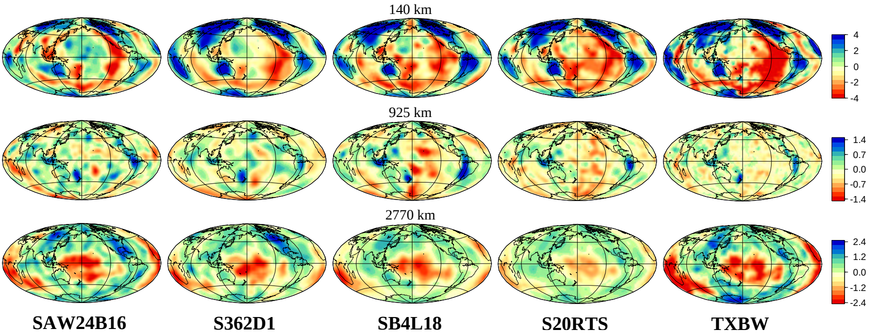
# Least-squares criterion: variations on a theme

Éric Beucler

Laboratoire de Planétologie et Géodynamique
Université de Nantes - France

SPICE Research & Training workshop III
Kinsale (Ireland, July 2006)

Thanks to
J.-P. Montagner, A. Tarantola, J. H. Woodhouse,
and M. Cara, J. Trampert, and H. van Heijst.




Lateral variations in *S* velocity at three sample depths (Romanowicz, 2003).

- SAW24B16: Mégnin & Romanowicz (2000)
- S362D1: Gu, Dziewonski, Su & Ekström (2001)
- SB4L18: Masters, Bolton & Laske (1999)
- S20RTS: Ritsema, van Heijst & Woodhouse (1999)
- TXBW: Grand (2002)

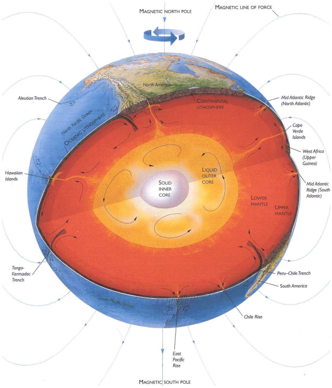
# Why such differences?

- The physical nature of the data sets (data space):
  - *P* and/or *S* body waves;
  - Rayleigh and/or Love surface waves;
  - normal-mode splitting measurements.
- The mathematical relation between the data and the model spaces (forward problem);
- The choice and the sensitivity of the parameters (model space);
- The inversion scheme.

# Why such differences?

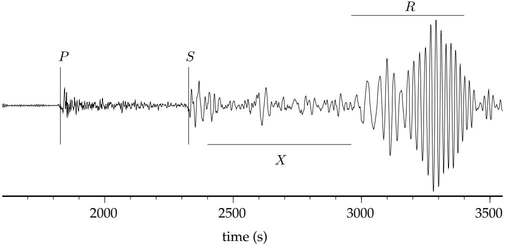
- The physical nature of the data sets (data space):
  - *P* and/or *S* body waves;
  - Rayleigh and/or Love surface waves;
  - normal-mode splitting measurements.
- The mathematical relation between the data and the model spaces (forward problem);
- The choice and the sensitivity of the parameters (model space);
- The inversion scheme.

# Surface waves: The physical insight




Lamb & Sington (1998).

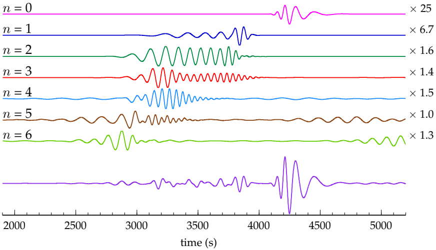
Vertical broad-band component recorded at DRV



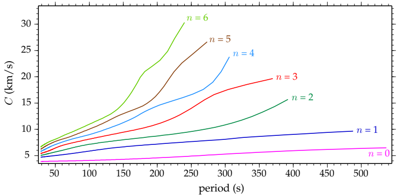

Event located in New-Guinea, May 16th 1999.
$\Delta \approx 6925\,\text{km}$.

- Good coverage at global scale;
- well suited for the detection of large-scale heterogeneities;
- dispersive waves.

# Surface waves: The mathematical insight

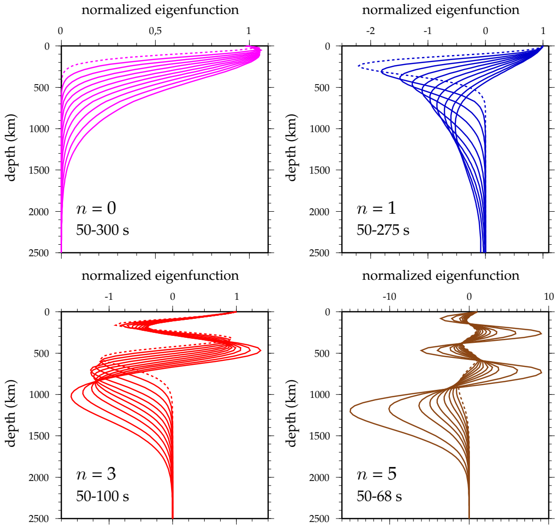



Contribution of each synthetic monomode computed in PREM.




Phase velocities of the first seven Rayleigh modes computed in PREM.

# Surface waves: The mathematical insight




Displacement as a function of depth for several spheroidal mode branches. Computations are made in PREM.

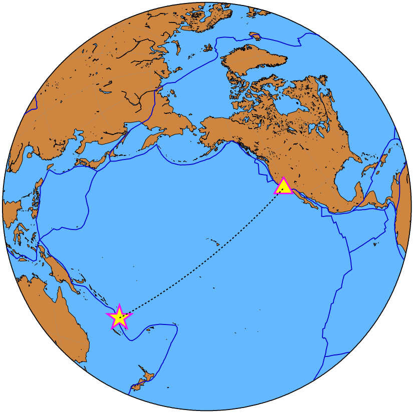
1. Introduction
   - 3D global models
   - Surface waves
2. Path averaged phase velocity measurements
   - The forward problem
   - The roller coaster approach
   - Results and comparisons
3. Phase velocity maps
   - The forward problem
   - The CLASH
   - Synthetic tests
   - Real data set inversions

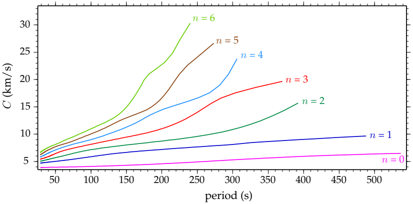
1. Introduction
   - 3D global models
   - Surface waves

2. Path averaged phase velocity measurements
   - The forward problem
   - The roller coaster approach
   - Results and comparisons

