A Bayesian Perspective on Seismic Tomography

────

With an application to the thermochemical structure beneath the North American continent

A. Khan

Institute of Geochemistry and Petrology, ETH Zürich, Switzerland

Charles University, Prague, March 2012

Outline

Introduction - Putting the problem into perspective

Methodology - An integrated approach to inverting geophysical data directly for composition and temperature

Application - Inversion of surface-wave phase-velocities for thermo-chemical and physical structure of the North American continent.

Conclusions

Lines of inquiry

Where does our information on the internal structure of the Earth and terrestrial planets come from ?

Geophysics:

- ▶ Seismology
- ▶ Electromagnetic sounding
- ▶ Gravity

Geochemistry:

- ▶ Geochemical and petrological analyses of terrestrial rocks and meteorites.

Geodynamics:

- ▶ Numerical modeling of mantle flow.

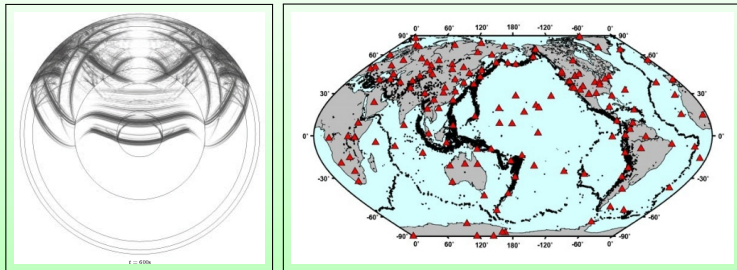# Seismology and Earth's mantle structure

Figure: Seismic wave propagation in the Earth, global distribution of seismometers and seismicity.
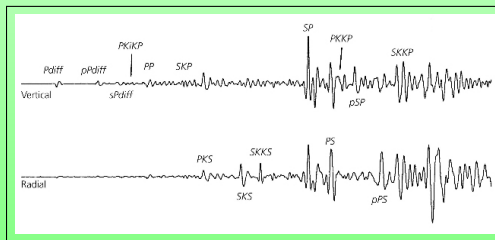

Figure: Seismogram from a station.

Seismic tomography and Earth's mantle structure

Seismic tomography has provided spectacular images of the large-scale structure of the Earth's interior, which in turn offers clues about mantle chemistry and thermal state,
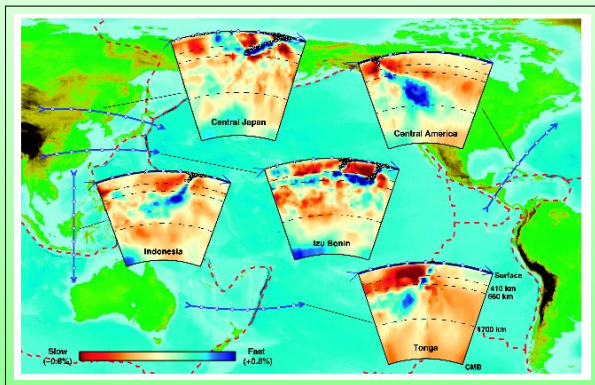

Figure: 3D tomographic structure of the mantle close to subduction zones in the Pacific and central America. From Karason & Van der Hilst (2000).

Seismic tomography and Earth's mantle structure

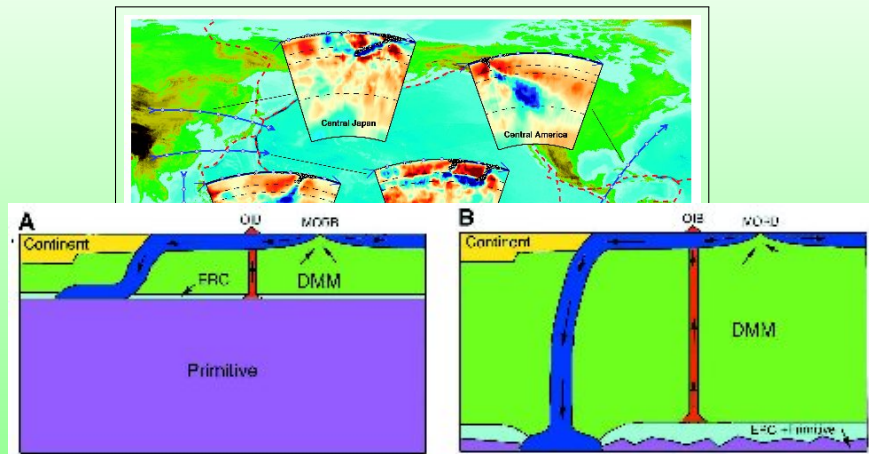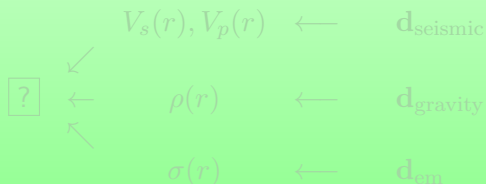and has changed our perception of mantle dynamics

Figure: Two end-member scenarios for the present-day constitution of the Earth's mantle. (A) layered mantle convection. (B) whole mantle convection. From Tackley (2000).

Limitations with current approaches and a step beyond

The problem is compounded by the fact that physical properties are not an end in themselves

$$
\begin{array}{rcl}
V_s(r), V_p(r) & \longleftarrow & \mathbf{d}_{\text{seismic}} \\
\rho(r) & \longleftarrow & \mathbf{d}_{\text{gravity}} \\
\sigma(r) & \longleftarrow & \mathbf{d}_{\text{em}}
\end{array}
$$

We need to look deeper for a set of parameters that describe rock properties at a more fundamental level

$$
\begin{array}{rcl}
V_s(r), V_p(r) & \longleftarrow & \mathbf{d}_{\text{seismic}} \\
\boxed{?} \;\leftarrow\; \rho(r) & \longleftarrow & \mathbf{d}_{\text{gravity}} \\
\sigma(r) & \longleftarrow & \mathbf{d}_{\text{em}}
\end{array}
$$

Limitations with current approaches and a step beyond

The problem is compounded by the fact that physical properties are not an end in themselves

$$
\begin{aligned}
V_s(r), V_p(r) &\longleftarrow \mathbf{d}_{\text{seismic}} \\
\rho(r) &\longleftarrow \mathbf{d}_{\text{gravity}} \\
\sigma(r) &\longleftarrow \mathbf{d}_{\text{em}}
\end{aligned}
$$

We need to look deeper for a set of parameters that describe rock properties at a more fundamental level

$$
\begin{array}{ccc}
& V_s(r), V_p(r) & \longleftarrow \quad \mathbf{d}_{\text{seismic}} \\
\swarrow & & \\
\boxed{?} \leftarrow & \rho(r) & \longleftarrow \quad \mathbf{d}_{\text{gravity}} \\
\nwarrow & & \\
& \sigma(r) & \longleftarrow \quad \mathbf{d}_{\text{em}}
\end{array}
$$

A novel approach

The underlying fundamental parameters are composition (c) and temperature (T):

$$
\begin{array}{ccc}
 & V_s(r), V_p(r) & \longleftarrow & \mathbf{d}_{\text{seismic}} \\
\swarrow & & & \\
\boxed{c(r), T(r)} \quad \leftarrow & \rho(r) & \longleftarrow & \mathbf{d}_{\text{gravity}} \\
\nwarrow & & & \\
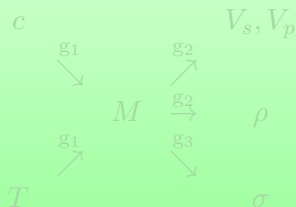 & \sigma(r) & \longleftarrow & \mathbf{d}_{\text{em}}
\end{array}
$$
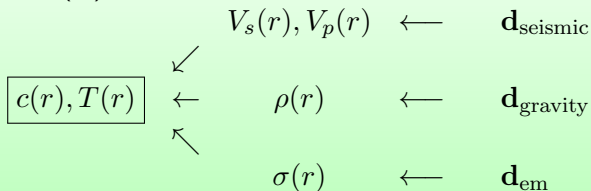
A more definitive approach would have us test geophysical data directly for composition and temperature !

