SES3D.inv

PART III

ITERATIVE SOLUTION OF THE TOMOGRAPHIC INVERSE PROBLEM

1. French cheese, non-linearity and the multi-scale approach
2. Misfit functionals
3. Towards quantitative resolution analysis

French cheese,

non-linearity

and the

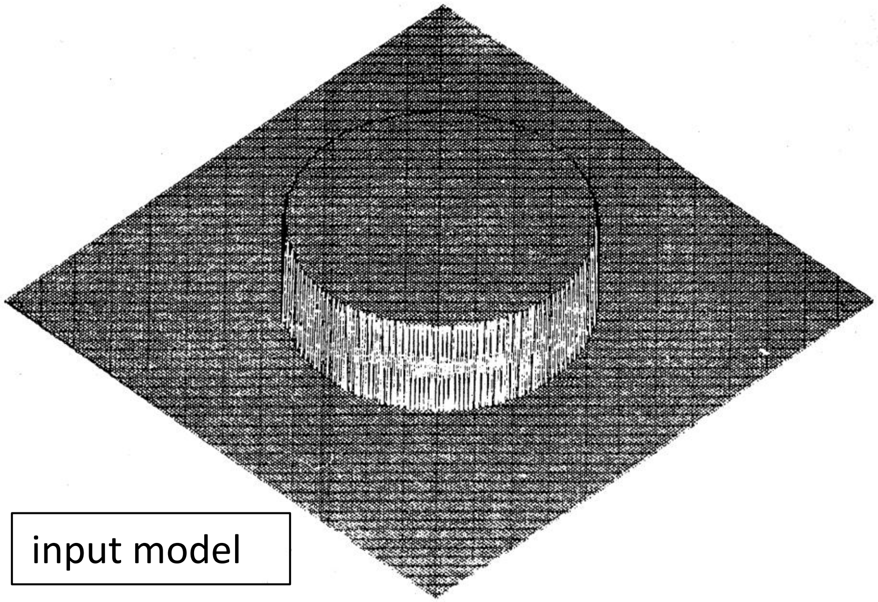
multi-scale approach

The Camembert experiment (Gauthier, Virieux & Tarantola, 1986)

- Synthetic full waveform inversion experiment.
- The misfit was the most obvious L_2 difference:

$$\boxed{X(m) = \frac{1}{2}\int \left[\underbrace{u}_{\text{synthetic waveform}} - \underbrace{u_0}_{\text{observed waveform}}\right]^2 dt}$$

- The goal was to recover an input model with the shape of a Camembert.



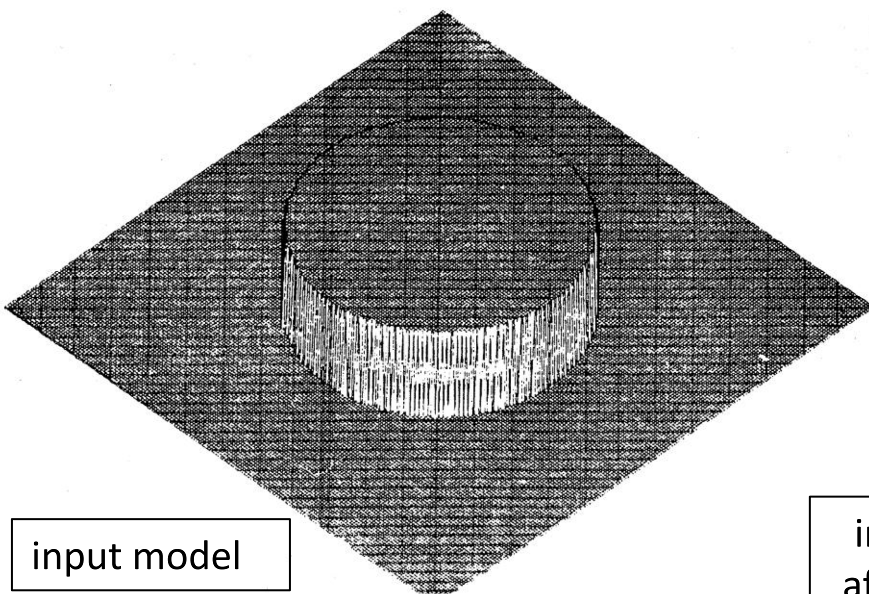
input model

The Camembert experiment (Gauthier, Virieux & Tarantola, 1986)

