

# Earthquake tomography on volcanoes with a Bayesian regularization approach

Jean-Luc GOT and Bernard VALETTE  
ISTerre, CNRS, Université de Savoie

# EARTHQUAKE TOMOGRAPHY: PROBLEM SETTING

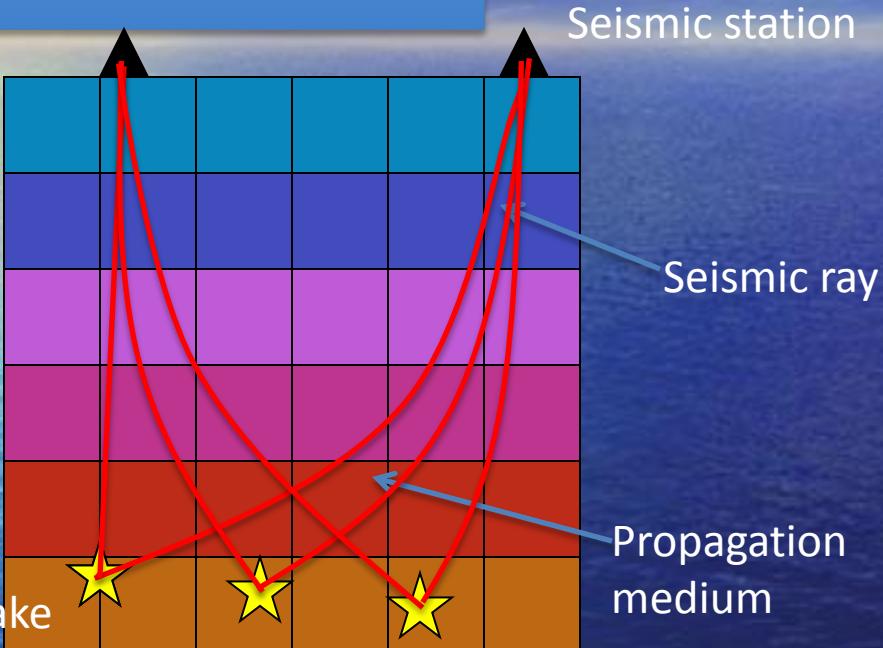
Data: EQ arrival times

Parameters: EQ hypoc. param.

Velocity param.

1. Model discretization:

Cells/nodes → Velocity parameters



RAYS:

2. EQ and stations: discrete distrib.



Incomplete sampling by seismic rays

3. Travel times:  $\Delta t = \int_0^l \frac{ds}{V}$  « Direct problem »

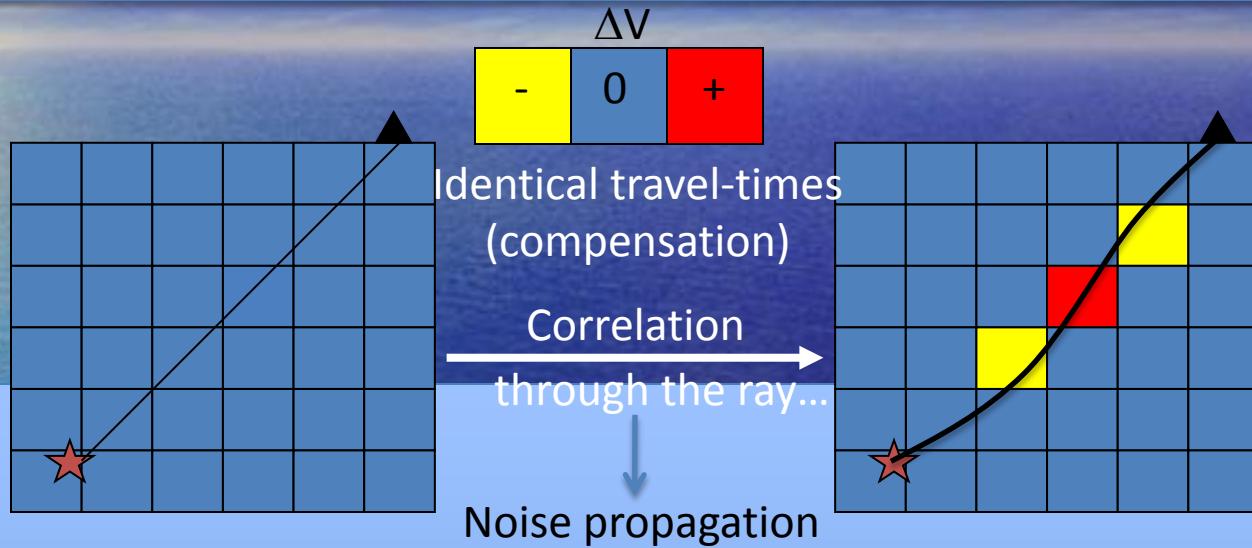
4. Uncertainty and non-linearity:

- Data :  $d$
- Theoretical travel times :  $g(m)$
- $g$  is non-linear in heterogeneous media

Tomography:

Non-linear parametric estimation  
→ inversion  
(velocity, EQ hyp.)

- Model discretization
- Incomplete sampling by rays
- Integral nature of travel-times → intrinsic correlation between parameters



Correct solution needs:

- bounding the error on parameters by realistic physical intervals;
- taking into account the intrinsic correlation between parameters;  
→ Regularization...

Probabilistic (« bayesian ») approach using « a priori » knowledge

# Bayesian approach

(Tarantola and Valette, 1982;  
Tarantola, 1987, 2005)

(A posteriori) knowledge on the model =  
a priori knowledge on the model AND knowledge from the data

A posteriori pdf:

$$\sigma_m(m) = \rho_m(m) f(d - g(m))$$

↑                      ↑  
A priori pdf      Data+modelling error pdf

Case of gaussian pdf for data and model parameters

$$\rho_m(m) = (2\pi \det C_m)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(m - m_0)^T C_m^{-1} (m - m_0)\right)$$
$$\rho_\varepsilon(d - d) = (2\pi \det C_d)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(d - g(m))^T C_d^{-1} (d - g(m))\right)$$

Maximize the a posteriori pdf (« MAP estimation »)

$$\sigma_m(m) \propto \exp\left(-\frac{1}{2}(m - m_0)^T C_m^{-1} (m - m_0) - \frac{1}{2}(d - g(m))^T C_d^{-1} (d - g(m))\right)$$

Minimize

$$(\mathbf{d} - g(\mathbf{m}))^T \mathbf{C}_{\mathbf{d}}^{-1} (\mathbf{d} - g(\mathbf{m})) + (\mathbf{m} - \mathbf{m}_0)^T \mathbf{C}_{\mathbf{m}}^{-1} (\mathbf{m} - \mathbf{m}_0)$$

Gauss-Newton

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \left( \mathbf{G}_k^T \mathbf{C}_{\mathbf{d}}^{-1} \mathbf{G}_k + \mathbf{C}_{\mathbf{m}}^{-1} \right)^{-1} \left( \mathbf{G}_k^T \mathbf{C}_{\mathbf{d}}^{-1} (\mathbf{d} - g(\mathbf{m}_k)) + \mathbf{C}_{\mathbf{m}}^{-1} (\mathbf{m}_k - \mathbf{m}_0) \right)$$

Implementation

$$\begin{pmatrix} \mathbf{C}_{\mathbf{d}}^{-1/2} \mathbf{G} \\ \mathbf{C}_{\mathbf{m}}^{-1/2} \end{pmatrix} \Delta \mathbf{m}_{k+1} = \begin{pmatrix} \mathbf{C}_{\mathbf{d}}^{-1/2} (\mathbf{d} - g(\mathbf{m}_k)) \\ \mathbf{C}_{\mathbf{m}}^{-1/2} (\mathbf{m}_k - \mathbf{m}_0) \end{pmatrix}$$

(Tarantola and Valette, 1982; Tarantola, 1987, 2005)

Physical interpretation:  
Fitting the data with the most simple (smooth) model

Understanding regularization... Unsufficient data:

- Classical approach: Unsufficient data →
  - \* very high variance and model perturbation;
  - \* Asymptotically unbiased solution;
- Bayesian approach: Unsufficient data →
  - \* null information : a posteriori knowledge = a priori knowledge  
→ no model perturbation;
  - \* solution is asymptotically biased ;

... → bias – variance compromise :

introducing controlled bias from optimal a priori knowledge  
for reducing noise variance.

... → bias – variance compromise :

introducing controlled bias from optimal a priori knowledge  
for reducing noise variance.

Needs:

- Good physical a priori model :  
→ Gradient or averaged models, std
- Correlation between parameters :  
→ Correlation kernel in the a priori covariance matrix

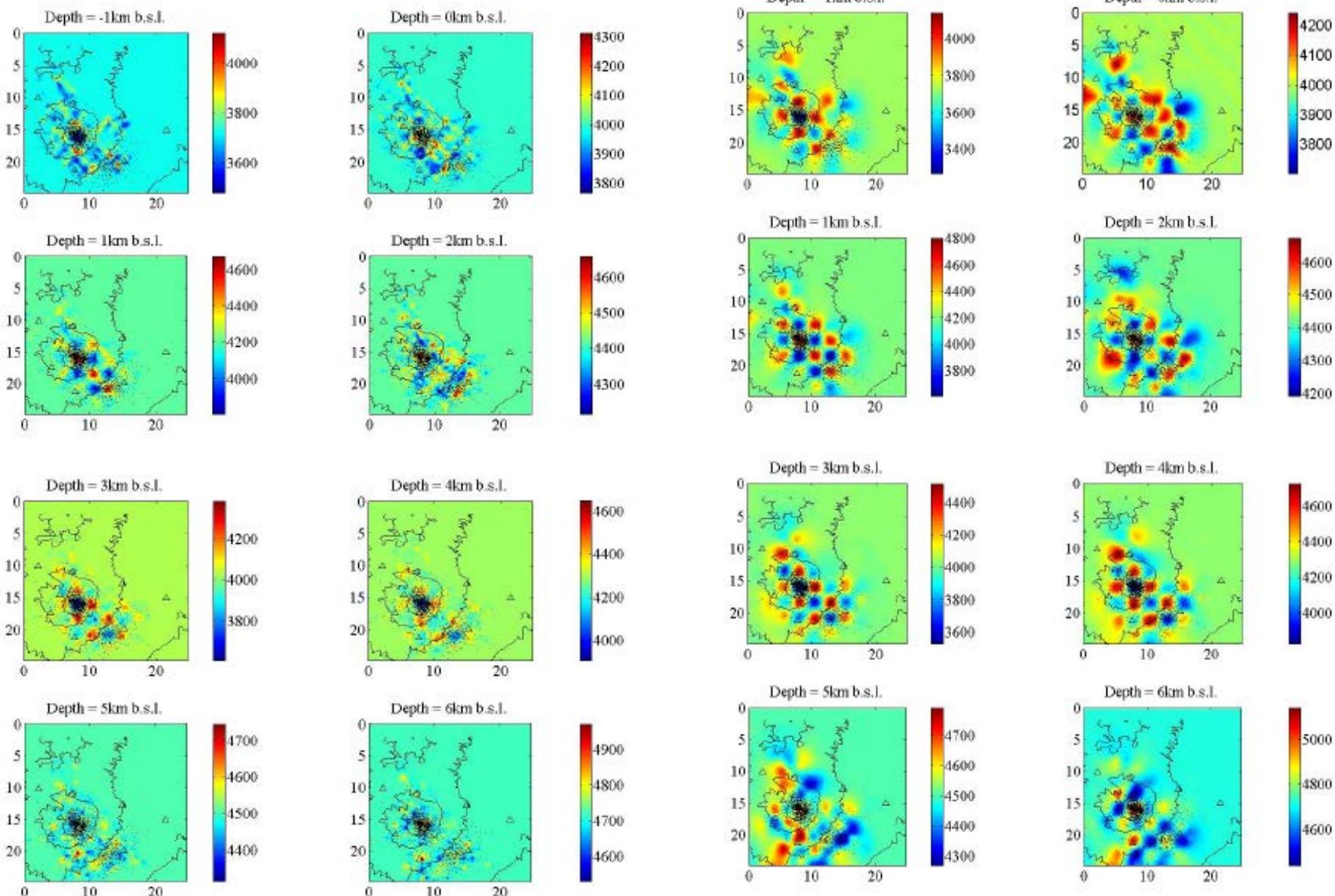
$$\mathbf{C}_v(\mathbf{p}, \mathbf{p}') = \sigma_v^2 e^{-\frac{| \mathbf{p} - \mathbf{p}' |}{\lambda}}$$