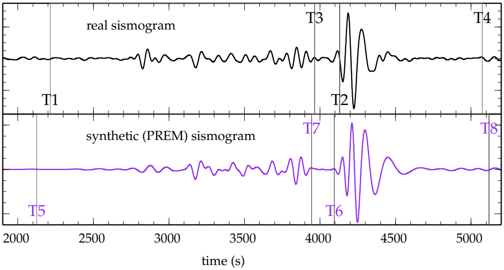
3. Phase velocity maps
   - The forward problem
   - The CLASH
   - Synthetic tests
   - Real data set inversions

# Path averaged phase velocity measurements



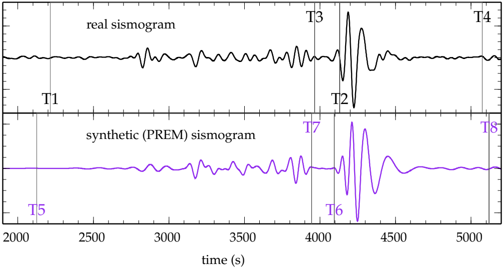


# The forward problem



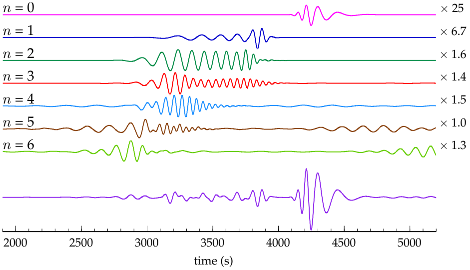

$$
\begin{aligned}
\mathbf{d} &= g(\mathbf{p}), \\
A^{(R)}(\vec{r},\omega)\exp\left[i\phi^{(R)}(\vec{r},\omega)\right] &= \sum_{j=0}^{n} A_j^{(S)}(\vec{r},\omega)\exp\left[i\phi_j^{(S)}(\vec{r},\omega)\right] \\
&\quad \times \exp\left[\frac{i\omega a\Delta}{C_j^{(S)}(\omega)}\,\frac{\delta C(\vec{r},\omega)}{C^{(R)}(\vec{r},\omega)}\right].
\end{aligned}
$$

# The forward problem




$$
\begin{aligned}
\mathbf{d} &= g(\mathbf{p}), \\
A^{(R)}(\vec{r},\omega)\exp\left[i\phi^{(R)}(\vec{r},\omega)\right] &= \sum_{j=0}^{n} A_j^{(S)}(\vec{r},\omega)\exp\left[i\phi_j^{(S)}(\vec{r},\omega)\right] \\
&\quad \times \exp\left[\frac{i\omega a\Delta}{C_j^{(S)}(\omega)}\,\frac{\delta C(\vec{r},\omega)}{C^{(R)}(\vec{r},\omega)}\right].
\end{aligned}
$$

# The forward problem




$$
\begin{aligned}
\mathbf{d} &= g(\mathbf{p}), \\
A^{(R)}(\vec{r},\omega)\exp\left[i\phi^{(R)}(\vec{r},\omega)\right] &= \sum_{j=0}^{n} A_j^{(S)}(\vec{r},\omega)\exp\left[i\phi_j^{(S)}(\vec{r},\omega)\right] \\
&\quad \times \exp\left[\frac{i\omega a\Delta}{C_j^{(S)}(\omega)}\mathbf{p}_j(\vec{r},\omega)\right].
\end{aligned}
$$

# The forward problem

$$A^{(R)}(\vec{r},\omega)\exp\left[i{\color{blue}\phi^{(R)}}(\vec{r},\omega)\right] = \sum_{j=0}^{n} A_j^{(S)}(\vec{r},\omega)\exp\left[i\phi_j^{(S)}(\vec{r},\omega)\right] \times \exp\left[\frac{i\omega a\Delta}{C_j^{(S)}(\omega)}{\color{red}\mathbf{p}}_j(\vec{r},\omega)\right].$$

- **Under-determination**
  - to impose correlations between parameters

and/or

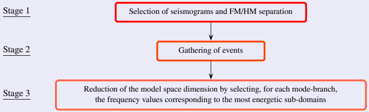
  - to increase the amount of independant data.

- **Non-linearity**
  - to find a way to explore the model space

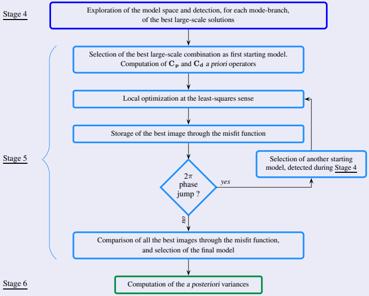
and

  - to detect the best model.

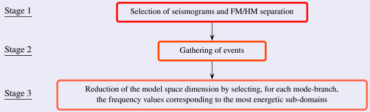
# Decreasing the under-determination



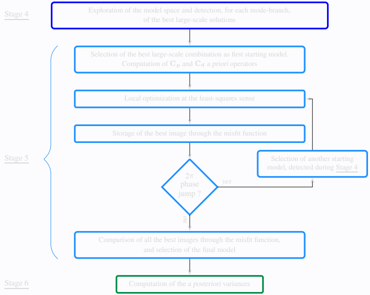

# Addressing the non-linearity



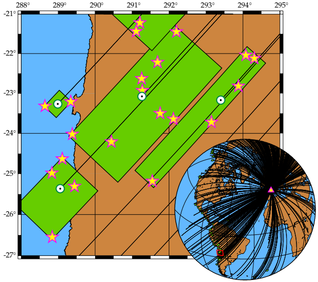
# Decreasing the under-determination



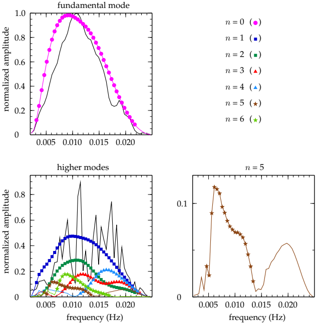

# Addressing the non-linearity



- Sufficiently close events are gathered;
- selection of the most energetic lobes for each mode-branch.

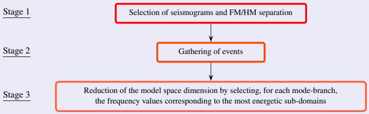



Events recorded at SSB (GEOSCOPE) between 24/08/1988 and 01/09/2001.

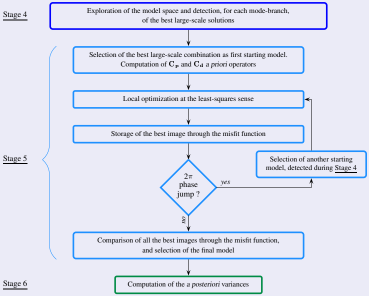



Normalized spectral amplitudes of data and synthetic mode-branches seismograms.