Thermodynamic modeling

Is it possible to compute physical properties directly for a given chemical composition (c), pressure (P) and temperature (T) ?

Yes, this can be achieved through the use of thermodynamic methods based on Gibbs free energy minimisation (e.g. Perple_X, Connolly, 2005):

$$c \qquad\qquad\qquad V_s, V_p$$

$$\searrow^{g_1} \qquad \nearrow^{g_2}$$

$$M \xrightarrow{g_2} \rho$$

$$\nearrow^{g_1} \qquad \searrow^{g_3}$$

$$T \qquad\qquad\qquad \sigma$$

where c is NCFMAS composition, comprising the oxides of the elements Na_2O-CaO-FeO-MgO-Al_2O_3-SiO_2.

Thermodynamic modeling

Is it possible to compute physical properties directly for a given chemical composition (c), pressure (P) and temperature (T) ?

Yes, this can be achieved through the use of thermodynamic methods based on Gibbs free energy minimisation (e.g. Perple_X, Connolly, 2005):
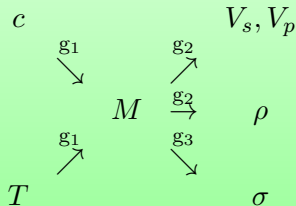
$$
\begin{array}{ccc}
c & & V_s, V_p \\
\;\;\searrow^{g_1} & \nearrow^{g_2} & \\
& M \xrightarrow{g_2} & \rho \\
\;\;\nearrow^{g_1} & \searrow^{g_3} & \\
T & & \sigma
\end{array}
$$

where c is NCFMAS composition, comprising the oxides of the elements Na_2O-CaO-FeO-MgO-Al_2O_3-SiO_2.

Thermodynamic Modeling - Equation of state

- Gruneisen model for Helmholtz energy:

$$A(V,T) = A_0 - A_c(V,T_0) + [A_{th}(V,T) - A_{th}(V,T_0)]$$

- Birch-Murnaghan "cold" part:

$$A_c = \frac{9}{2}K_0 V_0 [f^2 + (K_0 - 4)f^3], f = \frac{1}{2}\left[\left(\frac{V_0}{V}\right)^{2/3} - 1\right]$$

- Debye "thermal" part: $A_{th} = 9nRT\frac{T}{\theta^3}\int_0^{\theta/T}\ln[1 - \exp(-t)]t^2 dt$,

$$\theta = \theta_0 \exp\left[\frac{\gamma_0[1 - \left(\frac{V}{V_0}\right)^q]}{q}\right]$$

Model contains 10 species-specific parameters: $A_0, V_0, K_0, K_0^{'},$
$\theta_0, \gamma_0, q_0, G_0, G_0^{'}, \eta_{S0}$. EoS (Stixrude & Bukowinski, 1990; Stixrude &
Lithgow-Bertelloni, 2005a,b) and thermodynamic data from Xu et al.
(2008).

Thermodynamic Modeling - Physical Properties

Density

$$\rho = N \left(\frac{\partial G}{\partial P} \right)^{-1}$$

Bulk modulus

$$K_S = -\frac{\partial G}{\partial P} \left[\frac{\partial^2 G}{\partial P^2} + \frac{\left(\frac{\partial}{\partial P} \frac{\partial G}{\partial T} \right)^2}{\frac{\partial^2 G}{\partial T^2}} \right]^{-1}$$

Shear modulus is obtained from finite strain model EoS of Stixrude & Lithgow-Bertelloni (2005a,b)

$$\mu_S = f[G(\rho, T)]$$

Phase relations are most sensitive to integration constants and low-order derivatives, whereas seismic velocities are most sensitive to higher-order derivatives.

Thermodynamic modeling - phase equilibria and physical properties

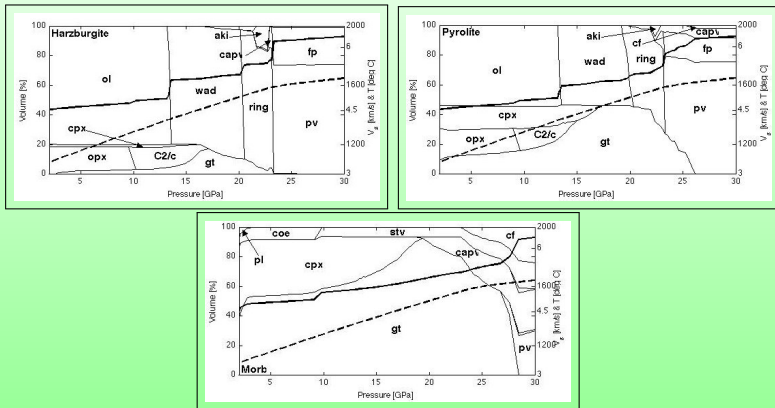$$c, T \xrightarrow{\ g_1\ } M \xrightarrow{\ g_2\ } V_P, V_S, \rho$$
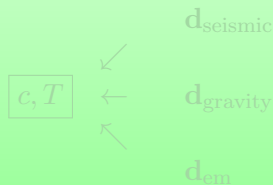

Figure: Phase equilibria and physical properties (S-wave velocities, solid line) for three different mantle compositions along a mantle adiabat (dashed line).

Putting it all together

Now, that we can compute physical properties for a given c, T (& P)

$$c, T \overset{\text{Perple\_X}}{\longrightarrow} V_P, V_S, \rho$$

it straightforward to define a quantitative approach to inferring composition and temperature from geophysical data

$$\mathbf{d}_{\text{seismic}}$$

$$\boxed{c, T} \quad \leftarrow \quad \mathbf{d}_{\text{gravity}}$$

$$\mathbf{d}_{\text{em}}$$

Putting it all together

Now, that we can compute physical properties for a given c, T (& P)

$$c, T \xrightarrow{\text{Perple\_X}} V_P, V_S, \rho$$

it straightforward to define a quantitative approach to inferring composition and temperature from geophysical data
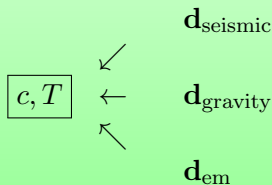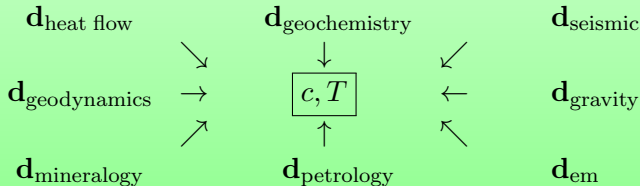
$$\boxed{c, T} \begin{array}{c} \swarrow \\ \leftarrow \\ \nwarrow \end{array} \begin{array}{l} \mathbf{d}_{\text{seismic}} \\ \\ \mathbf{d}_{\text{gravity}} \\ \\ \mathbf{d}_{\text{em}} \end{array}$$

Putting it all together

Now, that we can compute physical properties for a given c, T (& P)

$$c, T \overset{\text{Perple\_X}}{\longrightarrow} V_P, V_S, \rho$$

it straightforward to define a quantitative approach to inferring composition and temperature from geophysical data *as well as incorporating data/information from other fields of the geosciences*

$$
\begin{array}{ccccc}
\mathbf{d}_{\text{heat flow}} & & \mathbf{d}_{\text{geochemistry}} & & \mathbf{d}_{\text{seismic}} \\
 & \searrow & \downarrow & \swarrow & \\
\mathbf{d}_{\text{geodynamics}} & \rightarrow & \boxed{c, T} & \leftarrow & \mathbf{d}_{\text{gravity}} \\
 & \nearrow & \uparrow & \nwarrow & \\
\mathbf{d}_{\text{mineralogy}} & & \mathbf{d}_{\text{petrology}} & & \mathbf{d}_{\text{em}}
\end{array}
$$

Bayesian inversion of surface-wave phase-velocities: Application to the thermochemical and physical structure beneath North America
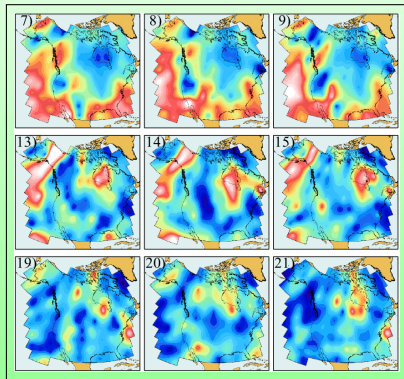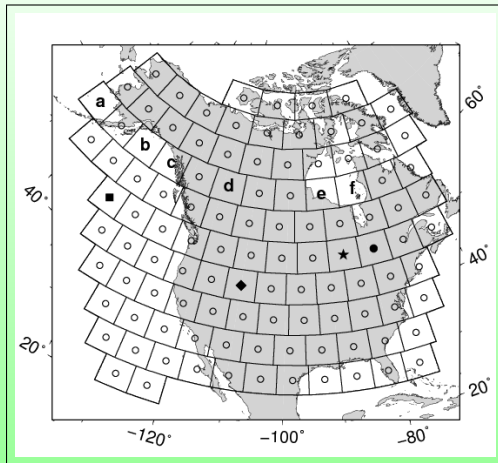
Parameterization I

Figure: Parameterization of the model. Dots at the center of each pixel denote the locations at which properties are defined laterally. Grid spacing is 5°. Radially the model is parameterized in terms of layers.

