What do We Learn from the Exercise of the SIV BlindTest I?

Guangfu Shao & Chen Ji (UCSB)

Acknowledgement

Ralph Archuleta, Daniel Lavallee, and folks in the UCSB earthquake discussion group

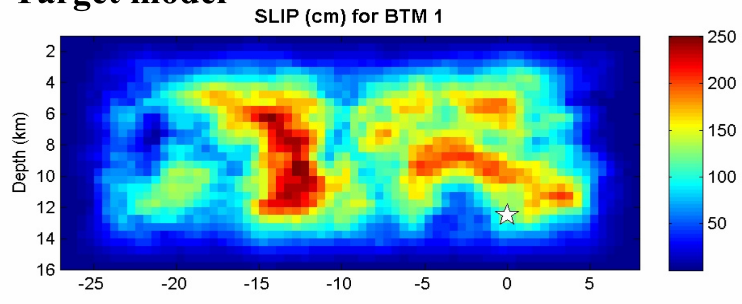
BlindTest I (Mai et al., 2007)

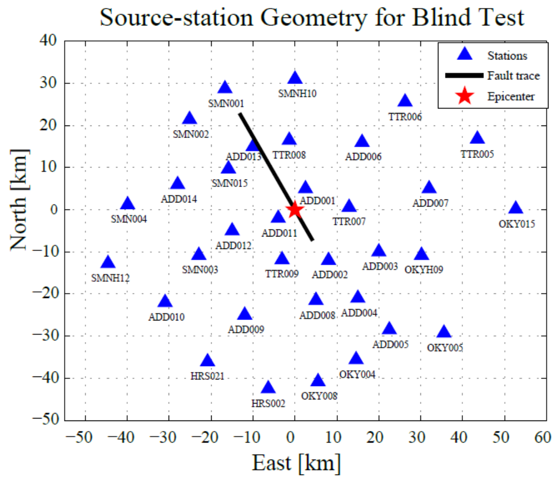
- **Data**
  1: Seismic data in velocity (*fmax* ~ 3 Hz)
  2: Static displacements

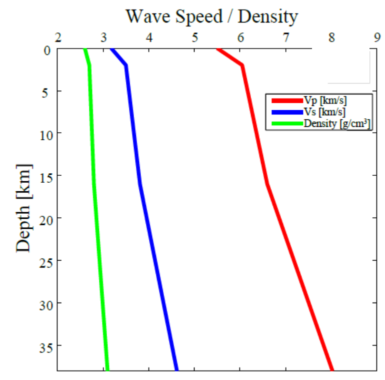
- **Available information**
  1: Fault geometry & Hypocentral location
  (strike, dip, rake: 150°, 90°, 0°)
  2: Total seismic moment:
  1.43×10^{26} dyne.cm
  3: Velocity structure
  4: Rupture does not break the surface

- **To be resolved:**
  1. Slip distribution on the fault plane
  2. Rupture velocity & rise time (both are constant; the investigators were given this information but not the values)

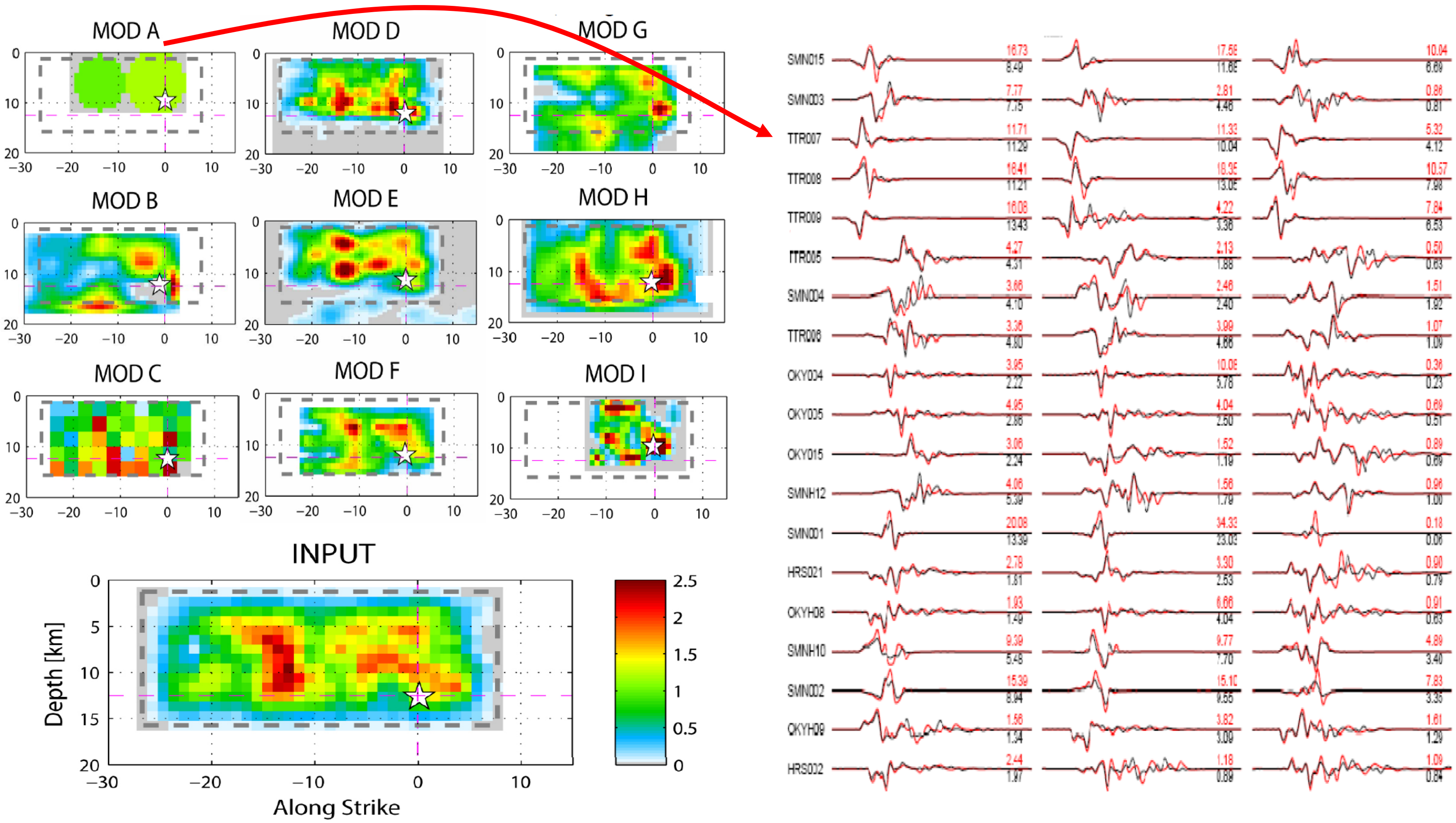
Target model





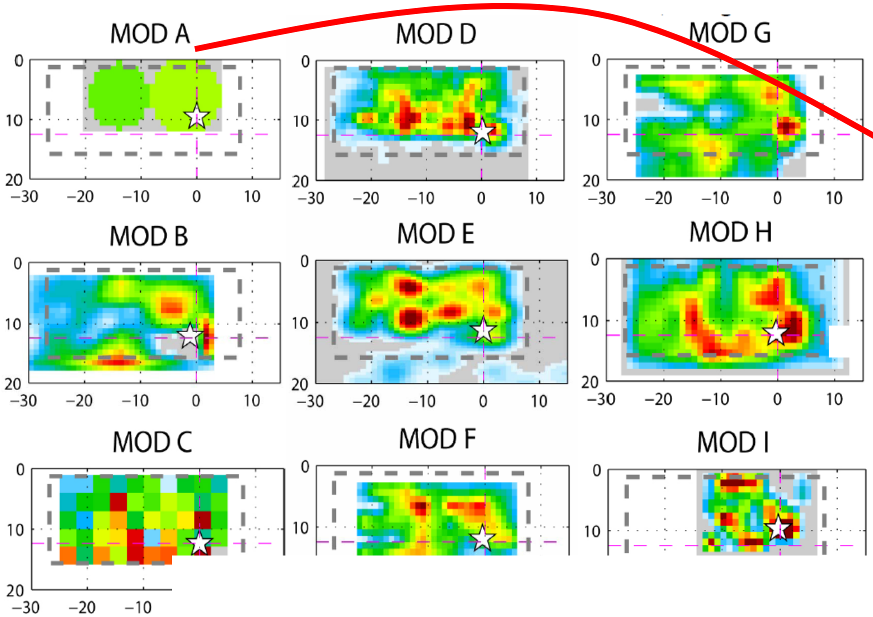


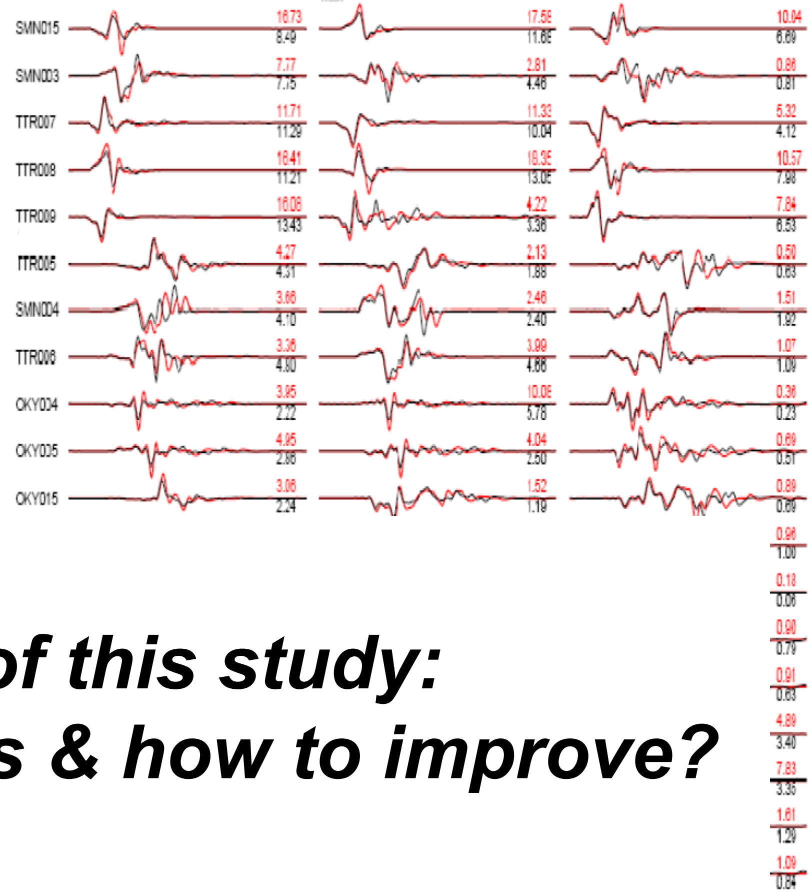
Previous results (Mai et al., 2007)

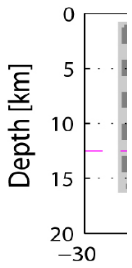


All the nine models provided reasonably good waveform fit but the correlations between the input and inverted slip models were not as good as we expected.