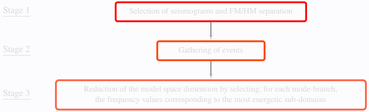
# Decreasing the under-determination



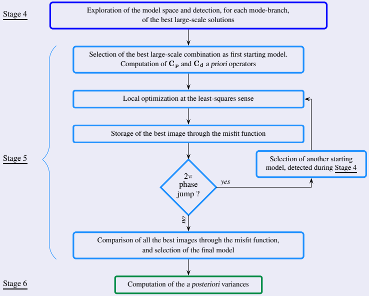

# Addressing the non-linearity



## Decreasing the under-determination

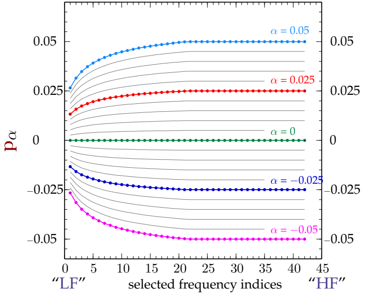



## Addressing the non-linearity

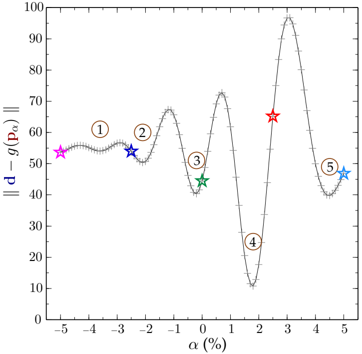


# Non-linearity (1)

Exploration of the model space and detection of large-scale solutions




Tested configurations of the FM parameter vector during large-scale exploration.




Result of the FM large-scale exploration (Beucler *et al.*, 2003).

# Non-linearity (2)

For each large-scale solution, the problem is now weakly non-linear
$\rightarrow$ **least-squares optimization** (Tarantola & Valette, 1982).

The parameter (or model), at the *k*th iteration, is given by

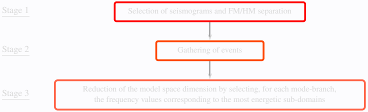
$$\mathbf{p}_k = \mathbf{p}_0 + \mathbf{C_p}\mathbf{G}_{k-1}^T \left[\mathbf{C_d} + \mathbf{G}_{k-1}\mathbf{C_p}\mathbf{G}_{k-1}^T\right]^{-1} \left[\mathbf{d} - g(\mathbf{p}_{k-1}) + \mathbf{G}_{k-1}(\mathbf{p}_{k-1} - \mathbf{p}_0)\right],$$

with $\mathbf{C_p}$ and $\mathbf{C_d}$ the *a priori* operators on parameters and data, respectively, $\mathbf{p}_0$ the starting model (*ie.* large scale solution), and $\mathbf{G}$ the partial derivative matrix: $\mathbf{G}_k = \frac{\partial g(\mathbf{p}_k)}{\partial \mathbf{p}_k}$.

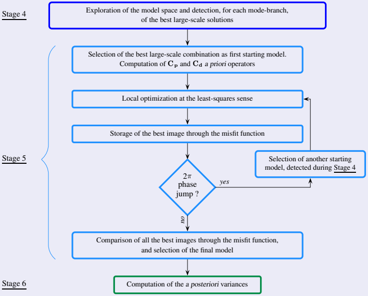
**Convergence threshold** $\rightarrow \| S(\mathbf{p}_k) - S(\mathbf{p}_{k-1}) \| \leqslant \varepsilon$,
where $S$ denotes the misfit function whose image is defined by

$$S(\mathbf{p}_k) = \frac{1}{2}\left[(g(\mathbf{p}_k) - \mathbf{d})^T\, \mathbf{C_d}^{-1}\,(g(\mathbf{p}_k) - \mathbf{d}) + (\mathbf{p}_k - \mathbf{p}_0)^T\, \mathbf{C_p}^{-1}\,(\mathbf{p}_k - \mathbf{p}_0)\right].$$