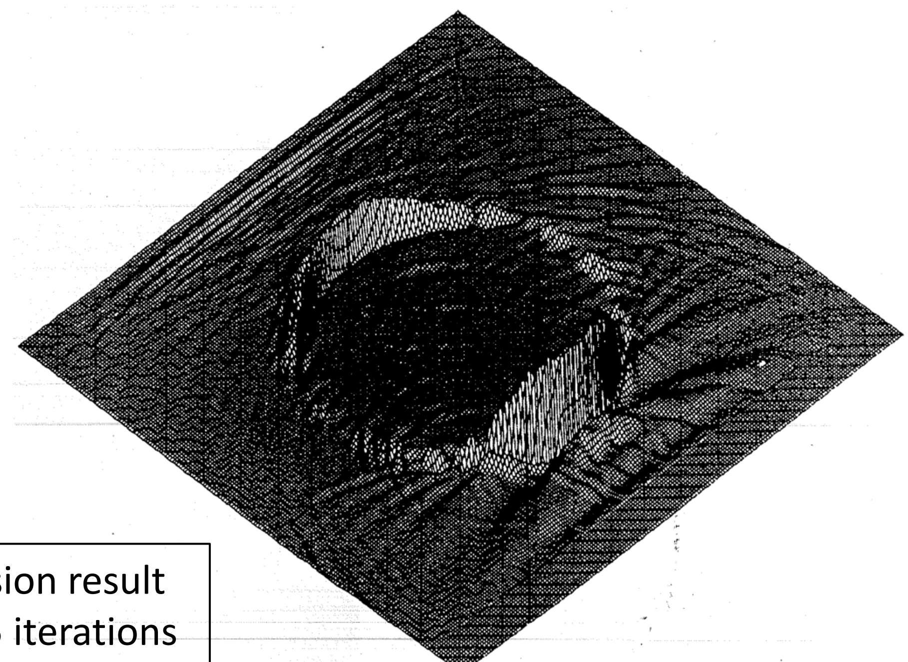
- Synthetic full waveform inversion experiment.
- The misfit was the most obvious L_2 difference:

$$\boxed{X(m) = \frac{1}{2}\int \left[\underset{\text{synthetic waveform}}{u} - \underset{\text{observed waveform}}{u_0} \right]^2 dt}$$

- The goal was to recover an input model with the shape of a Camembert.