Previous results (Mai et al., 2007)



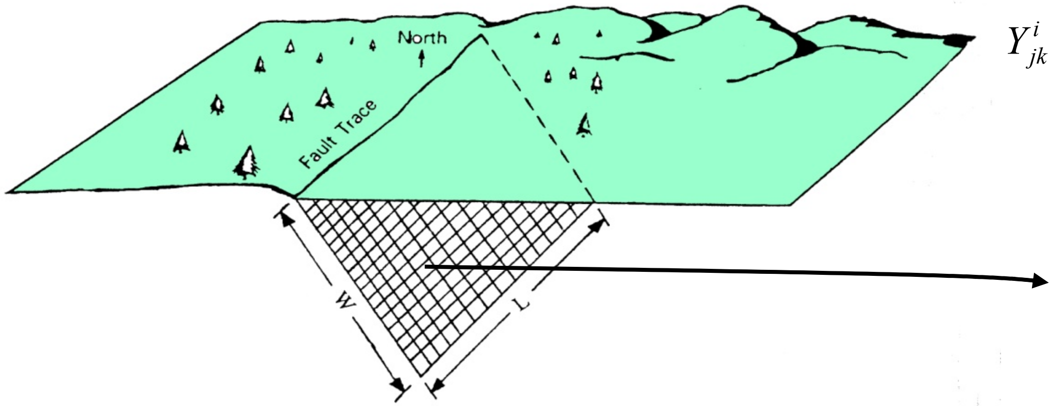




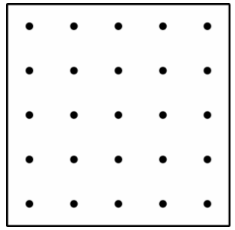
Motivation of this study:
What are the causes & how to improve?

"4 out of 9 inversion results are, statistically speaking, not better than a random model with somehow correlated slip " (Mai et al., 2007)

Finite fault approximation (Ji et al., 2003)



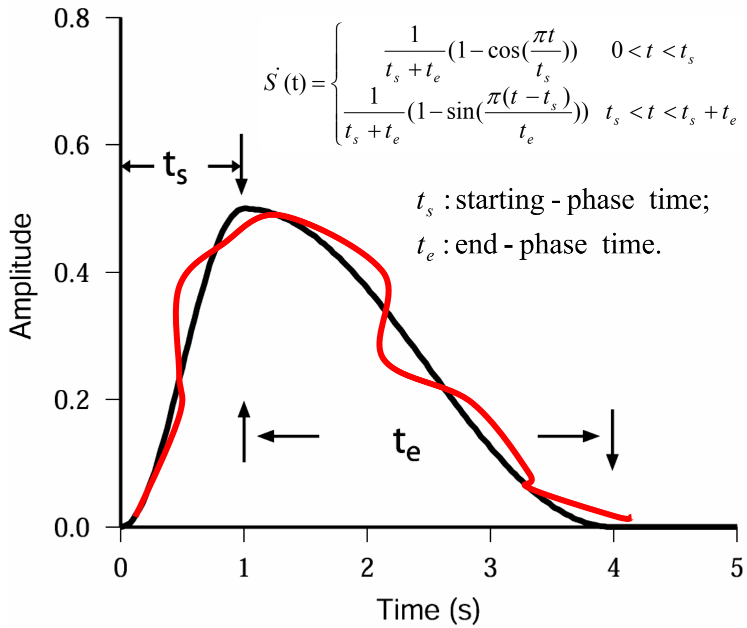
$$Y^i_{jk}(t,t',\vec{x}) = \sum_p G^i_{jk}(\vec{x}'_p,\vec{x},t) * \delta(t-\Delta t^p_{jk} - t')$$



$$u(t,\vec{x}) = \int_{-\infty}^{\infty} d\tau \iint_{\Sigma} [u_i(\xi,\tau)] c_{ijpq} \partial G_{np}(\vec{x},t-\tau;\xi,0)/\partial\xi_q \, d\Sigma$$

$$u(t,\vec{x}) \approx \sum_{j=1}^{n}\sum_{k=1}^{m} D_{jk}[\cos(\lambda_{jk}) Y^1_{jk}(t,t',\vec{x}) + \sin(\lambda_{jk}) Y^2_{jk}(t,t',x)] * \dot{S}_{jk}(t)$$

D_{jk}	Slip amplitude
λ_{jk}	Rake angle
$\dot{S}_{jk}(t)$	Derivative rise time function
t'	Rupture initiation time
$Y^i_{jk}(t,t',\vec{x})$	Subfault Green's functions



Rise time function S(t)

Geophysical Inversion 101

$$\vec{\mathrm{M}} \text{ is the vector of unknowns;}$$

$$\vec{D} \text{ is the vector of observations}$$

$$\vec{D}_{predicted} = G(\vec{M})$$

$$\text{An inversion is searching for :}$$

$$\vec{M}_{\text{target}} : \min(\vec{D}_{predicted} - \vec{D}_{observed});$$

How to improve a geophysical inversion?

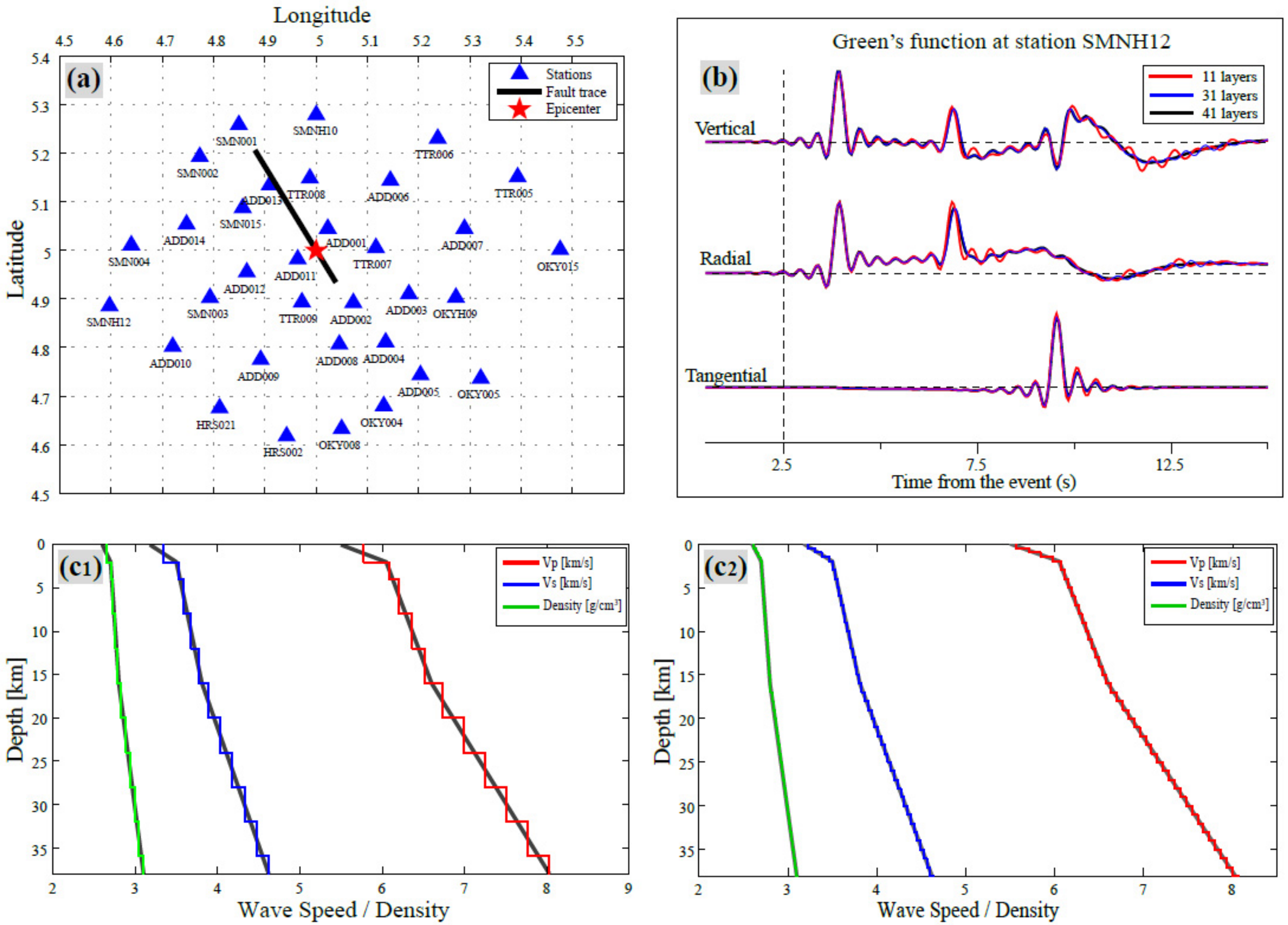
1. Increasing the number of independent observations
2. Improving the criteria for the minimization.
   i.e, minimum (Syn-Obs) == minimum (M_invert-M_target) ?
3. Reducing the number of unknowns
4. Avoiding the local minima

Pre-processes: Quality Control

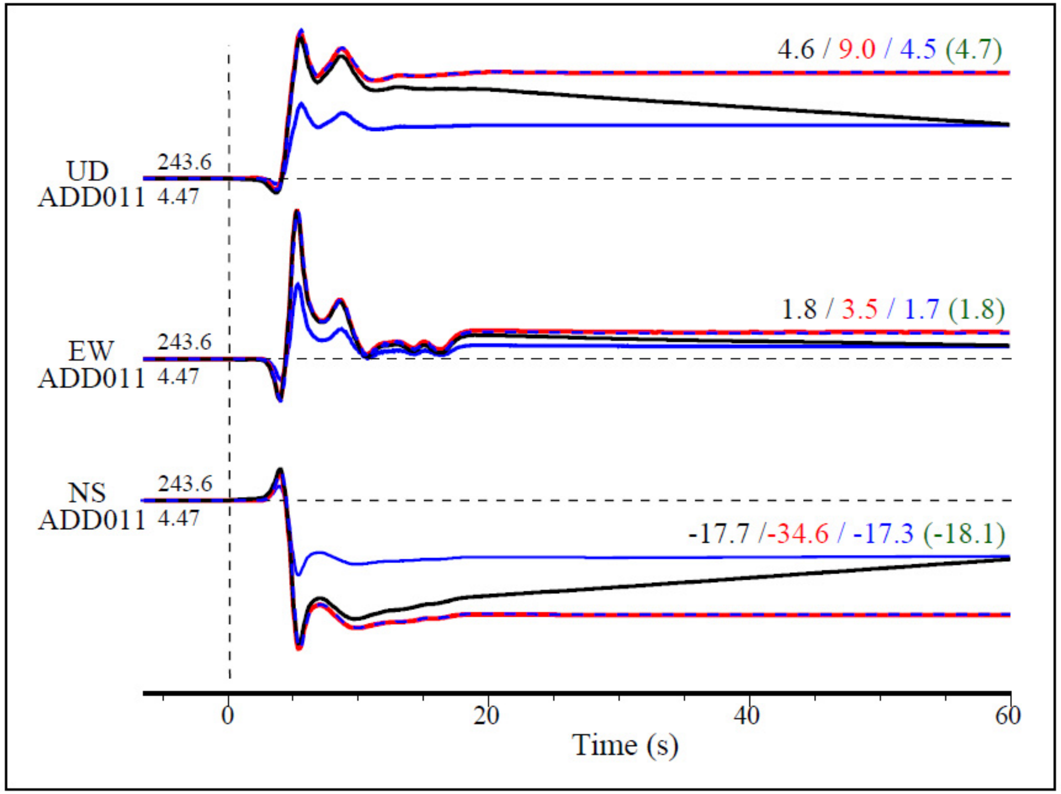
Is this
"garbage in and garbage out ?"