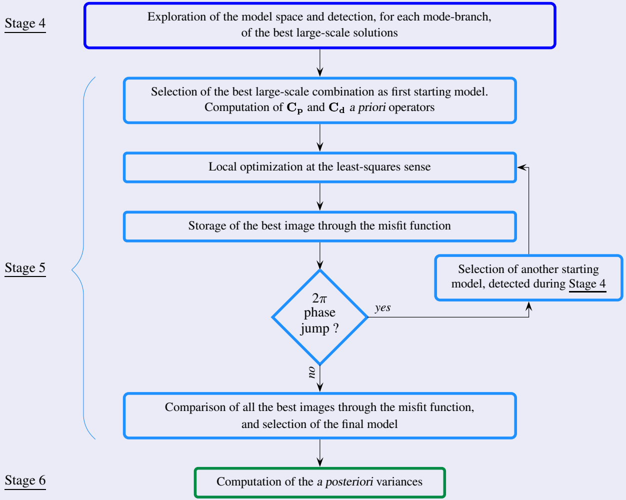
# Decreasing the under-determination



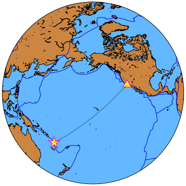

# Addressing the non-linearity

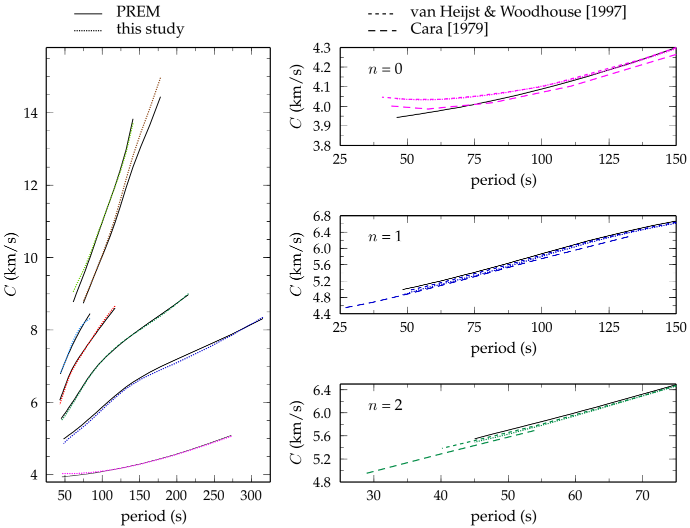


# Addressing the non-linearity



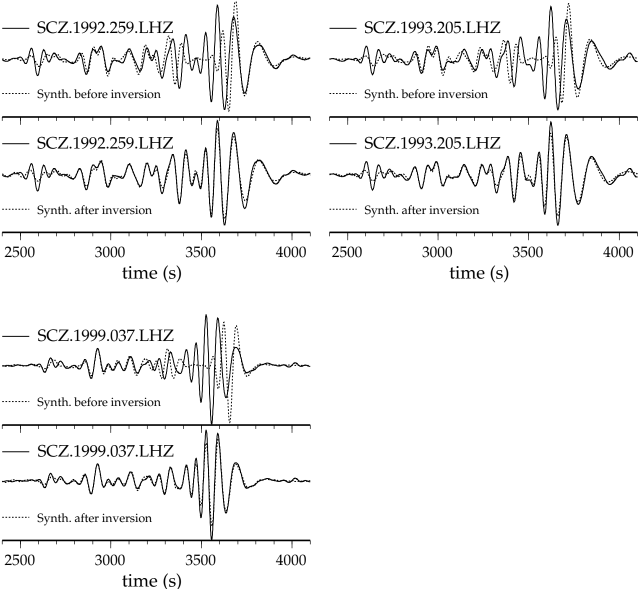
# Results and comparisons



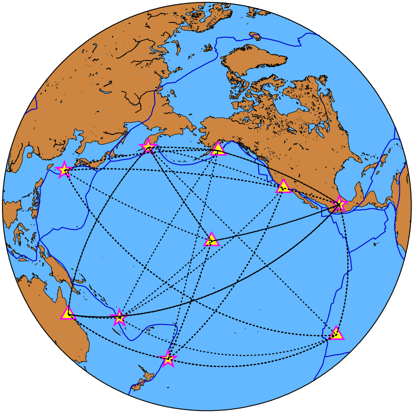


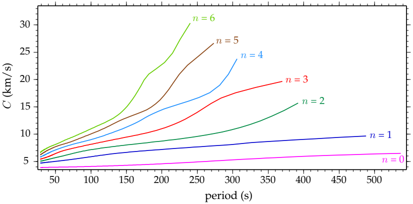

Comparison with previous results along a Vanuatu Islands-California path.

# Results and comparisons



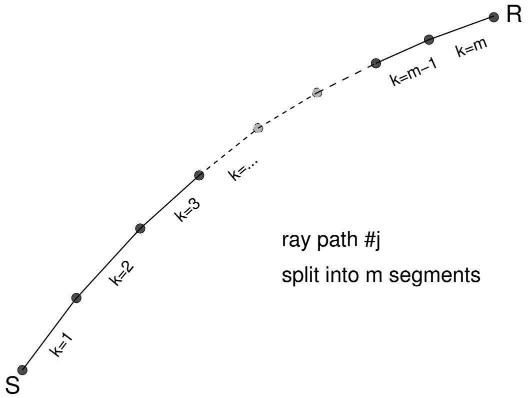
# Phase velocity maps





# The forward problem

$$\langle T_j(\omega)\rangle = \oint_0^{\Delta_j} \frac{dx_j}{C_j(\omega,x)} \iff \frac{\Delta_j}{\langle C_j(\omega)\rangle} = \sum_{k=1}^{m} \frac{l_{jk}}{C_{jk}(\omega)}.$$



# The forward problem

$$\langle T_j(\omega)\rangle = \oint_0^{\Delta_j} \frac{dx_j}{C_j(\omega, x)} \iff \frac{\Delta_j}{\langle C_j(\omega)\rangle} = \sum_{k=1}^{m} \frac{l_{jk}}{C_{jk}(\omega)}.$$

The phase velocity is defined, as a perturbation of a reference one (Smith & Dahlen, 1973; Montagner & Nataf, 1986), as

$$\begin{aligned} C(\theta,\phi,\omega,\psi) &= C_0(\theta,\phi,\omega) + \delta C(\theta,\phi,\omega,\psi), \\ \delta C(\theta,\phi,\omega,\psi) &= \frac{1}{2C_0}\left[A_1(\theta,\phi,\omega) + A_2(\theta,\phi,\omega)\cos 2\psi + A_3(\theta,\phi,\omega)\sin 2\psi \right. \\ &\quad \left. + A_4(\theta,\phi,\omega)\cos 4\psi + A_5(\theta,\phi,\omega)\sin 4\psi\right], \end{aligned}$$

where $\omega$ is the angular frequency, $\psi$ the azimuth (measured clockwise from the North) and $A_p (p = 1 \ldots 5)$ the different parameters (related to the physical properties in depth).

# The forward problem

$$
\begin{aligned}
\mathbf{d} &= g(\mathbf{p}), \\
\text{and}\ldots &\quad \text{for the } j\text{th ray path}, \\
\frac{C_0(\omega)}{\langle C_j(\omega)\rangle} - 1 &= \frac{1}{2\Delta_j C_0^2(\omega)} \sum_{k=1}^{m} l_{jk} \sum_{n=1}^{\infty} (-1)^n \left[ A_{1i}(\theta_i, \phi_i, \omega) \right. \\
&\quad + A_{2i}(\theta_i, \phi_i, \omega) \cos 2\psi_{jk} + A_{3i}(\theta_i, \phi_i, \omega) \sin 2\psi_{jk} \\
&\quad \left. + A_{4i}(\theta_i, \phi_i, \omega) \cos 4\psi_{jk} + A_{5i}(\theta_i, \phi_i, \omega) \sin 4\psi_{jk} \right]^n ,
\end{aligned}
$$

the $\oplus$'s surface is discretized in $q$ grid points $(\theta_i, \phi_i)$, $i \in \{1, 2, \ldots, q\}$.

- **Under-determination**
  - to impose correlations between parameters

and/or

  - to increase the amount of independant data.
- **Non-linearity**
  - negligeable.

# The CLASH (**C**omputation of **L**arge **A**nisotropic **S**eismic **H**eterogeneities)

$\dashrightarrow$ to compute isotropic and anisotropic values of phase velocities at various locations.

- Regular discretization of the model space (*ie.* the $\oplus$'s surface);
- azimuthal anisotropy is accounted in its comprehensive form;
- computations are based on crossing points between all ray paths;
- partial derivatives are assigned to the individual grid points;
- parameter prior operators with variable correlation lengths according to the ray path coverage;

# The CLASH (**C**omputation of **L**arge **A**nisotropic **S**eismic **H**eterogeneities)

$\dashrightarrow$ to compute isotropic and anisotropic values of phase velocities at various locations.

- Regular discretization of the model space (*ie.* the $\oplus$'s surface);
- azimuthal anisotropy is accounted in its comprehensive form;
- computations are based on crossing points between all ray paths;
- partial derivatives are assigned to the individual grid points;
- parameter prior operators with variable correlation lengths according to the ray path coverage;

# The CLASH (**C**omputation of **L**arge **A**nisotropic **S**eismic **H**eterogeneities)

$\dashrightarrow$ to compute isotropic and anisotropic values of phase velocities at various locations.

