MCMC-MTpy: A Python Package for Simultaneous Inversion of Source Location, Focal Mechanism and Rupture Directivity

Poster Time: .Pacific 3:45 PM - 4:45 PM on Thursday, April 22

Yin Fu, Wang Baoshan

School of Earth and Space Sciences, University of Science and Technology of China, Hefei, China

Key Points

- **We propose a new Bayesian inversion method MCMC-MTpy, which takes into account the uncertainty of source location and focal mechanism parameters.**
- **MCMC-MTpy takes phase (both P and S) travel times, first arrival P-wave polarity and the waveform as input to invert the source location, focal Mechanism, and rupture directivity iteratively.**
- **We introduce a weighting factor to measure the weight of waveform fitting error and travel time error, which can make the inversion process converge quickly.**

Motivation

The source parameters (Source Location, Focal Mechanism, and Rupture Directivity) are of key importance in seismic source studies and seismic hazard assessments. Generally, different source parameters are estimated separately. Separate inversion of source parameters may induce some inconsistence. And how to incorporate the uncertainty of model parameters into the inversion process is another problem we need to solve.

In this study, we proposed a new Bayesian inversion method MCMC-MTpy, which takes into account the uncertainty of source location and moment tensor parameters and it can invert the source location, focal Mechanism, and rupture directivity iteratively.

MCMC-MTpy: how to manage large amounts of Green Function data?

We used ASDF seismic data format to manage Green Function database, which is convenient to locally build up a database of pre-processed waveforms. ASDF greatly reduces the number of green functions files, and a single ASDF file can replace thousands of green function files which may threw us into confusion in the management process. We store the Green's functions corresponding to Event_ID (QuakeML) in "Waveforms" of ASDF data. In "Labels" of "Waveforms" , we store the phase (P and S) travel times which can be computed in advance by a given velocity model.

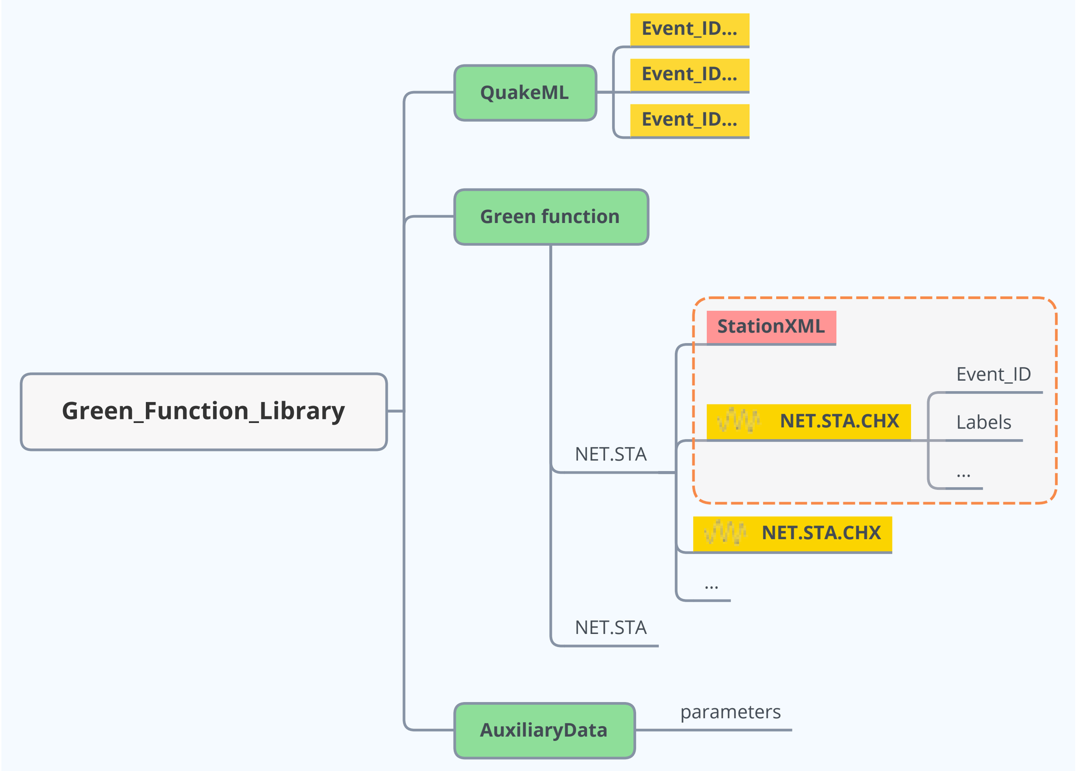

Fig. 1: The general structure of MCMC-MTpy stores Green's function with ASDF file format