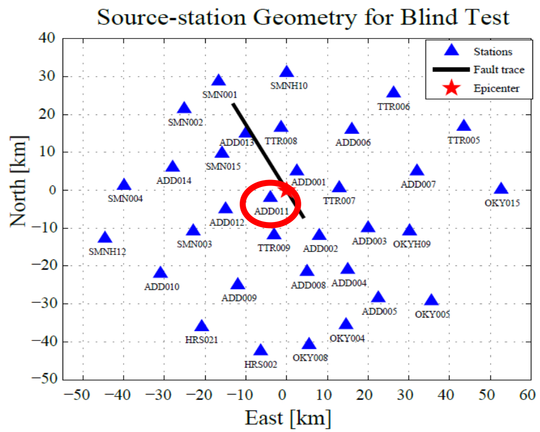
Comment from a respect seismologist

Preprocesses 1: fault & synthetics



Preprocesses 2: Data correction





Data

Corrected Data

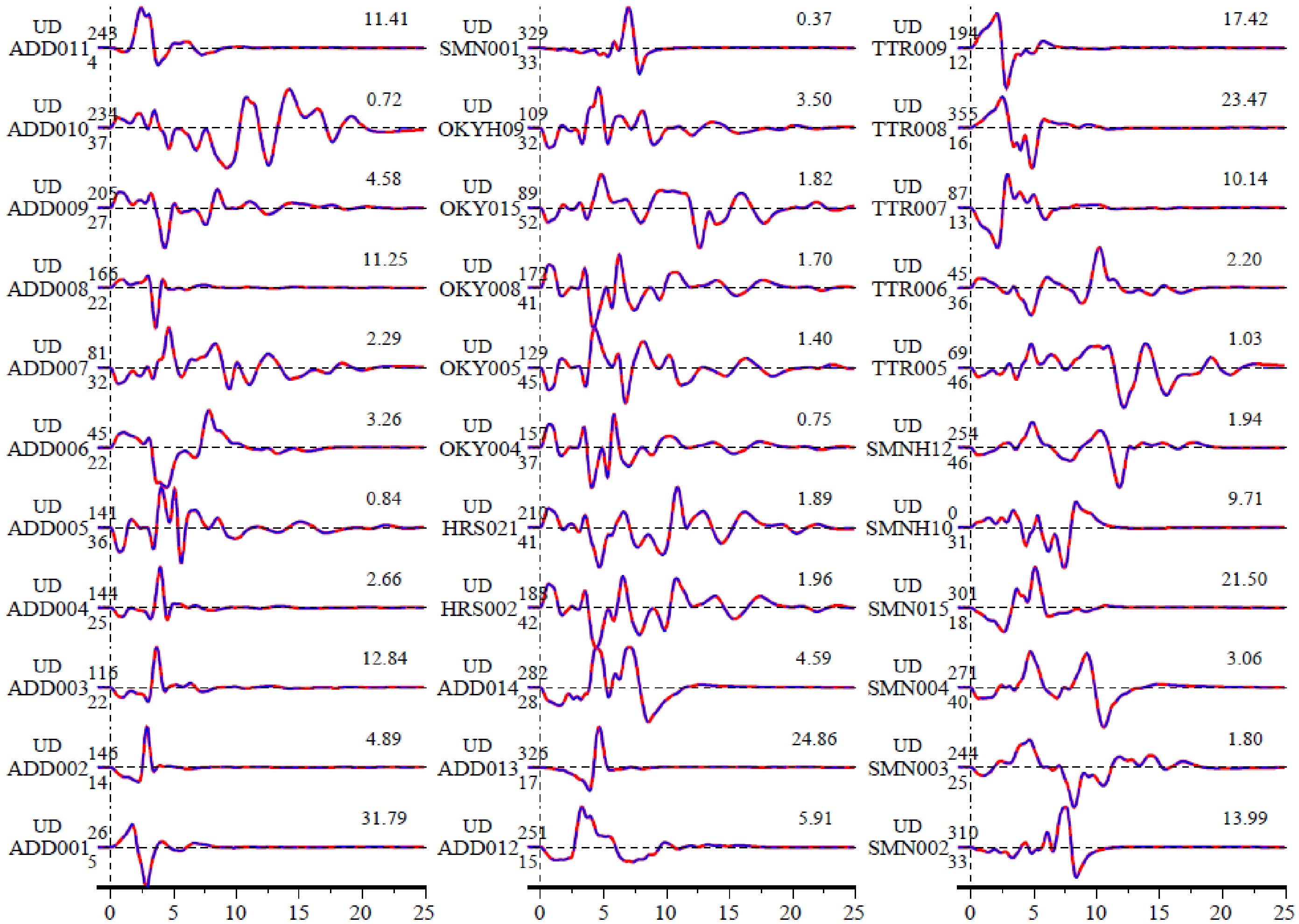
Our synthetics

Double of Our synthetics

Note:

We have corrected the constant offsets in synthetic data and double the **Target** slip as the new “**Target**” model

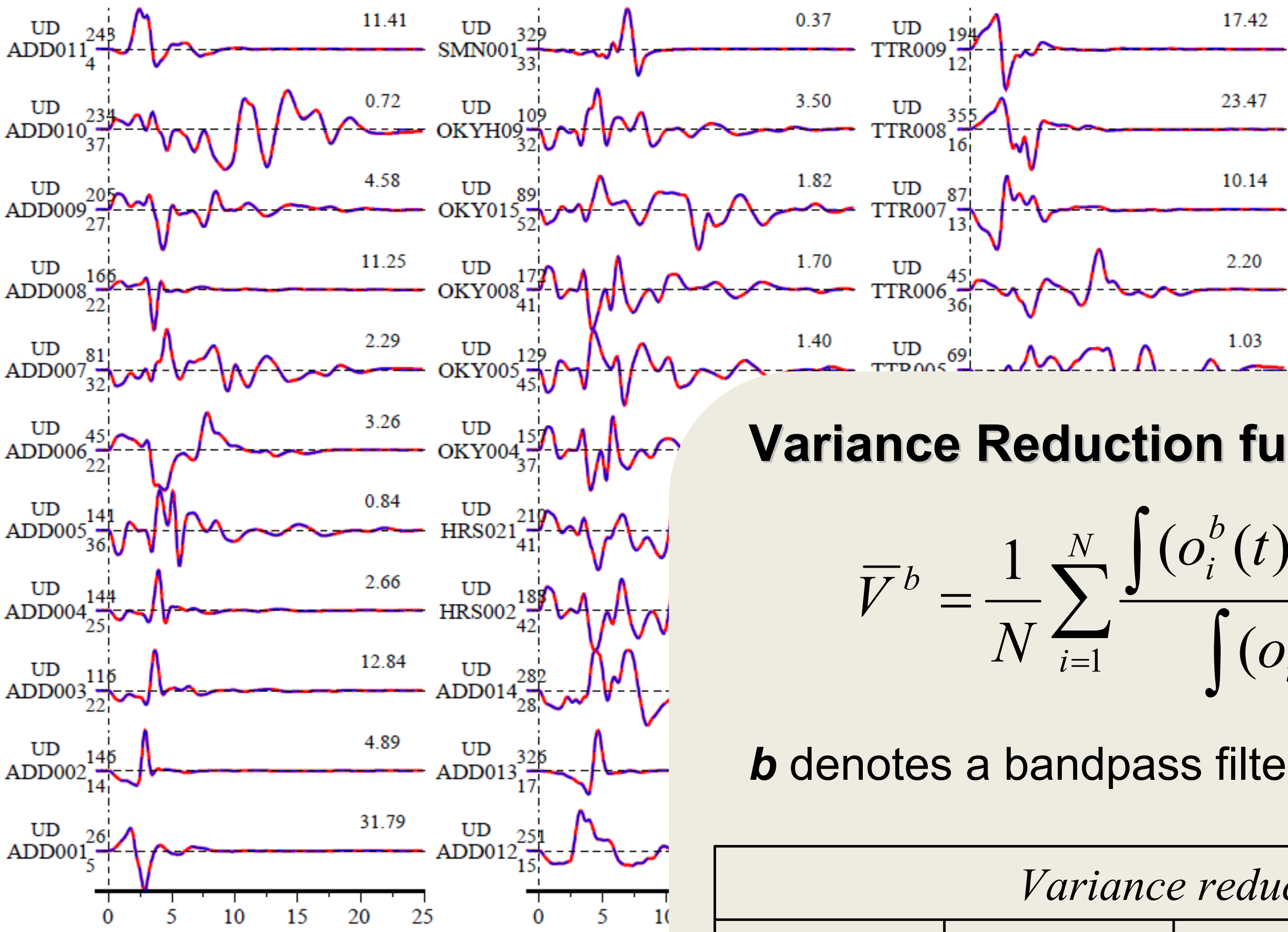
Comparison of data (red) and synthetics (dashed blue)



Vertical components

Visually indistinguishable !

Comparison of data (red) and synthetics (dashed blue)



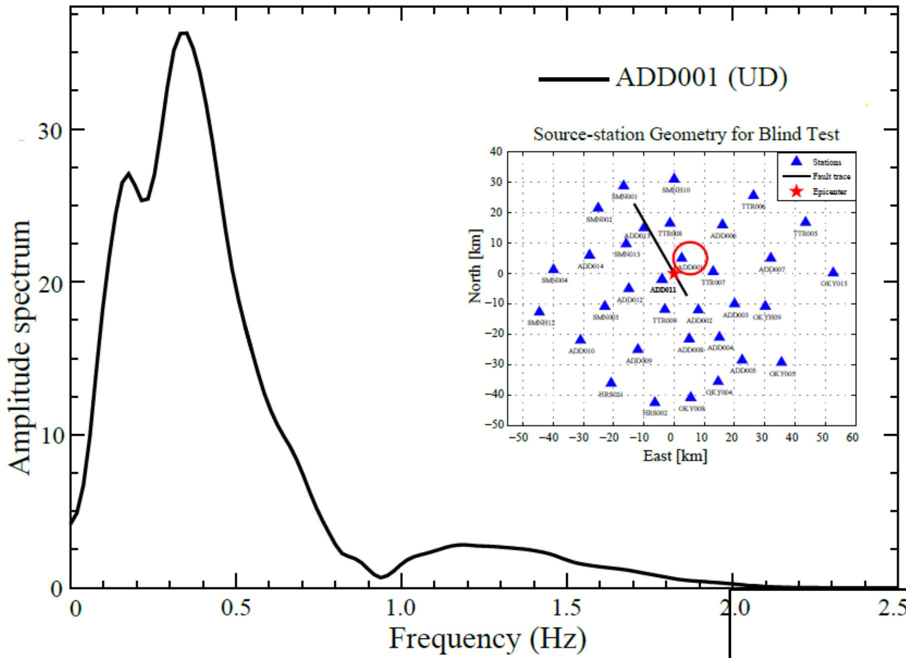
Variance Reduction function

$$\overline{V}^{b} = \frac{1}{N}\sum_{i=1}^{N}\frac{\int (o_i^b(t) - s_i^b(t))^2 dt}{\int (o_i^b(t))^2 dt}$$

b denotes a bandpass filter

Variance reductions			
0-2.0 (Hz)	*0-0.1 (Hz)*	*0.1-1.0 (Hz)*	*1.0-2.0 (Hz)*
99.91%	**99.98%**	**99.92%**	**97.53%**

Spectrum: Energy Ratio



Average relative energy

$$\overline{R}^{b} = \frac{100}{N} \sum_{i=1}^{N} \frac{\int (o_i^b(t))^2 dt}{\int (o_i(t))^2 dt}$$

Relative Energy			
0-2.0 (Hz)	*0-0.1 (Hz)*	*0.1-1.0 (Hz)*	*1.0-2.0 (Hz)*
100%	15.04%	86.02%	2.73%

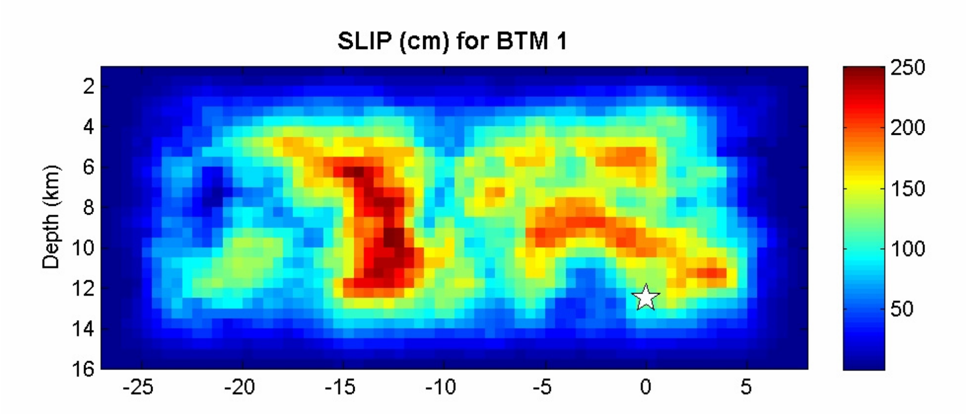
Note:

Misfit functions, such as variance reduction, are designed to catch the difference in amplitude (or energy). Therefore, for our case, it is dominated by the signals from 0.1 to 1 Hz.