- Regular discretization of the model space (*ie.* the $\oplus$'s surface);
- azimuthal anisotropy is accounted in its comprehensive form;
- computations are based on crossing points between all ray paths;
- partial derivatives are assigned to the individual grid points;
- parameter prior operators with variable correlation lengths according to the ray path coverage;

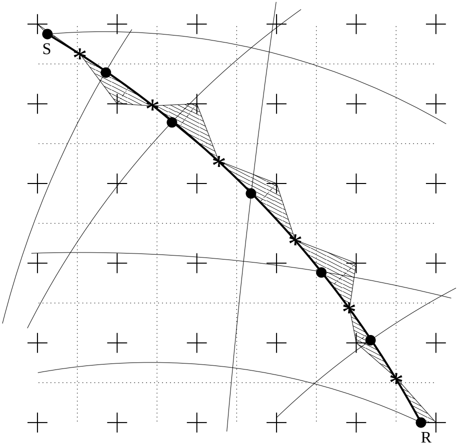
$$
\begin{aligned}
C(\theta,\phi,\omega,\psi) &= C_0(\theta,\phi,\omega) + \delta C(\theta,\phi,\omega,\psi), \\
\delta C(\theta,\phi,\omega,\psi) &= \frac{1}{2C_0}\left[A_1(\theta,\phi,\omega) + A_2(\theta,\phi,\omega)\cos 2\psi + A_3(\theta,\phi,\omega)\sin 2\psi\right. \\
&\quad \left. + A_4(\theta,\phi,\omega)\cos 4\psi + A_5(\theta,\phi,\omega)\sin 4\psi\right].
\end{aligned}
$$

# The CLASH (**C**omputation of **L**arge **A**nisotropic **S**eismic **H**eterogeneities)

⇢ to compute isotropic and anisotropic values of phase velocities at various locations.

- Regular discretization of the model space (*ie.* the $\oplus$'s surface);
- azimuthal anisotropy is accounted in its comprehensive form;
- computations are based on crossing points between all ray paths;
- partial derivatives are assigned to the individual grid points;
- parameter prior operators with variable correlation lengths according to the ray path coverage;

# The CLASH (Computation of Large Anisotropic Seismic Heterogeneities)



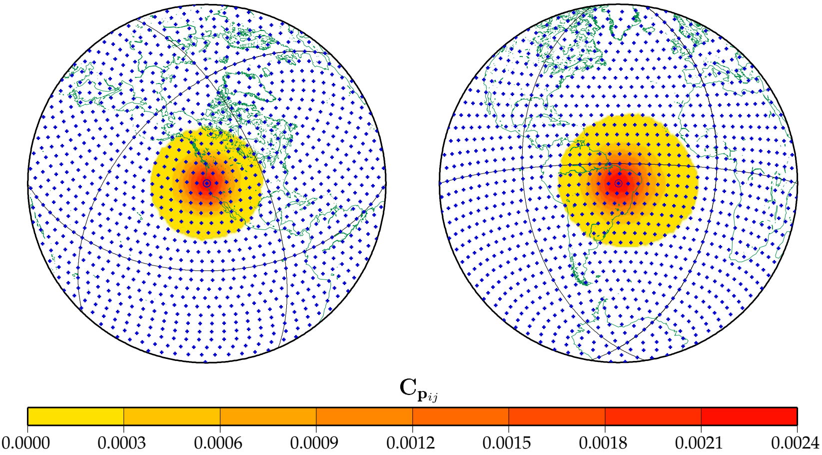

Sketch illustrating the relationships between the ray paths and the grid points (+). The hatched triangles represent the S-to-R ray path segments attributed to the grid points.

# The CLASH (Computation of Large Anisotropic Seismic Heterogeneities)

⇢ to compute isotropic and anisotropic values of phase velocities at various locations.

- Regular discretization of the model space (*ie.* the $\oplus$'s surface);
- azimuthal anisotropy is accounted in its comprehensive form;
- computations are based on crossing points between all ray paths;
- partial derivatives are assigned to the individual grid points;
- parameter prior operators with variable correlation lengths according to the ray path coverage;

# The CLASH (**C**omputation of **L**arge **A**nisotropic **S**eismic **H**eterogeneities)

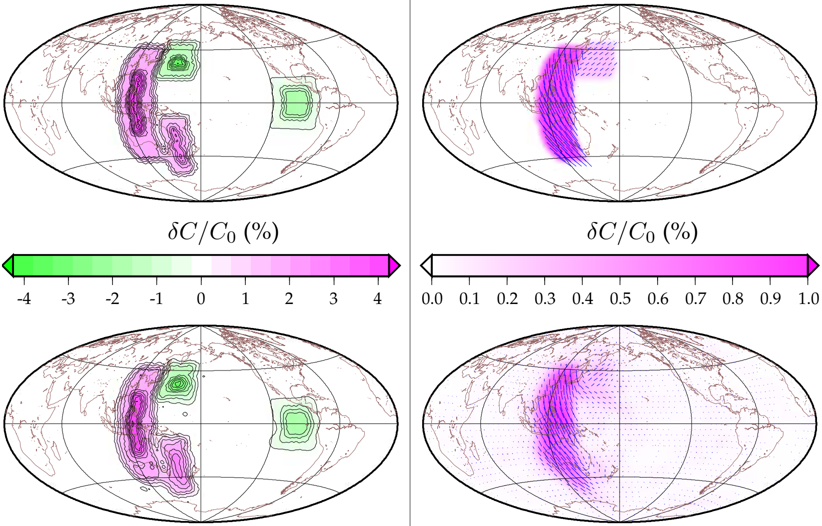



Laterally variable *a priori* covariances on parameters (Beucler & Montagner, 2006).

# Inversion in an unrealistic anisotropic model

Input Model (isotropic perturbations)

Input Model (anisotropic perturbations)

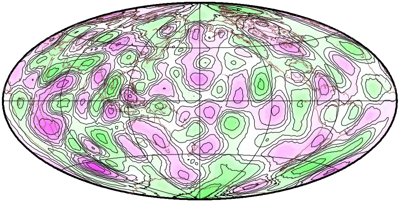



Output Model (isotropic perturbations)

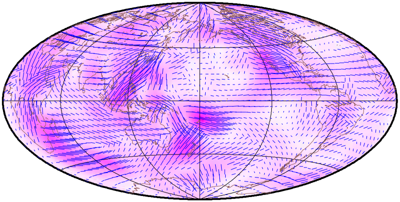
Output Model (anisotropic perturbations)

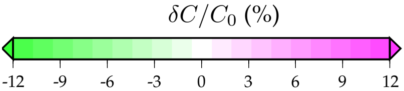
# Inversion in a realistic anisotropic model

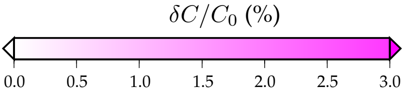
Input Model (isotropic perturbations)

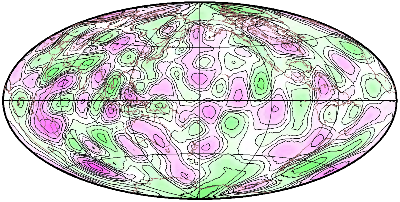


Input Model (anisotropic perturbations)