Parameterization II

At each geographical location the model is divided into crust and mantle layers, which are parameterized further using the following parameters

- ▶ Crust is divided into a 4-layer model of V_p, V_s, ρ based on CRUST2.0.
- ▶ Moho thickness
- ▶ NCFMAS composition in two mantle layers
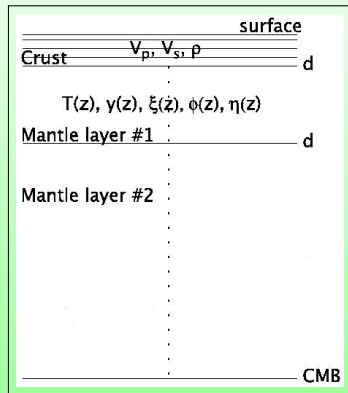- ▶ mantle temperature T in 30 layers.
- ▶ anisotropy parameters ξ, ϕ, η in 25 layers.
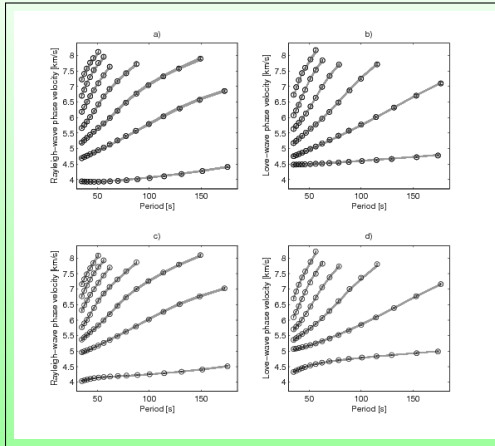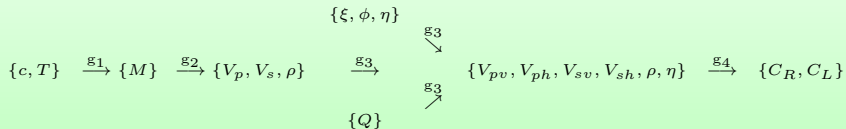

Figure: Radial parameterization.

Datafit

Figure: Comparison of calculated (gray lines) and observed Rayleigh and Love-wave phase-velocities (circles), including uncertainties (error bars) at two different locations, which are shown in figure 1 (a,b - filled square and c,d - filled circle)

Summary of forward model

A schematic illustration of the forward problem showing connection between model parameters and data via physical theory:

$$\{c,T\} \xrightarrow{g_1} \{M\} \xrightarrow{g_2} \{V_p, V_s, \rho\} \xrightarrow{g_3} \begin{array}{c} \{\xi, \phi, \eta\} \\[4pt] \searrow^{g_3} \\ \{V_{pv}, V_{ph}, V_{sv}, V_{sh}, \rho, \eta\} \\ \nearrow^{g_3} \\[4pt] \{Q\} \end{array} \xrightarrow{g_4} \{C_R, C_L\}$$

where

- $c, T, \xi, \phi, \ldots$ are model parameters,
- $g_1, g_2, \ldots$ denote physical theories, and
- C_R, C_L data.

Anisotropy

An anisotropic elastic medium in the general case is defined by 21 independent elements of the 4th order elastic tensor, whereas an isotropic solid is described by only two. In the case of transverse anisotropy (symmetry axis in vertical direction), the number of independent unknowns reduces to 5. These could be either

$$A = \rho V_{ph}^2, C = \rho V_{pv}^2, L = \rho V_{sv}^2, N = \rho V_{sh}^2, F = \frac{\eta}{A - 2L}.$$

(Love, 1927), or reparameterize the above coefficients using the Voigt average of the isotropic P and S wave velocities (Babuska & Cara, 1991)

$$V_s^2 = \frac{2V_{sv}^2 + V_{sh}^2}{3}, V_p^2 = \frac{V_{pv}^2 + 4V_{sh}^2}{5},$$

in addition to the following three anisotropy parameters

$$\xi = \frac{V_{sh}^2}{V_{sv}^2}, \phi = \frac{V_{pv}^2}{V_{ph}^2}, \eta = \frac{F}{A - 2L}. \tag{1}$$

Attenuation and dispersion

It has been found experimentally that attenuation Q is temperature dependent (e.g., Jackson et al., 2002)

$$Q = Q_0 \exp\left(\frac{\alpha(E + PV)}{RT}\right)$$

where α describes frequency dependence of Q ($\sim$0.25).

A consequence of a dissipative system is the presence of velocity dispersion due to attenuation

$$V(P, T, X, \omega) = V(P, T, X, \infty)\left[1 - \frac{2Q^{-1}}{\tan(\alpha\pi/2)}\right]$$

Solving the inverse problem

The discretized inverse problem can be written as

$$\mathbf{d} = g(\mathbf{m})$$

However,

- we do not have a closed-form mathematical expression for $g(\mathbf{m})$,
- but instead a method(s) that allows to evaluate $g(\mathbf{m})$ for a given $\mathbf{m}$:

$$\mathbf{m} \xrightarrow{g_1} \mathbf{m}' \xrightarrow{g_2} \mathbf{m}'' \xrightarrow{g_3} \ldots \xrightarrow{g_n} \mathbf{d}$$

This is a typical characteristic of non-linear inverse problems.

Solving the inverse problem

The discretized inverse problem can be written as

$$\mathbf{d} = g(\mathbf{m})$$

However,

- we do not have a closed-form mathematical expression for $g(\mathbf{m})$,
- but instead a method(s) that allows to evaluate $g(\mathbf{m})$ for a given $\mathbf{m}$:

$$\mathbf{m} \xrightarrow{g_1} \mathbf{m}' \xrightarrow{g_2} \mathbf{m}'' \xrightarrow{g_3} \ldots \xrightarrow{g_n} \mathbf{d}$$

This is a typical characteristic of non-linear inverse problems.

Solving the inverse problem

In the case of non-linear inverse problems the model space is typically also multi-modal
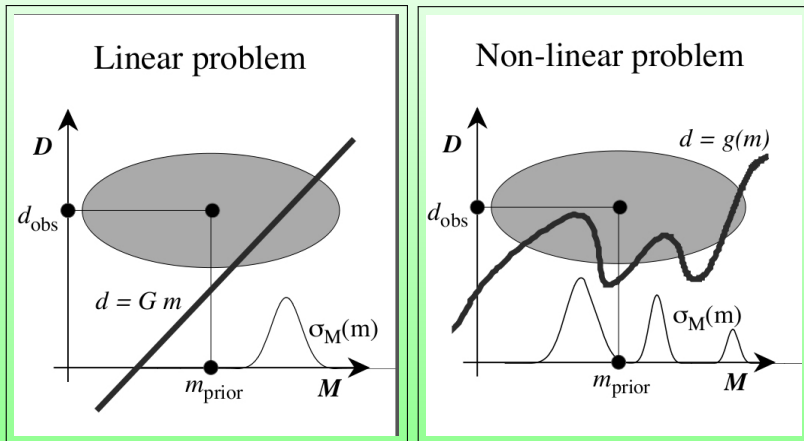
We employ a stochastic algorithm to sample the model space (=space of solutions).