Inversion setup ➔ model space (*M*)

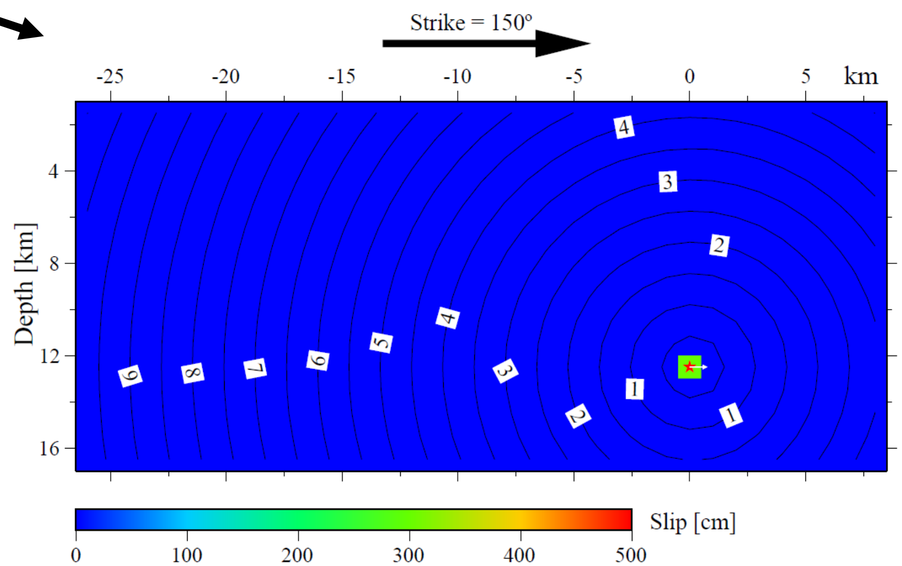
Origin Target model:

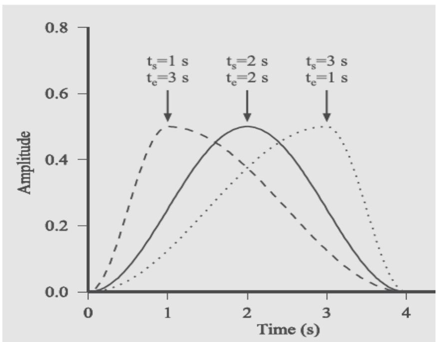
Fault dimensions: 34km by14.5km
Grid size: 0.5km by 0.5km
Rise time: 0.8 sec (symmetric triangle)



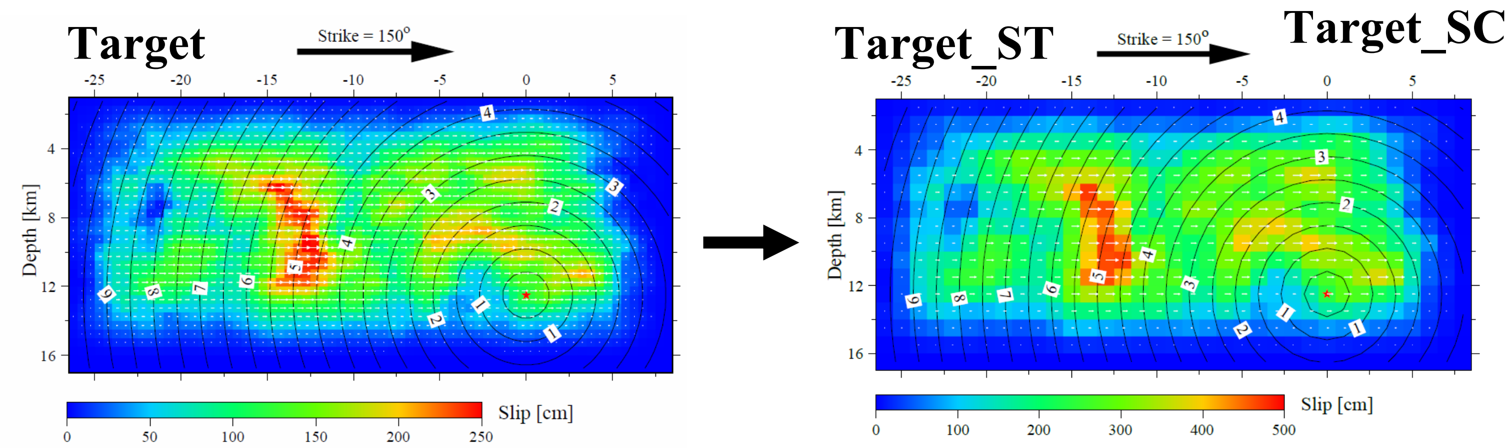
Inverted Models:

Fault dimension: 35km by 16km
Subfault size : 1km by 1km
Rupture velocity: 2.65 – 2.75 km/s
Rise time: starting time: 0.1 s -0.8 s
ending time: 0.1 s -0.8 s





What is the best model we expect to get?



Inside the model space ***M*** defined previously**,**

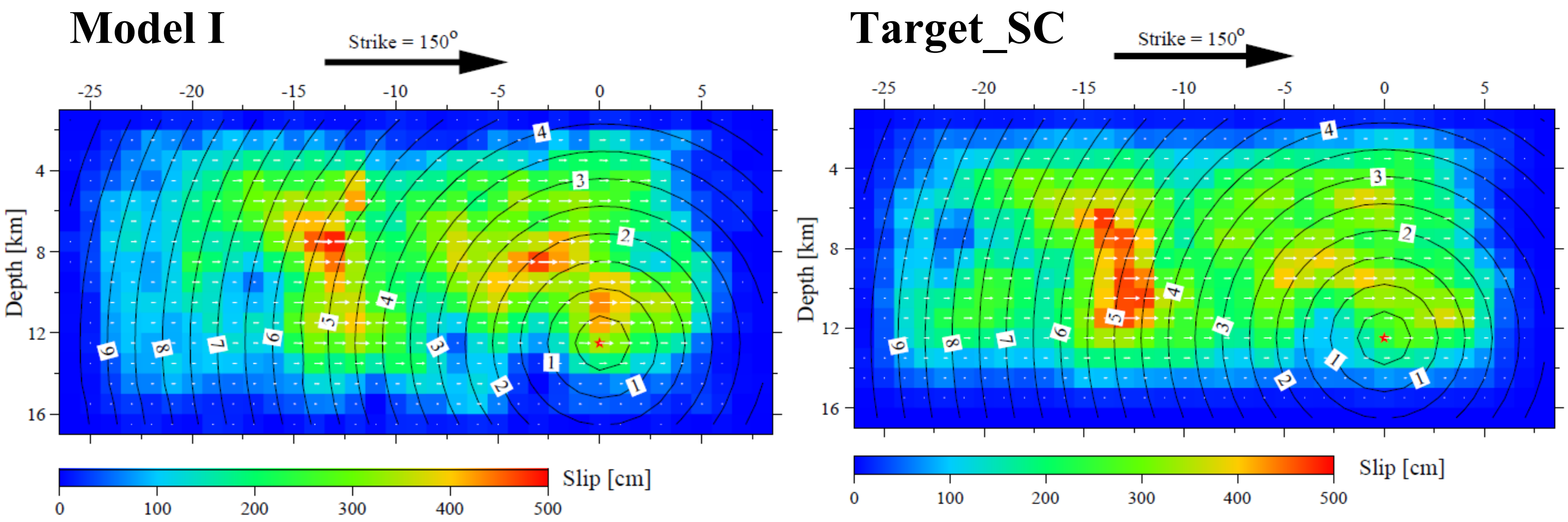
Target_SC is the model which best approximates the **Target** model, but whether it's also the model which matches the data best?

Target_ST	99.32%	99.72%	99.42%	86.12%
Target_SC	99.17%	99.71%	99.40%	80.29%

Model **Target_ST**: averaging the slip of **Target's** model into 1 X 1 km subfaults

Model **Target_SC** : further modified from the model **Target_ST** by replacing the triangle slip rate function with a symmetric cosine function

Q1: Can we reconstruct the rupture process?

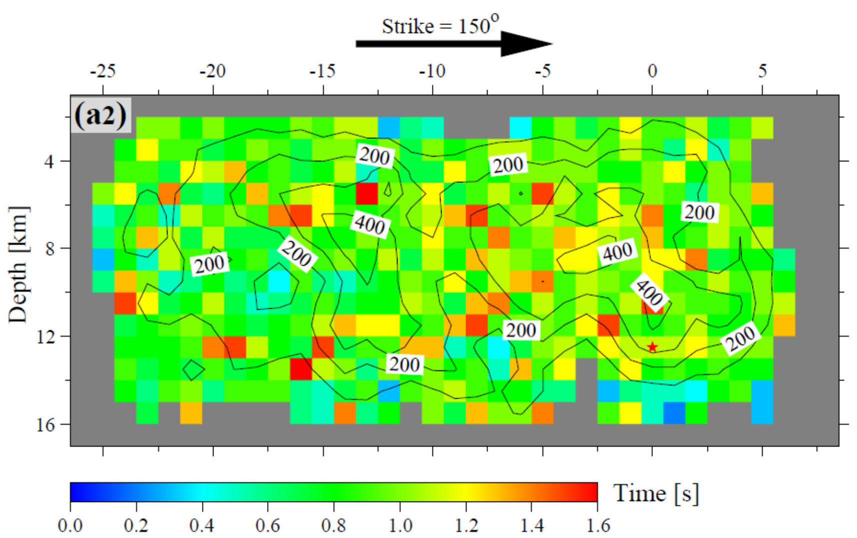


Slip Distribution

	Model I		**Target_SC**
Total moment:	2.72×10^{26} dyne.cm	vs.	2.86×10^{26} dyne.cm
Peak slip:	4.8 m	vs.	4.7 m

Q1: Can we reconstruct the rupture process?

Model I : Temporal variation



Rise time distribution (slip > 25 cm)

Good but definitely not prefect !!!

	Model I		Target_SC
Average starting time:	0.38 s	vs.	0.4 s
Average ending time:	0.50 s	vs.	0.4 s
Average rise time:	0.88 s	vs.	0.8 s

Can we further improve the result?

