.44800   41.00970  6.30   2429.82731   1.00000   1.00000   0.00000
 5646  132.58320   35.17803  6.10   2504.06425   0.00546   0.00724  -0.73698
 5733  135.62170   34.84430  6.10   2574.33234   1.00000   1.00000   0.00000
 5760  140.59200   35.16530  6.00   2593.84987   1.00000   1.00000   0.00000
 5840  142.00670   42.69080  6.70   2654.13055   0.02837   0.06975   0.52842
```

contains the centroid locations and normalized ellipsoidal coefficients for all event with magnitude not less than the cutoff magnitude, except for the first row that is the number of the additional M ≥ 6 data, ranges of longitudes and latitudes (see `hist-etas-mk-forecast.conf`). The other outputs, `etas2aniso.out2`, `etas2aniso.out4`, `etas2aniso.out8`, and `etas2aniso.out9` are also explained in §4.1.

We then use the input configuration file `hist-etas-mk-forecast.conf`:

`21.0  17.0  14012.0  308` !longitude span, latitude span, time span, starting time (days) of forecasting (= end time of the estimated period) for the `hist-etas-mk` model, and number of M ≥ 6 earthquakes to forecast.
`128.0  30.0  6.0   0.0 730.0  2.0` ! origin of the target rectangular region, cutoff magnitude, origin of time and end time of the short-term forecasting period. for the ranges of spatial rectangular region, time span, magnitude cutoff, etc.

We also use the Delaunay tessellation of the precursory period [`delone2.out`] in §6.4 to obtain [`work.para`] above by interpolating the `hist-etas-mk` coefficients [`hist-etas-mk.upda`] for each node; these coefficients are unchanged for the data.
Then we apply the program `workpara7pa` to make the summarized file [`work.para`]:

```
 -0.195903E+01  -0.770966E-01   143.8332 37.3025 6.1 14014.37850   1.0000   1.0000   0.0000
 -0.114274E+01   0.493784E-01   143.5827 37.8117 6.0 14025.92144   1.0000   1.0000   0.0000
  0.143420E+01   0.690773E-01   141.8213 37.6177 6.0 14029.85491   1.0000   1.0000   0.0000
  0.252823E+00   0.274905E+00   142.5908 39.9478 6.9 14034.28531   1.0000   1.0000   0.0000
 -0.156295E+00   0.200900E+00   143.2985 38.0631 7.3 14051.41467   1.0000   1.0000   0.0000
  0.180846E+01   0.622751E+00   142.0912 38.8737 6.4 14064.56555   1.0000   1.0000   0.0000
  0.167737E+01  -0.147276E-01   141.6267 37.7087 6.3 14066.16071   1.0000   1.0000   0.0000
 -0.204882E-01   0.155726E+00   141.2213 36.9032 6.5 14072.16239   1.0000   1.0000   0.0000
 -0.497493E+00  -0.208164E+00   138.5488 34.7070 6.2 14073.99874   1.0000   1.0000   0.0000
 -0.391050E-01   0.153219E+00   141.1610 36.9688 6.1 14084.14033   1.0000   1.0000   0.0000
                                 . . .
 -0.205721E+01   0.309927E-01   144.4887 38.0360 6.3 16316.06758   1.0000   1.0000   0.0000
  0.325441E+00   0.119500E+00   142.4553 40.2667 6.1 16322.22374   1.0000   1.0000   0.0000
 -0.203715E+01  -0.457134E-01   143.9483 37.4353 6.3 16331.70802   1.0000   1.0000   0.0000
```

```
-0.215702E+01  -0.979918E-01   144.8058 38.0062 6.0 16369.30843   1.0000   1.0000   0.0000
-0.392149E+00  -0.275451E+00   140.7453 32.3520 6.0 16372.78026   1.0000   1.0000   0.0000
 0.453226E+00  -0.209358E+00   142.4480 41.0097 6.3 16441.82731   1.0000   1.0000   0.0000
-0.312373E+01  -0.554472E+00   132.5832 35.1780 6.1 16516.06425   0.0055   0.0072  -0.7370
-0.179847E+01  -0.457204E+00   135.6217 34.8443 6.1 16586.33234   1.0000   1.0000   0.0000
 0.471500E+00  -0.303005E+00   140.5920 35.1653 6.0 16605.84987   1.0000   1.0000   0.0000
-0.148886E+01  -0.325565E+00   142.0067 42.6908 6.7 16666.13055   0.0284   0.0698   0.5284
```

for the additional earthquake in each raw, and the first two columns represent location dependent deviations from the baseline parameters log(μ_0) and log(K_0) of the `hist-etas-mk` model, respectively; 3 - 6 column stands for longitudes, latitudes, magnitudes and occurrence times in days unit, respectively. The last three columns indicate the anisotropic information of triggered descendants (same as those of `etas2aniso.out3` in the above); and `hist-etas-mk.upda` is the estimated coefficients of μ and K by the program `hist-etas-mk` in §9 for each M ≥ 6 earthquakes and some added points including those of boundaries from precursory period for the estimation.

Finally, given the time of the snapshot image `time.conf`:

`1780.05  ! one-hour after M6.5;` time of intensity in days = see `work.etas` for the time in days

in addition to the program `node` set coordinates of pixel node on which predicted intensity rate are given where the input configuration file is `node.conf`:

```
128. 149. 30. 47.     ! longitude and latitude ranges for all Japan Area
210 170               ! number of pixels for image,
```

which means that the resolution degree of the intensity image is unit pixel of 0.1^2 $\deg^2$ and unit time of 1 day, so that the each probability of M ≥ 6 earthquake in the space-time unit is 100^{-1} times of the intensity value; note that the estimated intensity values are per 1.0 $\deg^2$ and per day.

The output file is given in such a way that `node.dat`:

```
   128.0500   30.0500
   128.0500   30.1500
   128.0500   30.2500
   128.0500   30.3500
   128.0500   30.4500
   128.0500   30.5500
   128.0500   30.6500
   128.0500   30.7500
   128.0500   30.8500
   128.0500   30.9500
          . . .
   148.9500   46.0500
   148.9500   46.1500
   148.9500   46.2500
   148.9500   46.3500
   148.9500   46.4500
   148.9500   46.5500
```

```
  148.9500   46.6500
  148.9500   46.7500
  148.9500   46.8500
  148.9500   46.9500
```

for the locations at which the intensity is calculated.

For the snapshot at the time instances in `time.conf`, the program `intensityday2magnod` provides the location-dependent seismicity rates on the given node locations as the output [`img1.data`]:

```
 1780.05000 128.0500  30.0500  -4.48152
 1780.05000 128.0500  30.1500  -4.35851
 1780.05000 128.0500  30.2500  -4.22291
 1780.05000 128.0500  30.3500  -4.07589
 1780.05000 128.0500  30.4500  -3.92018
 1780.05000 128.0500  30.5500  -3.76075
 1780.05000 128.0500  30.6500  -3.60562
 1780.05000 128.0500  30.7500  -3.46630
 1780.05000 128.0500  30.8500  -3.35696
 1780.05000 128.0500  30.9500  -3.29134
                   . . .
 1780.05000 148.9500  46.0500  -5.00901
 1780.05000 148.9500  46.1500  -5.04269
 1780.05000 148.9500  46.2500  -5.14049
 1780.05000 148.9500  46.3500  -5.17804
 1780.05000 148.9500  46.4500  -5.21446
 1780.05000 148.9500  46.5500  -5.24642
 1780.05000 148.9500  46.6500  -5.27794
 1780.05000 148.9500  46.7500  -5.31176
 1780.05000 148.9500  46.8500  -5.34829
 1780.05000 148.9500  46.9500  -5.38806
```

for the forecasting based on `hist-etas-mk` model, where the last column represents the ordinary logarithm of the intensity values.

13.2 hist-etas-5pa forecast

The shell script `japan.sh` provides the same procedure as the above `japan.sh` except for using `workpara5pa` instead of `workpara7pa`, and `intensityday2magnod5pa` instead of `intensityday2magnod`. The program `workpara7pa` make the summarized file [`work.para`]:

```
-0.173143E+01 -0.235180E+00 -0.177008E-01  0.525061E-01  0.450011E-02  143.8332 37.3025 6.1 14014.37850  1.0000  1.0000  0.0000
-0.968881E+00 -0.969090E-01 -0.150673E-01  0.472810E-01  0.520681E-02  143.5827 37.8117 6.0 14025.92144  1.0000  1.0000  0.0000
 0.141514E+01 -0.370253E-01 -0.112767E-01  0.830377E-02  0.581182E-02  141.8213 37.6177 6.0 14029.85491  1.0000  1.0000  0.0000
 0.308055E+00  0.249292E+00 -0.574826E-01  0.147485E-01  0.628059E-02  142.5908 39.9478 6.9 14034.28531  1.0000  1.0000  0.0000
-0.105439E+00  0.816887E-01 -0.764281E-02  0.320480E-01  0.580723E-02  143.2985 38.0631 7.3 14051.41467  1.0000  1.0000  0.0000
 0.178974E+01  0.628283E+00 -0.426934E-01  0.234688E-01  0.707487E-02  142.0912 38.8737 6.4 14064.56555  1.0000  1.0000  0.0000
 0.156352E+01 -0.108644E+00 -0.226194E-01  0.775780E-02  0.544244E-02  141.6267 37.7087 6.3 14066.16071  1.0000  1.0000  0.0000
 0.174968E+00  0.100107E+00 -0.824457E-02  0.477913E-01  0.566260E-02  141.2213 36.9032 6.5 14072.16239  1.0000  1.0000  0.0000
-0.457258E+00 -0.190316E+00 -0.559044E-01 -0.398870E-02  0.304299E-02  138.5488 34.7070 6.2 14073.99874  1.0000  1.0000  0.0000
 0.149485E+00  0.993812E-01 -0.990510E-02  0.416871E-01  0.569163E-02  141.1610 36.9688 6.1 14084.14033  1.0000  1.0000  0.0000
```

```
. . .
-0.187238E+01  -0.790851E-01  -0.247511E-01   0.697094E-01   0.493365E-02   144.4887  38.0360  6.3  16316.06758   1.0000   1.0000   0.0000
 0.330056E+00   0.121612E+00  -0.695166E-01   0.102073E-01   0.592056E-02   142.4553  40.2667  6.1  16322.22374   1.0000   1.0000   0.0000
-0.180773E+01  -0.188937E+00  -0.187992E-01   0.553893E-01   0.458487E-02   143.9483  37.4353  6.3  16331.70802   1.0000   1.0000   0.0000
-0.202072E+01  -0.218767E+00  -0.327643E-01   0.612950E-01   0.439912E-02   144.8058  38.0062  6.0  16369.30843   1.0000   1.0000   0.0000
-0.463733E+00  -0.276171E+00  -0.344978E-01   0.223125E-02   0.136756E-02   140.7453  32.3520  6.0  16372.78026   1.0000   1.0000   0.0000
 0.375490E+00  -0.212455E+00  -0.715324E-01  -0.159353E-01   0.450233E-02   142.4480  41.0097  6.3  16441.82731   1.0000   1.0000   0.0000
-0.300655E+01  -0.604138E+00  -0.335803E-01   0.646442E-02   0.977202E-03   132.5832  35.1780  6.1  16516.06425   0.0055   0.0072  -0.7370
-0.180156E+01  -0.483693E+00  -0.572613E-01   0.122522E-01   0.188226E-02   135.6217  34.8443  6.1  16586.33234   1.0000   1.0000   0.0000
 0.486403E+00  -0.386306E+00  -0.298320E-01   0.209163E-01   0.284362E-02   140.5920  35.1653  6.0  16605.84987   1.0000   1.0000   0.0000
-0.140339E+01  -0.302433E+00  -0.730526E-01  -0.172558E-01   0.320788E-02   142.0067  42.6908  6.7  16666.13055   0.0284   0.0698   0.5284
```

for the additional earthquakes in each raw, where the first 5 columns represent location dependent deviations from the logarithm of reference values μ_0, K_0, α_0, p_0 and q_0 (the top five numbers in `hist-etas5pa.upda`), respectively, at each hypocenter location of longitudes, latitudes, magnitudes and occurrence times in days unit, respectively, given in 6 - 9 columns. The last three columns indicate the anisotropic information of triggered descendants (same as `etas2aniso.out3`). The input files are:
[`hist-etas5pa-forecast.conf`]