Solving the inverse problem

The inverse problem formulated as a combination of independent states of information:

1. prior information on $\mathbf{m}$ obtained independently of data $\rho_m(\mathbf{m})$.

2. information obtained from (uncertain) observations $\rho_d(\mathbf{d})$.

3. the joint prior $\rho(\mathbf{d}, \mathbf{m}) = \rho_d(\mathbf{d})\rho_m(\mathbf{m})$.

4. a distribution $\theta(\mathbf{d}, \mathbf{m})$ describing an uncertain theory $\mathbf{d} \approx \mathrm{g}(\mathbf{m})$ over the joint data/model space.

combining 1-4 using the conjunction of states of information results in

$$\sigma(\mathbf{d}, \mathbf{m}) = \frac{\rho(\mathbf{d}, \mathbf{m})\theta(\mathbf{d}, \mathbf{m})}{\mu(\mathbf{d}, \mathbf{m})}$$

Solving the inverse problem

The inverse problem formulated as a combination of independent states of information:

1. prior information on $\mathbf{m}$ obtained independently of data $\rho_m(\mathbf{m})$.

2. information obtained from (uncertain) observations $\rho_d(\mathbf{d})$.

3. the joint prior $\rho(\mathbf{d}, \mathbf{m}) = \rho_d(\mathbf{d})\rho_m(\mathbf{m})$.

4. a distribution $\theta(\mathbf{d}, \mathbf{m})$ describing an uncertain theory $\mathbf{d} \approx g(\mathbf{m})$ over the joint data/model space.

combining 1-4 using the conjunction of states of information results in

$$\sigma(\mathbf{d}, \mathbf{m}) = \frac{\rho(\mathbf{d}, \mathbf{m})\theta(\mathbf{d}, \mathbf{m})}{\mu(\mathbf{d}, \mathbf{m})}$$

Solving the inverse problem

The inverse problem formulated as a combination of independent states of information:

1. prior information on $\mathbf{m}$ obtained independently of data $\rho_m(\mathbf{m})$.
2. information obtained from (uncertain) observations $\rho_d(\mathbf{d})$.
3. the joint prior $\rho(\mathbf{d}, \mathbf{m}) = \rho_d(\mathbf{d})\rho_m(\mathbf{m})$.
4. a distribution $\theta(\mathbf{d}, \mathbf{m})$ describing an uncertain theory $\mathbf{d} \approx g(\mathbf{m})$ over the joint data/model space.

combining 1-4 using the conjunction of states of information results in

$$\sigma(\mathbf{d}, \mathbf{m}) = \frac{\rho(\mathbf{d}, \mathbf{m})\theta(\mathbf{d}, \mathbf{m})}{\mu(\mathbf{d}, \mathbf{m})}$$

Solving the inverse problem

The inverse problem formulated as a combination of independent states of information:

1. prior information on $\mathbf{m}$ obtained independently of data $\rho_m(\mathbf{m})$.

2. information obtained from (uncertain) observations $\rho_d(\mathbf{d})$.

3. the joint prior $\rho(\mathbf{d}, \mathbf{m}) = \rho_d(\mathbf{d})\rho_m(\mathbf{m})$.

4. a distribution $\theta(\mathbf{d}, \mathbf{m})$ describing an uncertain theory $\mathbf{d} \approx \mathbf{g}(\mathbf{m})$ over the joint data/model space.

combining 1-4 using the conjunction of states of information results in

$$\sigma(\mathbf{d}, \mathbf{m}) = \frac{\rho(\mathbf{d}, \mathbf{m})\theta(\mathbf{d}, \mathbf{m})}{\mu(\mathbf{d}, \mathbf{m})}$$
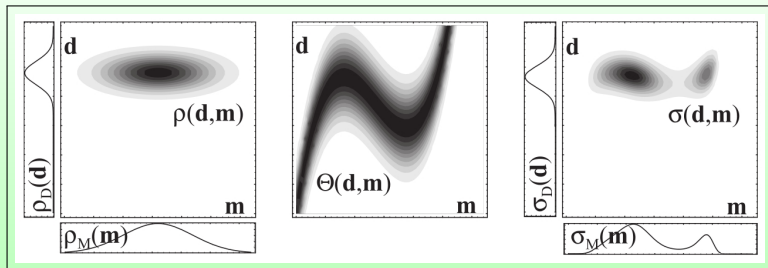
Solving the inverse problem

The inverse problem formulated as a combination of independent states of information:

1. prior information on $\mathbf{m}$ obtained independently of data $\rho_m(\mathbf{m})$.
2. information obtained from (uncertain) observations $\rho_d(\mathbf{d})$.
3. the joint prior $\rho(\mathbf{d}, \mathbf{m}) = \rho_d(\mathbf{d})\rho_m(\mathbf{m})$.
4. a distribution $\theta(\mathbf{d}, \mathbf{m})$ describing an uncertain theory $\mathbf{d} \approx \mathrm{g}(\mathbf{m})$ over the joint data/model space.

combining 1-4 using the conjunction of states of information results in

$$\sigma(\mathbf{d}, \mathbf{m}) = \frac{\rho(\mathbf{d}, \mathbf{m})\theta(\mathbf{d}, \mathbf{m})}{\mu(\mathbf{d}, \mathbf{m})}$$

Solving the inverse problem

The inverse problem formulated as a combination of independent states of information:

1. prior information on $\mathbf{m}$ obtained independently of data $\rho_m(\mathbf{m})$.
2. information obtained from (uncertain) observations $\rho_d(\mathbf{d})$.
3. the joint prior $\rho(\mathbf{d}, \mathbf{m}) = \rho_d(\mathbf{d})\rho_m(\mathbf{m})$.
4. a distribution $\theta(\mathbf{d}, \mathbf{m})$ describing an uncertain theory $\mathbf{d} \approx \mathrm{g}(\mathbf{m})$ over the joint data/model space.

combining 1-4 using the conjunction of states of information results in

$$\sigma(\mathbf{d}, \mathbf{m}) = \frac{\rho(\mathbf{d}, \mathbf{m})\theta(\mathbf{d}, \mathbf{m})}{\mu(\mathbf{d}, \mathbf{m})}$$

Solving the inverse problem

Figure: Combination of states of information. From Tarantola (2005).

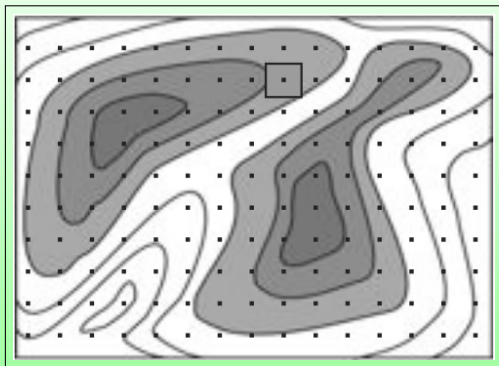which for many applications is typically written

$$\sigma_m(\mathbf{m}) = k\rho_m(\mathbf{m})L(\mathbf{m})$$

$$L(\mathbf{m}) = k\exp\left[-S(\mathbf{m})\right]$$

$$S(\mathbf{m}) = \frac{1}{2}\left[(\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs})^t\mathbf{C}_D^{-1}(\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs})\right]$$

Solving the inverse problem
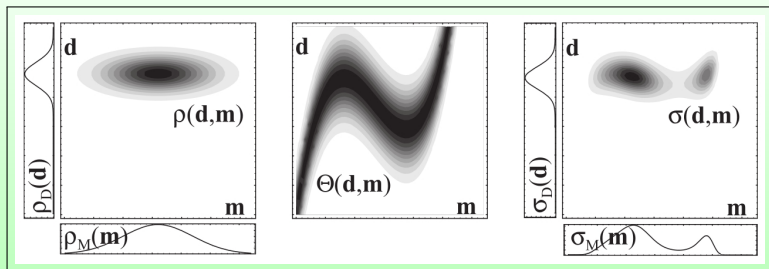
Figure: Combination of states of information. From Tarantola (2005).

which for many applications is typically written

$$\sigma_m(\mathbf{m}) = k\rho_m(\mathbf{m})L(\mathbf{m})$$

$$L(\mathbf{m}) = k \exp\left[-S(\mathbf{m})\right]$$

$$S(\mathbf{m}) = \frac{1}{2}\left[(\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs})^t\mathbf{C}_{\mathrm{D}}^{-1}(\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs})\right]$$

Sampling the posterior

Let us perform a grid search over the entire model space.