Models	*Variance reductions*				
	0-2.0 (Hz)	*0-0.1 (Hz)*	*0.1-0.5 (Hz)*	*0.5-1.0 (Hz)*	*1.0-2.0 (Hz)*
Target	99.91%	99.98%	99.94%	99.81%	97.53%
Target_SC	99.17%	99.71%	99.60%	98.67%	80.29%
Model I	99.35%	99.28%	99.71%	99.20%	77.02%

Statement I

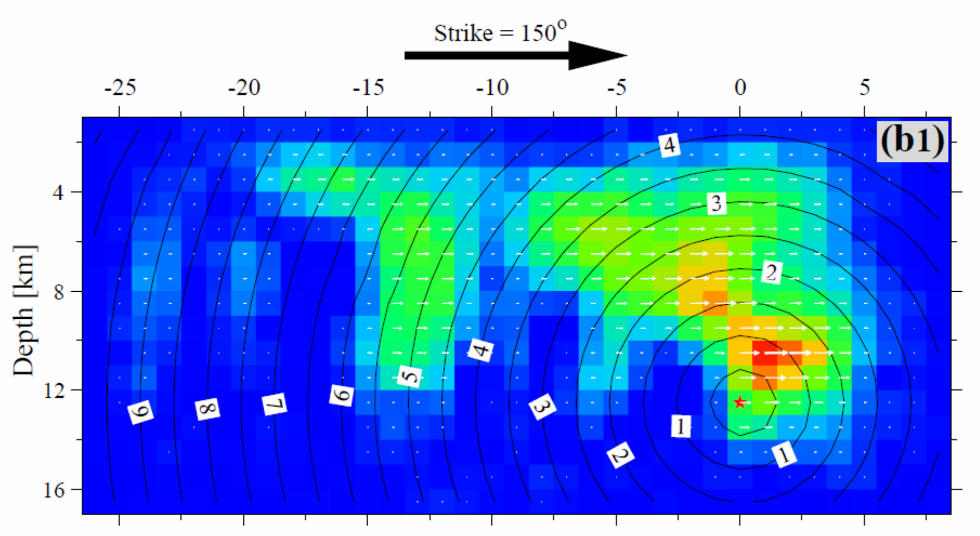
Inside the model space ***M*** defined by our source representation, **Target_SC** is the model which is most close to the **Target** model, but it is **NOT** the model which fits the data best in **a term of variance reduction**. Therefore**, the model cannot be further improved unless we use a different objective function.**

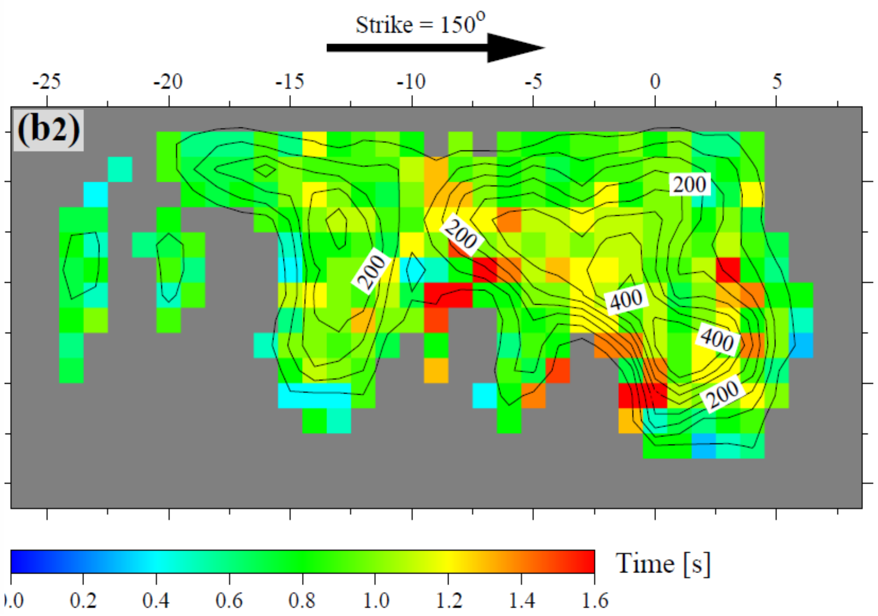
Two Restricted inversion exercises

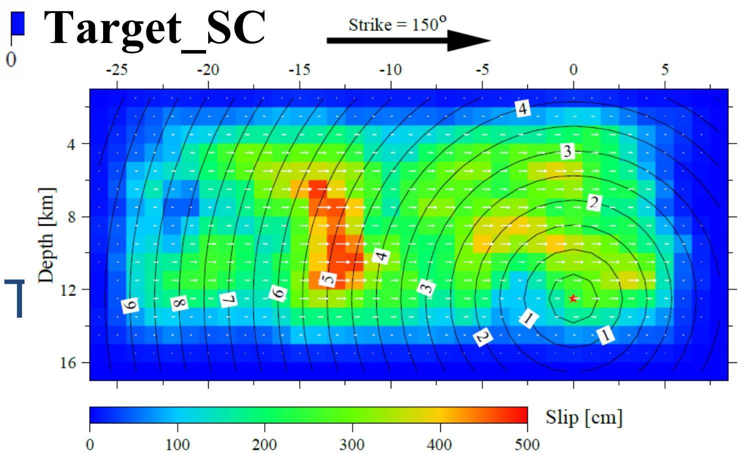
In a realistic world, it is impossible to achieve 99% variance reduction. So we further test whether the seismic data are sensitive to the following two parameters

1) **Model II**: Total seismic moment is only a **HALF** of the Target model.
2) **Model III**: Peak slip is only a **HALF** of the Target Model

Model II: (Total moment is only a half of the Target)







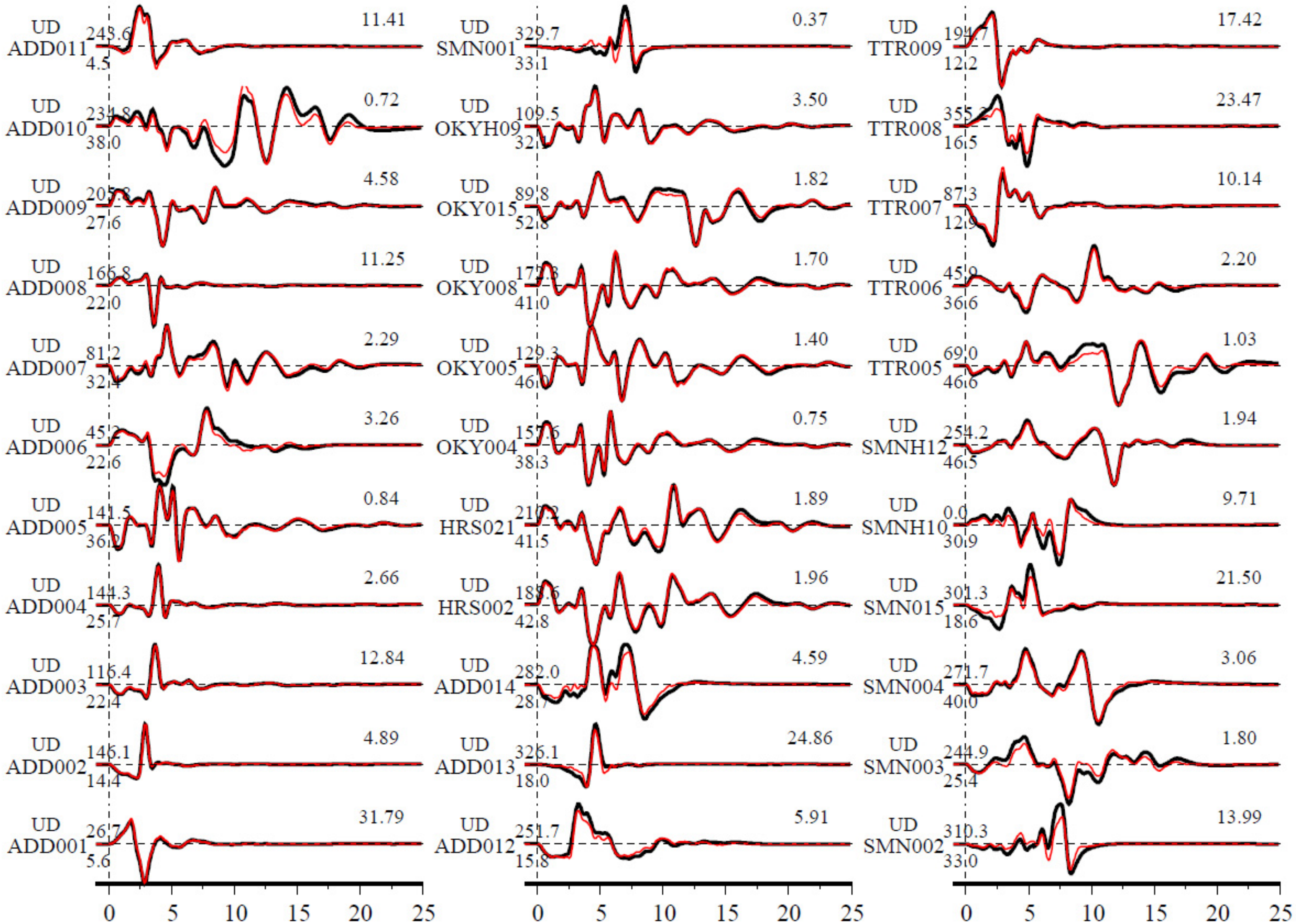
Rise time distribution (slip > 25 cm)

Average starting time: 0.38 s

Average ending time: 0.44 s

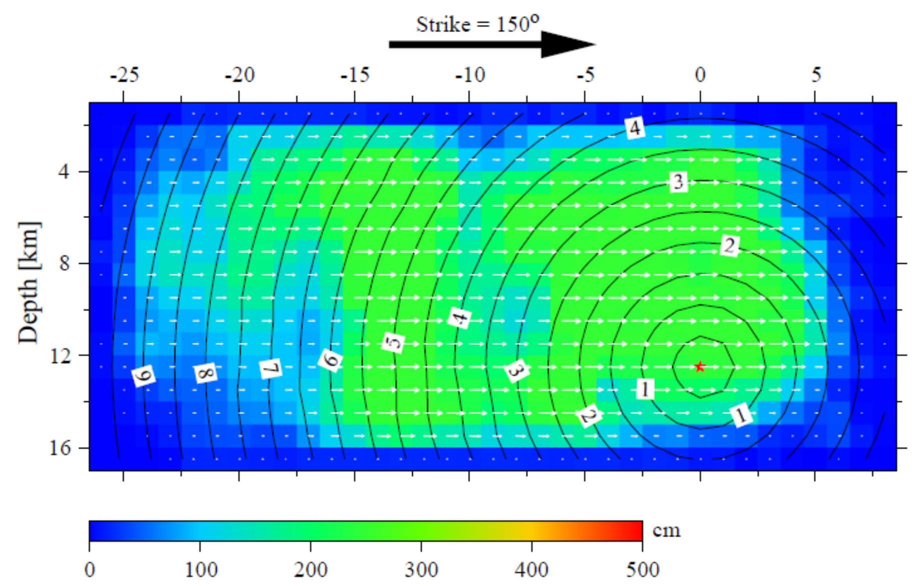
Average rise time: 0.82 s

Waveform fit (vertical components)

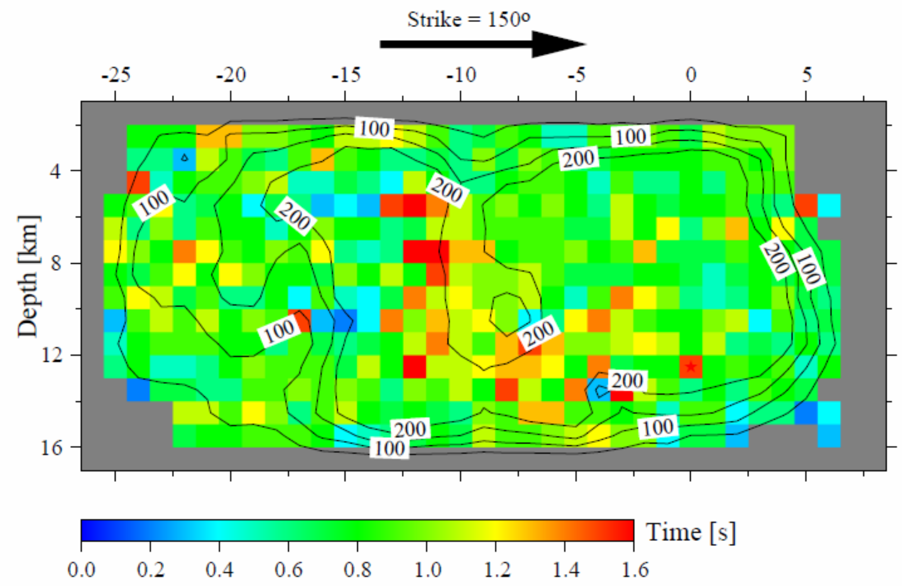


The average variance reduction of Model II is 93.15%

Model III: (Peak slip is only a half of the Target Model)