```
21.0     17.0     14012.   308
128.0    30.0     6.0      0.0     730.0    2.0
```

for the ranges of spatial rectangular region, time span, magnitude cutoff, etc., as explained for `hist-etas7pa-forecast.conf` in the above.

The program `intensityday2magnod` provides the location-dependent seismicity rates on the given node locations as the output [`img1.data`]:

```
 1780.05000 128.0500  30.0500  -4.47764
 1780.05000 128.0500  30.1500  -4.35814
 1780.05000 128.0500  30.2500  -4.22522
 1780.05000 128.0500  30.3500  -4.08007
 1780.05000 128.0500  30.4500  -3.92552
 1780.05000 128.0500  30.5500  -3.76672
 1780.05000 128.0500  30.6500  -3.61197
 1780.05000 128.0500  30.7500  -3.47310
 1780.05000 128.0500  30.8500  -3.36457
 1780.05000 128.0500  30.9500  -3.30022
                   . . .
 1780.05000 148.9500  46.0500  -4.90040
 1780.05000 148.9500  46.1500  -4.93188
 1780.05000 148.9500  46.2500  -5.01747
 1780.05000 148.9500  46.3500  -5.05222
 1780.05000 148.9500  46.4500  -5.08627
 1780.05000 148.9500  46.5500  -5.11626
 1780.05000 148.9500  46.6500  -5.14587
 1780.05000 148.9500  46.7500  -5.17757
 1780.05000 148.9500  46.8500  -5.21170
 1780.05000 148.9500  46.9500  -5.24869
```

for the forecasting based on `hist-etas-5pa` model.

13.3 Magnified forecasting image in a localized region

The shell script `kumamoto.sh` provides the same procedure as the above `japan.sh` except for the restriction of regions as given by `node.conf`:

```
130. 132. 32. 34.     !longitude and latitude ranges for Kumamoto Area
200 200               ! number of pixels for image
```

where the number of pixels adjust the resolution of image.

13.4 Plotting Snapshots of Short-Term Forecast Images

The example output images with relevant maps are given by `R` language:

```
Program: japan.r
Reads: img1.data, work.para, node.conf
Requires: drawmap.r, f2.r > filled2contour.r > see
     help(filled.contour) in R command.
```

`Writes: Rplots.pdf;` see the left-hand-side figure below (Fig. 13).

The example output images for the magnified region can be seen in:

```
Program:   kumamoto.r
Reads:     img1.data, work.para, nodekuma.conf
Requires:  drawmap.r, f2.r
```

`Writes: Rplors.pdf;` see the right-hand-side of the below figure (Fig.13).

We get the following panels of `Rplot.pdf` that delineates snapshots of the short-term probability forecast at the time of one-hour after the M6.5 Kumamoto Earthquake (see `time.conf` above). These are conditional intensity function $\lambda(t, x, y \mid H_t)$ as mathematically defined in §A.5

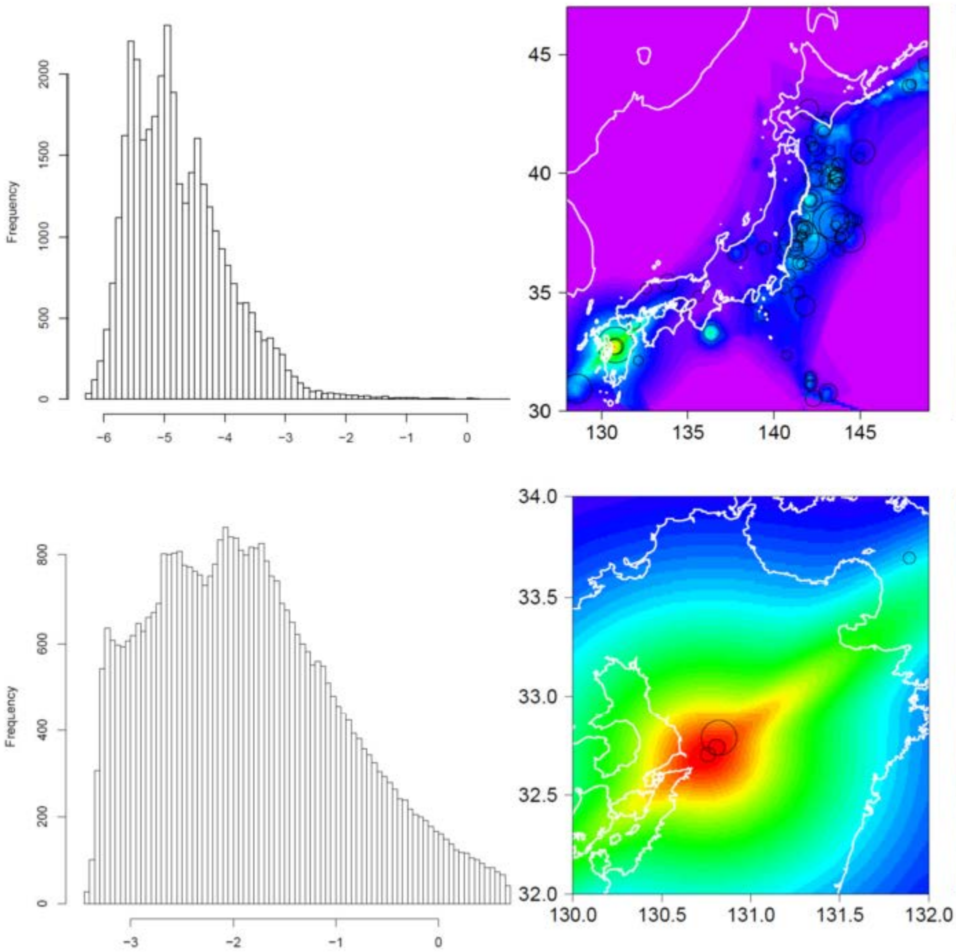


Fig. 13: Snapshots of probability forecasts of M ≥ 6 earthquakes at one-hour after the largest M6.4 foreshock before the 2016 M7.3 Kumamoto earthquake; image in the main Japan area and enlarged image in Kyushu area. The circles indicate actual M ≥ 6 earthquakes occurring during the forecast periods. The histograms show the frequency of intensity values at each pixel against the ordinary logarithm of the intensity. Color scale of the image shows expected number of M ≥ 6 earthquakes per one square degree (~ 100km^2) per day.

Part V. Simulations

This chapter provides the simulation of hypocenters using the nonhomogeneous Poisson model, spatial magnitudes using by space-dependent b-values, HIST-ETAS-mk model and HIST-ETAS-5pa model. Examples here use the intensity b-values and conditional intensities estimated in §7 ~ §10.

14. Nonhomogeneous Poisson simulation by spatial intensity rate function

This program fits a nonhomogeneous spatial Poisson model with stationary Poisson time component to the location of earthquakes. The simulation is done using the 2D spatial Poisson intensity given by coefficients at the nodes of the Delaunay tessellations (§6) and their interpolations can be found §12. All the used files in this section are in the program directory of `Section14files/` in the program package.

Mathematical explanation of Poissonian spatial intensity is described in §A.3.

14.1 File Names

For the example we use the intensity estimated in §8:

```
Program: simNHPoisson.f
Object:  simNHPoisson
Configuration: poisson.conf
Reads: delone2.out, delo2d-poisson.omap
Writes: fort.2 (= simNHPoi.hypo)
```

For the spatial plot, done in R:

```
Program: fort2.R
Reads: fort.2, drawmap.r, ../MapsData/jp.br.dat & jp.pp.dat
Writes: Rplots.pdf (= 1993.1119.1046.pdf)
```

14.2 Configuration File Format

The configuration file poisson.conf includes the following three lines:

`128.0 21.0 30.0 17.0` ! xmin, ymin, tx, ty

`1993 1119 1046` !4 digit seeds of for a series of uniform random numbers, where different seeds are expected to provide mutually independent random number series.

14.3 Program Execution

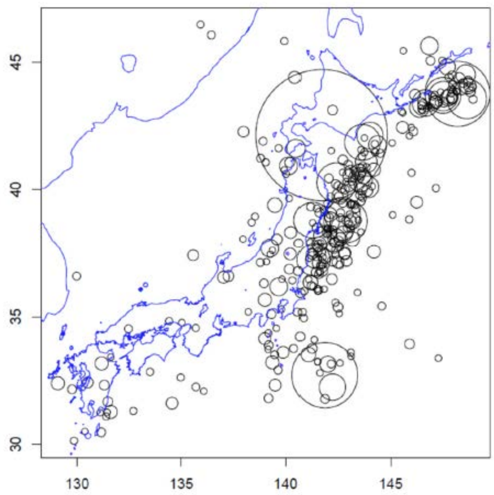
For the simulations, done in FORTRAN:

`./simNHPoisson |tee simNHPoisson.prt` !which is given in `Section14files/`.

For the spatial plot, done in R:

```
> source('r. fort2’)
```

`Writes:  Rplots.pdf (= 1993.1119.1046.pdf);` which shows the following plot (Fig. 14):



`delo2d-poisson.pdf`:

Fig. 14. Simulated epicenter coordinates by the nonhomogeneous Poisson intensity §8.3 in and around mainland Japan. Sizes of circle radii are proportional to exponential of the same magnitude series (M ≥ 6) of the original JMA data.

After simulation we can make reestimation of nonhomogeneous Poisson intensity, starting from constructing the new Delone tessellation of the simulated data.

15. Magnitude simulation given spatially varying b-values of G-R law

These programs simulate magnitudes given the *b*-value over a spatial region. Magnitude are simulated by GR-law at any location based on *b*-values interpolated on the Delaunay tessellations (§6). All used files in this section are selected in the program directory of `Section15files/` in the program package.

15.1 File Names

For the simulations, done in FORTRAN:

```
Program:       bvalue2magsim.f
Object:        bvalue2magsim
Configuration: delo2d-bvalues.conf  !same as poisson.conf in §14.2
Reads:         delone2.out
Writes:        fort.2 (= fort.2Mc595, fort.2.Poiconfig)
```

For the spatial plot, done in R:

```
Program: fort2.R
Reads:   fort.2, drawmap.r,../MapsData/jp.br.dat & jp.pp.dat
Writes:  Rplots.pdf (= original2magsim.pdf, binterpo2magsim.pdf)
```

15.2 Configuration File Format

The configuration file delo2d-bvalues.conf includes the following three lines:

`128. 30. 5.95` !xmin, ymin, threshmag = magnitude threshold

`6.0d0` !w1, which used in §8, but not used here.

7 !ipr, which used in §8, but not used here.

Magnitude rounding issue: if magnitude data are rounded to 0.1 units, the threshold magnitude here should be modified to 5.95 (= $M_c - 0.05$) to avoid the *b*-value MLE bias. This is because a rounded value of 6.0 may have been as small as 5.95 or large as 6.05. This applies to the traditional catalogs such as the JMA, NEIC-PDE, and ISC catalogs. Otherwise, namely, less than 0.01 magnitude unit, we can keep threshmag = 6.0 .

15.3 Program Execution

FORTRAN execution command:

`./delo2d-bvalues |tee delo2d-bvalues.prt` !which is given in `Section16files/.`

For the spatial plot, done in R:

`> source('delo2d-bvalues.R') ;` which shows the following two plots (Fig. 15):

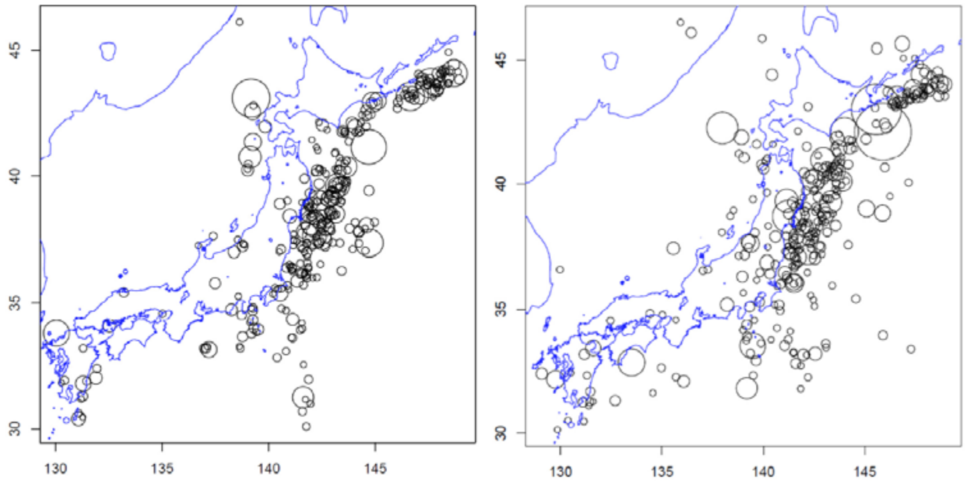


Fig. 15. Simulated earthquake magnitudes on the epicenter coordinates of the JMA data (left panel) and on the locations (right panel) simulated by the nonhomogeneous Poisson intensity. Sizes of circle radius is proportional to exponential of simulated magnitudes (M>5.95) by the interpolated b-values in §7.3.

After simulation we can make re-estimation of *b*-values, starting from constructing the new Delaunay tessellation of the simulated data.

16. HIST-ETAS simulation

The programs in this section produce simulated data files for given sets of parameters in the point process model used in HIST-ETAS models (See Ogata, 1981, 1998) for theoretical basis. It is noted that the intensity defined by a combination of parameter values should be well-defined; due to some combinations of parameter values, the simulated data can be explosive (Zhuang and Ogata, 2006).

There are two options; either using magnitudes in `delone2.out` or simulating magnitude by (modified) Gutenberg-Richter's Law. The first option simulates the same number of events that are not less than threshold magnitude in the data, this is the present option, and therefore the parameter *b*-value is not used in this particular example. For the second option, you have to provide *b*-value of G-R law and number of events to be simulated; you can simply modify the FORTRAN program `histetasim.f` below by changing the commented line to execute for simulating magnitude sequence.

Furthermore, simulation can start based on an occurrence history of precursory period; the users may also extend these program.

Finally, the program `histetasim.f` here support only the case of isotropic clustering that ignores the last three columns of delone2.out, but, if necessary, this can be extended by modifying `histetasim.f` in reference of `subroutine func17` of the optimization programs `hist-etas-mk.f` or `hist-etas5pa.f` in sections 9 and 10 or forecasting programs in Section 13, with the same the format of the current delone2.out.

The FORTRAN program `histetasim.f` needs configuration `histetasim.conf` as explained below. The example of input file is the same as `hist-etas-mk.upda` or `hist-etas5pa.upda` which was the output in sections 9 and 10, respectively. All used files in this section are selected in the program directory of `Section16files/` in the program package.

16.1 File names

For the simulation, done in FORTRAN:

```
Program: histetasim.f
Object: histetasim
Configuration: histetasim.conf
Reads: delone2.out,
       hist-etas-mk.upda, or
       hist-etas5pa.upda
Writes: histetasim.prt, fort.7, fort.2
```

For the spatial plot, done in R:

```
Program: histetasim.R
Reads: fort.2, fort.7,
       drawmap.r, ../MapsData/jp.br.dat & jp.pp.dat
Writes:   Rplots.pdf (= 7pa1993,1119,1046.pdf or 5pa1993,1119,1046.pdf)
```

16.2 Configuration File Format

Explanation of the configuration file `histetasim.conf` consists of:

5 !Choose either of `simulation` model 7 or 5 for `hist-etas-mk` or `hist-etas5pa`, respectively

1.1 6.0 128.0 21.0 30.0 17.0 14012.0 !bmg,cm0,tx0,tx,ty0,ty,tend

1993 1119 1046 !Seeds of uniform random number series; triplet four digits.

The variable `bmg` and `cm0` stand for *b*-value, lower cutoff magnitude, respectively; `tx0` and `ty0` stand for the longitude and latitude origin of the focal region, respectively; `tx` and `ty` stand for the length of the rectangular region, respectively; and `tend` stands for the time length.

Different random number seeds are assumed working independent simulation experiments.

16.3 Executing the Program

The following command executes the compiled FORTRAN code.

`./histetasim |tee histetasim.prt` (= `histetasimuk.prt` or `histetasim5pa.prt`),

where all output files listed below are selected in the program directory of `Section17files`.

16.4 Output Produced by Program with configuration file of different first line

16.4.1 `hist-etas-mk` case:

If the number in the first line of `histetasim.conf` is 7, representing `hist-etas-mk` model simulation, then the output files are:

`histetasim.prt (= histetasimuk.prt)`, `fort.2(= fort.2.muk)`, `fort.7` (= `fort.7.muk`) which are all selected in the program directory of `Section17files/`, and they have the same format as those by the simulation of `hist-etas-mk` model. Calculated record of the program `histetasim` is stored by the name `histetasim.prt (= histetasimuk.prt)` which shows some key parameters to compare with the key parameters for checking consistency together with hypocenter data that are same as `fort.7.`

`fort.2` includes:

```
  308    21.00000    17.00000     6.00000  1701.00000
    1   146.37236    43.22112     7.70000    33.96782
    2   147.45493    43.47246     6.00000    34.02402
    3   145.96078    43.80458     7.10000    34.12867
    4   140.58100    36.19143     6.60000    37.06475
    5   143.57109    41.56049     6.00000    39.57438
                                ……
  304   142.19733    35.97812     6.10000  1680.84593
  305   141.96287    40.80838     6.20000  1683.40119
  306   140.64088    33.22854     6.00000  1684.12163
  307   143.43240    39.97664     6.10000  1692.37786
  308   148.13354    44.18092     6.10000  1700.54208
```

where the first line shows the number of events, rectangular side lengths in degrees, cutoff magnitude and the entire time span. The rest lines indicate the serial number of events, epicenter coodinates, magnitude that are same as those in delone2.out in §13.1.

`fort.7` (= `fort.7.muk`) includes:

```
        308
    1   146.372   43.221  7.70     33.96782    0 0.00     1
    2   147.455   43.472  6.00     34.02402    1 7.70     1
    3   145.961   43.805  7.10     34.12867    1 7.70     1
    4   140.581   36.191  6.60     37.06475    0 0.00     2
    5   143.571   41.560  6.00     39.57438    0 0.00     3
    6   147.592   43.674  6.50     40.59455    0 0.00     4
    7   146.909   44.236  6.10     40.61836    6 6.50     4
    8   143.568   41.716  6.00     41.64508    5 6.00     4
    9   140.782   35.176  6.70     47.89052    0 0.00     5
   10   140.718   35.286  6.10     48.30271    9 6.70     5
                                   ……
  299   142.177   36.975  6.00   1613.69128  298 6.10   183
  300   142.531   38.418  7.10   1651.89759  128 7.30   183
  301   145.575   43.010  6.60   1656.22287    0 0.00   184
  302   141.519   34.460  6.20   1671.84874    0 0.00   185
  303   146.391   43.451  6.00   1680.62511    0 0.00   186
  304   142.197   35.978  6.10   1680.84593  241 9.00   186
  305   141.963   40.808  6.20   1683.40119    0 0.00   187
  306   140.641   33.229  6.00   1684.12163    0 0.00   188
  307   143.432   39.977  6.10   1692.37786    0 0.00   189
  308   148.134   44.181  6.10   1700.54208    0 0.00   190
```

where the first to five columns are same as those of fort.2, sixth and seventh columns represent shows the identification of parent and its magnitude, where 0 represents the 0-generation event that is simulated by the contribution of background intensity $\mu(x,y)$; and the last columns show cluster number of the same family trees.

For the plot, done in R, then `R.plots.pdf(7pa1993,1119,1046.pdf)` shows below plots (Fig. 16):

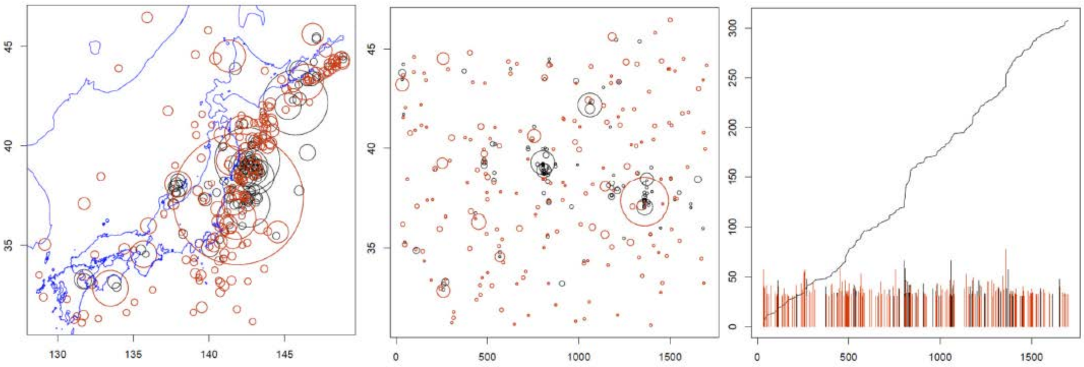


Fig. 16: Simulated data by the `HIST-ETAS-mK` model. Left panel shows epicenters with sizes of circle radii are proportional to exponential of the same magnitude series (M>5.95) of the original JMA data. Middle panel shows latitude versus time plots. Right panel shows the cumulative number curve and magnitude versus time plots. In all panels, red ones indicate 0-th generation earthquake events generated by the background intensity.

After simulation we can make reestimation starting from constructing 2D Delaunay tessellation for the simulated data sets.

`16.4.2 hist-etas-5pa` case:

If the number in the first line of `histetasim.conf` is `5` representing `hist-etas-5pa` model simulation, then the output files are:

`histetasim.prt (= histetasim5pa.prt), fort.2(= fort.2.5pa)`, and `fort.7` (= `fort.7.5pa`) which are selected in the program directory of `Section16files/`, and have the same format as those by the simulation of `hist-etas-mk` model.

For the plot, done in `R`, then R.plots.pdf (5pa1993,1119,1046.pdf) show below plots (Fig. 17):

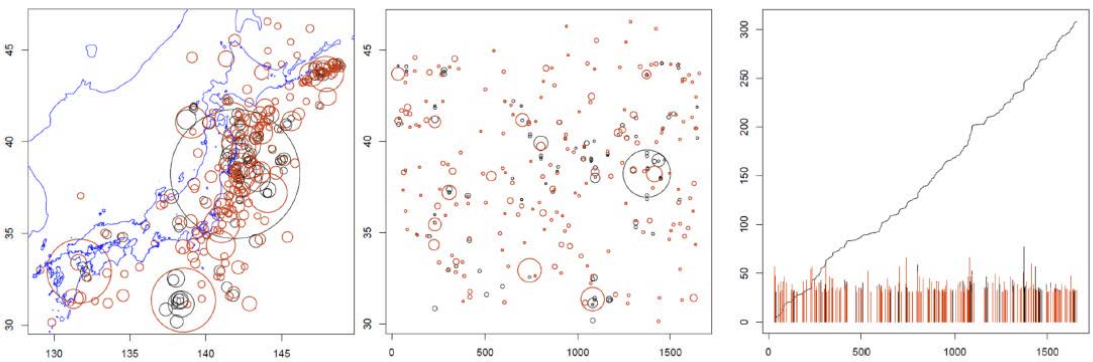


Fig. 17: Simulated data by the `HIST-ETAS-5pa` model. Left panel shows epicenters with sizes of circle radii are proportional to exponential of the same magnitude series (M>5.95) of the original JMA data. Middle panel shows latitude versus time plots. Right panel shows the

cumulative number curve and magnitude versus time plots. In all panels, red ones indicate 0-th generation earthquake events generated by the background intensity.

After simulation we can make re-estimation, but we need to start from constructing new 2D Delaunay tessellation for the simulated data sets.

APPENDICES

A. Mathematical Outline of Models

The ETAS model (Ogata, 1985, 1988, 1989) was extended for space-time data, and among the possible modelings for the space component, the best form described in §A.3 (Ogata, 1998) is selected by the goodness-of-fit comparison using the Akaike information criterion (*AIC*: Akaike, 1974). Incidentally, see Zhuang *et al*. (2005) and Ogata and Zhuang (2006) for further improvement of the space-time ETAS model, but we do not consider this for the hierarchical extensions of the parameters.

We give a brief outline here of the space-time ETAS models that are fitted by this software. We initially define the space-time ETAS model in a general way that encompasses all of the specific models fitted by this software. We then describe what constraints are imposed by specific models. Further details are available in Ogata (2010) for an example.

A.1 Determination of Anisotropic Clusters

Before fitting the space-time ETAS models with anisotropic spatial clustering effect, we aim at compiling similar solution as the centroid Moment tensor solution (Dziewonski *et al*.1981) using early aftershocks activity, which was first investigated by Utsu and Seki (1955) and Utsu (1969). Also, see Ogata *et al*. (1995) and Ogata (1998).

The large earthquakes of $M \geq M_m$ in the catalogue are selected, and their immediate aftershocks are determined. The threshold magnitude M_m of the main shocks is determined appropriately, taking account of the cutoff magnitude M_c of the earthquakes in the catalog, such as $M_m = M_c + 1.0$. For example, the space window is a square centered at the epicenter of the main shock, with sides of length $3.33 \times 10^{0.5M-2} + \varepsilon$ centered at the epicenter location, where M is the magnitude of the main shock. The last term ε is to quantify the error of epicenter estimates, usually takes 0 but we take ε =66.6 km (0.3 degree in latitude) in early days in offshore Japanese events. For the time span for estimation purpose, we can set one day (24 hours) or the shorter. The time window can be longer than 1 day in a low detection region or during an earlier period. On the other hand, from a forecasting perspective nowadays, one might set “0.05”, i.e. about one hour, to quickly determine the centroid location and orientation characteristics of the impending aftershock sequence after a main shock event using all detected earthquakes.

For each main shock and its aftershock sequence, a bivariate normal distribution is fitted to the spatial values. In particular, for each, the covariance matrix

$$S = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

and the centroid of the main shock and its aftershock sequence are estimated.

The null model assumes that S is the identity matrix, and the cluster center is at the location of the main shock. There are three possible alternative models:

1. S is different to the identity matrix but the cluster center is not different to that of the main shock;
2. S is not different to the identity matrix but the cluster centre is located at the centroid;
3. S is different to the identity matrix and the cluster center is located at the centroid.

Cases 2 and 3 are achieved by relocating the main shock to the centroid location. For each of the four models of a given cluster, the AIC is calculated. That model with the smallest AIC is selected for each cluster.
See Ogata (1998, 2010), Ogata (2004, Appendix B) and Ogata and Zhuang (2006, Appendix A) for more details. This procedure is executed by the program `aniso2etas`.

The procedure is illustrated below (Fig. 18):

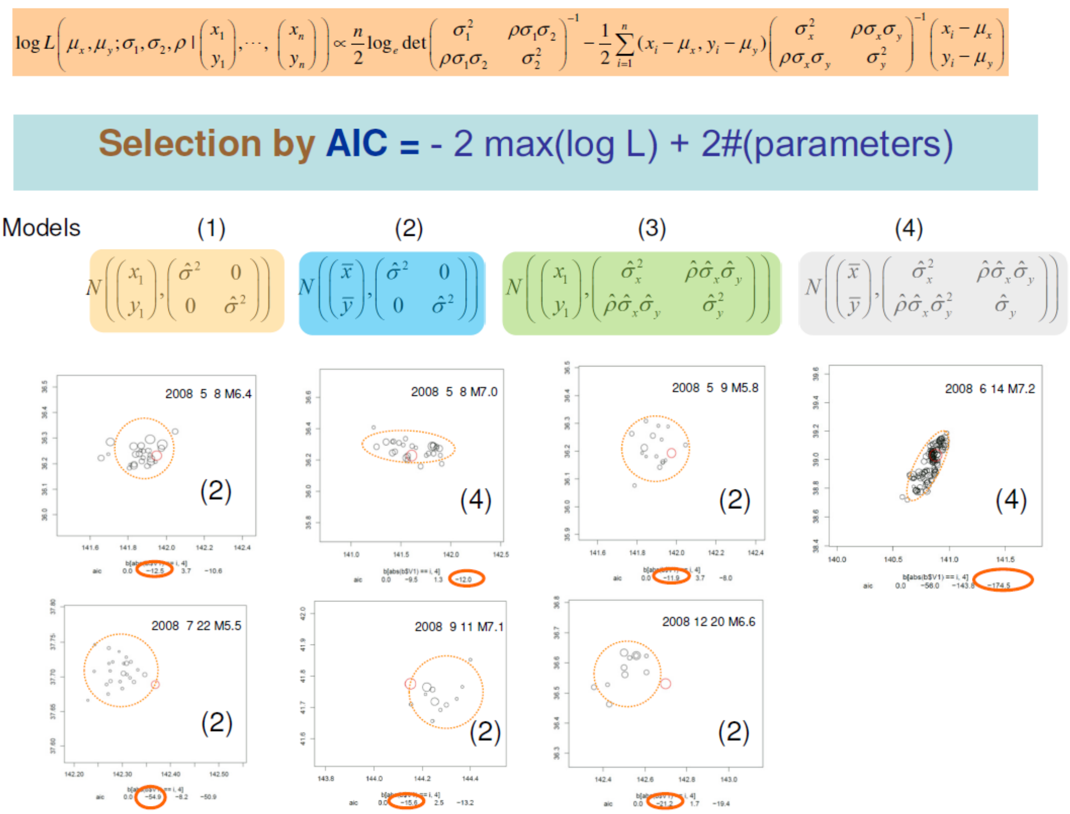


Fig. 18: These panels show aftershocks occurring during the first hour after the main shock that is indicated by a small red circle (x_1, y_1). The occurrence date and magnitude of the main shock are printed. The AIC values of Models (1) ~ (4) relative to the largest one are listed in each panel, where the model of the smallest value is adopted for the forecast of the aftershock cluster

anisotropy. Namely, we compare the goodness-of-fit of the following four 2-dimensional Normal distributions by the AIC. The model (1) stands for isotropic cluster with the centroid as the original epicenter. The model (2) stands for isotropic cluster, but the centroid coordinates are different from the original epicenter. The model (3) stands for anisotropic cluster with the centroid as the original epicenter. And the model (4) stands for anisotropic cluster but the centroid coordinates are different from original epicenter. The model with the smallest AIC value is adopted, and each panel illustrates a contour of the selected model.

The isotropic Space-Time Epidemic-Type Aftershock Sequence (ST-ETAS) model

$$\lambda(t,x,y \mid H_t) = \mu(x,y) + \sum_{\{i;\, t_i<t\}} \frac{K_0}{(t-t_i+c)^p} \left[\frac{(x-x_i)^2+(y-y_i)^2}{e^{\alpha(M_i-M_0)}} + d \right]^{-q}$$

can be extended to non-isotropic clusters for the earthquakes indicated by the output aniso2etas.out3, aiming at a better fit of the models to an earthquake catalog. For this, each response function is extended in such a way that the isotropic term in the response functions is replaced by

$$\frac{1}{\sqrt{1-\rho^2}} \left(\frac{\sigma_2}{\sigma_1} x^2 - 2\rho xy + \frac{\sigma_1}{\sigma_2} y^2 \right),$$

so that the corresponding iso-circle and iso-ellipse as a cross-section of the function at the same height have the same area as each other. Namely, the corresponding circle and ellipse as a cross-section of $z = z_c$ have the same area to each other. Then the integral of the above conditional intensity function remains the same (cf., Ogata, 1998).

A.2 Delaunay Tessellation

The Delaunay tessellation is a rather elegant method that can be used to estimate background seismicity or, in fact, to get estimates of anything that may vary in space where we have values of the entity of interest at any given points. It involves drawing triangles where the vertices are points, and no point falls within any of the circumcircles of the drawn triangles. Algorithms for the implementation of the techniques can be found in the `Wikipedia`, for example.

In the case of a two-dimensional surface, each triangle provides a flat surface where the height of the surface is known at the three vertices. At any other point on the surface within a triangle, the height of the surface can be estimated using linear interpolation. The program `interpolated.f` performs such an interpolation. In regions where point density is large, the triangles will be very small and hence the interpolation error will be small, and conversely, where the point density is small the interpolation error will be relatively larger. Further, the rate at which points occur in a given region will be inversely proportional to the area of the triangles within that region.

Consider the Delaunay triangulation (e.g., Green and Sibson, 1978); that is to say, the whole rectangular region *A* is tessellated by triangles with the vertex locations of earthquakes and some additional points $\{(x_i, y_i),\ i=1,\ldots,N+n\}$, as given in Fig. 19, where *N* is the number of earthquakes and *n* is the number of the additional points on the rectangular boundary including the corners. Here, for successfully fulfilling a Delaunay tessellation, we sometimes need very small perturbation of epicenters to avoid lattice structure or duplicated locations in a local domain. The panel below shows such a tessellation based on the epicenters of a JMA dataset and the additional points on the boundaries. Then, define the piecewise linear function $\phi(x, y)$ on the tessellated region such that its value at any location (x, y) in each triangle is linearly interpolated by the three values at the vertices. Specifically, consider a Delaunay triangle and the coordinates of its vertices (x_i , y_i), $i = 1, 2, 3$. Then, for the values $\phi_i = \phi(x_i, y_i)$, $i = 1, 2, 3$, the function value at any location inside the triangle is given as follows: Consider the linear equations