$\sigma_v$  : a priori velocity standard deviation  
 $\lambda$  : correlation length

- Optimization strategy :  
→ defining and optimizing hyper-parameters:
  - a priori standard deviation
  - correlation length

# Understanding correlation length (1)

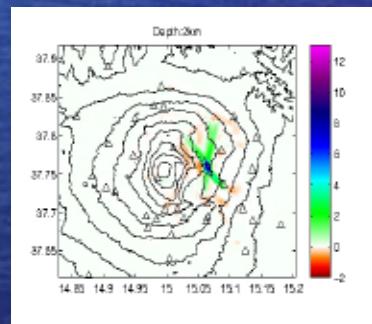
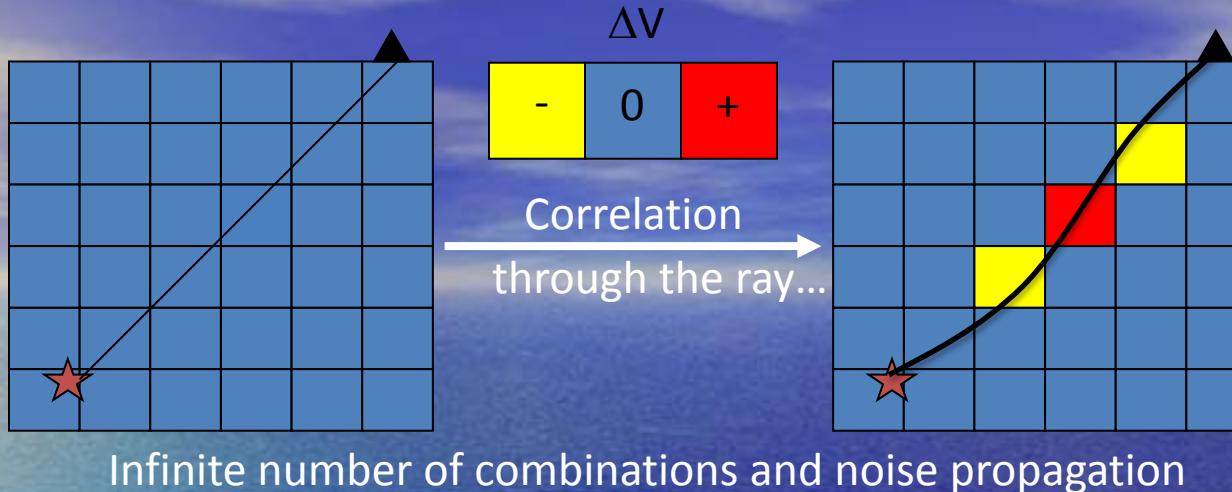
Berger, Got, Valdes, Monteiller, 2011



Checkerboard test: (1) under-regularization ( $\lambda = 0.5$  km)  
High-frequency noise

(2) Optimal regularization ( $\lambda = 3$  km)  
→ Low-pass filter

## Understanding correlation length (2) Ray propagation through model cells



Resolution

For each cell:

Parameter value = true value + noise (neighbouring cells)

Optimal filtering : bias-variance OPTIMAL compromise

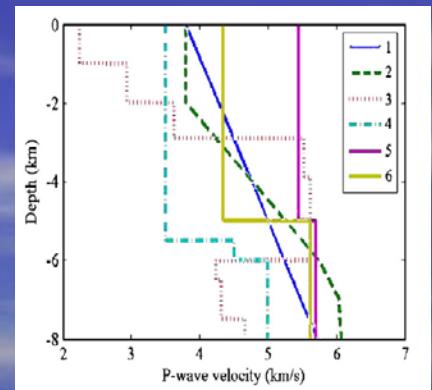
Correlating parameters with a length :

- not too large to bias physical values
- but sufficiently large to filter out the noise and reduce variance

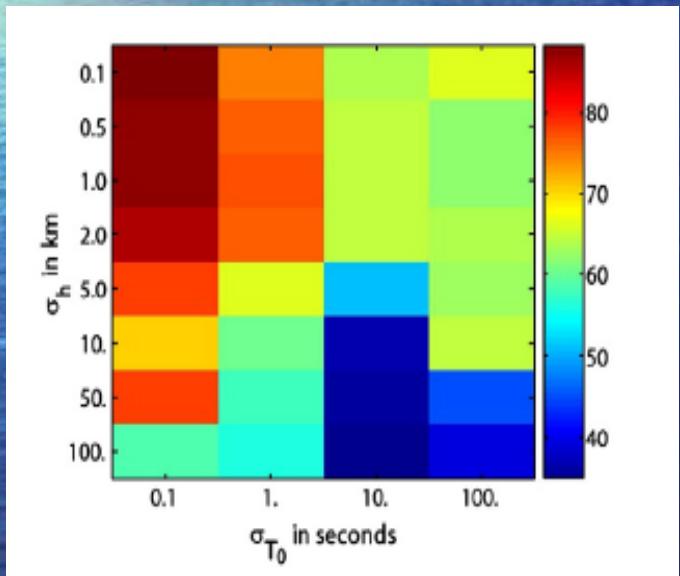
LOW-PASS FILTERING OF HIGH-FREQUENCY NOISE

## Choosing the regularization:

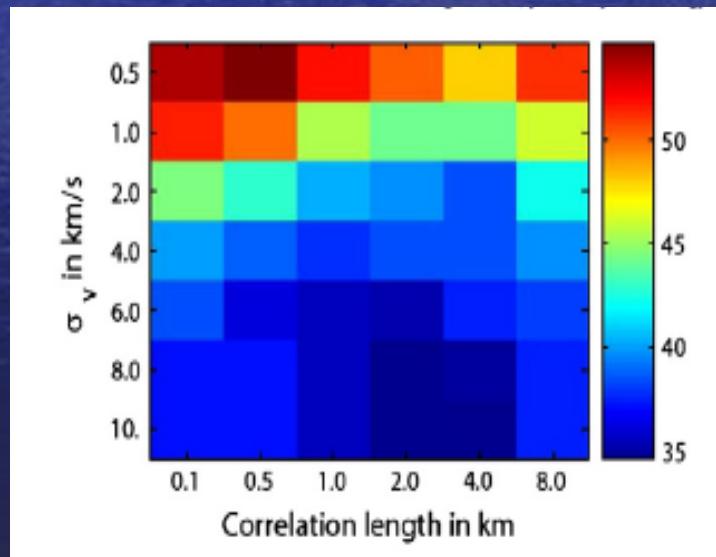
1. A priori average model ( $m_0$ ): constant gradient