Sampling the posterior

Instead, let us try a sampling-based method such as importance sampling (Metropolis, Gibbs).
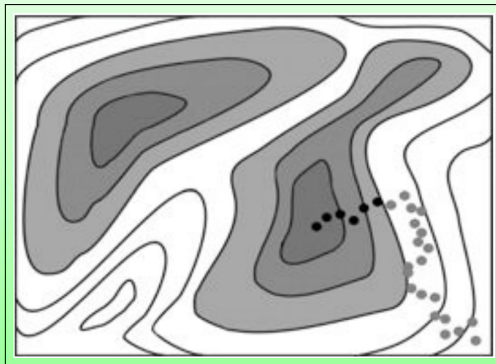
There are several algorithms available: crude Monte Carlo (random search), genetic algorithms, Importance sampling methods (Gibbs, Metropolis), Neighbourhood algorithm, simulated annealing, etc.

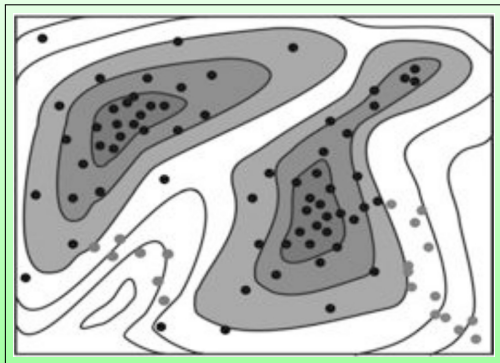# Sampling the posterior

Importance sampling using Metropolis algorithm
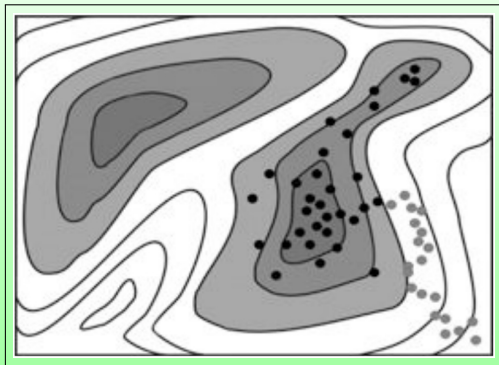
- if $L(\mathbf{m}_j) \geq L(\mathbf{m}_i)$, accept proposed transition $i \to j$.
- if $L(\mathbf{m}_j) < L(\mathbf{m}_i)$, accept proposed transition with probability
$P_{i \to j} = \frac{L(\mathbf{m}_j)}{L(\mathbf{m}_i)}$

Sampling the posterior

Importance sampling

Sampling the posterior

This became Albert's preferred idea - representing probability densities with samples from the probability distribution.
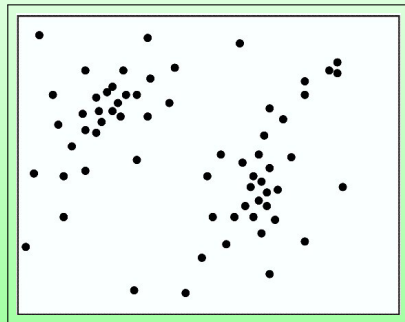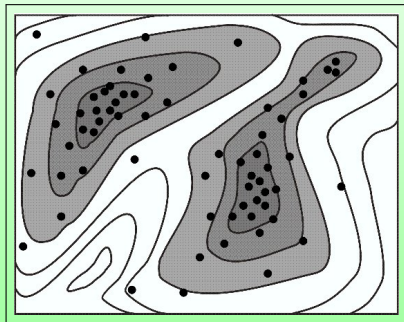
The solution is not one model but a collection of models that are consistent with both prior information and data

Analysis of the posterior distribution

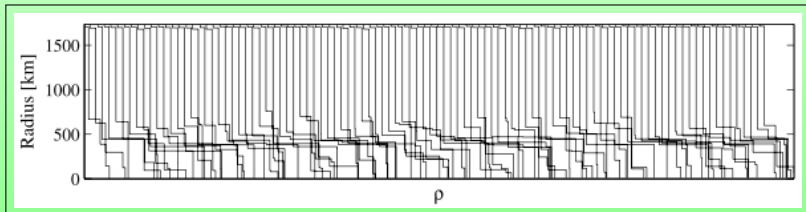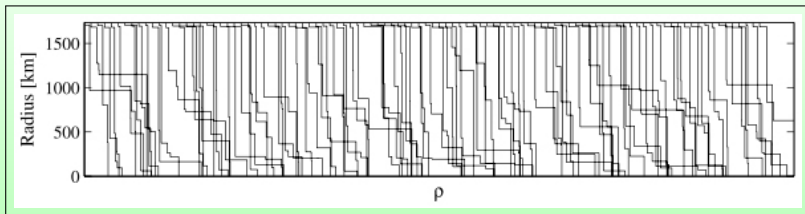1) displaying prior and posterior models (showing movies)

Figure: 100 lunar density models from the prior (top) and posterior (bottom) pdf. From Khan et al. (2004).

Analysis of the posterior distribution

2) calculation of resolution measures

$$\mathcal{R}(\Lambda, f) = \int_\Lambda f(\mathbf{m})\sigma(\mathbf{m})d\mathbf{m} \approx \frac{1}{N} \sum_{\{n|m_n \in \Lambda\}} f(\mathbf{m}_n) \qquad (2)$$

3) Bayesian hypothesis testing

Given two hypotheses $\mathcal{H}_i$, $\mathcal{H}_j$, the Bayes factor, $\mathcal{B}_{ij}$ in favour of $\mathcal{H}_i$ (and against $\mathcal{H}_j$) is given by the posterior to prior odds ratio.

$$\mathcal{B}_{ij}(\mathbf{d}) = \frac{\mathcal{P}(\mathbf{d}|\mathcal{H}_i)}{\mathcal{P}(\mathbf{d}|\mathcal{H}_j)} = \frac{\mathcal{P}(\mathcal{H}_i|\mathbf{d})/\mathcal{P}(\mathcal{H}_j|\mathbf{d})}{\mathcal{P}(\mathcal{H}_i)/\mathcal{P}(\mathcal{H}_j)} \qquad (3)$$

Analysis of the posterior distribution

2) calculation of resolution measures

$$\mathcal{R}(\Lambda, f) = \int_{\Lambda} f(\mathbf{m})\sigma(\mathbf{m})d\mathbf{m} \approx \frac{1}{N} \sum_{\{n | m_n \in \Lambda\}} f(\mathbf{m}_n) \qquad (2)$$

3) Bayesian hypothesis testing

Given two hypotheses $\mathcal{H}_i$, $\mathcal{H}_j$, the Bayes factor, $\mathcal{B}_{ij}$ in favour of $\mathcal{H}_i$ (and against $\mathcal{H}_j$) is given by the posterior to prior odds ratio.

$$\mathcal{B}_{ij}(\mathbf{d}) = \frac{\mathcal{P}(\mathbf{d}|\mathcal{H}_i)}{\mathcal{P}(\mathbf{d}|\mathcal{H}_j)} = \frac{\mathcal{P}(\mathcal{H}_i|\mathbf{d})/\mathcal{P}(\mathcal{H}_j|\mathbf{d})}{\mathcal{P}(\mathcal{H}_i)/\mathcal{P}(\mathcal{H}_j)} \qquad (3)$$
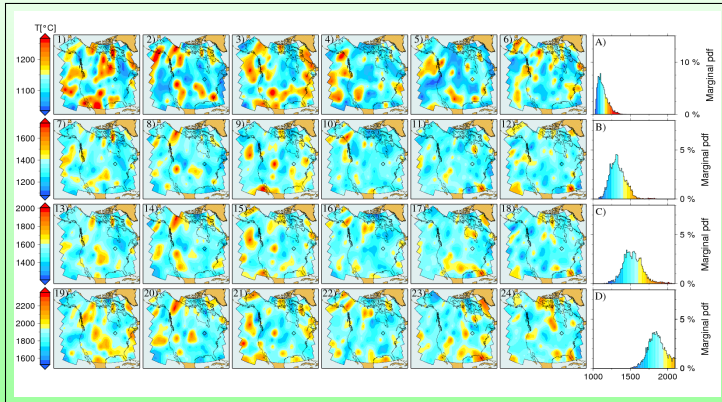
Prior thermal movie