$$a_1 x_1 + a_2 x_2 + a_3 x_3 = x$$
$$a_1 y_1 + a_2 y_2 + a_3 y_3 = y$$
$$a_1 + a_2 + a_3 = 1$$

to obtain the non-negative solution $\hat{a}_1$, $\hat{a}_2$ and $\hat{a}_3$ so that we have

$$\phi(x, y) = \hat{a}_1 \phi_1 + \hat{a}_2 \phi_2 + \hat{a}_3 \phi_3 .$$

Such a function suitably represents the variation of the samples on a highly non-homogeneous or clustered point pattern. That is to say, we can estimate detailed changes of rate in a region where the observations are densely populated.

For further details on Delaunay tessellations, see the `Wikipedia`, Tanemura *et al*. (1983), Ogata (2004), Ogata *et al*. (2003), and Green and Sibson (1978).

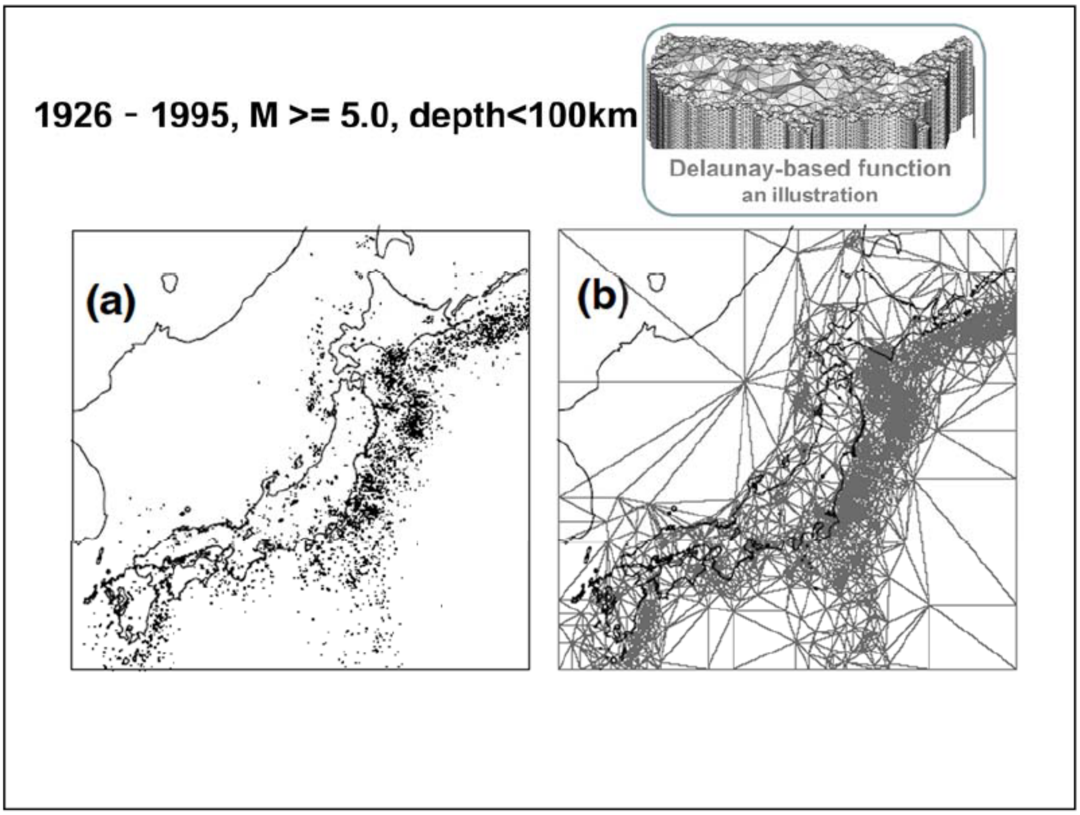


Fig. 19: (a) Epicenter locations (dots) of earthquakes of M$\geq$5.0 in and around Japan for the target period 1926-1995 together with those of M$\geq$6.0 from the period 1885-1925 that are used as the history of the ETAS model, and (b) Delaunay tessellation connecting the

epicenters and some points on the boundary.

A.3 Spatial Non-homogeneous Poisson Model

An objective method is developed for the estimation of the spatial intensity of the point locations. Consider superimposed epicenters throughout a period. Let us estimate the spatial seismicity from the earthquake locations. Now, we can consider two possible parameterizations for an intensity function $\lambda_\theta(x, y)$ of the nonhomogeneous Poisson processes. The first one is a bi-linear cubic spline function (Ogata and Katsura, 1988). However, this does not work efficiently relative to the number of necessary coefficients unless the locations are rather uniformly distributed throughout the region. The alternative is the Delaunay triangulation of this region tessellated by the earthquake locations, namely, a 2-dimensional piecewise linear function defined on the tessellation where the function value at any location is determined by the values at the vertices of Delaunay triangles. The modelling using Delaunay tessellation is suited for observations of clustered points. Namely, we can see detailed changes in the region where the observations are densely populated while smoother changes are expected in the sparsely populated regions. For the random location data $\{(x_i, y_i); i = 1, 2, \cdots, n\}$ in a region *A*, we can write the log-likelihood function as

$$\ln L(\theta) = \sum_{i=1}^{n} \ln \lambda_\theta(x_i, y_i) - \iint_A \lambda_\theta(x, y)\, dxdy$$

where we have about the same number of parameters, or even more, as the number of earthquakes. Hence, we consider the penalized log likelihood

$$R(\theta \mid w) = \ln L(\theta) - Q(\theta \mid w),$$

where, in the case of a Delaunay piecewise function,

$$Q(\theta \mid w) = w \iint_A \left\{ \left(\frac{\partial \lambda_\theta(x, y)}{\partial x} \right)^2 + \left(\frac{\partial \lambda_\theta(x, y)}{\partial y} \right)^2 \right\} dxdy$$

$$= \sum_{j:\, Delaunay\, triangles} w \Delta_j \left(\begin{vmatrix} \phi_1^j & y_1^j & 1 \\ \phi_2^j & y_2^j & 1 \\ \phi_3^j & y_3^j & 1 \end{vmatrix}^2 + \begin{vmatrix} x_1^j & \phi_1^j & 1 \\ x_2^j & \phi_2^j & 1 \\ x_3^j & \phi_3^j & 1 \end{vmatrix}^2 \right) \Bigg/ \begin{vmatrix} x_1^j & y_1^j & 1 \\ x_2^j & y_2^j & 1 \\ x_3^j & y_3^j & 1 \end{vmatrix}^2 .$$

The objective tuning of the weight *w* is carried out by the Bayesian method described in §A6 below, hence we obtain unique solutions of both the optimal weight and then the maximum a posterior estimate (in short, the MAP estimate) of the intensity function.

A.4 *b*-value estimate and forecasting seismicity

Initially assume that the *b*-value of the Gutenberg-Richter's magnitude frequency law (Gutenberg and Richter, 1944) is location independent. Historically, based on the moment method, Utsu (1965) proposed the estimator $\hat{b} = N \log e / \sum_{i=1}^{N} (M_i - M_c)$ for the observation of magnitude sequence $\{M_i,\ i=1,\ldots, N\}$ where M_c is usually the lowest bound of the magnitudes above which almost all the earthquakes are detected. This is modified by Utsu (1970) to replace M_c by $M_c - 0.05$ for the unbiased estimate of the *b*–values in case when the given magnitudes are rounded into values with 0.1 unit, and hereafter we follow this modification for the JMA catalog. Aki (1965) showed that the Utsu's *b*-estimator is nothing but the maximum likelihood estimate (MLE) that maximizes the likelihood function

$$L(b) = \prod_{i=1}^{N} \beta e^{-\beta(M_i - M_c)}, \; M_i > M_c \text{ and } \beta = b \ln 10.$$

Here, we want to assume that the *b*-value, or coefficient of the exponential distribution of magnitude, is dependent on the location in such a way that $\beta_{\boldsymbol{\theta}}(x, y) = b_{\boldsymbol{\theta}}(x, y) \ln 10$ where $\boldsymbol{\theta}$ is a parameter vector characterizing the function (Ogata *et al.*, 1991). We will solve these problems by a Bayesian procedure. Having observed the magnitude data M_i for each hypocenter's coordinates (x_i, y_i) with $i = 1,2,..., N$, the current likelihood function of θ can be written by

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{N} \beta_{\boldsymbol{\theta}}(x_i, y_i) e^{-\beta_{\boldsymbol{\theta}}(x_i, y_i)(M_i - M_c)}$$

for $M_i > M_c$. Since β, or *b*, is positive valued, we make the re-parameterization of the function $\beta_{\boldsymbol{\theta}}(x, y) = e^{\phi_{\boldsymbol{\theta}}(x,y)} / \log_{10} e$, so that the estimate of the *b*-values in space is given by $b_{\boldsymbol{\theta}}(x, y) = e^{\phi_{\boldsymbol{\theta}}(x,y)}$, where the ϕ-function is piecewise linear on the Delaunay tessellation, as given above. For a set of clusters of earthquakes, the Delaunay-based function fits better than the bi-cubic B-spline function that was used in Ogata & Katsura (1988) and Ogata *et al.* (1991). The estimation of the coefficients is undertaken by the penalized log-likelihood,

$$R(\theta \mid w) = \ln L(\theta) - w \iint_A \left\{ \left(\frac{\partial \beta_\theta(x, y)}{\partial x} \right)^2 + \left(\frac{\partial \beta_\theta(x, y)}{\partial y} \right)^2 \right\} dxdy$$

where the penalty weight *w* is tuned by a similar Bayesian procedure based on the ABIC (see Appendix B).

A.5 Space-Time ETAS Models: General Model Formulation

Denote the history of the process up to but not including time *t* as H_t where

$$H_t = \{(t_i, x_i, y_i, M_i) : t_i < t\}$$

and where (t_i, x_i, y_i, M_i) represents the time-space-magnitude outcome of the *i*-th event. The model parameters are μ, K_0, *c*, α, *p*, *d*, and *q*. In the fitted models, some or all of these parameters will vary in space, and will be denoted as $\mu(x,y)$, $K(x,y)$, *c*, $\alpha(x,y)$, $p(x,y)$, *d*, and $q(x,y)$.

Let

$$f_j(t, x, y) = \left[t - t_j + c \right]^{-p(x,y)}$$

and

$$g_j(x, y) = \left[\frac{(x - x_j, y - y_j) S_j (x - x_j, y - y_j)^t}{e^{\alpha(x,y)(M_j - M_0)}} + d \right]^{q(x,y)} \qquad \text{(a1)}$$

where M_0 is a reference magnitude (xmg0) that can be usually a threshold magnitude of completely detected (cutm in §4.2), (x_j, y_j) is the centroid location of the main shock-aftershock sequence associated with the jth event, and S_j describes the major and minor axes of the spatial intensity associated with the jth event. Note that in many cases, S_j will just be the identity matrix and (x_j, y_j) will be the location of the epicenter in the original catalog. Alternative spatial response functions to (a1) are examined in Ogata (1998) to show the predominance of (a1) in and around Japan.

The conditional intensity function can now be written as

$$\lambda(t,x,y \mid H_t) = \mu(x,y) + K_0(x,y) \sum_{\{j:t_j<t\}} g_j(x,y) f_j(t,x,y)$$

Using the Delaunay tessellations, the spatial versions of the model parameters can be expressed as

$$\mu(x,y) = \bar{\mu}\, e^{\phi_1(x,y)} \quad \text{(a2)}$$

$$K_0(x,y) = \bar{K}_0\, e^{\phi_2(x,y)} \quad \text{(a3)}$$

$$\alpha(x,y) = \bar{\alpha}\, e^{\phi_3(x,y)} \quad \text{(a4)}$$

$$p(x,y) = \bar{p}\, e^{\phi_5(x,y)} \quad \text{(a5)}$$

$$q(x,y) = \bar{q}\, e^{\phi_7(x,y)} \quad \text{(a6)}$$

In the programs, we assume that the temporal scaling parameter c and the scaling parameter d are location independent. See Ogata *et al*. (2003) and Ogata (2004).

A.5.1 Anisotropic space-time ETAS model (etas2aniso)

The simplest model (st-etas) is where no model parameters vary in space, i.e.

$$\phi_1(x,y) = \phi_2(x,y) = \phi_3(x,y) = \phi_5(x,y) = \phi_7(x,y) = 0,$$

for all x and y for functions in (a2) – (a6).

This model includes an approximate version to shorten the long computation time by considering a range within a certain prescribed distance for each earthquake that is useful for the application to seismicity in wide regions. For this version, we need to indicate a spatial distance bound of the triggering range. The input parameter is how many times of the Utsu Spatial Distance $USD = 3.33 \times 10^{0.5M-2}$ km (cf., §A.1). As the default value, it is set to be 2 times of USD in the configuration file. Hence, for the exact calculation, we put the parameter bi2 such that bi2 x USD exceeds the largest distance between earthquakes in the region.

A.5.2 HIST-ETAS model of location dependent μ and K_0-parameters

In this model (hist-etas-mk), we assume that only μ and K_0 vary over space, i.e.

$$\phi_3(x,y) = \phi_5(x,y) = \phi_7(x,y) = 0$$

and $\phi_1(x,y)$ and $\phi_2(x,y)$ are not zero for all x and y for functions in (a2) – (a6). The model is fitted by using the values of the other parameters as estimated by the model in §A.5.1, as the initial values to start, and fitting the two spatial functions given by Eqs. 1 and 2. Here, all seven baseline parameters $\bar{\mu}, \bar{K}_0, c, \bar{\alpha}, \bar{p}, d$, and $\bar{q}$ are re-estimated along with $\phi_1(x,y)$ and $\phi_2(x,y)$; i.e., Eqs. (a2) ~ (a6).

A.5.3. HIST-ETAS model (`hist-etas5pa`)

In this model, we assume that five of the parameters vary in space: μ, K_0, α, p and q, i.e. Equations 1~ 5, respectively. The values of the two constant parameters (c and d) are those as estimated by the model in §A.5.1. In addition to the parameters c and d, the baseline parameters of α, p and q as estimated by the model in §A.5.1 are fixed throughout the computation. Namely, those are same as obtained in `hist-etas-mk.` Effectively we are fixing the parameter values to those estimated in A.5.2 and only estimating the ϕ_i's, $i = 1, 2, 3, 5, 7$.

A.5.4. Forecasting by HIST-ETAS models