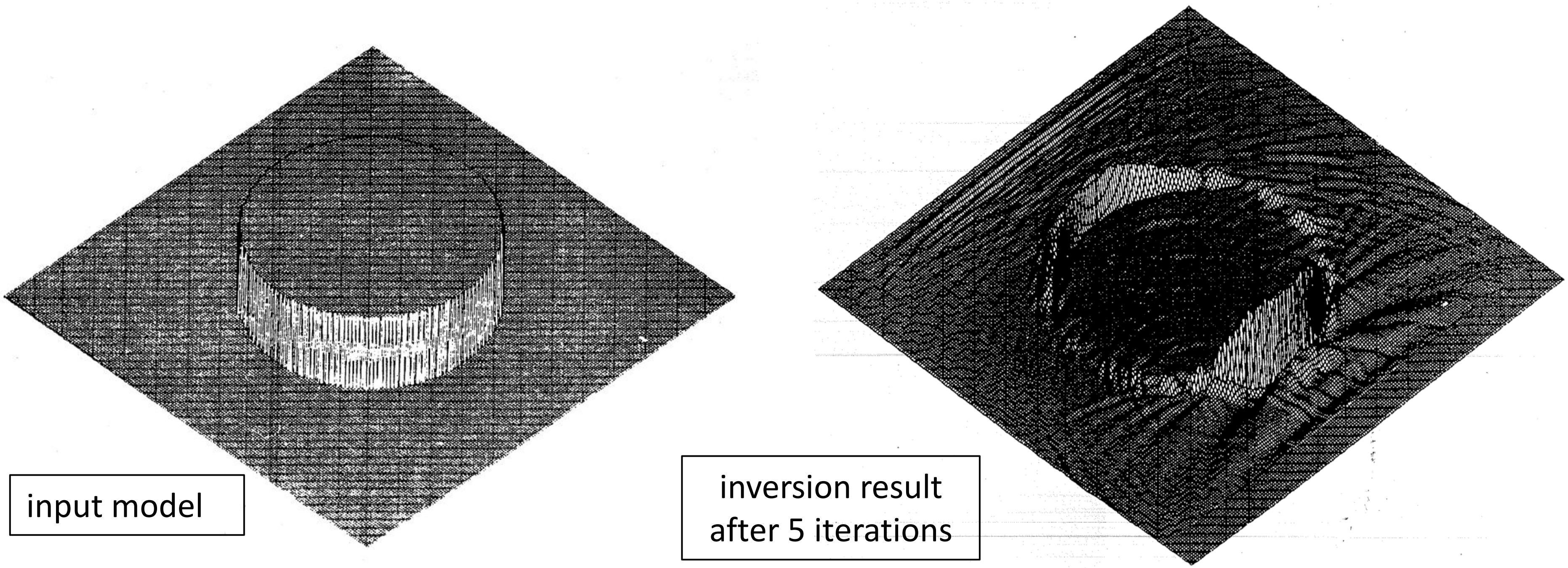
input model



inversion result after 5 iterations

The Camembert experiment (Gauthier, Virieux & Tarantola, 1986)

The inversion was trapped in a local minimum of the misfit function.



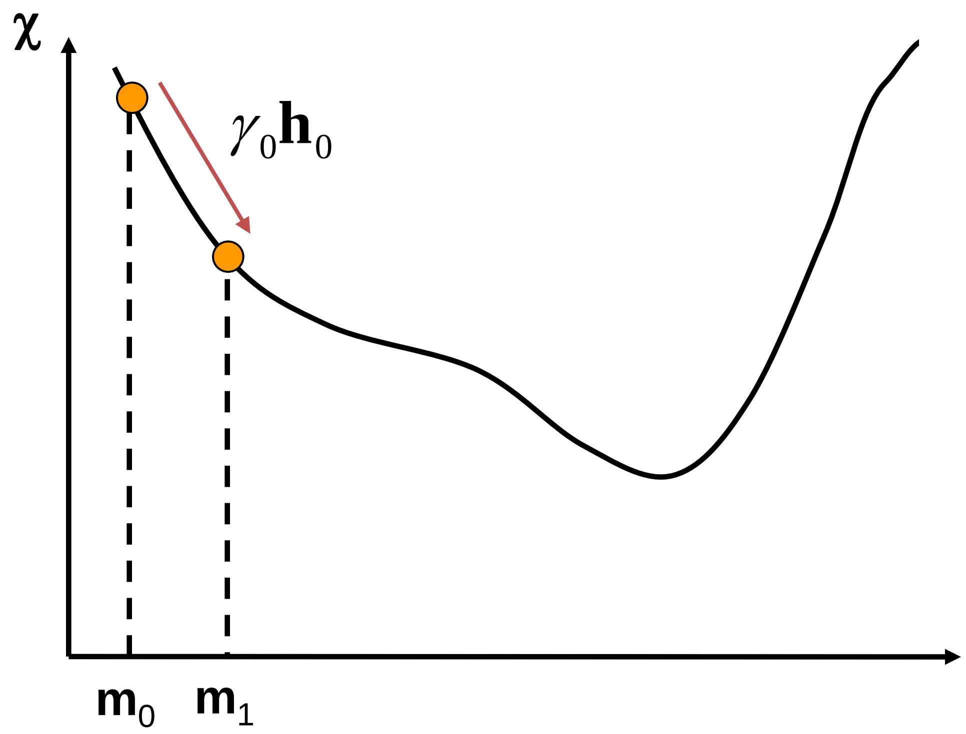
- Seismic waveforms depend non-linearly on the structure of the Earth.
- This can result in the presence of multiple local minima of the misfit.