Figure: Figure: Prior mantle temperatures at 100, 300, 500 and 1000 km depth. Plots A-D show marginal prior 1D pdf's for the location indicated by the circle on the maps.
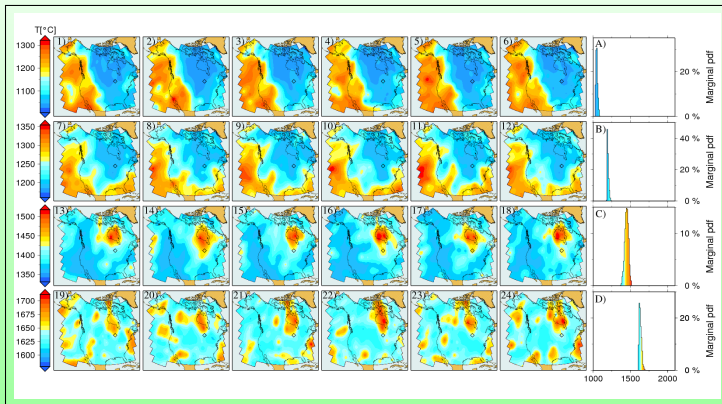
Posterior thermal movie

Figure: Posterior mantle temperatures at 100, 300, 500 and 1000 km depth. Plots A-D show posterior 1D pdf's for the location indicated by the circle on the maps (Khan et al., 2011).
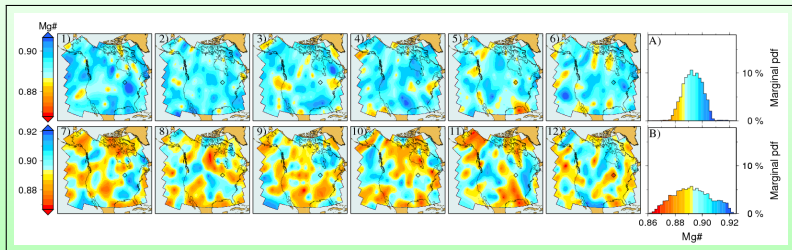
Prior compositional movie

Figure: Prior mantle compositions in the upper (first panel) and lower mantle (second panel). Plots A,B show marginal prior 1D pdf's for the location indicated by the circle on the maps.
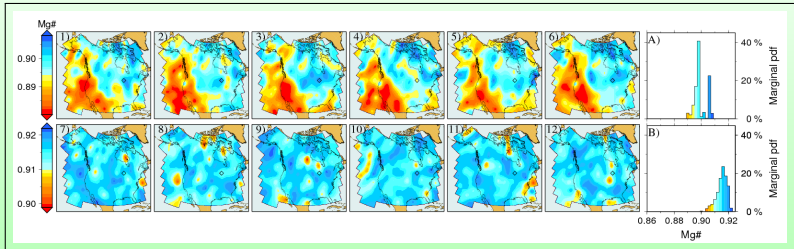
Posterior compositional movie

Figure: Posterior mantle compositions in the upper (first panel) and lower mantle (second panel). Plots A,B show marginal posterior 1D pdf's for the location indicated by the circle on the maps.

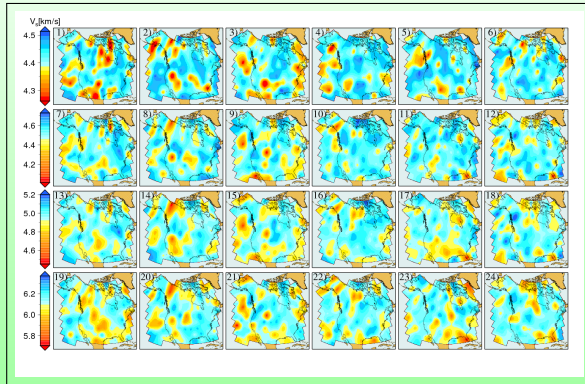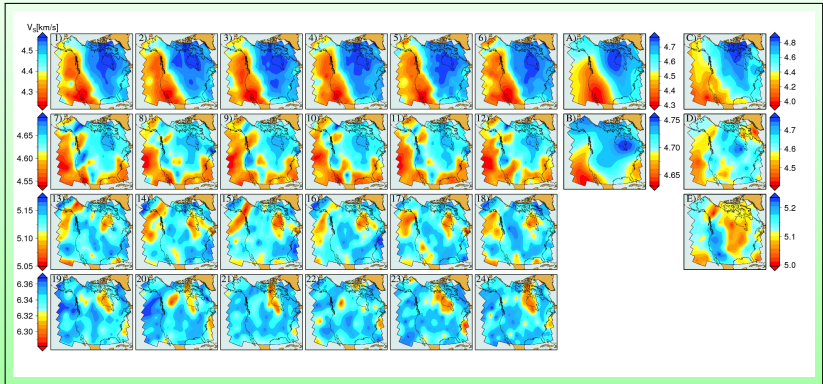# Prior shear-wave velocity movie

Figure: Prior isotropic shear-wave velocity movie. In each panel the six maps represent six shear-wave velocity models that are picked randomly from the prior distribution at depths of 100 km (1-6), 300 km (7-12), 500 km (13-18) and 1000 km (19-24), respectively (from Khan et al., 2011a).

Posterior shear-wave velocity movie

Posterior isotropic shear-wave velocity movie at depths of 100 km (1-6), 300 km (7-12), 500 km (13-18) and 1000 km (19-24). Plots A (at 100 km depth) and B (at 300 km depth) - Nettles & Dziewonski (2008), plots C (100 km), D (300 km) and E (500 km) - Yuan et al. (2010)).
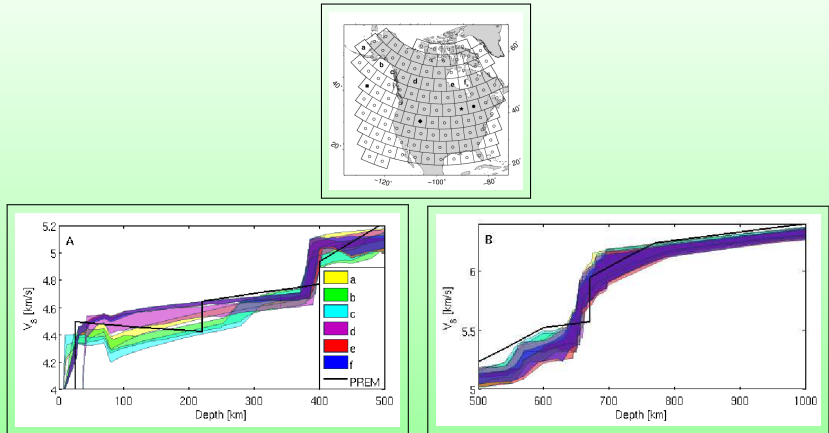
Posterior shear-wave velocity profiles

Figure: Selected shear-wave velocity models beneath different tectonic settings in the upper mantle and transition zone (A) and lower transition zone and mantle (B) : Oceanic (a-c), young continent (d) and old stable continent (e,f). Profiles encompass all sampled models. Geographic location of letters are indicated in figure 1.