In a short-term span after a large earthquake j, we can make space-time forecast of aftershock activity. First, we only make a real time forecast using the isotropic matrix S_j (see §A1) within one hour after the occurrence of the earthquake j; but during the same period, a cluster analysis for the S_j is carried out. Specifically, the centroid hypocenter and variance-covariance matrix of a spatial cluster of aftershocks are formed using all detected and located earthquakes during the first hour, say, after the large earthquake. Then, based on this, the general non-isotropic space-time forecasting is performed after that.

Then, in principle, the short-term probability forecast in space-time-magnitude bin is calculated, by the simple joint distribution of the separable combination between seismicity and magnitude, given by:

$$\lambda(t, x, y; M \mid H_t)\, dt dx dy = \lambda(t, x, y \mid H_t) \cdot \hat{\beta}(x, y)\, e^{-\hat{\beta}(x,y)(M - M_c)}\, dt dx dy,$$

where the estimation procedure of the location-dependent parameter $\hat{\beta}(x, y) = \hat{b}(x, y) \ln 10$ for magnitude frequency could be applied.

However, the $\hat{b}(x, y)$ -values represent the frequency feature near the small earthquake near the threshold magnitude, but the magnitude distribution in many local regions do not follow the GR law for larger magnitudes such as taking shapes of tapering or characteristic earthquake type. For example, maximum likelihood estimates are obtained for many modified Gutenberg-Richter magnitude frequency distributions (see Utsu, 1999). Another issue is that b-values for the mainshocks and aftershocks can be significantly different (Utsu, 1971). Also, Ogata et al. (2018) did not confirm that the magnitude forecasts by location dependent b-value throughout Japan region outperform the baseline G–R law with the b value of 0.9. Hence, at this moment, we may rather assume generic magnitude frequency $\hat{\beta} = \hat{b} \ln 10$ with $\hat{b} = 0.9$ throughout the entire target region, instead of location-dependent estimate $\hat{\beta}(x, y)$, for a stable forecasting.

A.6 Likelihoods and Penalized Likelihoods

A.6.1 log-likelihood function and its maximization

Now we start with the simplest space-time ETAS model in which all the parameters $\boldsymbol{\theta} = (\mu. K, c, \alpha, p, d, q)$ of the ASTETAS model in §A3.1 are constant throughout the whole region, equivalently, all the functions $\phi_k(x, y)$, $k = 1, 2$, 3, 5, 7 defined are equal to zero. The maximum likelihood estimates (MLE) are obtained by the maximizing the log-likelihood function

$$\ln L(\boldsymbol{\theta}) = \sum_{\{i;\, S<t_i<T\}} \ln \lambda_\theta(t_i, x_i, y_i \mid H_{t_i}) - \int_S^T \iint_A \lambda_\theta(t, x, y \mid H_t)\, dxdydt, \qquad \text{(a7)}$$

for the earthquakes in the target period $[S, T]$, where H_t is the history of earthquake occurrences before time t including those from the precursory period $[0, S]$. For the detailed numerical description of the log-likelihood function, especially of the second integral term in (a7), the reader is referred to Ogata (1998). Then we use a quasi-Newton method (Fletcher and Powell, 1963; Kowalik and Osborne, 1968, etc.) for the numerical maximization.

When the number of earthquakes (say, n) in the data is large, the computing take a substantial time due to the double sum of $n^2/2$ terms in the first part of the log likelihood (a7). Unlike the computation using the Markovian recursive relation in the conditional intensity of the ETAS model (Ogata *et al*., 1993), such a recursive calculation of the conditional intensity of the space-time ETAS is not available. Instead, one may be interested in a quicker spatially approximate computation by only taking the double sum of the earthquake pairs closer than a certain distance, such as 2 times the Utsu Spatial Distance $3.33 \times 10^{0.5M-2}$ km (cf., §A.1). The HIST-ETAS models in A5 and A6 use this restriction.

A.6.2 Penalised log-likelihood function and its optimization

Here we consider the hierarchical models with location dependent parameters in §A.3 to describe spatial heterogeneity. These models require a large number of further parameters for the coefficients of functions $\phi_k(x, y)$, $k = 1, 2, \ldots, 5$. Let such coefficients be described by the parameter set $\{\boldsymbol{\theta} = (\theta_i) \in \Theta\}$, and let the likelihood function be given by $L(\boldsymbol{\theta} | \text{data})$. To estimate the parameters, we frequently use the penalised log likelihood (Good and Gaskins, 1971)

$$R(\boldsymbol{\theta}, \boldsymbol{\tau} \mid \text{data}) = \ln L(\boldsymbol{\theta} | \text{data}) - Q(\boldsymbol{\theta} | \boldsymbol{\tau}), \qquad \text{(a8)}$$

where the function Q represents a positive valued penalty function, and $\boldsymbol{\tau} = (w_1, w_2)$ or $\boldsymbol{\tau} = (w_1, \cdots, w_5)$ is a vector of the hyper-parameters that control the strength of some constraints between the parameters bundled by $\boldsymbol{\theta}$. Greater constraints will impose more smoothness in $\phi_k(x, y)$, less constraints allows greater roughness. For the penalties, besides the simplest penalty in §A3 and §A4, we can consider

$$Q(\theta \mid \tau) = w \iint_A \left\{ \left(\frac{\partial \phi}{\partial x} \right)^2 + \left(\frac{\partial \phi}{\partial y} \right)^2 \right\} dxdy$$

for b-values of the location-dependent G-R law and non-homogeneous Poisson processes, we use

$$Q(\theta \mid \tau) = \sum_{k=1}^{2} w_k \iint_A \left\{ \left(\frac{\partial \phi_k}{\partial x} \right)^2 + \left(\frac{\partial \phi_k}{\partial y} \right)^2 \right\} dxdy \qquad \text{(a9)}$$

for the HIST-ETAS with location dependent μ and K parameters, and

$$Q(\theta \mid \tau) = \sum_{k=1}^{5} w_k \iint_A \left\{ \left(\frac{\partial \phi_k}{\partial x} \right)^2 + \left(\frac{\partial \phi_k}{\partial y} \right)^2 \right\} dxdy \qquad \text{(a10)}$$

for the HIST-ETAS with location dependent μ, K, α, p and q parameters. Furthermore, in addition to each penalty, we sometimes need damping constraints for ϕ_1 and ϕ_2 corresponding to μ and K_0, $\sum_{k=1}^{2} w_0 \iint_{\partial A} \phi_k(x,y)^2\, dxdy$, only on the boundary of the region ∂A, where w_0 is fixed throughout the optimization procedure of other hyperparameters (weights).

The penalized log-likelihood in (a8) defines a trade-off between the goodness of fit to the data and the uniformity of each function, namely, the facets of the piecewise linear function being as flat as possible. A smaller weight leads to a higher regional variability of the ϕ-functions. The crucial point here is the tuning of the vector $\boldsymbol{\tau}$. From the Bayesian viewpoint, the penalty function is related to the prior probability density

$$\pi(\boldsymbol{\theta} \mid \boldsymbol{\tau}) = e^{-Q(\boldsymbol{\theta} \mid \boldsymbol{\tau})} / \int_{\Theta} e^{-Q(\boldsymbol{\theta} \mid \boldsymbol{\tau})} \mathbf{d\theta},$$

and the exponential to the penalized log likelihood function R is proportional to the posterior function. For determining suitable values of the hyper-parameters $\boldsymbol{\tau}$, consider the posterior probability density function

$$p(\boldsymbol{\theta} \mid \text{data}; \boldsymbol{\tau}) = L(\boldsymbol{\theta} \mid \text{data}) \pi(\boldsymbol{\theta} \mid \boldsymbol{\tau}) \big/ \Lambda(\boldsymbol{\tau} \mid \text{data})$$

with normalizing factor

$$\Lambda(\boldsymbol{\tau} \mid \text{data}) = \int_{\Theta} L(\boldsymbol{\theta} \mid \text{data}) \pi(\boldsymbol{\theta} \mid \boldsymbol{\tau}) \mathbf{d\theta}. \qquad \text{(a11)}$$

The maximization of this normalizing factor or its logarithm with respect to the hyper- parameters τ is called the method of the Type II maximum likelihood due to Good (1965). Given a set of data, one seeks to compare the goodness-of-fit of Bayesian models that have distinct likelihoods or distinct priors and to search for the optimal hyper-parameter values. For instance, Ogata *et al*. (1991) compared the use of different priors for isotropic and anisotropic smoothness constraints, which need two and five hyper-parameters, respectively. For such a purpose, Akaike (1980) justified and developed Good’s method based on the entropy maximization principle (Akaike, 1978) and defined

$$\mathrm{ABIC} = -2\max_{\boldsymbol{\tau}} \ \ln \ \Lambda(\boldsymbol{\tau}| \,\mathrm{data}) \ +2\dim(\boldsymbol{\tau}) \qquad \text{(a12)}$$

for consistent use with the Akaike Information Criterion (AIC; Akaike, 1974). Here, $\dim(\boldsymbol{\tau})$ is the number of the hyper-parameters . Both ABIC and AIC are to be minimized for the comparison of Bayesian and ordinary likelihood-based models, respectively, for better fit to the data. The normalizing factor $\Lambda(\boldsymbol{\tau}|$ data) in (a11) is called the likelihood of the Bayesian model with respect to the hyper-parameters $\boldsymbol{\tau}$.

For practical computation of the normalizing factor $\Lambda(\boldsymbol{\tau}|$ data) in (a11), see the §B.2 below.

B Background to Computation Algorithms

This Appendix gives a description of the computing algorithms that are used to fit the models.

B.1 Nonlinear optimization for the maximum likelihood estimates (MLE)

For the maximum likelihood procedure of a space-time ETAS model (etasSelectAniso) in §A3.1, we use a quasi-Newton optimization for non-linear functions called Davidon-Fletcher- Powell algorithm (Fletcher and Powell, 1963). Also see Kowalik and Osborne (1968) or *Wikipedia* for an introduction.

To get the optimal parameters, we repeat the following steps (A) - (D):

(A) For a given fixed τ, calculate the negative log-likelihood and its gradient vector $\boldsymbol{u}$ at an initially given parameter vector $\boldsymbol{\theta}_0$.

(B) Search the smallest negative log likelihood function (a7) with respect to $\boldsymbol{\theta}$ on the one-dimensional straight line determined by the initial parameter vector $\boldsymbol{\theta}_0$ and the gradient vector $\boldsymbol{u}$ (Linear Search; e.g., Kowalik and Osborne, 1968).

(C) Replace the minimizing parameter $\hat{\boldsymbol{\theta}}$ in step (B) by $\boldsymbol{\theta}_0$. Then, compute the gradient vector $\boldsymbol{u}_0$ at $\boldsymbol{\theta}_0$. Solve the equation $H_T \boldsymbol{u} = \boldsymbol{u}_0$ by an estimated Hessian to get a vector $\boldsymbol{u}$ for the direction of the next linear search in step (B).

(D) Repeat A-C until the negative log-likelihood function T attains the minimum overall $\boldsymbol{\theta}$, which is the maximum likelihood estimate (MLE).

In quasi-Newton methods the Hessian matrix (second derivatives of the function) need not be computed. An estimated inverse Hessian matrix is calculated by using the gradients during the steps of searching for the minimum of the negative log-likelihood function.

B.2 Computations of Bayesian models through Gaussian approximations