2. Cost function as a function  
of regularization hyper-parameters

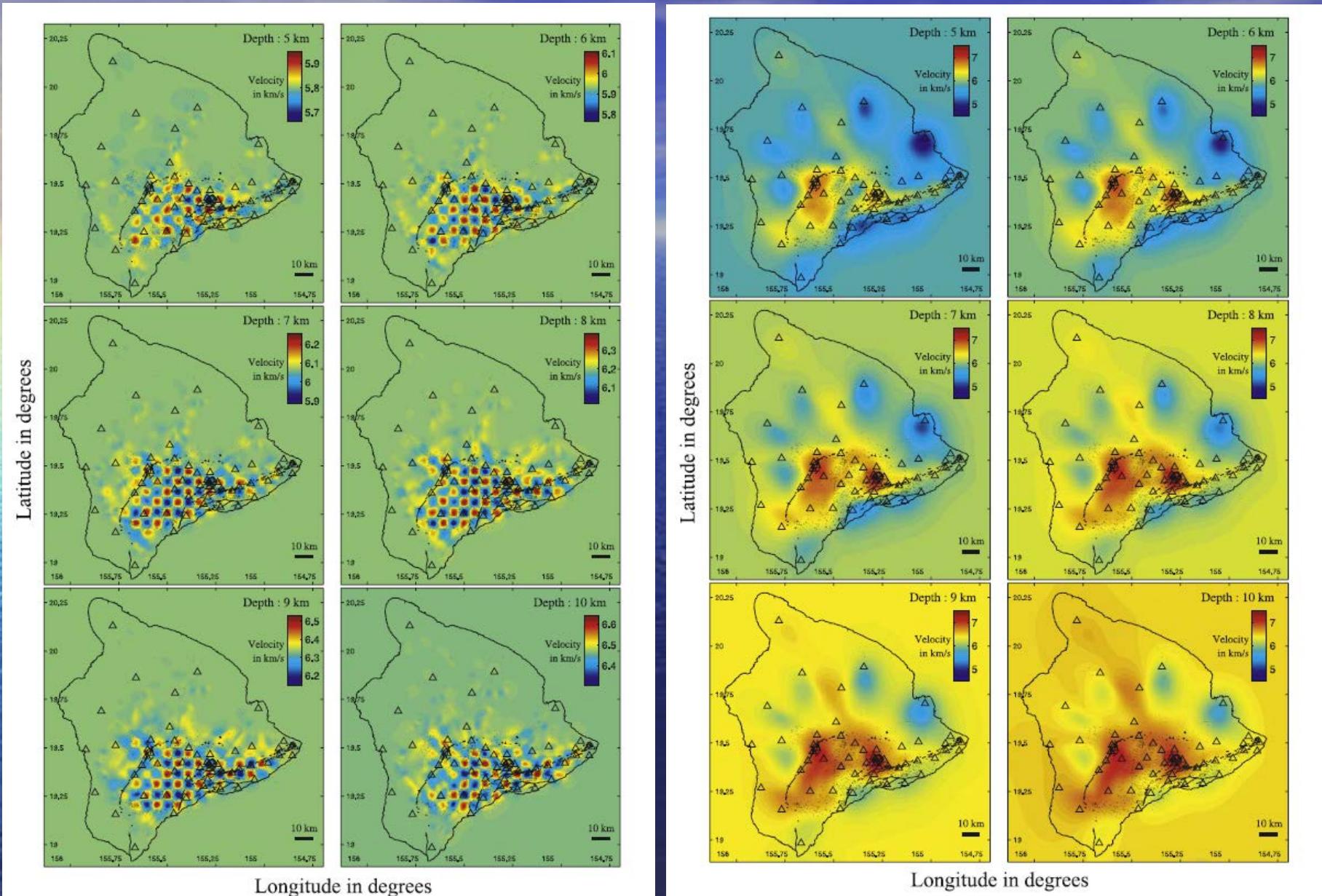


Hypocentral hyper-parameters



Velocity hyper-parameters

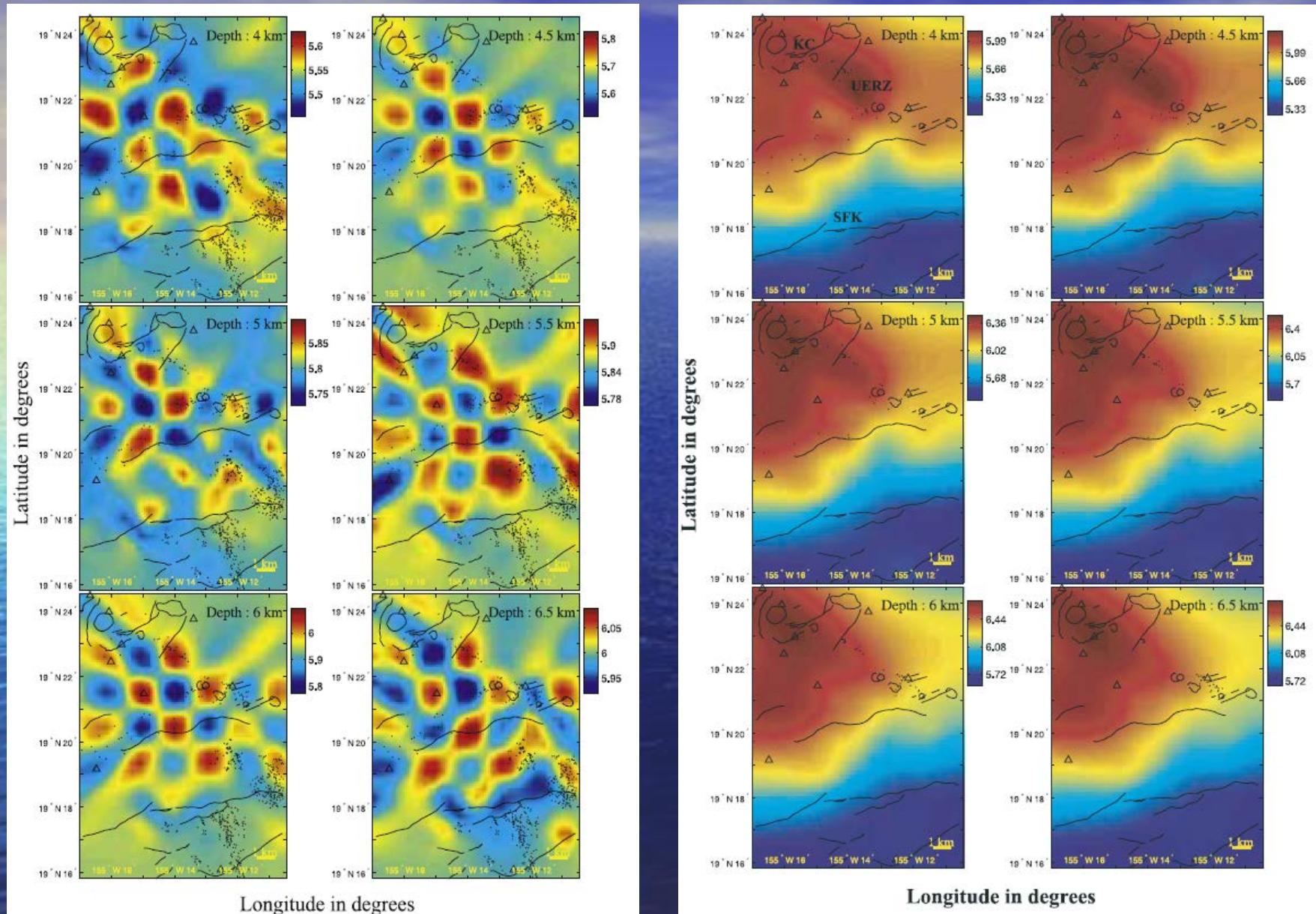
# Travel-time earthquake tomography results, Hawaii Monteiller, Got, Virieux, Okubo, JGR, 2005



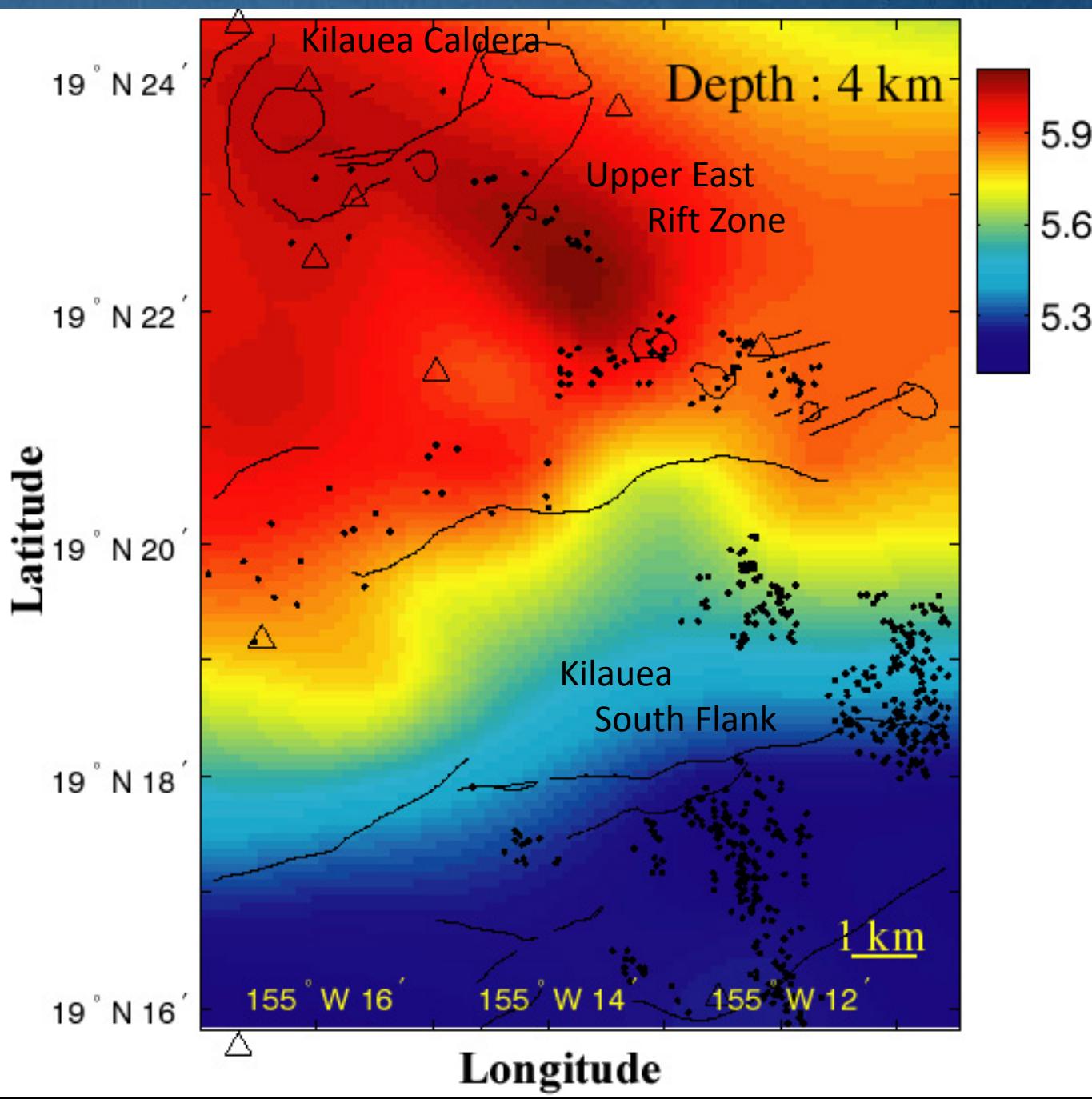
41886 P travel-time, 1358 evts, 31 arr. times/evt (av.) , 959077 nodes

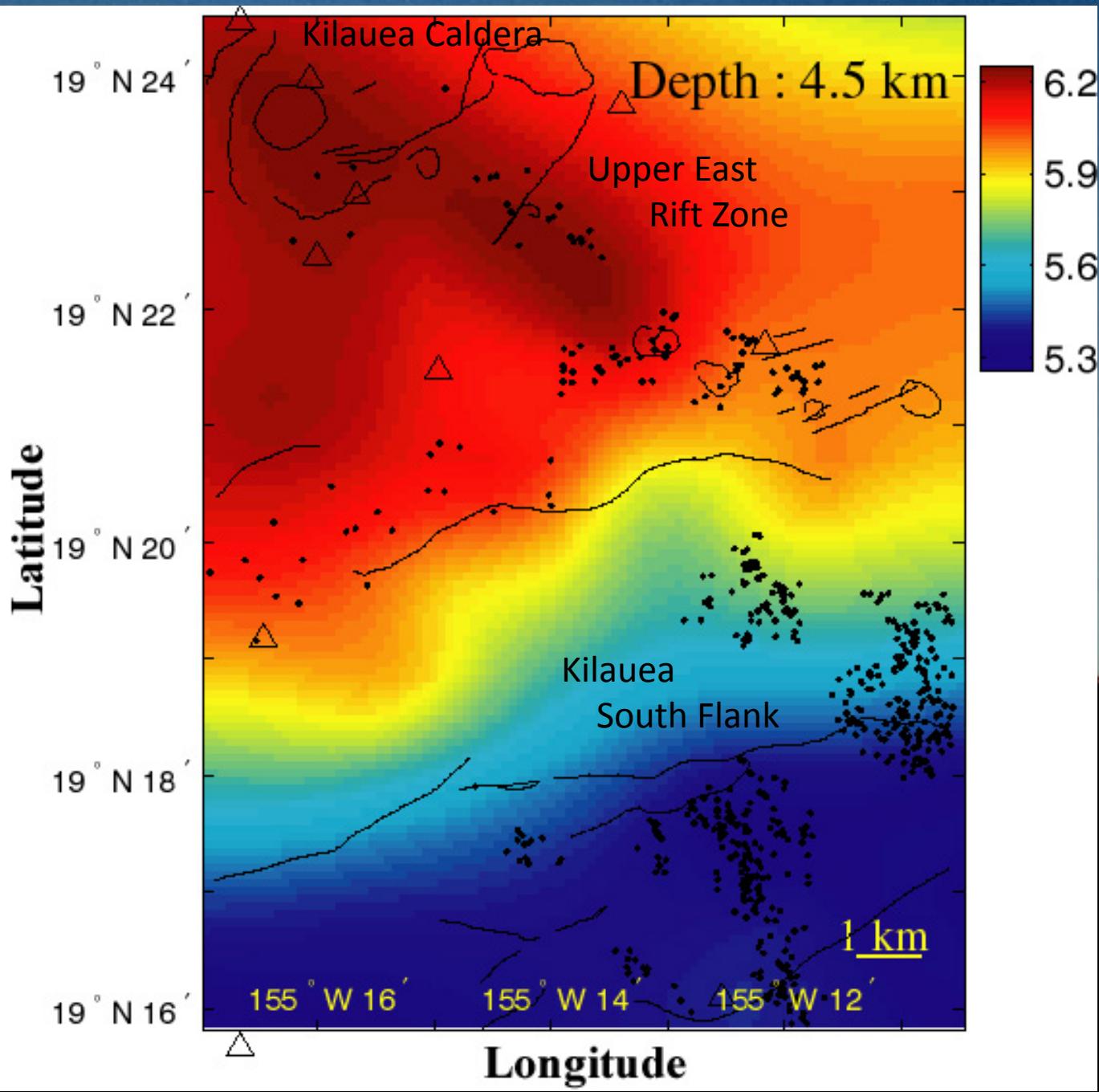
$$\lambda = 5 \text{ km}, \sigma_v = 1 \text{ km/s}$$