Posterior filtering: geoid anomalies

Additional geophysical data can be employed as a tool to refine and narrow the collection of tomographic models.
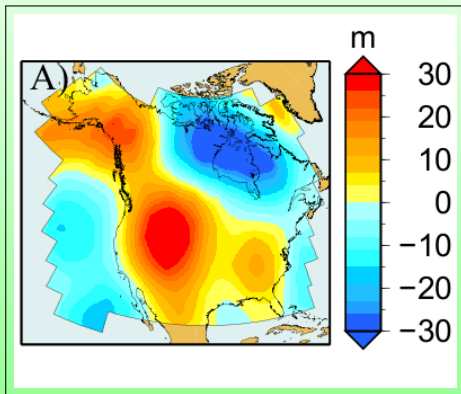

Figure: Observed geoid GGM02 of Tapley et al. (2005) for harmonic degrees 6-20.
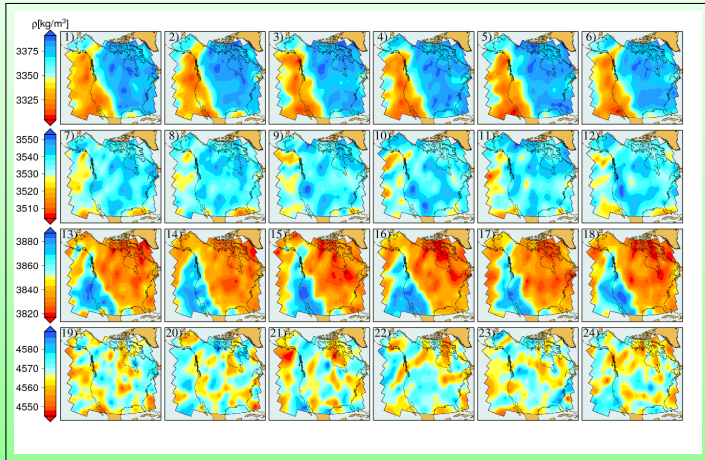
Posterior density movie

Figure: Posterior density movie at depths of 100 km (1-6), 300 km (7-12), 500 km (13-18), 1000 km (19-24).

Posterior filtering: geoid anomalies

Geoid anomalies are computed from

$$\delta N(\theta, \phi) = \frac{3}{4\pi\rho_m} \int_{r_{\mathrm{CMB}}}^{R} \int_{0}^{2\pi} \int_{-\pi/2}^{\pi/2} K_g(\Delta, r)\delta\rho(r, \theta', \phi') \sin\theta' d\theta' d\phi' dr \tag{4}$$

whereas geoid kernels, expressing the relationship between surface observables and internal density loads that drive the flow, are given by

$$K_g(\Delta, r) = \sum_{l=l_1}^{l_2} G_l(r)\mathrm{P}_l^0(\cos\Delta) \tag{5}$$

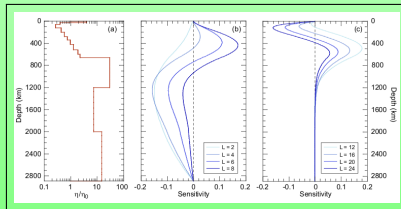where $\mathrm{P}_l^0(\cos\Delta)$ and $G_l(r)$ are Legendre Polynomials and radial geoid kernels. The latter are computed using the viscous flow theory of Forte (2000) and thus depend on the radial viscosity profile.

Posterior filtering: geoid anomalies

Radial viscosity profiles using $\eta = \eta_0 \exp(-(E + PV)/RT)$
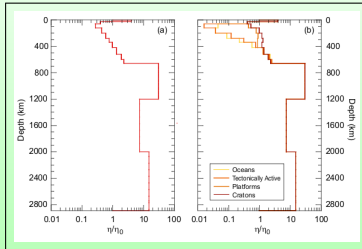
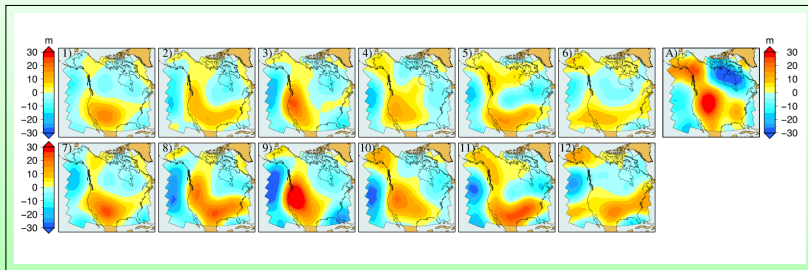Computed geoid kernels

Posterior filtering: geoid anomalies

Figure: Reconstructed geoid anomalies for the six posterior density models using a continental average viscosity profile (plots 1-6) and regionally averaged viscosity profiles (plots 7-12). Only harmonic degrees 6 to 20 are used. For comparison, plot A shows the observed geoid GGM02 of Tapley et al. (2005) for the same harmonic degrees.

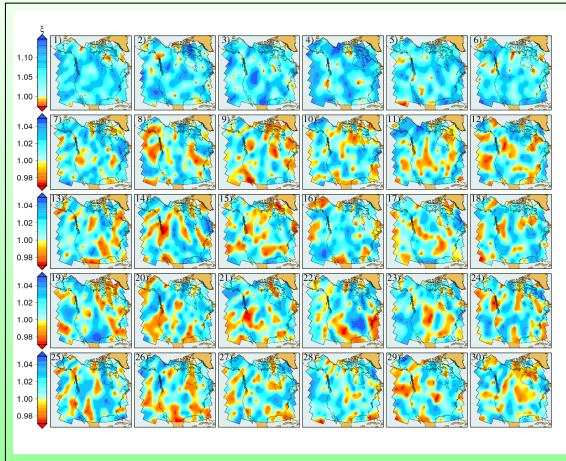# Prior anisotropic shear-wave velocity movie

Figure: Prior shear-wave anisotropy movie. In each panel the six maps represent six different shear-wave anisotropy models that are picked randomly from the prior distribution at depths of 100 km (1-6), 300 km (7-12), 500 km (13-18) and 1000 km (19-24), respectively.

Posterior anisotropic shear-wave velocity movie
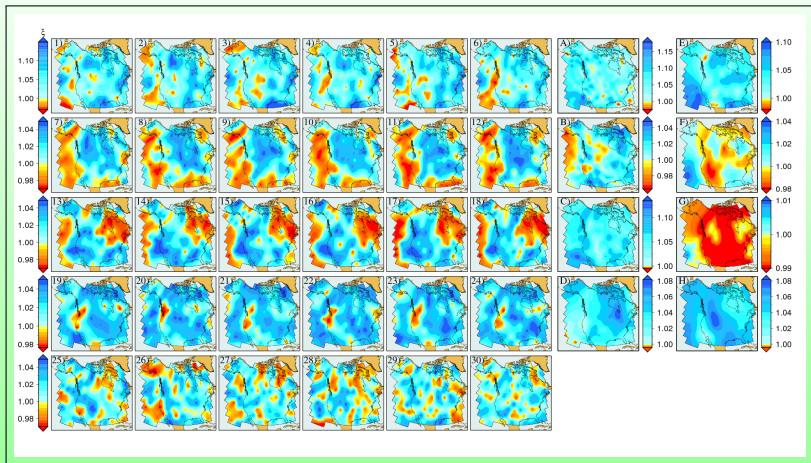
Figure: Posterior shear-wave anisotropy movie at depths of 100 km (1-6), 300 km (7-12), 500 km (13-18), 1000 km (19-24) and 1800 km (25-30). Plots A (at 100 km depth) and B (at 300 km depth) - Nettles & Dziewonski (2008), plots C (100 km depth) and D (300 km depth) - Boschi & Ekström (2002), plots E (100 km), F (300 km) and G (500 km) - Yuan et al. (2010) and plot (H) - Marone et al. (2007) at 100 km depth. Note differences in colourbars.