In general, it is hard to get the high dimensional integration (a11) analytically unless the posterior distribution is Gaussian. This is because the likelihood function of the point-process model is not Gaussian distributed. Nevertheless, by virtue of the Gaussian prior distribution, Gaussian approximation of the posterior function is useful. Namely, we take the Gaussian approximation of the posterior distribution, utilising the quadratic form around the log-posterior maximum solution. That is to say, the penalized log-likelihood is well approximated by the quadratic form

$$T(\boldsymbol{\theta}|\,\boldsymbol{\tau}) \equiv \ln L(\boldsymbol{\theta}|\,\mathbf{Y}) + \ln \pi(\boldsymbol{\theta}|\,\boldsymbol{\tau}) \approx T(\hat{\boldsymbol{\theta}}|\,\boldsymbol{\tau}) - \frac{1}{2}(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})H_T(\hat{\boldsymbol{\theta}}|\,\boldsymbol{\tau})(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})^t \qquad \text{(b1)}$$

around $\hat{\boldsymbol{\theta}} = \arg\{\max_{\boldsymbol{\theta}} T(\boldsymbol{\theta}|\,\boldsymbol{\tau})\}$, and $H_T(\boldsymbol{\theta}|\,\boldsymbol{\tau})$ is the Hessian of $T(\boldsymbol{\theta}|\,\boldsymbol{\tau})$ consisting of its negative second-order partial derivatives with respect to $\boldsymbol{\theta}$.

We further assume that the Hessian matrix in (b1) is well approximated by a block diagonal matrix of five sub-matrices, H_T =diag{ $H_T^1, H_T^2, H_T^3, H_T^4, H_T^5$ }, relying on the Hessian of the prior where each block relates the model parameters μ. K_0, α, p, and q, respectively. Namely, we assume independency between the coefficients of the different ϕ_k -functions in the penalized log-likelihood (a8). Thus, the logarithm of the likelihood (11) of the Bayesian model is given by

$$\begin{aligned}\ln \Lambda(\mathbf{Y}) &= \log \int_{\Theta} L(\boldsymbol{\theta}\,|\,\mathbf{Y})\,\pi(\boldsymbol{\theta}\,|\,\boldsymbol{\tau})\,\mathbf{d}\boldsymbol{\theta} \\ &\approx T(\hat{\boldsymbol{\theta}}\,|\,\boldsymbol{\tau}) - \frac{1}{2}\ln\det\{H_T(\hat{\boldsymbol{\theta}}\,|\,\boldsymbol{\tau})\} + \frac{1}{2}\dim\{\boldsymbol{\theta}\}\log 2\pi \\ &= R(\hat{\boldsymbol{\theta}}\,|\,\boldsymbol{\tau}) - \frac{1}{2}\ln\det\{H_R(\hat{\boldsymbol{\theta}}\,|\,\boldsymbol{\tau})\} + \frac{1}{2}\ln\det\{H_Q(\hat{\boldsymbol{\theta}}\,|\,\boldsymbol{\tau})\},\end{aligned}$$

where H_R and H_Q is the block diagonal Hessian matrix of the function R and Q in (a8), respectively, and 'det{.}' indicates the determinant of the matrices.

Then, we implement the maximization of the penalized log-likelihood (a8) with respect to the coefficients of the ϕ-functions.

In the maximization with respect to the $2(N+n)$ dimensional coefficient vectors, we alternately adopt a linear search procedure and the incomplete Cholesky conjugate gradient (ICCG) method by inverting a block diagonal Hessian matrix $H_R(\hat{\boldsymbol{\theta}}|\,\boldsymbol{\tau})$ (see §B.2), where N is the number of earthquakes and n is the number of the additional points on the rectangular boundary including the corners (see §6.4 and the figure in §6.5). This procedure makes the convergence very rapid regardless of the high dimensionality of $\boldsymbol{\theta}$ if the Gaussian approximation at Equation (b1) is adequate for the posterior function.

Having attained such convergence for a given hyper-parameter $\boldsymbol{\tau}$, we further need to perform the maximization of $\Lambda(\boldsymbol{\tau})$ defined in (a11) with respect to $\boldsymbol{\tau}$ by a direct search such as the simplex method (e.g., Kowalik and Osborn, 1968) in either 2 or 7 dimensional space depending on the programs. Thus, we perform the double optimizations with respect the parameters (coefficients) $\boldsymbol{\theta}$ and the hyper-parameters (weights) $\boldsymbol{\tau}$. These are alternately repeated until the latter maximization converges (see the diagram in Fig. 20 below). The whole optimization procedure usually converges when initial vector values for $\boldsymbol{\tau}$ are set in such a way that the penalty is reasonably close to the correct value; otherwise, it may take very many steps to reach the solution, or it may even diverge. Eventually, we obtain the optimal maximum posterior (OMAP) solution $\hat{\boldsymbol{\theta}}$ for the maximum likelihood estimate $\hat{\boldsymbol{\tau}}$.

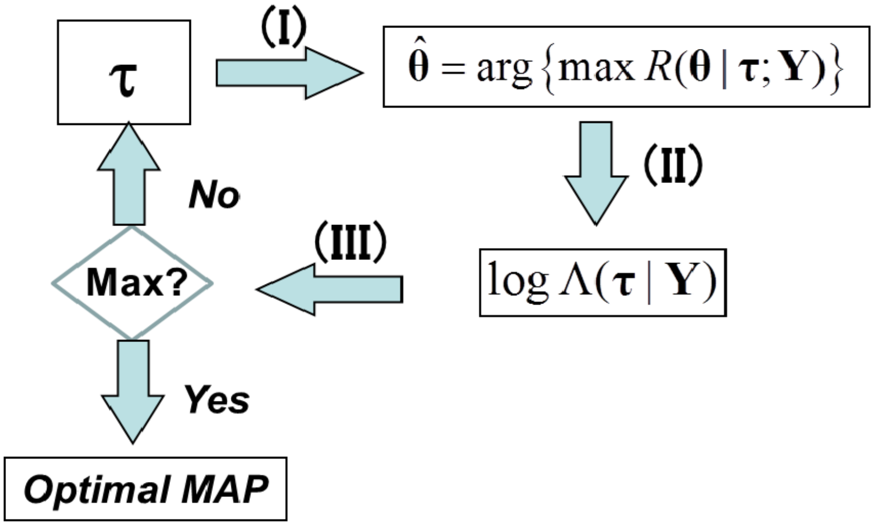


Fig. 20. Diagram of Double Optimizations. (I) performs the maximization of the function R with respect to θ. (II) calculates the log likelihood of the Bayesian model using the quadratic approximation expanded at $\hat{\boldsymbol\theta}$. (III) maximizes the log likelihood with respect to τ.

To get the optimal hyper-parameters, we repeat the following steps (A) - (D):

(A) For a given τ being fixed, set the gradient of the penalized log-likelihood, $\boldsymbol{u}$ = $\partial T / \partial \boldsymbol\theta$ at an initial parameter $\boldsymbol\theta_0$.

(B) Maximize T in (b1) with respect to $\boldsymbol\theta$, that is, on the one-dimensional straight line determined by the initial parameter vector $\boldsymbol\theta_0$ and the gradient vector $\boldsymbol{u}$ (Linear Search; e.g., Kowalik and Osborne, 1968).

(C) Replace the maximizing parameter $\hat{\boldsymbol\theta}$ in step (B) by $\boldsymbol\theta_0$. Then, compute the gradient vector $\boldsymbol{u}_0 = \partial T / \partial \boldsymbol\theta$ at $\boldsymbol\theta_0$. Solve the equation $H_T \boldsymbol{u} = \boldsymbol{u}_0$ by the Incomplete Cholesky Conjugate Gradient (ICCG) method (e.g., Mori, 1986) to get the vector $\boldsymbol{u}$ for the direction of the next linear search in step (B) until the function T attains the overall maximum $\boldsymbol\theta$, which is the maximum posterior (MAP) solution for the given $\boldsymbol\tau$.

(D) Calculate $\log \Lambda(\boldsymbol\tau)$ using the quadratic approximation around the MAP $\hat{\boldsymbol\theta}$, and go to step (A) with the other $\boldsymbol\tau$ to maximize $\log \Lambda(\boldsymbol\tau)$ by the direct-search maximizing method, such as the simplex method (e.g., Kowalik and Osborne, 1968; and Murata, 1992). The steps (A) ~ (D) are repeated in turn until $\log \Lambda(\boldsymbol\tau)$ converges.

According to our experience, the convergence rate in step (C) is very fast in spite of the very high dimensionality of $\boldsymbol\theta$. This is expected when the quadratic approximations of T are adequate in a region around the MAP solution, otherwise it is likely to take endless iterations or even diverge. After all, by assuming a uni-modal posterior function, we can get the optimal MAP solution $\hat{\boldsymbol\theta}$ for the maximum likelihood estimate $\hat{\boldsymbol\tau}$ of the hyper-parameters. The reader is referred to Ogata and Katsura (1988, 1993), Ogata *et al*. (1991, 2000, 2001), and related references therein which further describe computational details.

B.3 Notes on location-dependent μ and K_0 ETAS fitting (`hist-etas-mk`)

The penalized log-likelihood in (a8) defines a trade-off between the goodness of fit to the data and the uniformity of each parameter function. We obtain the optimal weights $\hat{\boldsymbol{\tau}} = (\hat{w}_1, \hat{w}_2)$ together with the maximizing baseline parameters ($\bar{\mu}$, $\bar{K}$) for the first two programs or ($\bar{\mu}$, $\bar{K}$, c, α, p, d, q) for the last two, by the principle of maximizing the integrated posterior function (a11). Here note that the baseline parameters $\bar{\mu}$ and $\bar{K}$ are automatically determined by the zero-sum constraint of the corresponding ϕ-function. This overall maximization can be eventually attained by repeating alternate procedures of the separated maximizations with respect to the parameters (coefficients) and hyper-parameters (weights) described as follows.

First of all, for the initial inputs, we use the MLEs $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{K}, \hat{c}, \hat{\alpha}, \hat{p}, \hat{d}, \hat{q})$ obtained by the primary space-time ETAS model (`st-etas`), for the baseline parameter, and also set all the coefficients of ϕ-functions to be zero such that $\phi_1(x, y) = \phi_2(x, y) = 0$.

Since the penalty functions already have the quadratic form with respect to the parameters θ, the prior density is of a multivariate Gaussian distribution, in which the Hessian matrix H_Q consists of the elements of the negative second order partial derivatives of the penalty function Q. Actually, the present penalty function implies that the Hessian is a block diagonal matrix of five sub-matrices corresponding to each ϕ_k -function in (a2)~(a6) such that H_Q =diag{ H_μ^1, H_κ^2 }. This is because we do not consider any restrictions a priori between the different ϕ_1 and ϕ_2 -functions. Here, all sub-matrices of H_Q^k are sparse, and have the same configuration of non-zero elements. Specifically, the $(i,\ j)$ -element is non-zero if and only if the pair of points i and j are vertices of the same Delaunay triangle; cf., §6.3.

B.4 Notes on location-dependent μ, K, α, p and q ETAS fitting (`hist-etas5pa`)

Having obtained the optimal weights $\hat{\boldsymbol{\tau}} = (\hat{w}_1, \hat{w}_2)$ and the MAP coefficients of $\hat{\phi}_1(x, y)$ and $\hat{\phi}_2(x, y)$ with the baseline parameters $\hat{\mu}, \hat{K}, \hat{c}, \hat{\alpha}, \hat{p}, \hat{d}$, $\hat{q}$ in the μK-HIST-ETAS model, we use all of these for initial inputs to stably estimate the HIST-ETAS model in §A.3 with five spatially varying parameters in (a2) - (a6). Also, set other coefficients of α, p and q parameter functions being zero such that $\phi_3(x, y) =$ $\phi_4(x, y) =$ $\phi_5(x, y) =$ 0 with the estimated baseline values $\hat{\mu}, \hat{K}, \hat{c}, \hat{\alpha}, \hat{p}, \hat{d}$ and $\hat{q}$ of the μK-HIST-ETAS model (`hist-etas-mk`).

Here, we consider the penalized log-likelihood function (a8) with the penalty function

$$Q(\theta \mid \tau) = \sum_{k=1}^{5} w_k \iint_A \left\{ \left(\frac{\partial \phi_k}{\partial x} \right)^2 + \left(\frac{\partial \phi_k}{\partial y} \right)^2 \right\} dxdy \qquad \text{(b2)}$$

of $\boldsymbol{\tau} = (w_1, ..., w_5)$. In addition, we need damping constraints for ϕ_1 and ϕ_2 corresponding to μ and K_0; $\sum_{k=1}^{2} w_k \iint_{\partial A} \{\partial\phi_k(x, y)/\partial x\}^2 + \{\partial\phi_k(x, y)/\partial y\}^2 \, dxdy$ only on the boundary of the region ∂A. For technical reasons, the baseline values

$\hat{\mu}$, $\hat{K}$, $\hat{c}$, $\hat{\alpha}$, $\hat{p}$, $\hat{d}$, $\hat{q}$ and w_0 in the programs are fixed throughout the whole computations. Thus the optimal weights $\hat{\boldsymbol{\tau}} = (\hat{w}_1, \hat{w}_2, \hat{w}_3, \hat{w}_4, \hat{w}_5)$ are obtained by the similar procedure of maximizing the integrated posterior function (see A.5.2) to that of the μK-HIST-ETAS model in §B.3.