Non-linearity and multiple minima

1. Start from initial Earth model $\mathbf{m}_0$

2. Update according to $\mathbf{m}_{i+1} = \mathbf{m}_i + \gamma_i \mathbf{h}_i \,, \quad \text{with} \quad \chi(\mathbf{m}_{i+1}) < \chi(\mathbf{m}_i)$

step length — γ_i; descent direction — $\mathbf{h}_i$

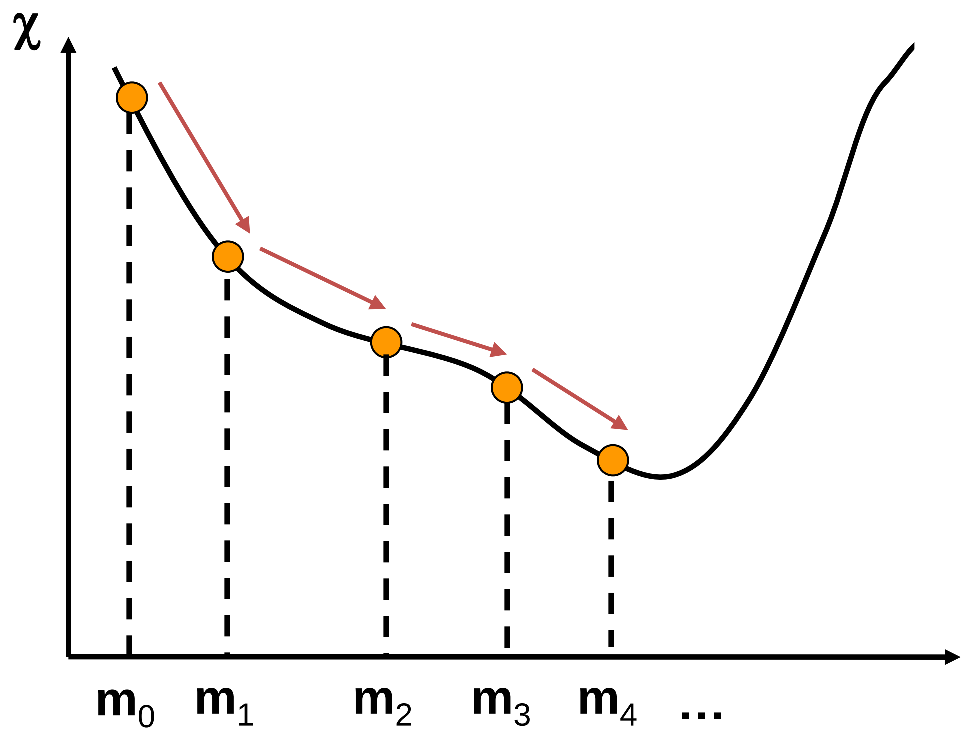


Non-linearity and multiple minima

1. Start from initial Earth model $\mathbf{m}_0$

2. Update according to $\mathbf{m}_{i+1} = \mathbf{m}_i + \gamma_i \mathbf{h}_i\,, \quad \text{with} \quad \chi(\mathbf{m}_{i+1}) < \chi(\mathbf{m}_i)$