Method

The Bayesian inference provides us a way to quantify uncertainties in the inverse problem by statistical inference. In the Bayesian formula m describes the model parameters. The solution to inverse problem is represented by posterior PDFs $\pi_{post}(m)$ when given observed data d_{obs} and the prior probability density $\pi_{prior}(m)$

$$\pi_{\text{post}}(\mathbf{m} \mid \mathbf{d}_{\text{obs}}) \propto \pi_{\text{prior}}(\mathbf{m})\pi_{\text{like}}(\mathbf{d}_{obs} \mid \mathbf{m}) \quad (1)$$

where π_{like} is the likelihood function describes the conditional probability of d_{obs} given m. F is the objective function, which can be related to π_{like} by the canonical distribution. F (formula 2), is composed of two parts, and the first part represents the error of phase arrival time (t) and the second part represents the error of waveform fitting (v). These two parts are connected by the weight factor a. Seismic waveforms are calculated by formulas 3, we have

$$F = a * \sum\left(t^{obs} - t^{syn}\right) + \sum\left(v^{obs} - v^{syn}\right) \quad (2)$$

$$V_i^n = \sum_{j=1}^{3}\sum_{k=1}^{3} m_{jk} G_{ij,k}^n(t) * s(t) \quad (3)$$

Synthetic Test

We test the Markov chain Monte Carlo method to source inversion with synthetic waveforms. In this test, we use fk method to calculate synthetic seismograms with 10 stations about 25 kilometers away the epicenter. The flow of our algorithm is illustrated in Fig.4. The misfit function decreases gradually with the increase of the sample number and finally converges showing in Fig.2.

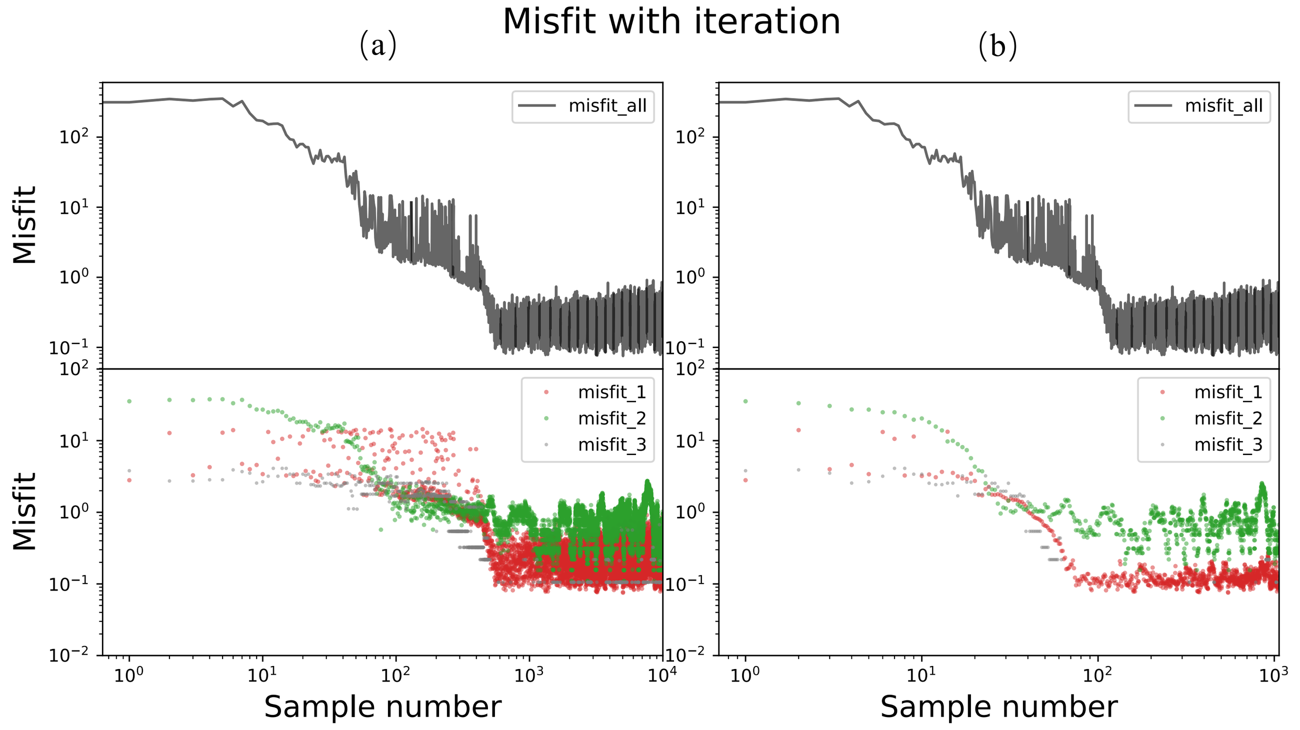

Fig. 2. (a) The misfit variation of generated models. (b) The misfit variation of accepted models. The gray line 'misfit_all' represents the overall misfit which is "F" in formula 2. The red line represents the phase time's misfit and the green line represents the waveform fitting's misfit. 'misfit_3' is the p-wave polarity's misfit.