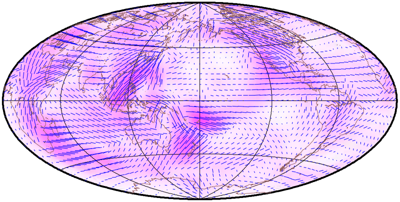




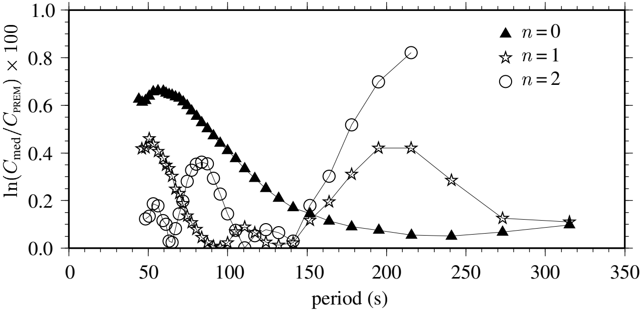



Output Model (isotropic perturbations)



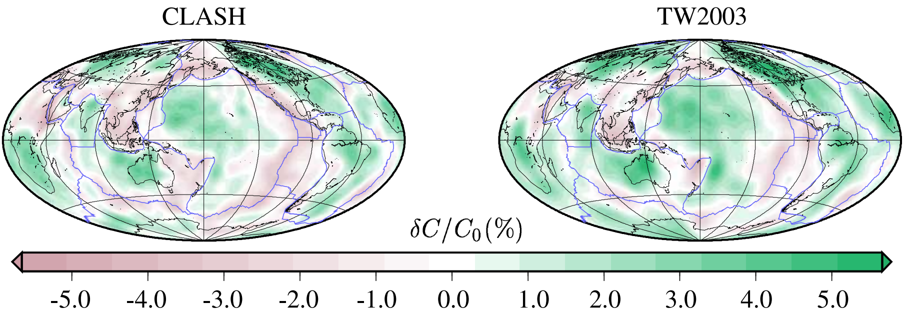
Output Model (anisotropic perturbations)

# Real data set inversions



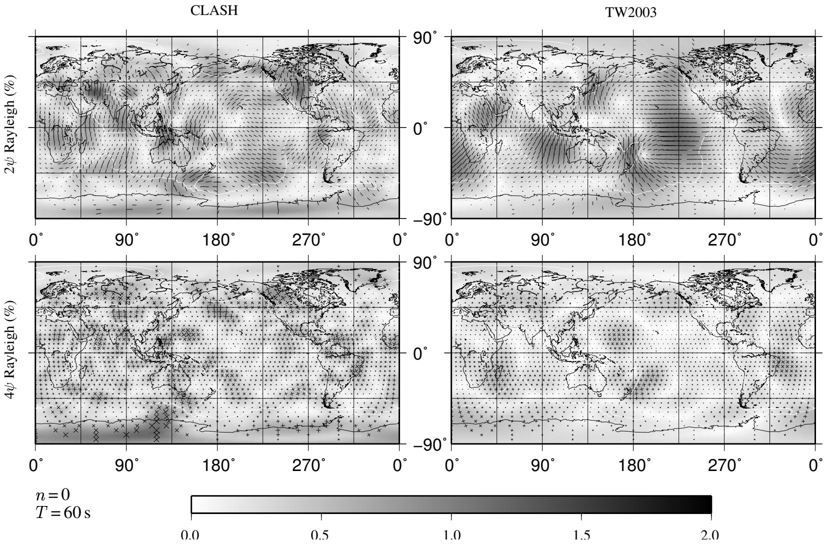

Data set (*ie.* S-to-R integrated phase velocities) perturbations with respect to the PREM.

- 229 422 seismograms manually inspected and 19 020 retained;
- the actual amount of ray paths varies for each mode-branch and each period;
- 2 036 grid points $\rightarrow$ 10 180 parameters;
- *a priori* variations: 5% for the 0$\psi$ and 1% for the 2- and 4$\psi$ parameters.




$n = 0$, T$= 60$ s

Isotropic phase velocity maps for Rayleigh waves - TW2003 = Trampert & Woodhouse (2003).




Anisotropic phase velocity maps for Rayleigh waves - TW2003 = Trampert & Woodhouse (2003).

# Summary

- Sometimes it is possible to use a least-squares criterion with a non-linear forward problem (*eg.* by mixing explorations and local optimizations);
- The issues of the under-determination and of the non-linearity worth to be addressed before choosing an inverse method.

Beucler, É. & J.-P. Montagner (2006). Computation of Large Anisotropic Seismic Heterogeneities (CLASH). *Geophys. J. Int. 165*, 447–468.

Beucler, É., É. Stutzmann & J.-P. Montagner (2003). Surface-wave higher mode phase velocity measurements using a roller coaster type algorithm. *Geophys. J. Int. 155*, 289–307.

Grand, S. P. (2002). Mantle shear-wave tomography and the fate subducted slabs. *Philos. Trans. R. Soc. London A. 360*, 2475–2491.

Gu, Y., A. M. Dziewonski, W.-J. Su & G. Ekström (2001). Models of mantle shear velocity and discontinuities in the pattern of lateral heterogeneties. *J. Geophys. Res. 106*, 11169–11199.

Lamb, S. & D. Sington (1998). *Earth Story: The Shaping of Our World*. Princeton Univ. Press, Princeton, New Jersey.

Masters, G., H. Bolton & G. Laske (1999). How pervasive are chemical anomalies in the mantle. *EOS, Trans. Am. geophys. Un. 80*, S14.

Mégnin, C. & B. Romanowicz (2000). The three-dimensional shear velocity structure of the mantle from the inversion of body, surface and higher-mode waveforms. *Geophys. J. Int. 143*, 709–728.

Montagner, J.-P. & H. C. Nataf (1986). A simple method for inverting the azimuthal anisotropy of surface waves. *J. Geophys. Res. 91*, 511–520.

Ritsema, J., H. J. van Heijst & J. H. Woodhouse (1999). Complex shear wave velocity structure imaged beneath Africa and Iceland. *Science 286*, 1925–1928.

Romanowicz, B. (2003). Global mantle tomography: progress status in the last 10 years. *Ann. Rev. of Earth and Planet. Sciences 31*, 303–328.

Smith, M. L. & F. A. Dahlen (1973). The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropic medium. *J. Geophys. Res. 78*, 3321–3333.

Tarantola, A. & B. Valette (1982). Generalized nonlinear inverse problems solved using the least squares criterion. *Rev. Geophys. Space Phys. 20*, 219–232.

Trampert, J. & J. H. Woodhouse (2003). Global anisotropic phase velocity maps for fundamental mode surface waves between 40 and 150 s. *Geophys. J. Int. 154*, 154–165.

van Heijst, H. J. & J. H. Woodhouse (1999). Global high-resolution phase velocity distributions of overtone and fundamental-mode surface waves determined by mode-branch stripping. *Geophys. J. Int. 137*, 601–620.