Slip Distribution



Rise time distribution (slip> 25 cm)

Total moment: 2.71×10^{26} dyne.cm

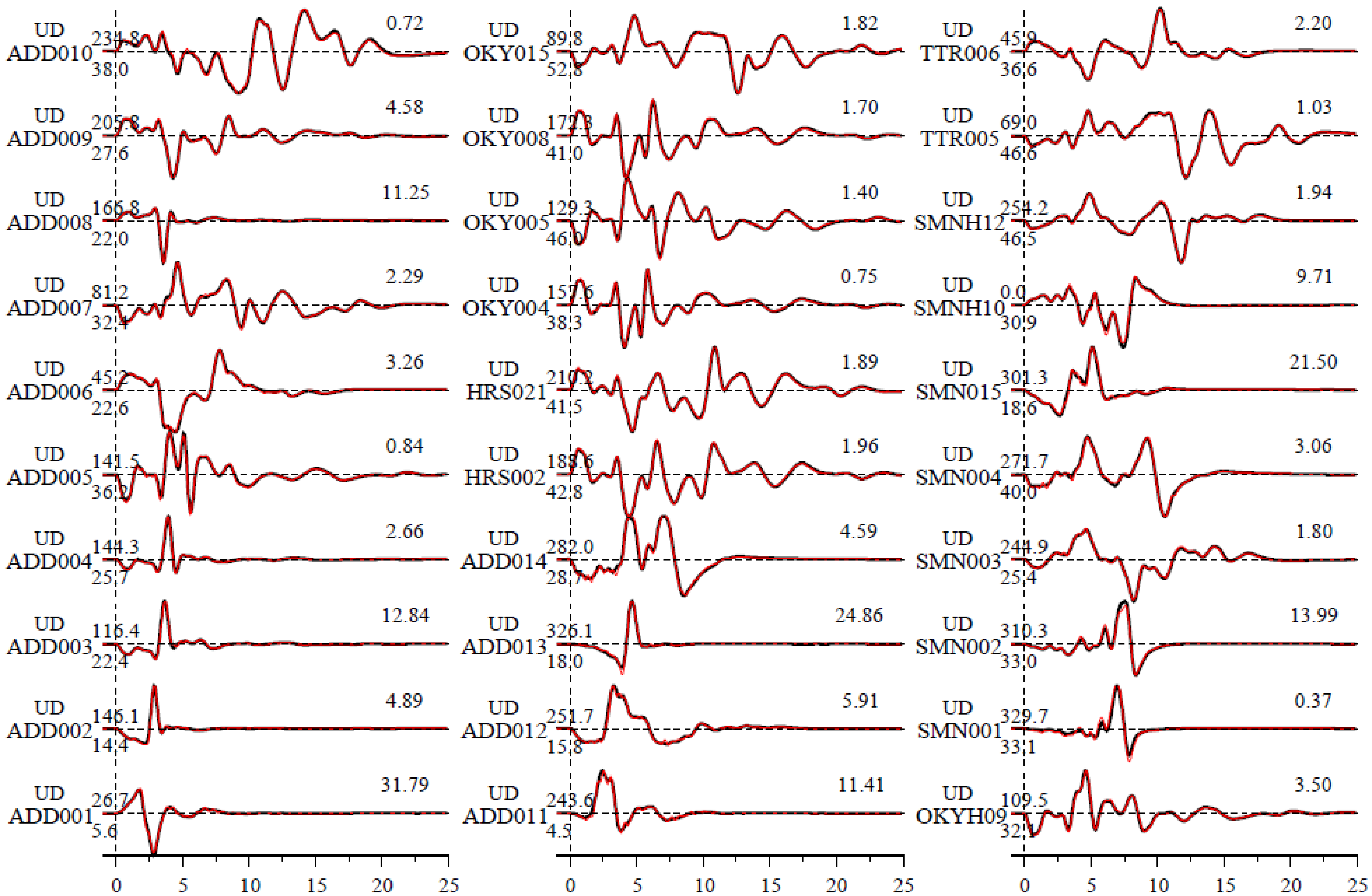
Peak slip: 2.5 m

Average starting time: 0.36 s

Average ending time: 0.46 s

Average rise time: 0.82 s

Waveform fit (vertical components)



The average variance reduction of Model III is 98.97%

Model II associates with a variance reduction of 93%. Then can we say it explains data well?

Models	*Variance reductions*			
	0-2.0 (Hz)	*0-0.1 (Hz)*	*0.1-1.0 (Hz)*	*1.0-2.0 (Hz)*
Target	99.91%	99.98%	99.92%	97.53%
Target_SC	99.17%	99.71%	99.40%	80.29%
Model I	99.35%	99.28%	99.61%	77.02%
Model II	93.15%	76.86%	**95.15%**	86.36%

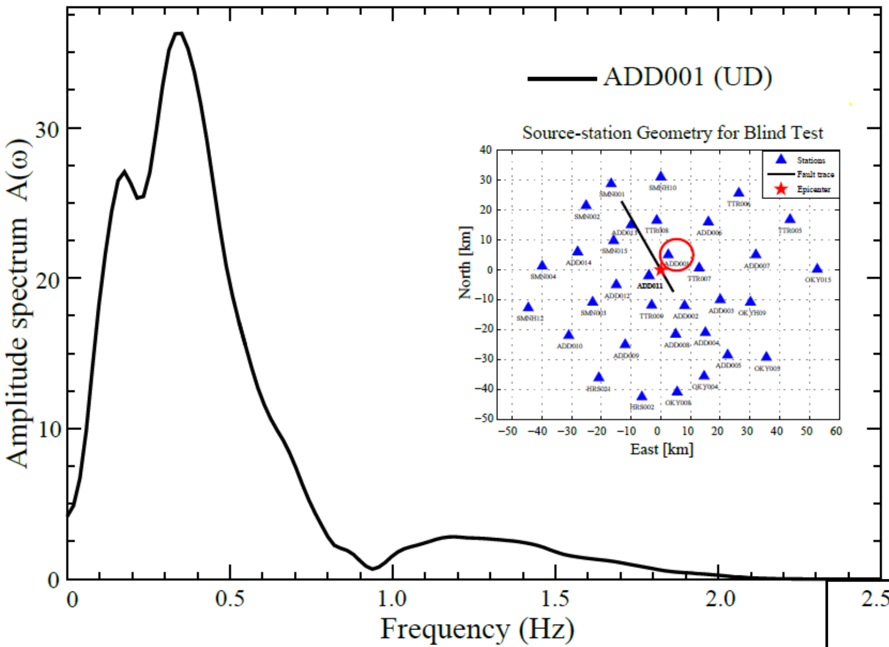
- **Model II** explains signals from 0.1 to 1 Hz well with a variance reduction around 95% and to the higher frequency bands from 1 to 2 Hz, this model matches the data better than **Model I** .
- **However, the model cannot well explain the data from 0 to 0.1 Hz.**

Model III associates with a variance reduction of 99%. Then can we say it explains data well?

Models	*Variance reductions*			
	0-2.0 (Hz)	*0-0.1 (Hz)*	*0.1-1.0 (Hz)*	*1.0-2.0 (Hz)*
Target	99.91%	99.98%	99.92%	97.53%
Target_SC	99.17%	99.71%	99.40%	80.29%
Model I	99.35%	99.28%	99.61%	77.02%
Model III	98.97%	98.90%	99.36%	64.26%

- **Model III** explains signals from 0 to 1 Hz well with a variance reduction around 99%
- **but it cannot match the signals from 1 to 2 Hz.**

Spectrum: Energy Ratio



Average relative energy

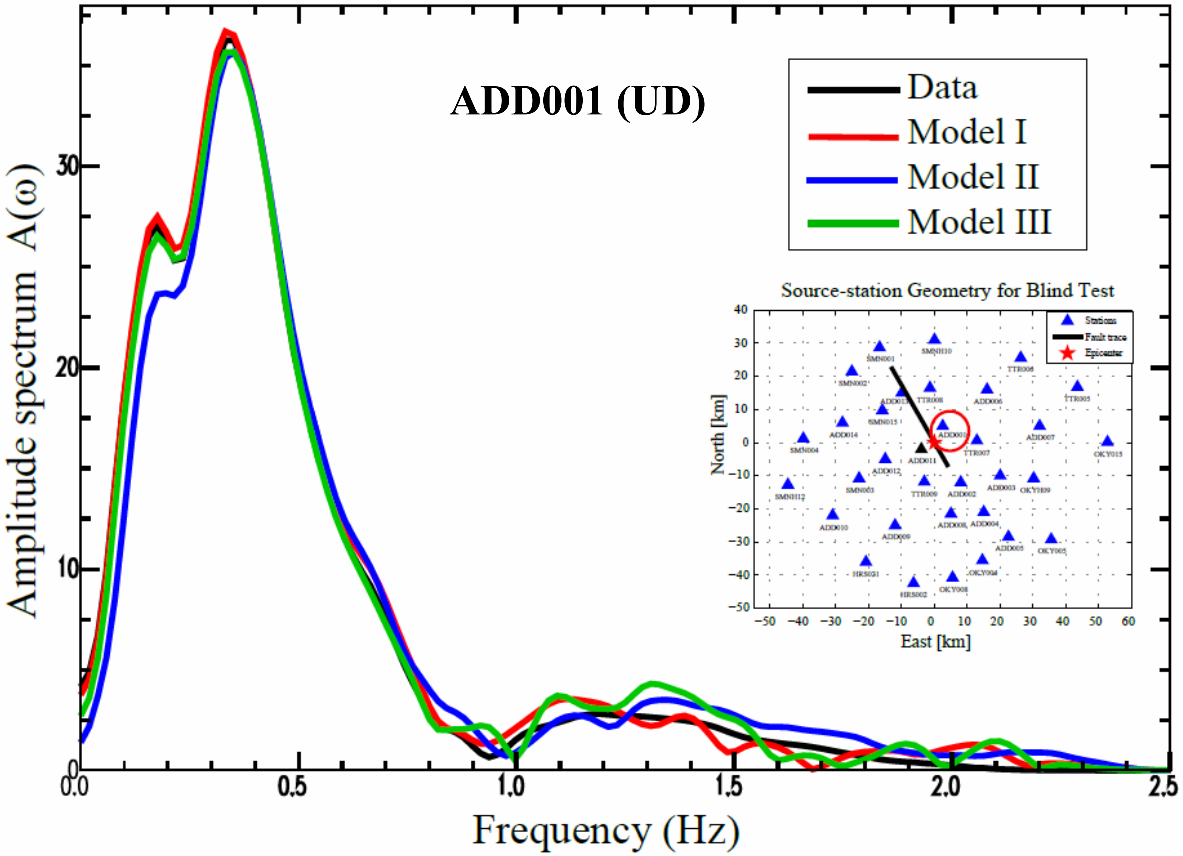
$$\overline{R}^b = \frac{100}{N}\sum_{i=1}^{N}\frac{\int (o_i^b(t))^2 dt}{\int (o_i(t))^2 dt}$$

Relative Energy			
0-2.0 (Hz)	*0-0.1 (Hz)*	*0.1-1.0 (Hz)*	*1.0-2.0 (Hz)*
100%	15.04%	86.02%	2.73%

Note:

Misfit functions, such as variance reduction, are designed to catch the difference in amplitude (or energy). Therefore, for our case, it is dominated by the signals from 0.1 to 1 Hz.