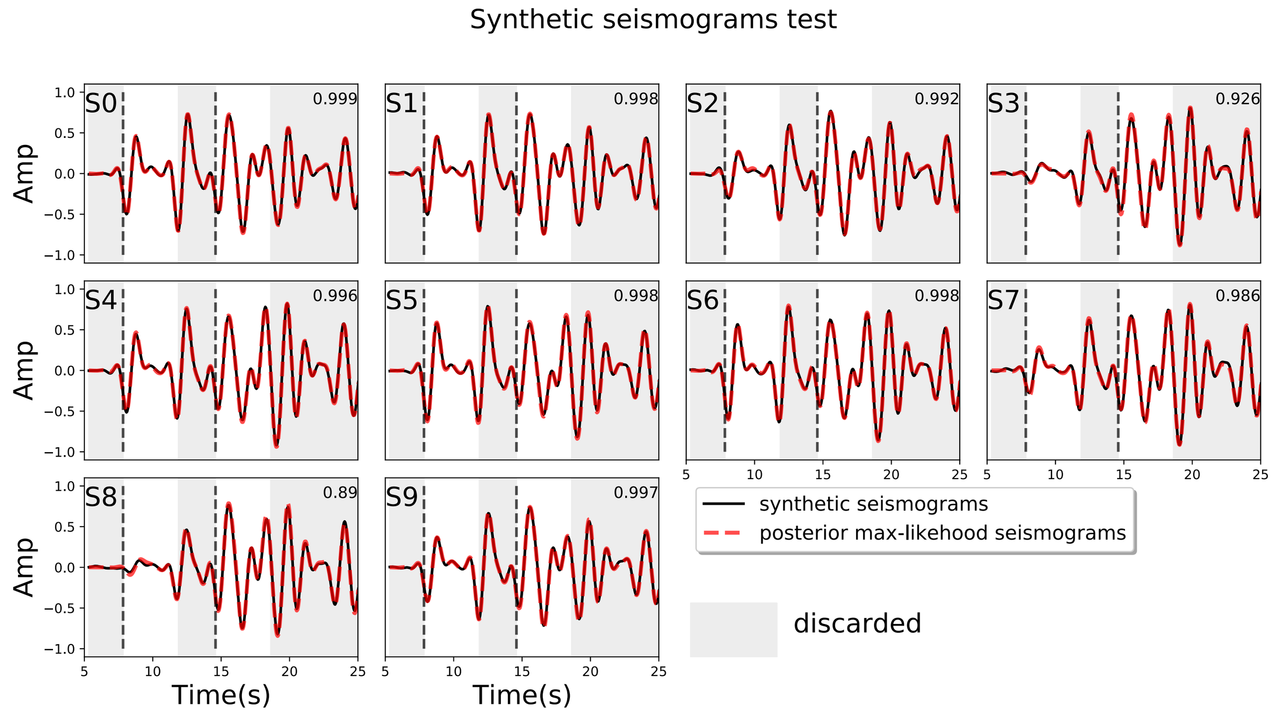

Fig. 3. Raw waveforms are plotted in black, synthetics for the posterior max-likelihood source in red dashed, The measurement windows used in sampling are marked in white. The number in the upper right corner represents the correlation coefficient between the waveforms.

```
Algorithm: MCMC-MTpy algorithm to sample
Define initial sample m_0
Compute π_post(m_0)
for i=1,..., samples do:
    m = m^{i-1}
    m* ~ randn()
    r = min( 1, π_post(m*) / π_post(m) )
    u ~ U[0,1]
    If u < r then:
        m^i = m*
    else:
        m^i = m
    end
end
```

Fig. 4. The algorithm in MCMC-MTpy.