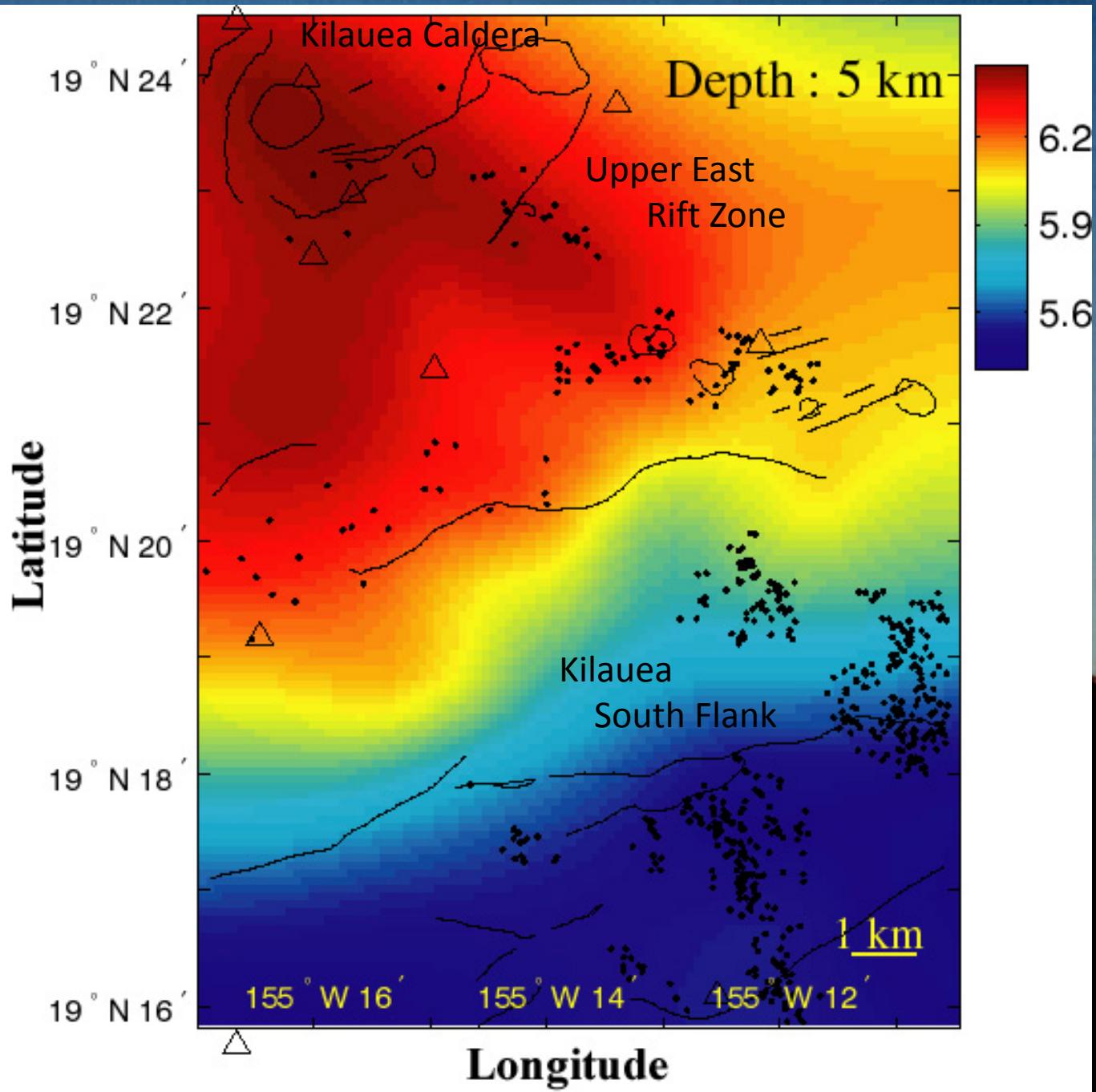
# Double-difference earthquake tomography results, Hawaii Monteiller, Got, Virieux, Okubo, JGR, 2005