Let's compare the data in the frequency domain



Model I: Best model

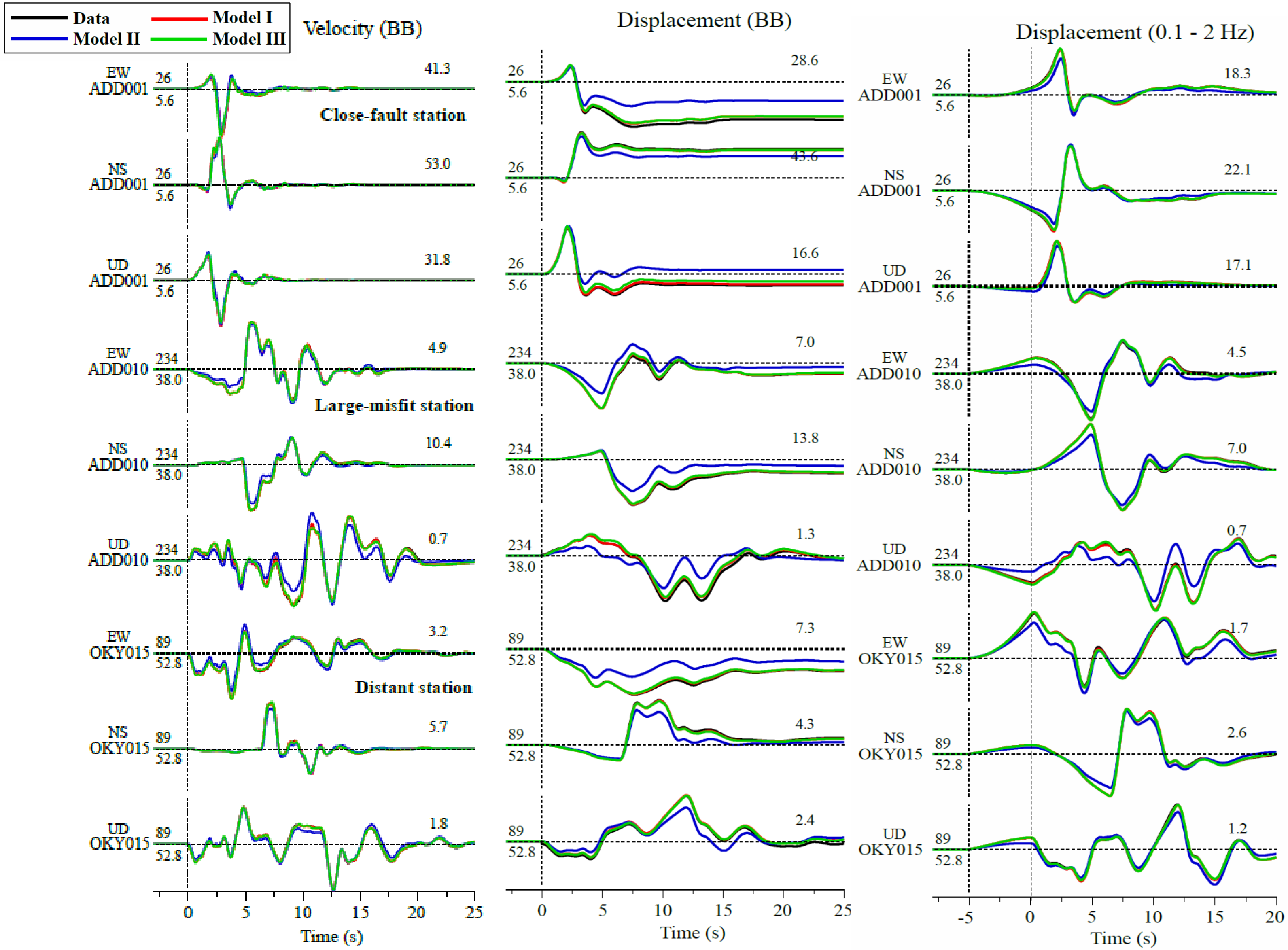
Model II: 0.5 * Mo

Model III: (slip<=2.5m)

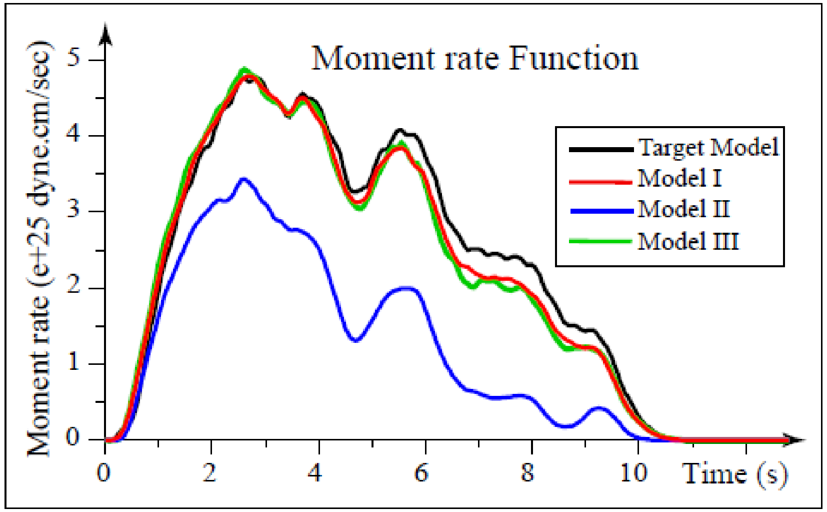
Statement II:

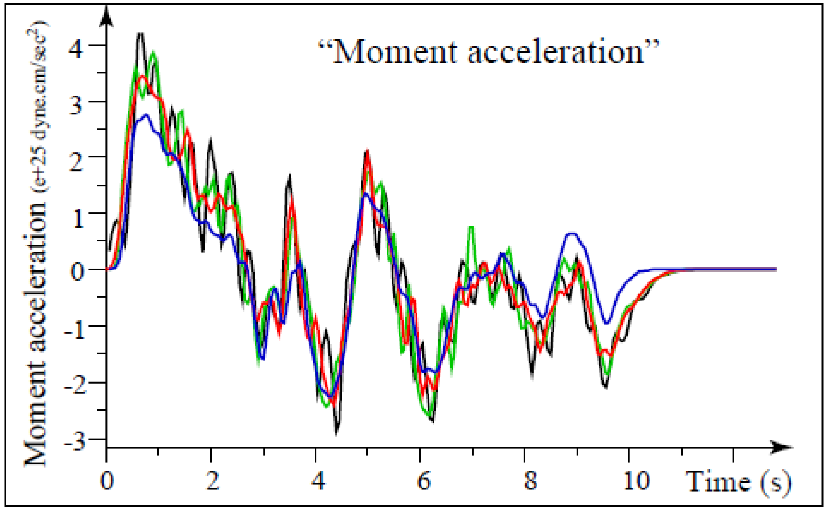
Most misfit functions used before, such as variance reduction, are sensitive more to the energy of signals rather than the amount of independent information.

Waveform comparison



Comparison of moment rate functions



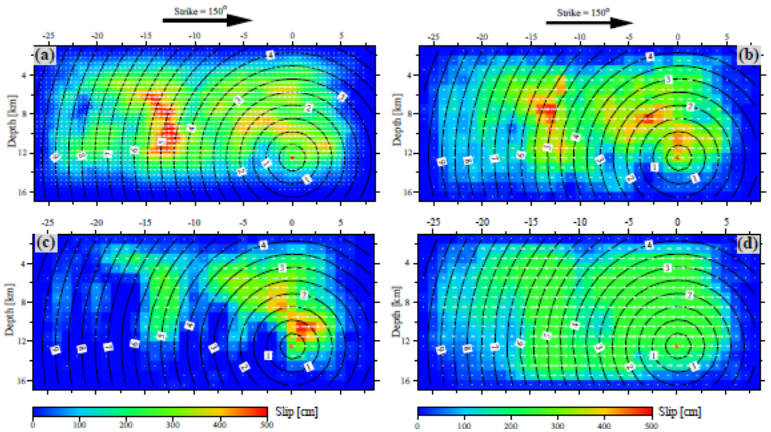


Far field body-wave

1: Displacement $$U(\vec{r},t) \approx \frac{1}{4\pi\rho v^3}\psi(\theta,\phi)\frac{1}{r}\dot{M}(t)$$

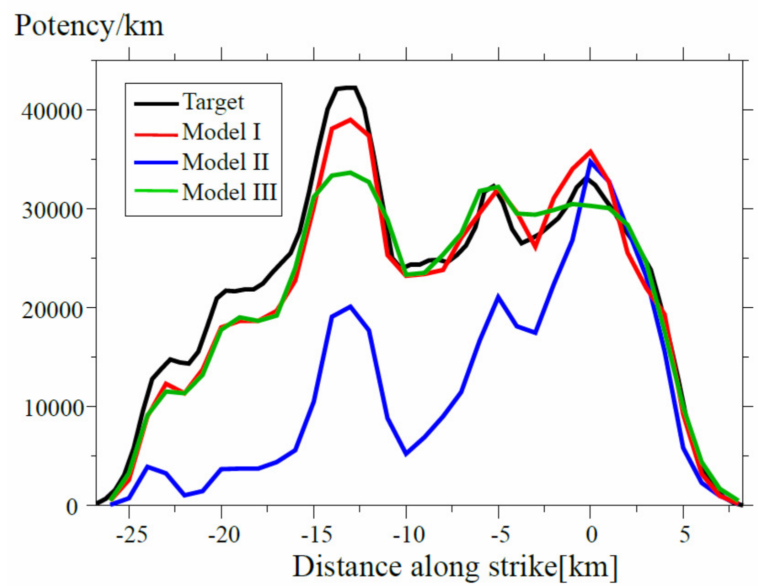
2: Velocity $$V(\vec{r},t) \approx \frac{1}{4\pi\rho v^3}\psi(\theta,\phi)\frac{1}{r}\ddot{M}(t)$$

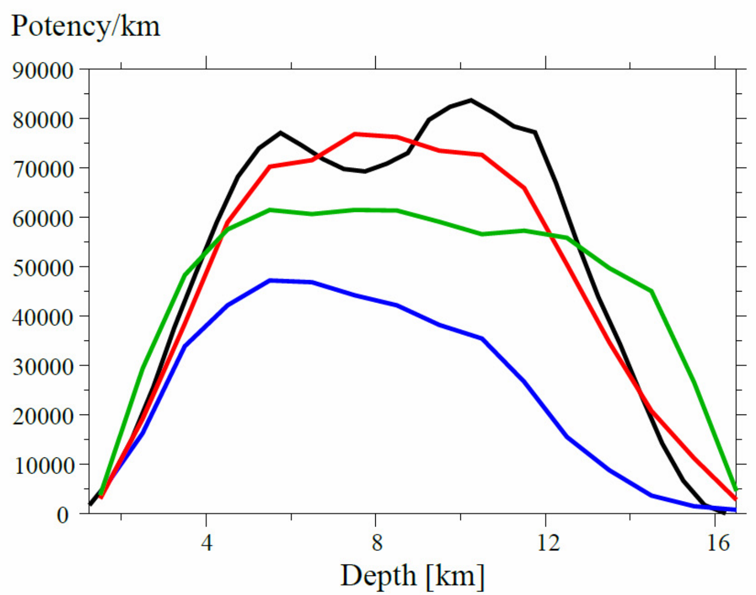
Comparison of Potency density profiles



Potency = slip amplitude * slip area

Along strike spatial variation is better resolved than the resolution along the depth





Conclusion I: Causes of bad results

1. When we use a larger subfault to simplify the problem, the model matches the data best (for instance, **Model I**) is not the model (**Model Target_SC**) which matches the Target model best.

2. For the finite fault study, the bandwidth of inverted signals is as important as the spatial coverage of stations. Practically, investigators must be aware that bandlimited seismic data can lead to erroneous results even if the synthetics have a good fit to the data.