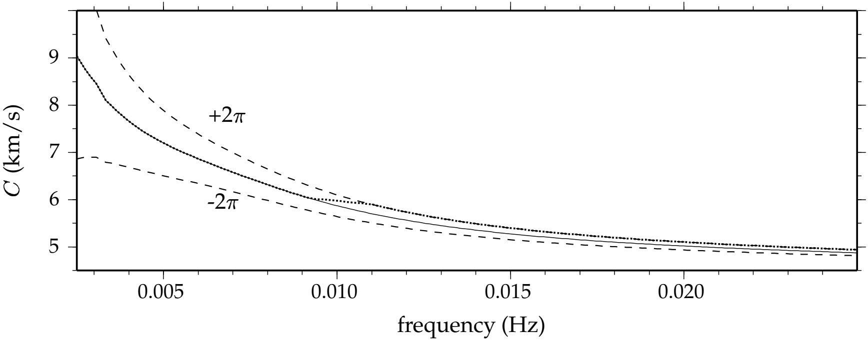
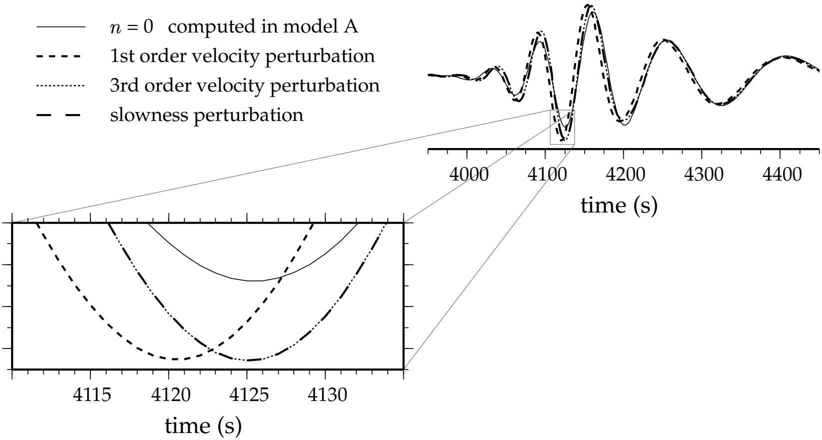
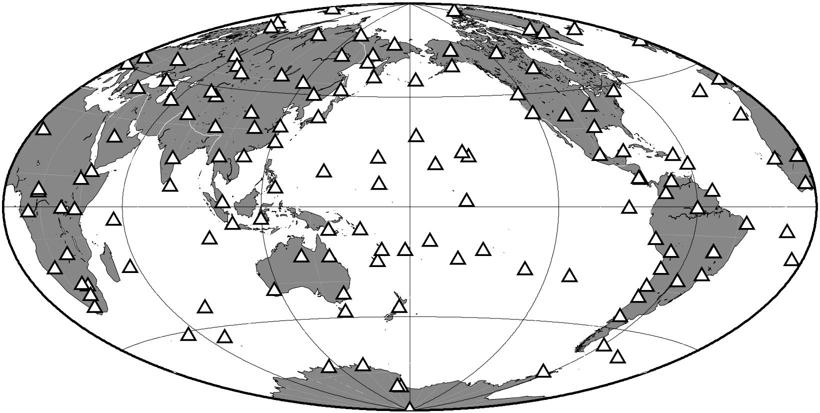


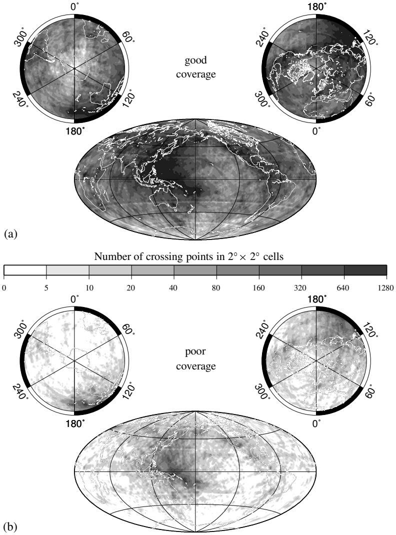

Illustration of possible $2\pi$ phase jumps over the whole frequency range (dashed lines) or localised around a given frequency (dotted line). The reference phase velocity used to compute these three curves is represented in solid line.



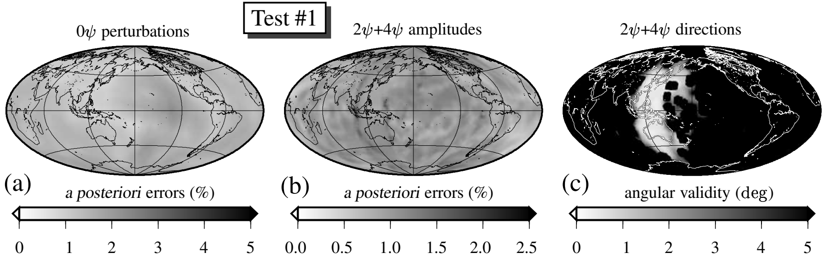

The dotted line represents a fundamental mode seismogram computed in model A. The model A phase velocity is directly used to phase shift the FM seismogram computed in PREM (not displayed here for clarity). When expanding the phase velocity perturbation to first order the reconstruted signal (small-dashed line) is 5 s phase shifted as shown in the zoomed part of the figure. It is correctly retrieved when expanding the phase velocity perturbation up to third order (represented in dotted line). The long-dashed line represents the recovered phase when using a slowness perturbation development.



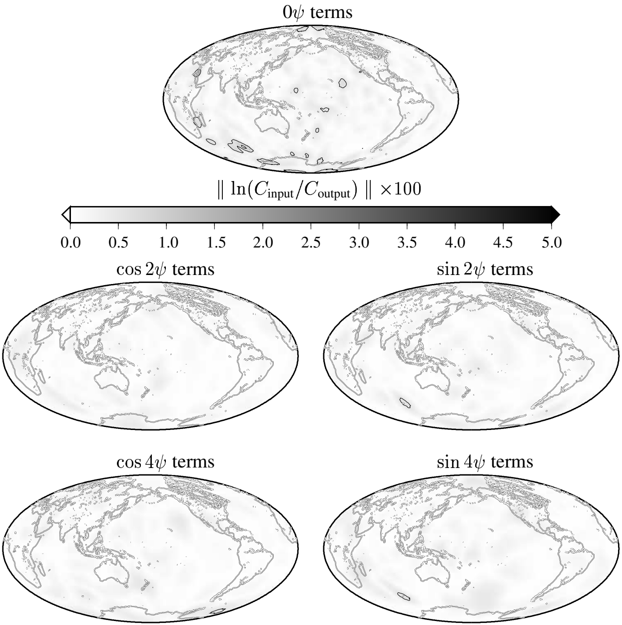
Locations of the 141 FDSN stations (IRIS and GEOSCOPE) processed for the construction of our Rayleigh wave phase velocity maps.