Conclusion

- ▶ Thermal as well as compositional variations are needed in order to fit data,

- ▶ The thermo-chemical and physical structure of the North American upper mantle follows the surface tectonic age-division closely,

- ▶ the old stable continental parts are cold and Fe-depleted, while the tectonically younger continental regions and oceanic lithosphere appear to be relatively hot and Fe-enriched,

- ▶ shear-wave velocity differences between oceans and continents disappear around 300 km depth,

- ▶ within the transition zone a decoupling of the structure (thermo-chemical and anisotropic) from that of the upper mantle is accompanied by a decrease in amplitude of velocity anomalies,

- ▶ absence of strong heterogeneities in the lower mantle, with evidence for compositionally distinct upper and lower mantles,

Conclusion

- ▶ Thermal as well as compositional variations are needed in order to fit data,

- ▶ The thermo-chemical and physical structure of the North American upper mantle follows the surface tectonic age-division closely,

- ▶ the old stable continental parts are cold and Fe-depleted, while the tectonically younger continental regions and oceanic lithosphere appear to be relatively hot and Fe-enriched,

- ▶ shear-wave velocity differences between oceans and continents disappear around 300 km depth,

- ▶ within the transition zone a decoupling of the structure (thermo-chemical and anisotropic) from that of the upper mantle is accompanied by a decrease in amplitude of velocity anomalies,

- ▶ absence of strong heterogeneities in the lower mantle, with evidence for compositionally distinct upper and lower mantles,

Conclusion

- ▶ Thermal as well as compositional variations are needed in order to fit data,

- ▶ The thermo-chemical and physical structure of the North American upper mantle follows the surface tectonic age-division closely,

- ▶ the old stable continental parts are cold and Fe-depleted, while the tectonically younger continental regions and oceanic lithosphere appear to be relatively hot and Fe-enriched,

- ▶ shear-wave velocity differences between oceans and continents disappear around 300 km depth,

- ▶ within the transition zone a decoupling of the structure (thermo-chemical and anisotropic) from that of the upper mantle is accompanied by a decrease in amplitude of velocity anomalies,

- ▶ absence of strong heterogeneities in the lower mantle, with evidence for compositionally distinct upper and lower mantles,

Conclusion

- ▶ Thermal as well as compositional variations are needed in order to fit data,

- ▶ The thermo-chemical and physical structure of the North American upper mantle follows the surface tectonic age-division closely,

- ▶ the old stable continental parts are cold and Fe-depleted, while the tectonically younger continental regions and oceanic lithosphere appear to be relatively hot and Fe-enriched,

- ▶ shear-wave velocity differences between oceans and continents disappear around 300 km depth,

- ▶ within the transition zone a decoupling of the structure (thermo-chemical and anisotropic) from that of the upper mantle is accompanied by a decrease in amplitude of velocity anomalies,

- ▶ absence of strong heterogeneities in the lower mantle, with evidence for compositionally distinct upper and lower mantles,

Conclusion

- ▶ Thermal as well as compositional variations are needed in order to fit data,

- ▶ The thermo-chemical and physical structure of the North American upper mantle follows the surface tectonic age-division closely,

- ▶ the old stable continental parts are cold and Fe-depleted, while the tectonically younger continental regions and oceanic lithosphere appear to be relatively hot and Fe-enriched,

- ▶ shear-wave velocity differences between oceans and continents disappear around 300 km depth,

- ▶ within the transition zone a decoupling of the structure (thermo-chemical and anisotropic) from that of the upper mantle is accompanied by a decrease in amplitude of velocity anomalies,

- ▶ absence of strong heterogeneities in the lower mantle, with evidence for compositionally distinct upper and lower mantles,

Conclusion

- ▶ Thermal as well as compositional variations are needed in order to fit data,
- ▶ The thermo-chemical and physical structure of the North American upper mantle follows the surface tectonic age-division closely,
- ▶ the old stable continental parts are cold and Fe-depleted, while the tectonically younger continental regions and oceanic lithosphere appear to be relatively hot and Fe-enriched,
- ▶ shear-wave velocity differences between oceans and continents disappear around 300 km depth,
- ▶ within the transition zone a decoupling of the structure (thermo-chemical and anisotropic) from that of the upper mantle is accompanied by a decrease in amplitude of velocity anomalies,
- ▶ absence of strong heterogeneities in the lower mantle, with evidence for compositionally distinct upper and lower mantles,

Conclusion continued

- anisotropic upper mantle structure is similar to previous studies, where old stable continents are characterized by positive shear-wave anisotropy, whereas younger areas are characterized by negative shear-wave anisotropy. In the transition zone anisotropy appears to reverse, which likely reflects the presence of distinct anisotropic layers in the mantle,

- testing posterior tomographic models using geoid anomalies, which are sensitive to density, presents a promising tool for refining the collection of sampled tomographic and thermo-chemical models. A current limitation, however, is the accuracy of the reconstructed geoid, which requires a good knowledge of the mantle viscosity structure.

Conclusion continued

- anisotropic upper mantle structure is similar to previous studies, where old stable continents are characterized by positive shear-wave anisotropy, whereas younger areas are characterized by negative shear-wave anisotropy. In the transition zone anisotropy appears to reverse, which likely reflects the presence of distinct anisotropic layers in the mantle,

- testing posterior tomographic models using geoid anomalies, which are sensitive to density, presents a promising tool for refining the collection of sampled tomographic and thermo-chemical models. A current limitation, however, is the accuracy of the reconstructed geoid, which requires a good knowledge of the mantle viscosity structure.

Acknowledgements

Computation of electrical conductivity structure

We can measure the electrical conductivity for a set of relevant mantle minerals in the laboratory, which is found to follow an Arrhenius type relation.
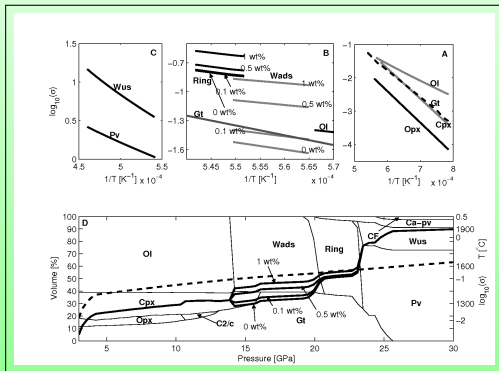

Upper mantle:
$$\sigma = \sigma_{\mathrm{o}} \exp(-\Delta H/kT)$$
Transition zone:
$$\sigma = AC_w \exp(-[\Delta H - \alpha C_w^{1/3}]/kT)$$
Lower mantle:
$$\sigma = \sigma_{\mathrm{o}} T^m \exp(-[E + PV]/kT)$$

Figure: Electrical conductivity profiles as a function of major element composition, water content and temperature (from Khan et al., 2011b).
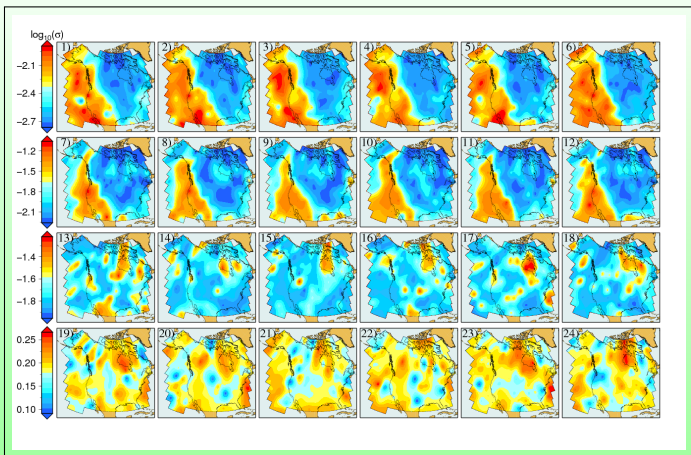
Posterior electrical conductivity movie

Figure: Posterior electrical conductivity movie at depths of 100 km (1-6), 300 km (7-12), 500 km (13-18), 1000 km (19-24).

Posterior filtering: 3D EM C-responses

3D EM response functions are computed using the integral-equation approach of Kuvshinov & Semenov (2012).
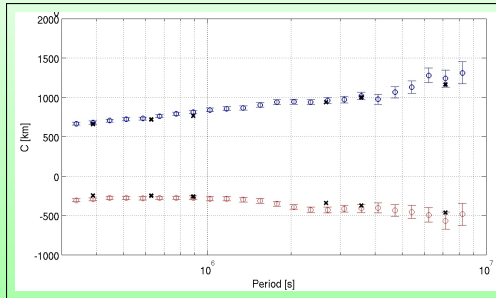

Figure: Comparison of computed and observed 3D EM C-responses for station TUC (North America).