Guide to run the code:

This guide explains how to RUN the code related to the seismicity forecasting based on a Bayesian spatio-temporal ETAS model. This code has been originally developed in MATLAB by Hossein Ebrahimian (University of Naples Federico II, UNINA, Italy) and Fatemeh Jalayer (Institute for Risk and Disaster Reduction, IRDR, University College London, UK & University of Naples Federico II) in the following two publications (the methodology is demonstrated by retrospective early forecasting of seismicity associated with the 2016 Central Italy and 2017-2019 Western Iran seismic sequence activities):

If you use this code, please cite the following two articles:

(1) Ebrahimian, H., Jalayer, F., Maleki Asayesh, B., Hainzl, S., Zafarani, H. Improvements to seismicity forecasting based on a Bayesian spatio-temporal ETAS model. *Scientific Reports*, (2022) https://doi.org/10.1038/s41598-022-24080-1.

(2) Ebrahimian, H. & Jalayer, F. Robust seismicity forecasting based on Bayesian parameter estimation for epidemiological spatio-temporal aftershock clustering models. *Scientific Reports* 7, 9803 (2017), https://doi.org/10.1038/s41598-017-09962-z.

The main script is called "**main_script.mat**". The main input parameters, defined in this script, are as follows:

- lonMin, lonMax, latMin, latMax: The minimum and maximum longitude and latitude boundaries of the Aftershock zone (AS).
- deltaGrid: the mesh grid size.
- catalog.txt: The text file containing the catalog of the ongoing seismic sequence which consists of 5 column vectors including [latitude [deg], longitude [deg], magnitude, time, time_T1]. It is to note "time" is the time from the origin (T_o) and "time_T1" is the time from the first event in the sequence. They should have the same unit (e.g., hour or day).
- background seismicity.txt: The text file containing the background seismicity data. In Reference [1] above, we have defined it as $\mu(x, y|M_l)$. Therefore, it should be a matrix with the size of [length (Ycgrid), length (Xcgrid)]. It is noted that "Ycgrid" and "Xcgrid" are latitude and longitude of the center of each cell unit within the grid mesh.
- Mmax: Upper-bound Magnitude of the aftershock zone.

- Mc: Lower cut-off magnitude (which should be equal to or greater than the completeness magnitude).
- vec_m: The vector containing the magnitude values *m* used for forecasting number of events with $M \geq m$. It starts with Mc.
- T_start, T_end: Start and end time of the forecasting time interval of interest.
- method_for_analysis: 'Slow' which is the *Proposed method* (see Reference [1], Figure 10), 'Semi-Fast' and 'Fast' methods.
- use_background: (=1) consider background seismicity that is in "background seismicity.txt" file; (=0) consider BGS=0 (see Reference [1]).
- use_kernel_distance_m: (=1) magnitude-dependent spatial kernel density; (=0) simple spatial kernel density (see Reference [1]).
- beta_ini, K_ini, alpha_ini, c_gen, p_gen, d_ini, q_ini, gamma_ini: Initial median values for prior distribution of ETAS model parameters [β, K, α, c, p, d, q, γ]. It is to note that the priors have a lognormal distribution.
- cov_beta, cov_K, cov_alpha, cov_c, cov_p, cov_d, cov_q, cov_gamma: COV (coefficient of variation) of the prior of ETAS parameters.
- output_Dir: Output directory address

- do_MCMC_updating: (=1) Perform MCMC and generate samples from the ETAS model parameters, (=0) Skip this operation and use the already existing samples in the output directory. The results will be saved in the file "**samples.mat**"
- do_find_Robust_estimate: (=1) Generate sequences within the forecasting interval and perform a robust estimate of seismicity (see Reference [1]).
- do_N_test_S_test (=1) do N-test and S-test.

In case that someone need to modify the MCMC parameters, the following parameter in the function
"**sample_posterior_MCMC_script_ETAS_spatialModel_new_Blockwise.mat**" should be modified accordingly.

- maxIterations_chain1: Number of posterior samples to draw in chain 1, default =520
- maxIterations_ chain2_end: Number of posterior samples to draw in chain 2 to the last chian, default =1000
- burnin: Burnin period, default =20
- numChain: Number of Markov chains, default =6

The outputs are as follows:

(1) Posterior samples for ETAS model parameters are in the file "**samples.mat**"

(2) The file "**sampleN_mgrM.mat**" includes all the seismicity estimates in terms of the number of events equal to or greater than the magnitude levels defined in vecM ($M \geq m$). It has the dimension of (total number of grid)×(number of generated samples)×(length of vecM).

(3) In case that "do_N_test_S_test=1", the data in the file "**sampleN_mgrM.mat**" will be postprocessed in order to provide the mean, median (p50), 2nd (p02), 16th (p16), 84th (p84) and 98th (p98) percentiles of the forecasted number of events with magnitudes $M \geq m$ in two format:

   (a) Across the whole aftershock zone as:
   "[N_robust_mean,N_robust_p50,N_robust_p16,N_robust_p84,N_robust_p02,N_robust_p98]", which is proper for N-test.

   (b) Over each spatial grid cell for drawing the seismicity maps:
   [N_robust_mean_grid,N_robust_p50_grid,N_robust_p16_grid,N_robust_p84_grid,N_robust_p02_grid,N_robust_p98_grid]

(4) "P_mean" is the probability that an event exceeding a given magnitude level $M \geq m$ occurs within the aftershock zone in the forecasting interval (see Equation 12 in Reference [1])

(5) [Prob1_Stest,Prob_new_Stest] are the two probabilities of the S-test based on the standard, and the simulation-based methods respectively as presented in Reference [1].