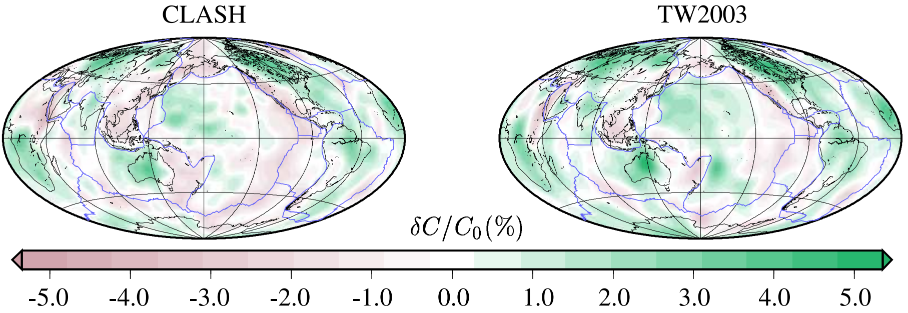

Ray path coverages used for both the computation of synthetic data sets and the inversions of real data sets. They correspond to (a) the fundamental mode at 100 s and (b) the second overtone at 150 s.




*A posteriori* uncertainties on isotropic and anisotropic (amplitude and azimuth) for the unrealistic input model test.

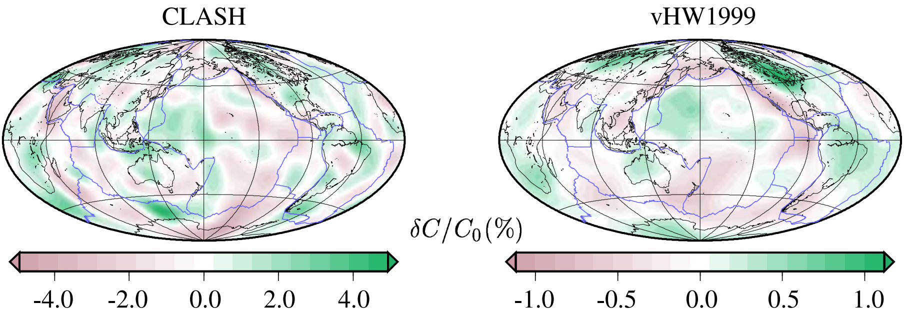



9 288 ray paths



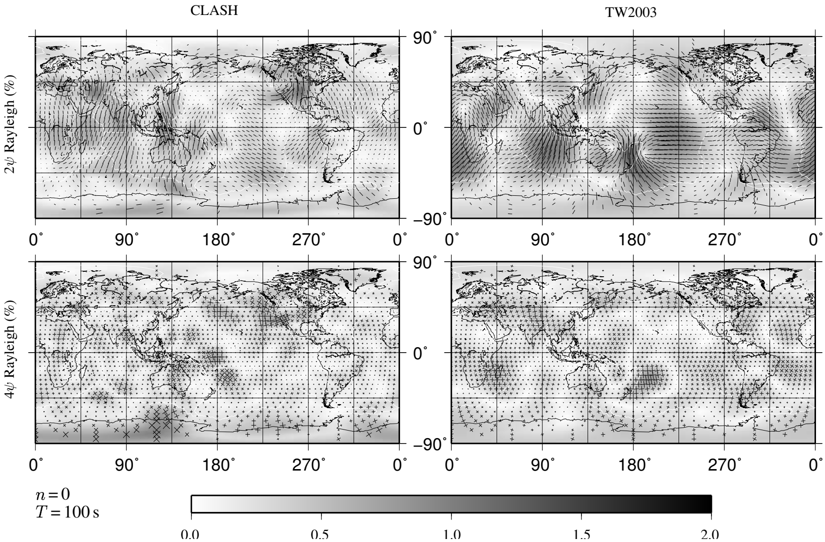

$n = 0$, T$=$ 100 s

Isotropic phase velocity maps for Rayleigh waves - TW2003 = Trampert & Woodhouse (2003).

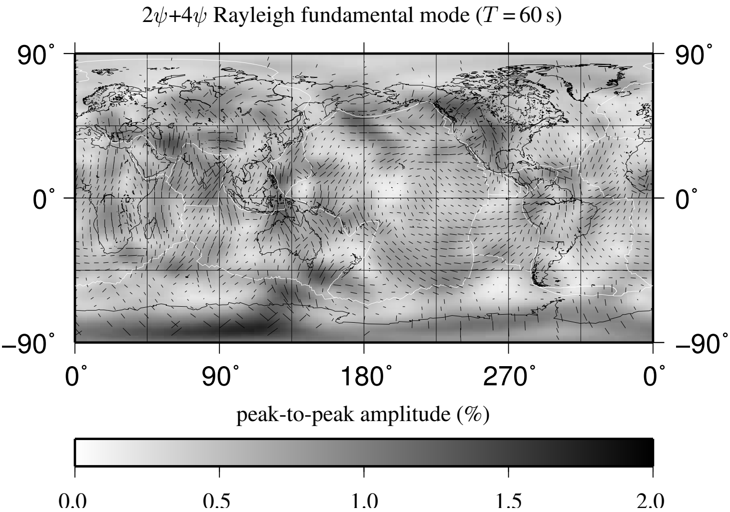



$n = 2$, T= 150 s

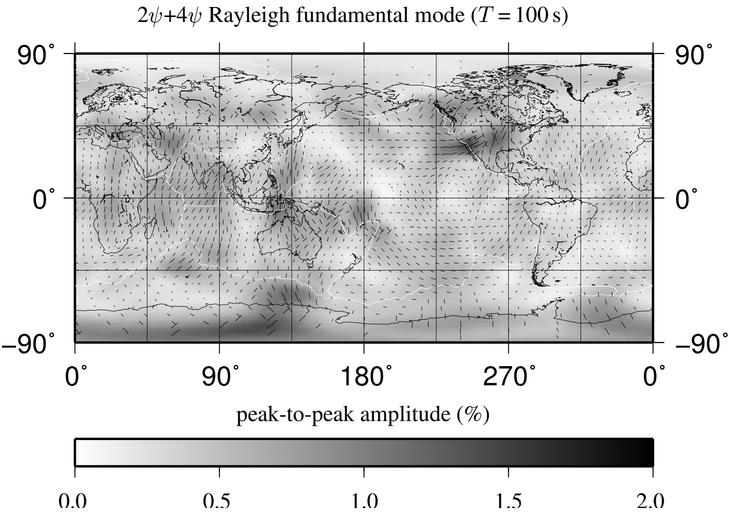
Isotropic phase velocity maps for Rayleigh waves - vHW1999 = van Heijst & Woodhouse (1999).




Anisotropic phase velocity maps for Rayleigh waves - TW2003 = Trampert & Woodhouse (2003).




$2\psi$+$4\psi$ model for the fundamental mode of Rayleigh waves at 60 s. The grey-scale in the background corresponds to the peak-to-peak amplitude of anisotropy, expressed with respect to the PREM phase velocity. The black segments represent the fast-axis directions which are also scaled to the amplitude shown in the background.




$2\psi$+$4\psi$ model for the fundamental mode of Rayleigh waves at 100 s. The grey-scale in the background corresponds to the peak-to-peak amplitude of anisotropy, expressed with respect to the PREM phase velocity. The black segments represent the fast-axis directions which are also scaled to the amplitude shown in the background.