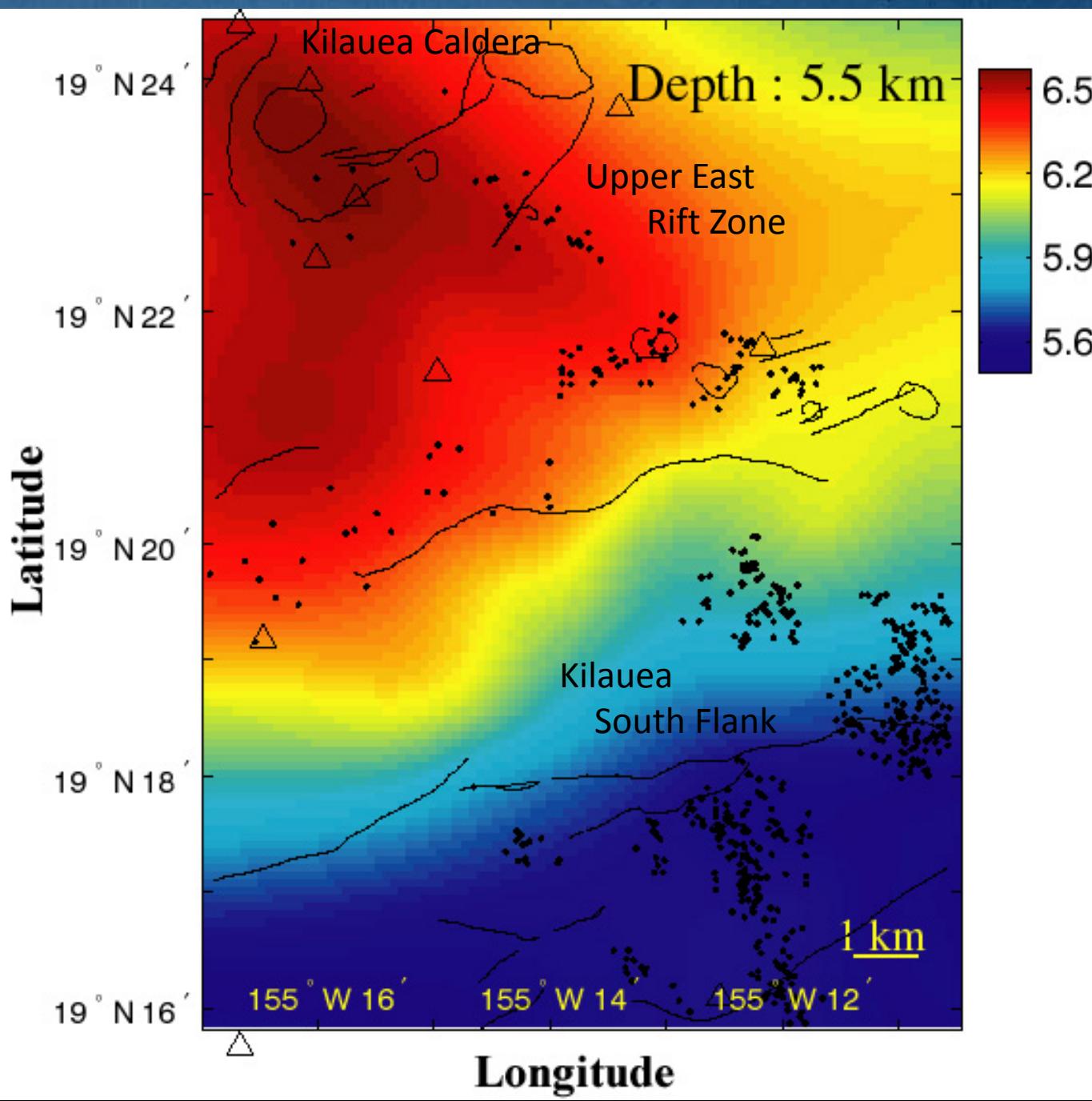


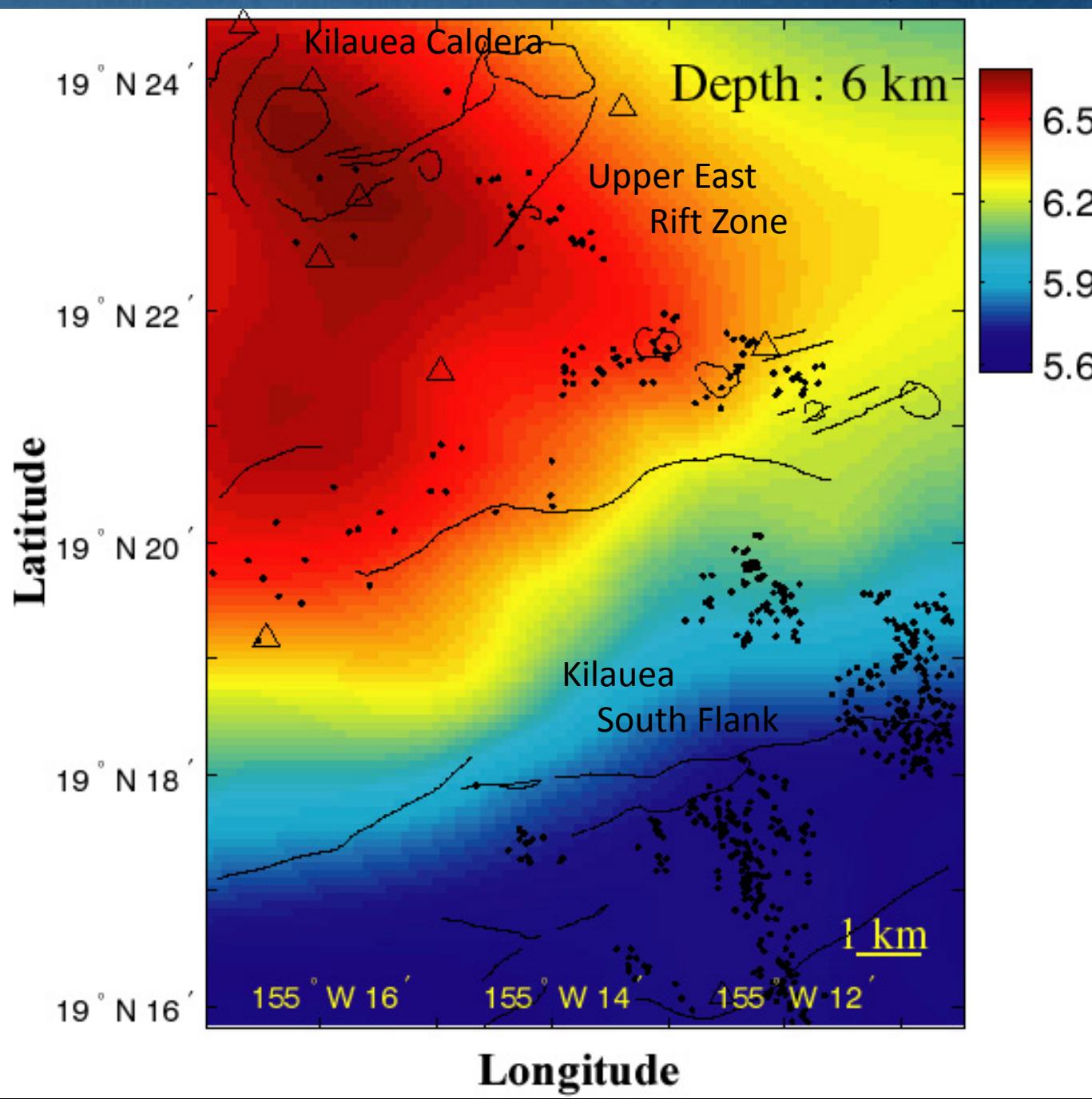
415436 time delays P, 614 evts, 25 time delays/cpl (avg), 20625 nodes  
 $\lambda = 1 \text{ km}$ ,  $\sigma_v = 0.5 \text{ km/s}$