Since the penalty function in (b1) already has the quadratic form with respect to the parameters θ, the prior density is of a multivariate Gaussian distribution, in which the Hessian matrix H_Q consists of the elements of the negative second order partial derivatives of the penalty function Q. Actually, the present penalty function implies that the Hessian is a block diagonal matrix of five sub-matrices corresponding to each ϕ_k -function in (a2)~(a6) such that H_Q =diag{ $H_Q^1, H_Q^2, H_Q^3, H_Q^4, H_Q^5$ }. This is because we do not consider any restrictions a priori between the different ϕ_k -functions. Here, all sub-matrices of H_Q^k are sparse, and have the same configuration of non-zero elements. Specifically, the (i, j) -element is non-zero if and only if the pair of points i and j are vertices of the same Delaunay triangle; cf., §6.3.

Specifically, this maximization is performed sequentially and alternately as follows. First, we implement the maximization of the penalized log-likelihood (a8) with respect to the coefficients of the ϕ-functions; see Eqs. (a2) - (a6). For the calculation, we adopt a linear search using the incomplete Cholesky conjugate gradient (ICCG) method for $5(N+n)$ dimensional coefficient vectors, where $N+n$ is the same number as given in §6.3. Alternately, we implement the simplex algorithm in the 5-dimensional space of $(\hat{w}_1, \hat{w}_2, \hat{w}_3, \hat{w}_4, \hat{w}_5)$ to maximize $\Lambda(\boldsymbol{\tau})$ until this converges. Here, before doing the 5-dimensional simplex search, we recommend to firstly make a lattice search of (w_3, w_4, w_5) in the logarithmic orders, such as $(10^i, 10^j, 10^k)$, for possible sets of integers i, j and k to compare the respective ABIC values h, while (w_1, w_2) remain fixed to those $(\hat{w}_1, \hat{w}_2)$ obtained in §9.3. It is a limitation of this procedure that this maximization may not converge for small sets of integers because the convergence relies on the quadratic approximation penalized log likelihood (see Appendix and the ICCG method). From our experience, selection from 2 or 3 or 4 for the above i, j and k, can be a good choice of the starting values. Then, using the set of weights with the smallest ABIC value, we can implement the 3-dimensional simplex search of (w_3, w_4, w_5) or even the 5-dimensional simplex search of $(w_1, w_2, w_3, w_4, w_5)$ for a global minimum. Here it is important to make use of the previously converged solutions of parameters (coefficients) for the next initial parameters of such large dimensions.

It is also useful to examine whether or not the characteristic parameters, particularly $\alpha(x,y)= \hat{\alpha}\exp\{\phi_3(x, y)\}$, $p(x,y)= \hat{p}\exp\{\phi_4(x, y)\}$ and $q(x,y) = \hat{q}\exp\{\phi_5(x, y)\}$ are significantly uniform (i.e., spatially invariant). For this we can calculate the Akaike Bayesian Information Criterion (ABIC; see Appendix) as a byproduct of the above simplex optimization. A model with a smaller ABIC value indicates a better fit. For example, we can compare the ABIC values of the HIST-ETAS model for the optimal weights $(\hat{w}_1, \hat{w}_2, \hat{w}_3, \hat{w}_4, \hat{w}_5)$ with the one for $(\hat{w}_1, \hat{w}_2, \hat{w}_3, \hat{w}_4, 10^8)$ to examine whether q-value is location dependent or not.

Acknowledgement

We are grateful for the comments by Yicun Guo (former Ph. D. student of Peking University), Taku Ueda (Ph. D. student of Earthquake Research Institute, University of Tokyo) based on the use of previous version of HIST-PPM program package; which were very useful for the revised version. The careful reading of this manual and many comments by anonymous reviewer for publishing in the ISM Computer Science Monographs the manual are greatly appreciated. Authors are supported by the JSPS KAKENHI grant 17H00727.

References

Akaike, H. (1974). A new look at the statistical model identification, *IEEE Trans. Autom. Control*, **AC-19**, 716– 723.

Akaike, H. (1978). A new look at the Bayes procedure, *Biometrika*, **65**, 53–59.

Akaike, H. (1980). Likelihood and Bayes procedure, in *Bayesian Statistics*, eds. J. M. Bernard, *et al*., 1 –13, Univ. Press, Valencia, Spain.

Aki, K. (1965). Maximum likelihood estimate of b in the formula log *N=a-bM* and its confidence limits, *Bull. Earthq. Res. Inst.*, **43**, 237-239.

Dziewonski, A.M., T.A. Chou, and J.H.Woodhouse (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res*., **86**, 2825-2852.

Fletcher, R. and M.J.D., Powell (1963). A rapidly convergent descent method for minimization, *Comput. J*., **6**, 163-168.

Good, I.J. (1965), *The Estimation of Probabilities*, M.I.T. Press, Cambridge, Massachusetts.

Good, I.J. and R.A. Gaskins (1971), Nonparametric roughness penalties for probability densities, *Biometrika* **58**, 255-277.

Green, P.J.and Sibson, R. (1978). Computing Dirichlet tessellations in the plane. *Computer Journal* **21(2)**, 168–173. DOI:10.1093/comjnl/21.2.168

Gutenberg, R. and C.F. Richter (1944). Frequency of earthquakes in California, *Bull. Seismol. Soc. Amer.*, **34**, 185-188.

Kowalik, J. and M.R. Osborne (1968). *Methods for Unconstrained Optimization Problems*, American Elsevier, New York.

Mori, M. (1986). *FORTRAN 77 Numerical Analysis Programming*, 342pp., Iwanami Publisher, Tokyo, (in Japanese).

Murata, Y. (1992). Estimation of optimum surface density distribution only from gravitational data: an objective Bayesian approach, *J. Geophys. Res*., **98**, 12097-12109.

Ogata, Y. (1981), On Lewis' simulation method for point processes, IEEE Information Theory, IT-27, 23-31.

Ogata, Y. (1985). Statistical models for earthquake occurrences and residual analysis for point processes, Research Memorandum, No. 288, The Institute of Statistical Mathematics, Tokyo, http://www.ism.ac.jp/editsec/resmemo-e.html.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes, *J. Amer. Statist. Assoc*., 83, 9-27.

Ogata, Y. (1989), Statistical model for standard seismicity and detection of anomalies by residual analysis, *Tectonophysics*, **169**, 159-174.

Ogata, Y. and Katsura, K. (1993). Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogues, *Geophys. J. Int*., **113**, 727-738.

Ogata, Y., Matsu'ura, R. S. and Katsura, K. (1993). Fast likelihood computation of epidemic type aftershock sequence model, *Geophys. Res. Lett.*, **20(**19), 2143-2146.

Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Ann. Inst. Statist. Math.* **50**, 379–402. DOI:10.1023/A:1003403601725

Ogata, Y. (2004). Space-time model for regional seismicity and detection of crustal stress changes. *J. Geophys. Res.* **109(B3)**, B03308. DOI:10.1029/2003JB002621. See corrections in *J. Geophys. Res.* **109(B6)**, B06308. DOI:10.1029/2004JB003123.

Ogata, Y. (2006). *Statistical Analysis of Seismicity, Updated Version* (*SASeis2006*),Computer Science Monograph, No 33, Institute of Statistical Mathematics, Tokyo.

Ogata, Y. (2008), Occurrence of the large earthquakes during 1978~2007 compared with the selected seismicity zones by the Coordinating Committee of Earthquake Prediction (in Japanese), *Report of the Coordinating Committee for Earthquake Prediction*, 79, 623-625.

Ogata, Y. (2010). Significant improvements of the space-time ETAS model for forecasting of accurate baseline seismicity. *Earth, Planets and Space* **63**, 217-229, doi:10.5047/eps.2010.09.001**.**

Ogata, Y., M. Imoto and K. Katsura, (1991). 3-D spatial variation of b-values of magnitude-frequency distribution beneath the Kanto District, Japan, *Geophys. J. Int*. **104**, 135-146.

Ogata, Y. and K. Katsura (1988). Likelihood analysis of spatial inhomogeneity for marked point patterns, *Ann. Inst. Statist. Math*., **40**, 29-39.

Ogata, Y. and K, Katsura (1993). Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogues, *Geophys. J. Int.*, **113**, 727-738.

Ogata, Y., Utsu ,T. and Katsura, K. (1995). Statistical features of foreshocks in comparison with other earthquake clusters, *Geophys. J. Int*., **121**, 233-254.

Ogata, Y., Katsura, K., Tanemura, M. (2003). Modelling heterogeneous space-time occurrences of earthquakes and its residual analysis. *Applied Statistics* **52(4)**, 499–509. DOI:10.1111/1467-9876.00420

Ogata, Y., Katsura, K., Zhuang, J.C. (2006). *TIMSAC84: Statistical Analysis of Series of Events (TIMSAC84-SASE), Version 2.* Computer Science Monograph, No 32. Institute of Statistical Mathematics, Tokyo.

Ogata, Y., Matsu’ura, R.S., Katsura, K. (1993). Fast likelihood computation of epidemic type aftershock sequence model. *Geophys. Res. Lett.* **20(19)**, 2143–2146. DOI:10.1029/93GL02142

Ogata, Y., Zhuang, J.C. (2006). Space-time ETAS models and an improved extension. *Tectonophysics* **413(1-2)**, 13–23. DOI:10.1016/j.tecto. 2005.10.016

Ogata, Y. and Katsura, K. (2012). Prospective foreshock forecast experiment during the last 17 years, *Geophys. J. Int.*, **191**, Issue3, 1237–1244, doi:10.1111/j.1365-246X.2012.05645.x

Rproject R Development Core Team. (2009). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna. ISBN: 3-900051-07-0, URL: http://www.R-project.org

Tanemura, M., Ogawa, T., Ogita, N. (1983). A new algorithm for three-dimensional Voronoi tessellation. *Journal of Computational Physics* **51(2)**, 191–207. DOI:10.1016/0021-9991(83)90087-6

Tsuruoka, H. (1996). Development of seismicity analysis software on workstation. *Tech. Res. Rep.* 2, 34–42, Earthquake Res. Inst., Univ. Tokyo (in Japanese).

Utsu, T. (1965). A method for determining the value of b in a formula log n = a-bM showing the magnitude frequency relation for earthquakes (in Japanese), *Geophys. Bull. Hokkaido Univ.*, **13**, 99-103.

Utsu, T. and A. Seki (1955). Relation between the area of aftershock region and the energy of the main shock, *Zisin* (*J. Seismol. Soc. Japan*), 2nd Ser., ii, **7**, 233-240 (in Japanese with English summary).

Utsu, T. (1969). Aftershocks and earthquake statistics (I): some parameters which characterize an aftershock sequence and their interaction, *J. Faculty Sci., Hokkaido Univ*., Ser. VII (geophysics) **3**, 129-195.

Utsu, T. (1970). Aftershocks and earthquake statistics (II): Further investigation of aftershocks and other earthquake sequences based on a new classification of earthquake sequences, *J. Faculty Sci., Hokkaido Univ*., Ser. VII (geophysics), **3**, 198-266.

Utsu, T. (1982), Catalog of Large Earthquakes in the Region of Japan From 1885 Through 1980, *Bull. Earthq. Res. Inst., Univ. Tokyo*, **57**, 401-463.

Utsu, T. (1985). Catalog of large earthquakes in the region of Japan from 1885 through 1980: Correction and supplement. *Bull. Earthq. Res. Inst., Univ. Tokyo*, **60**, 639-642.

Zhuang, J., C. Chang, Y. Ogata, and Y. Chen (2005). A study on the background and clustering seismicity in the Taiwan region by using point process models, *J. Geophys. Res.*, **110**, B5, B05S18, doi:10.1029/2004JB003157, 2005.

Zhuang J. and Ogata Y. (2006). Properties of the probability distribution associated with the largest event in an earthquake cluster and their implications to foreshocks. *Physical Review*, E. 73, 046134. doi: 10.1103/PhysRevE.73.046134