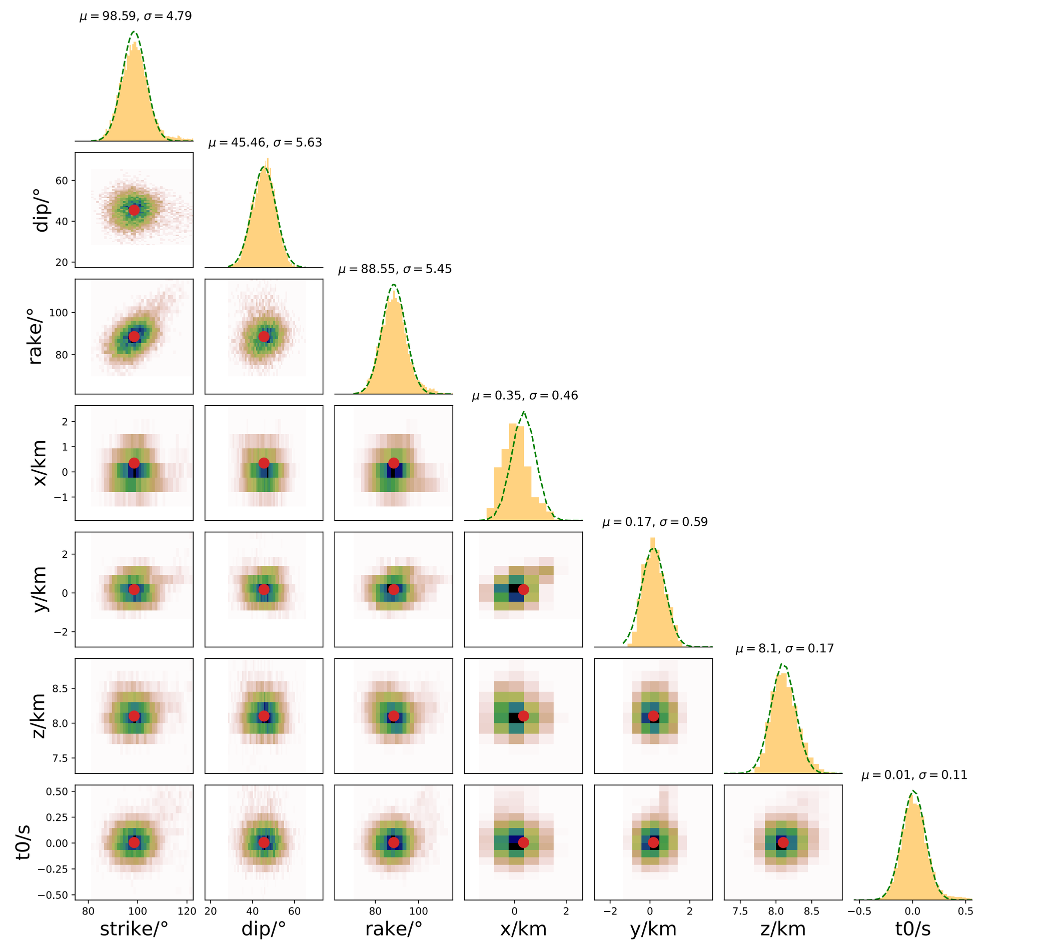

Fig. 5. The red dot in the 2D correlation maps mark the MAP solution. In the correlation subplots, blue colors are the high probability regions. Yellow histograms represents posterior distributions of the paraments, The posterior Gaussian distribution are shown for 10 thousand samples in green dotted line.

Fig. 3. shows a comparison of observed and raw synthetic seismograms. The results show that the waveform fitting correlation coefficient can reach 0.99, which because we added noise to the raw synthetic seismograms. Fig. 5. visualizes the posterior for the individual model parameters and the correlation between each two model parameters.

Discussion

- The phase arrival time misfit falls faster than the waveform fitting misfit. Because different weighting factors were used in different stage of the inversion. First the highest weight was assigned to the travel time objective function, which effectively constrains the source location. Then higher weight was put on the waveform and the polarity to estimate the focal mechanism.
- The new method only needs 100 iterations to find the optimal solution, but we still need a large number of samples to satisfy the stable distribution of model parameters.
- We did not test the rupture directivity in synthetic seismograms and we will test that in the following experiment. And we are confident that it will work in the actual data.

Conclusions

We propose a new Bayesian inversion method MCMC-MTpy, which work well in source inversion problem. The weighting factor measure the weight of waveform fitting error and travel time error, and it can make the inversion process converge quickly. We recommend using ASDF to manage Green's datebases, which is very convenient.

Reference:

[1]. Fichtner, A., & Simut˘e, S. (2018). Hamiltonian Monte Carlo inversion of seismic sources in complex media. Journal of Geophysical Research: Solid Earth, 123, 2984–2999. https://doi.org/10.1002/2017JB015249.

[2]. Lion Krischer, James Smith, Wenjie Lei, Matthieu Lefebvre, Youyi Ruan, Elliott Sales de Andrade, Norbert Podhorszki, Ebru Bozdağ and Jeroen Tromp (2016),An Adaptable Seismic Data Format,Geophysical Journal International, 207(2).http://dx.doi.org/10.1093/gji/ggw319.

[3]. Vasyura-Bathke, H.,J. Dettmer, A. Steinberg, S. Heimann,M. P. Isken, O. Zielke, P. M. Mai,H. Sudhaus, and S. Jónsson (2020). The Bayesian Earthquake Analysis Tool, Seismol. Res. Lett. 91, 1003–1018, doi: 10.1785/0220190075.

Wechat

Email: yinfu@mail.ustc.edu.cn