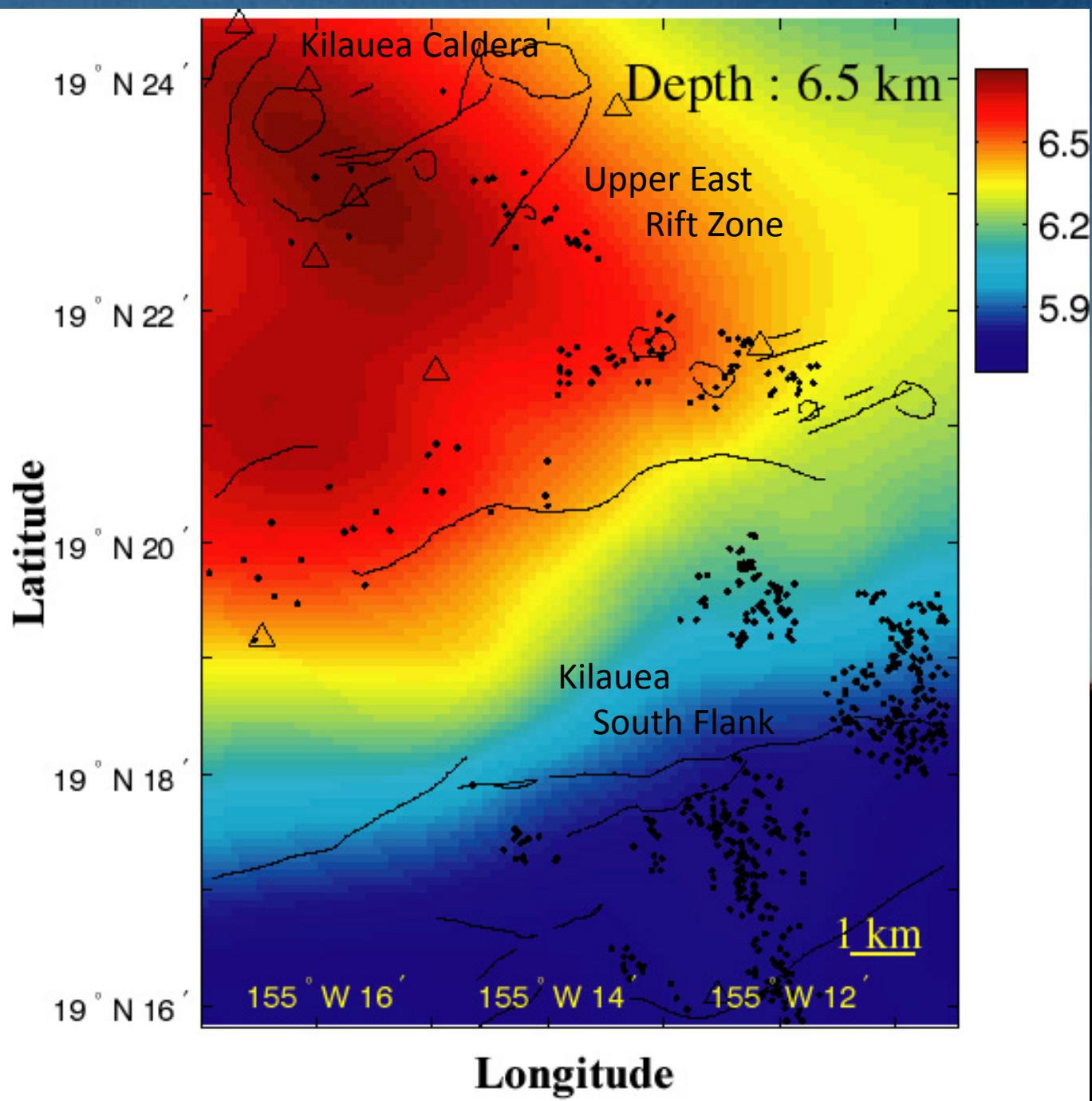


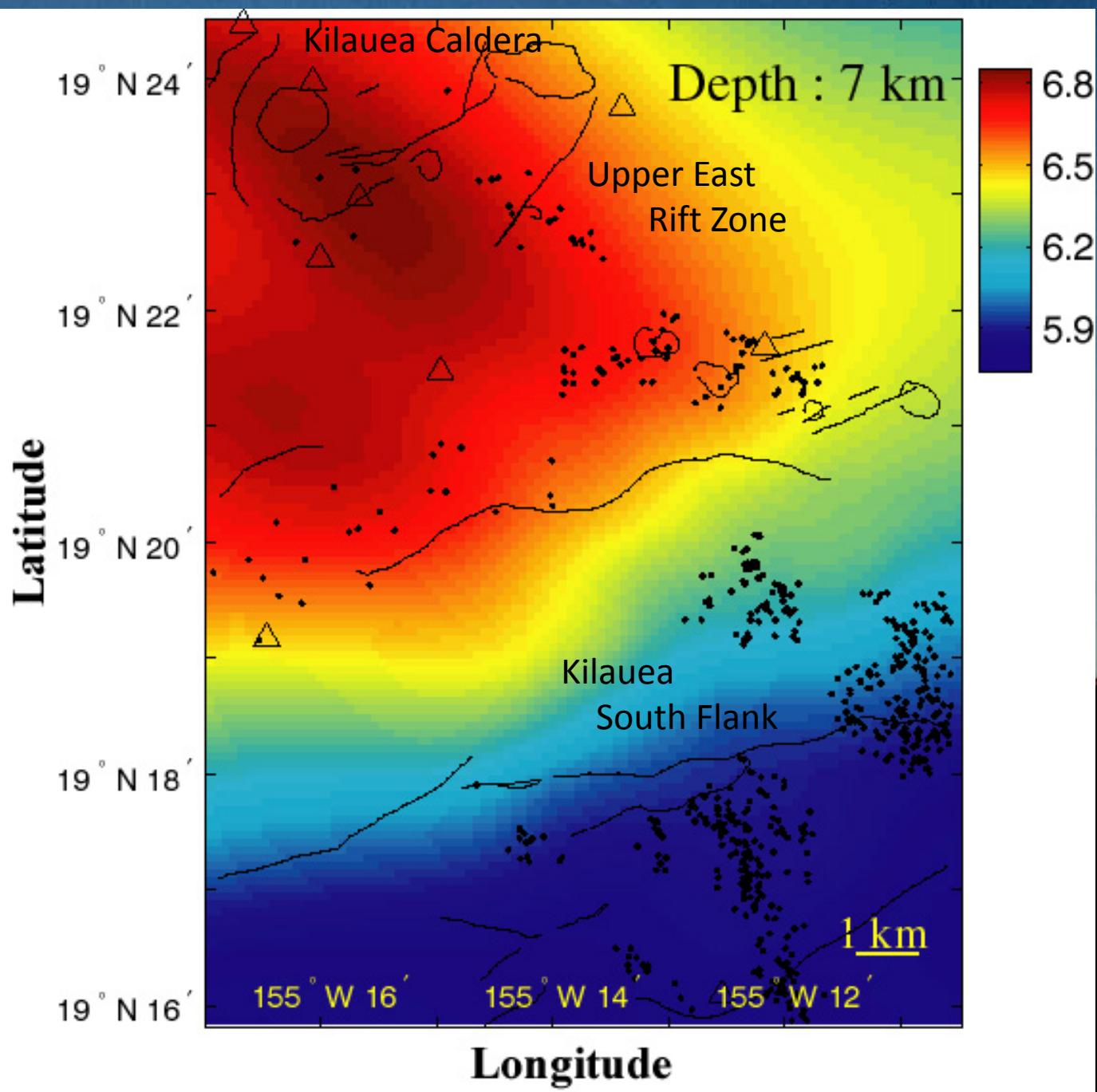


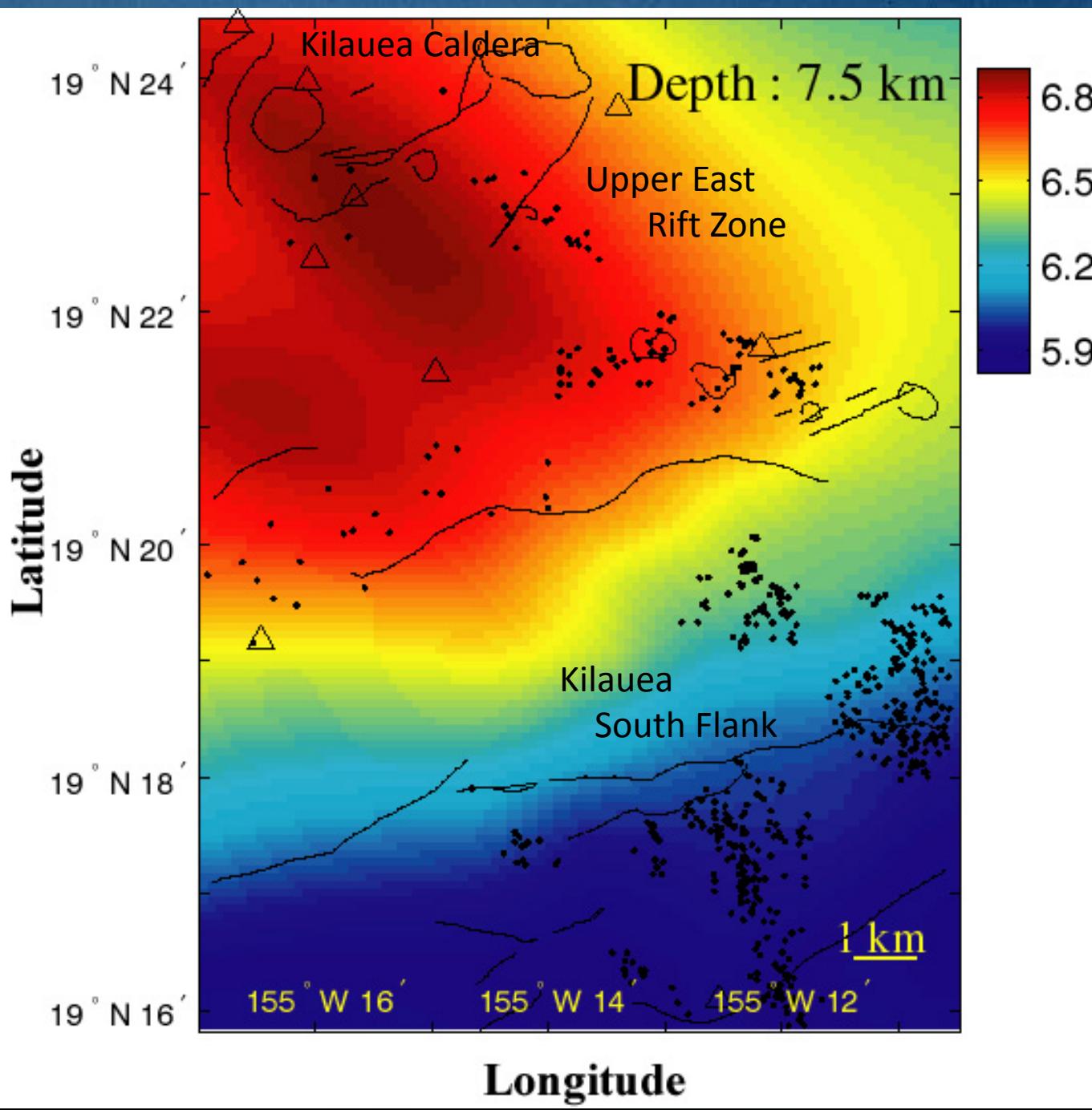


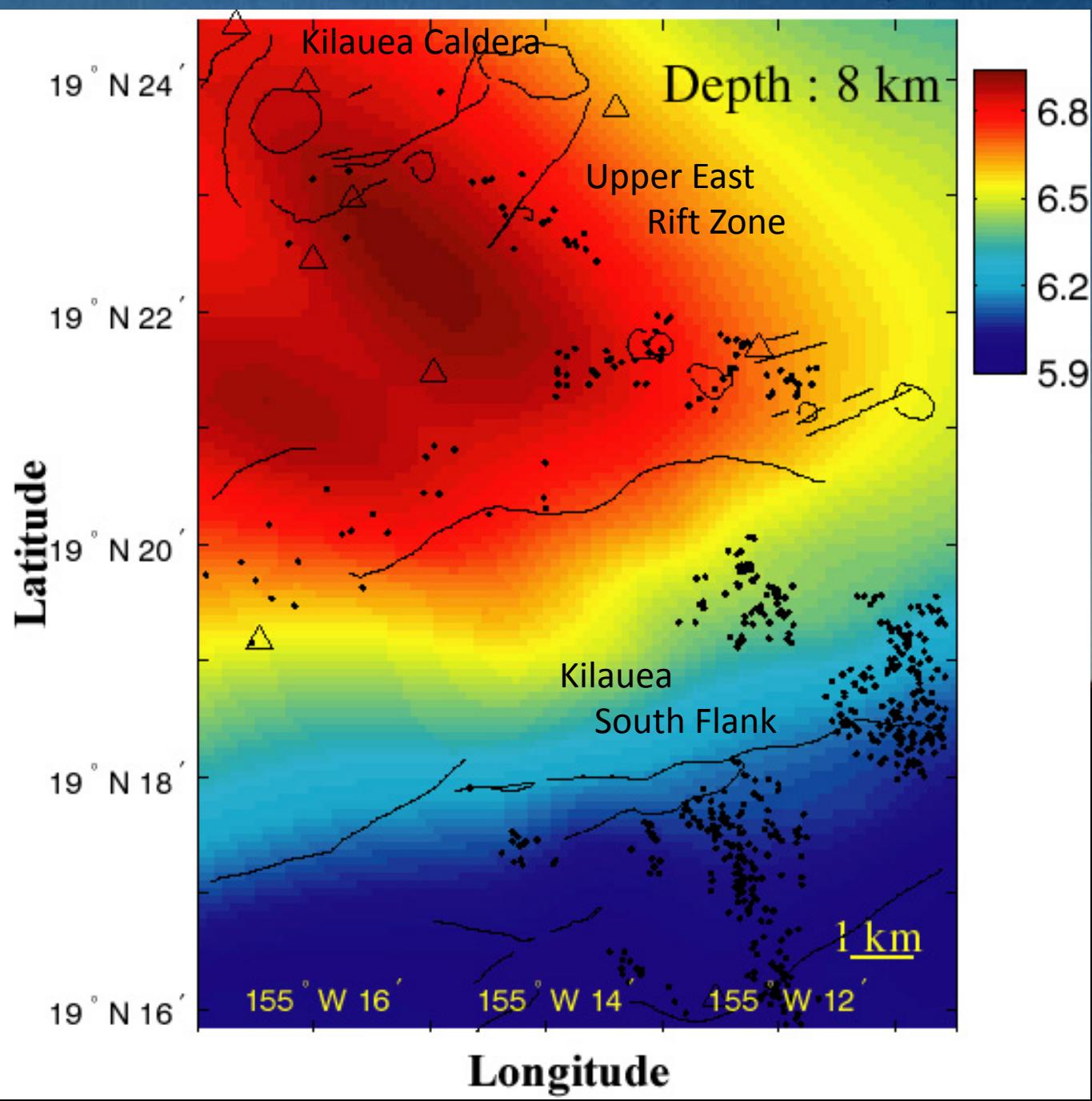


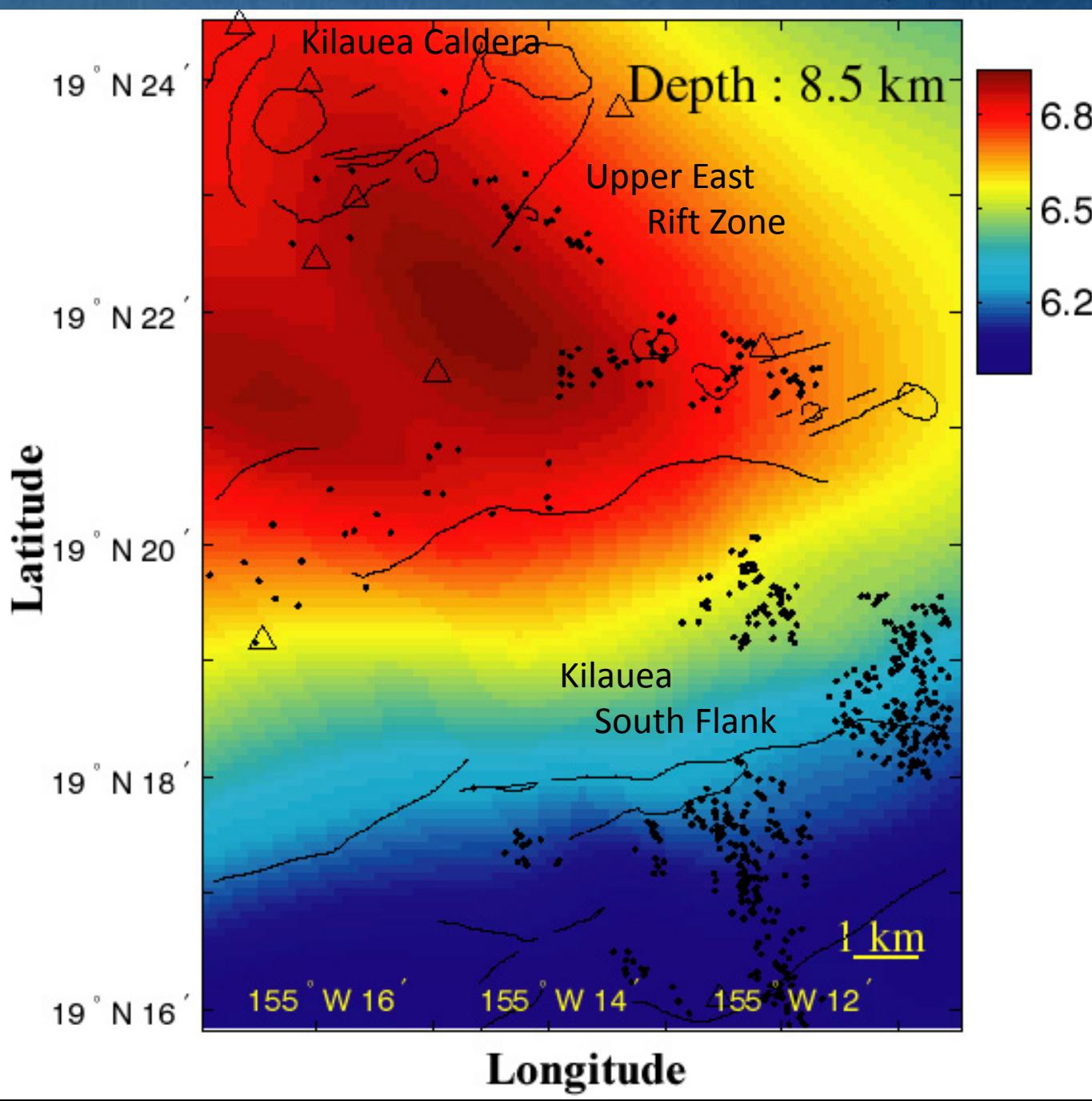


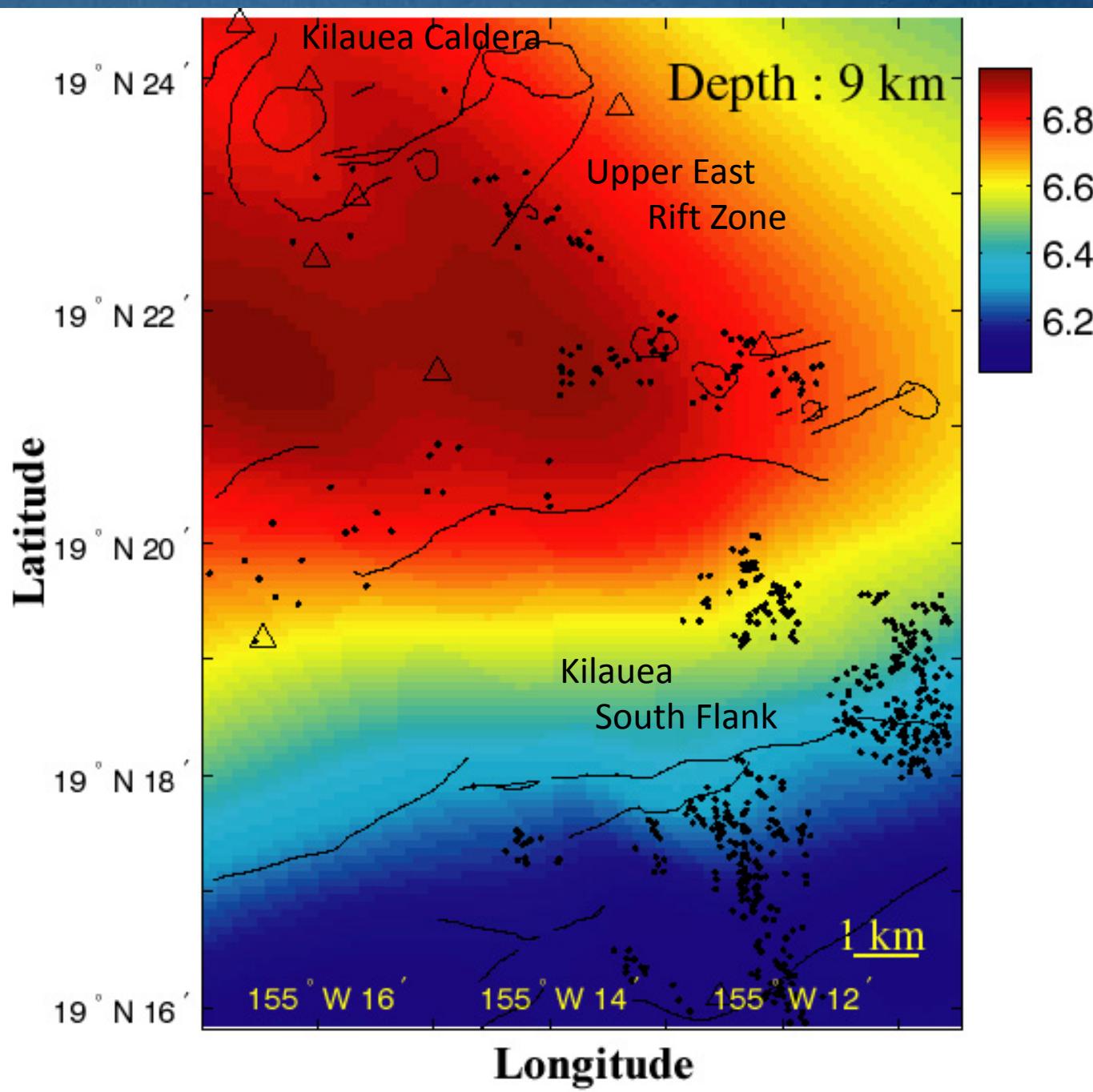


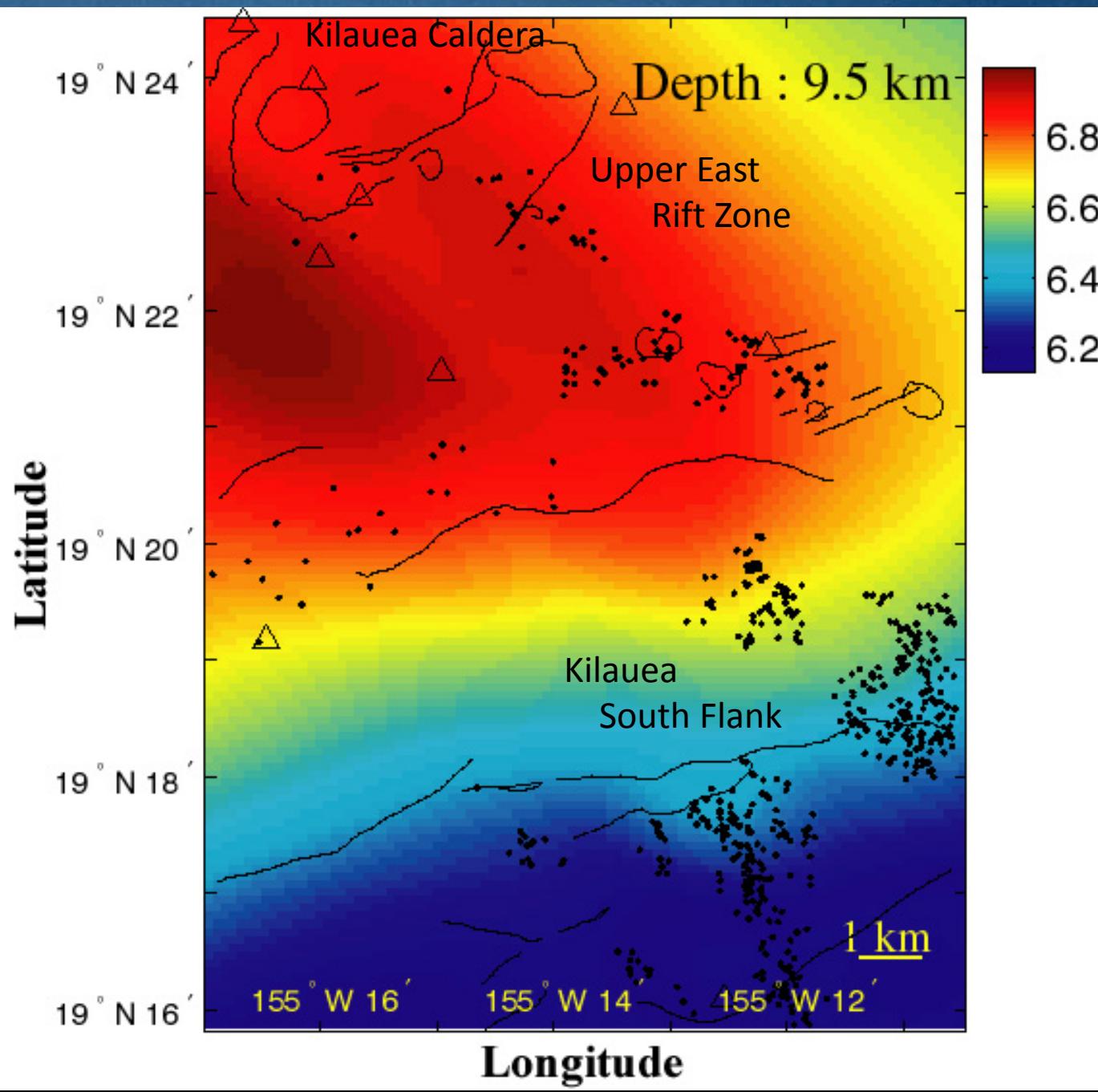




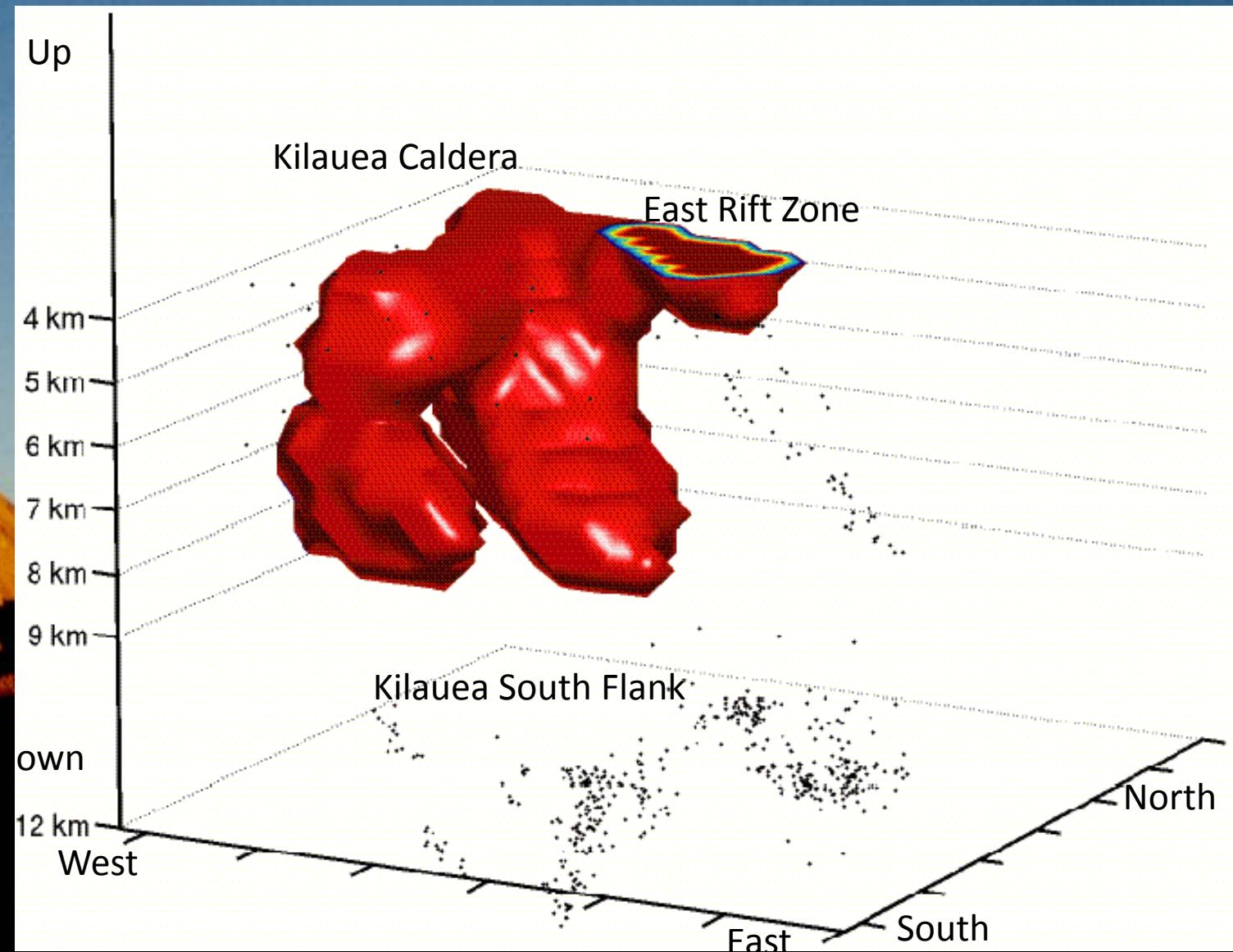








## Kilauea Volcano magma feeding system from double-difference tomography



The background of the image is a wide-angle photograph of a vast ocean. The water is a deep, dark blue, showing subtle ripples and reflections. Above the horizon, the sky is a lighter shade of blue, dotted with wispy, white clouds. In the upper left corner, there is a faint, vertical rainbow-like glow, possibly from a camera lens or a natural phenomenon.