3. The standard objective functions are **not equally** sensitive to all constraints embedded in the broadband seismic waveforms. Inversion codes then could be biased particularly when we match the data with “noise” due to instruments, Green’s functions, or even the model parameterization.

Conclusion II: How to improve?

1. Better objective functions, including a weight which varies with frequency or perhaps even wavelet as proposed by Ji et al. (2002), should be investigated.

2. Inversions using multiple datasets could extend the bandwidth of observations (e.g., Wald & Heaton, 1994; Hernandez et al., 1999; Ji et al., 2002; Custodio et al., 2009) and then should be advocated.

We also advocate that the source modelers include the misfits within different frequency bands in their papers.

Conclusion III: additional bonus

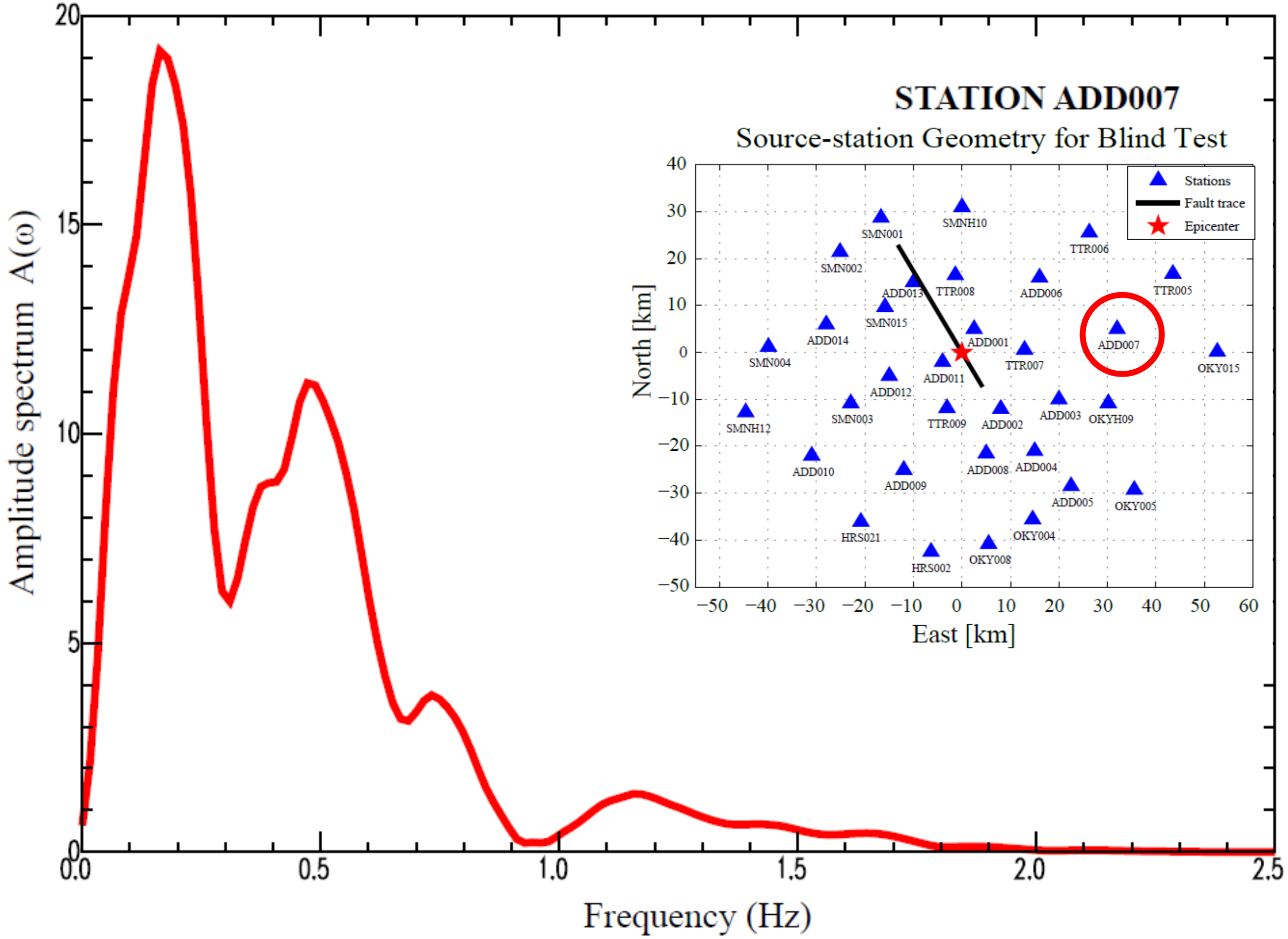
1. Velocity waveforms are more sensitive to the “moment acceleration” than the “moment rate”.

2. For a strike-slip event on a vertical fault, the spatial variation along the strike is better resolved than that along the depth.

Thank you

Thank you

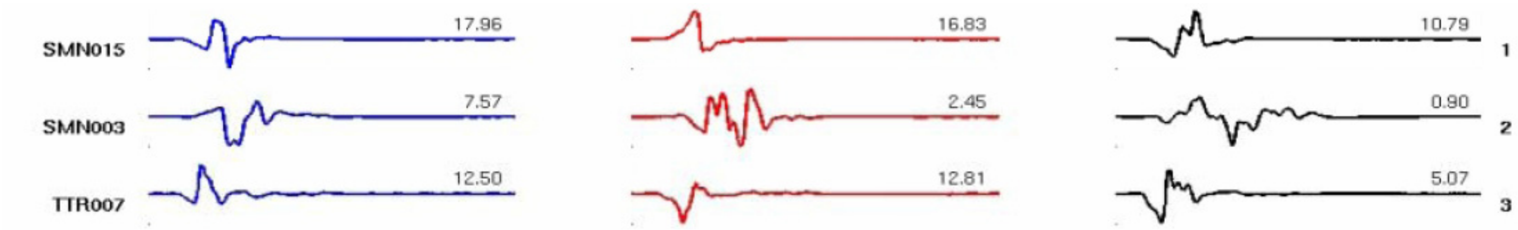
Appendix 1: Energy Ratio



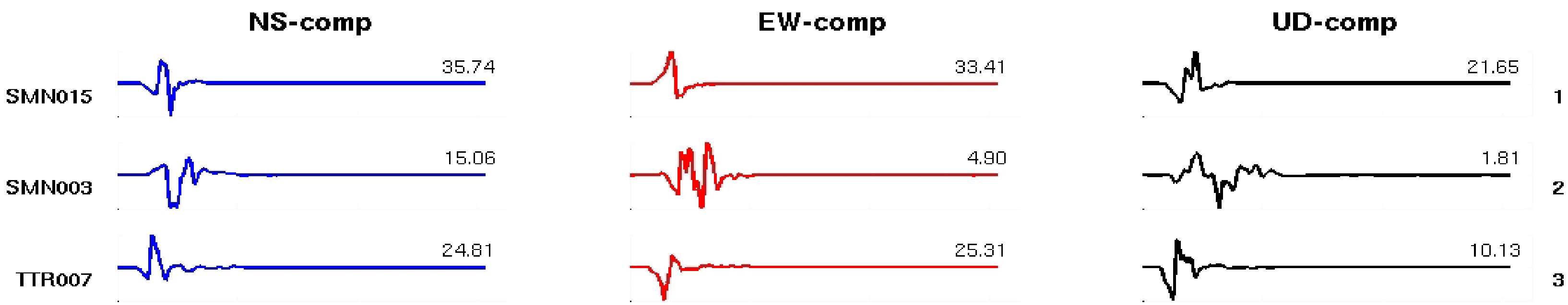
The amplitude spectrum of the velocity record (NS component) at the station ADD007.

Appendix 2: Data correction

Mai et al., 2007 AGU

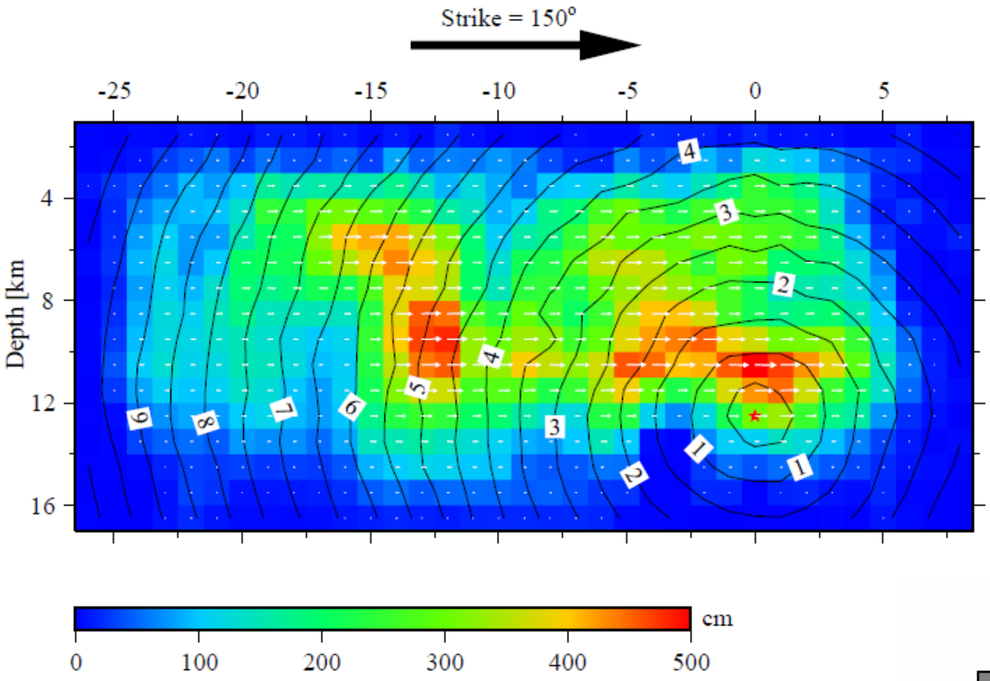


BlindTest Website



The amplitudes are different!

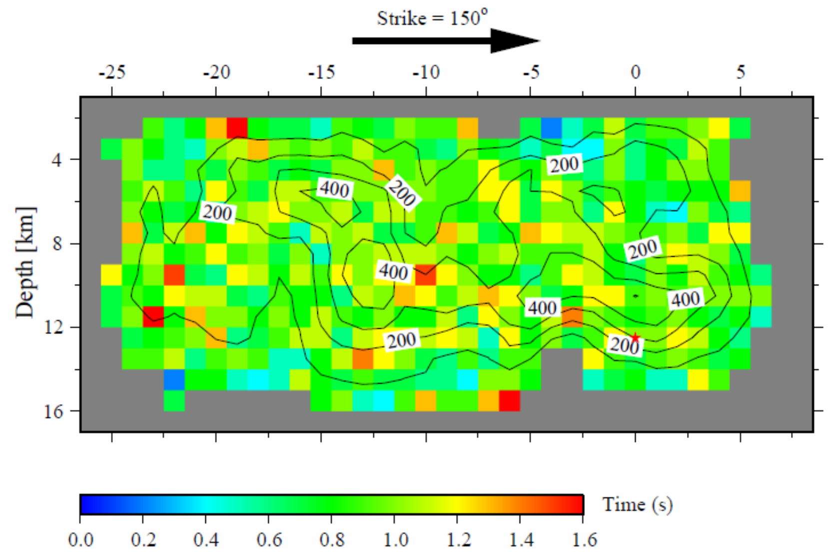
Appendix 3: Model IV: rupture velocity (2.2 -3.1 km/s)



Total moment: 2.71 × 10^{19} Nm

Average starting time: 0.39 s
Average ending time: 0.46 s
Average rise time: 0.85 s

	Variance reduction
Model IV	99.31%
Model I	99.35%



Appendix 3.2: Rupture velocity

Black contours: inverted rupture velocity of Model IV
White dashed contours: reference rupture velocity of 2.7 km/s