Thank you!!!

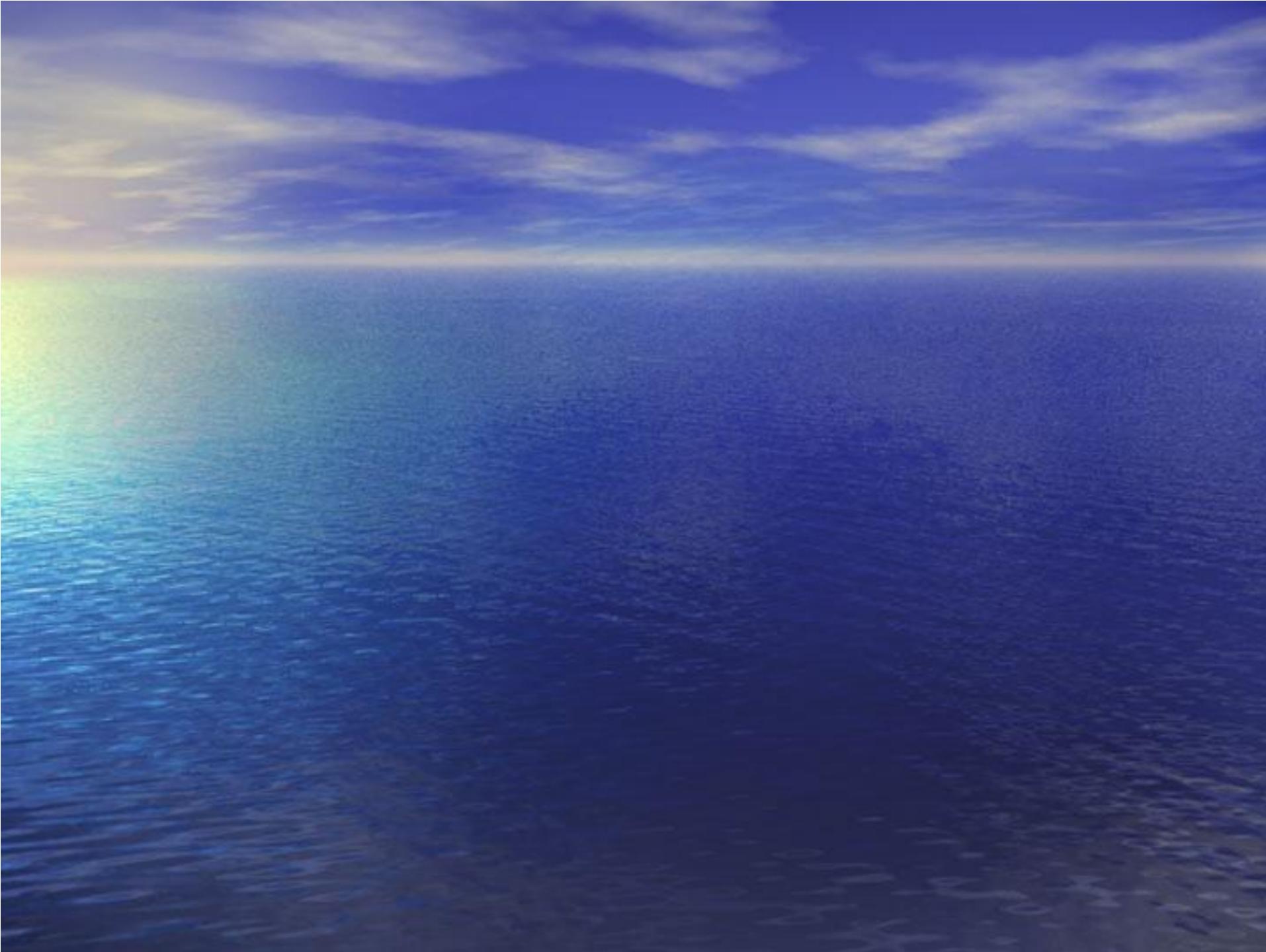
### **Further readings:**

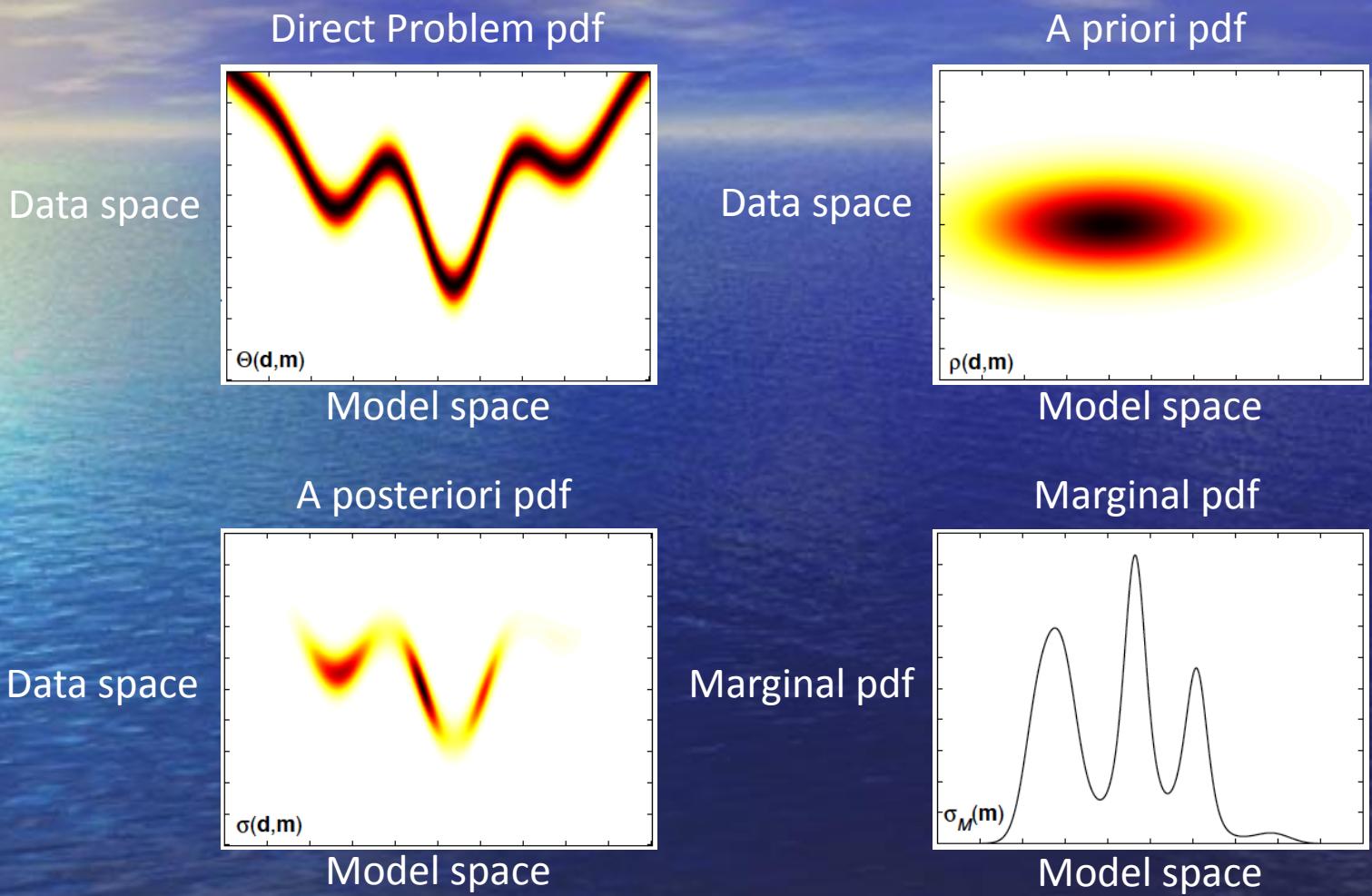
**Tarantola**, Inverse Problem Theory, 1987, Elsevier, or 2005, SIAM

**Tarantola and Valette, 1982**, Generalized nonlinear inverse problems solved using the least-square criterion, Rev. Geophys., 20, 219-232.

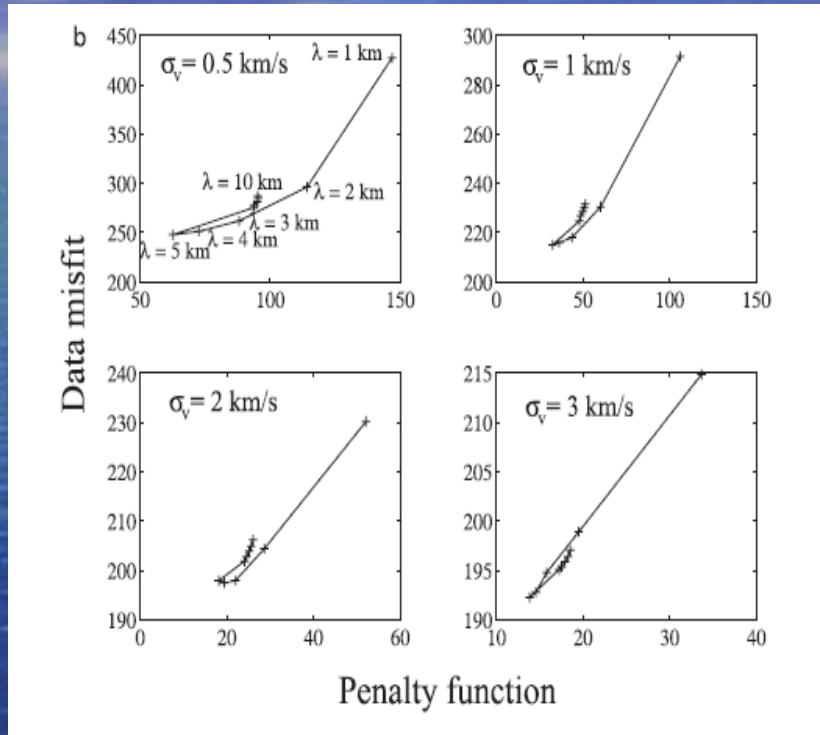
**Monteiller, Got, Virieux, Okubo, 2005**, An efficient algorithm for double-difference tomography and location in heterogeneous media, with an application to the Kilauea volcano, J. Geophys. Res., 110, B12306, doi:10.1029/2004JB003466.

**Berger, Got, Valdes, Monteiller, 2011**, Seismic tomography at Popocatepetl volcano, Mexico, J. Volc. Geoth. Res., 20, 234–244.





### 3. Data misfit – Penalty function curves



Monteiller, Got, Virieux, Okubo,  
JGR, 2005