step length — descent direction



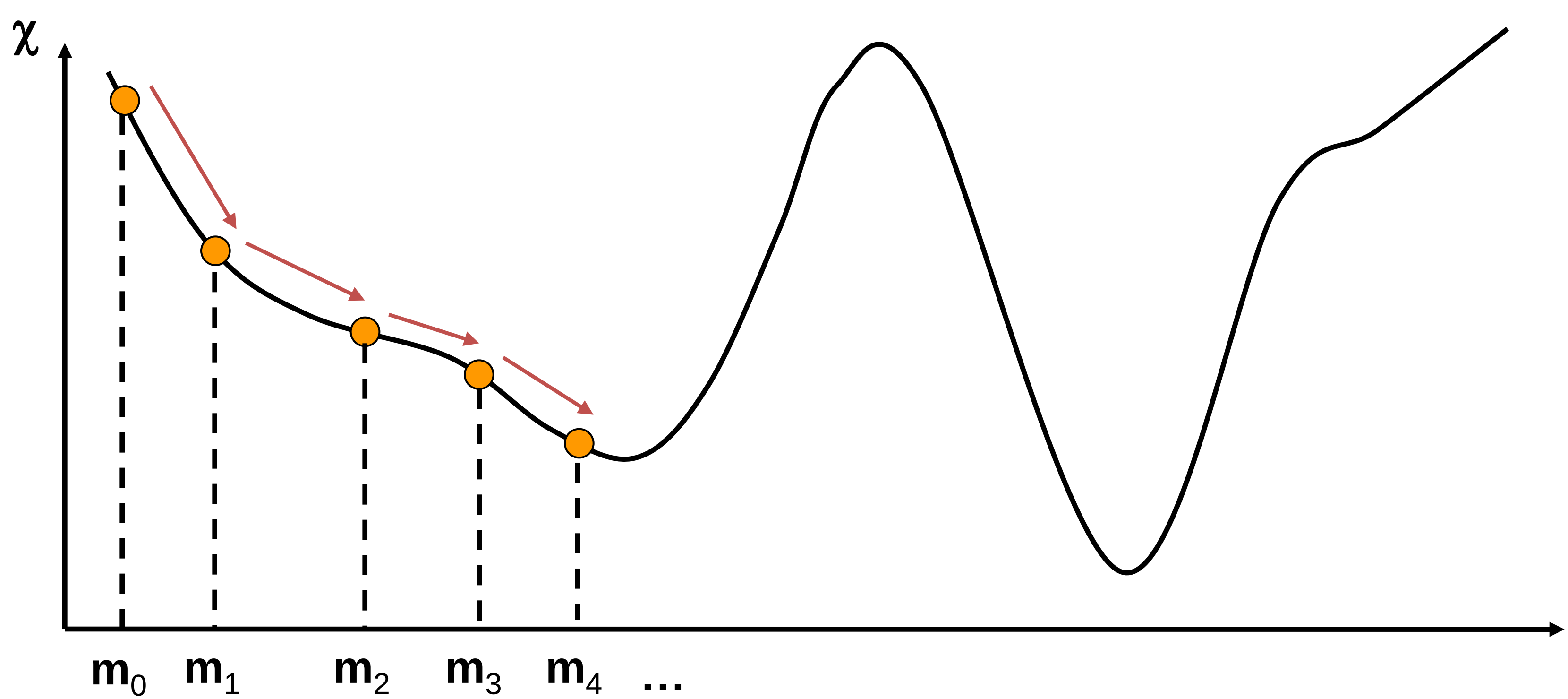
Iteratively approach the minimum misfit by following the local descent directions.

Non-linearity and multiple minima

1. Start from initial Earth model $\mathbf{m}_0$

2. Update according to $\mathbf{m}_{i+1} = \mathbf{m}_i + \gamma_i \mathbf{h}_i \,,\quad \text{with} \quad \chi(\mathbf{m}_{i+1}) < \chi(\mathbf{m}_i)$

step length — descent direction

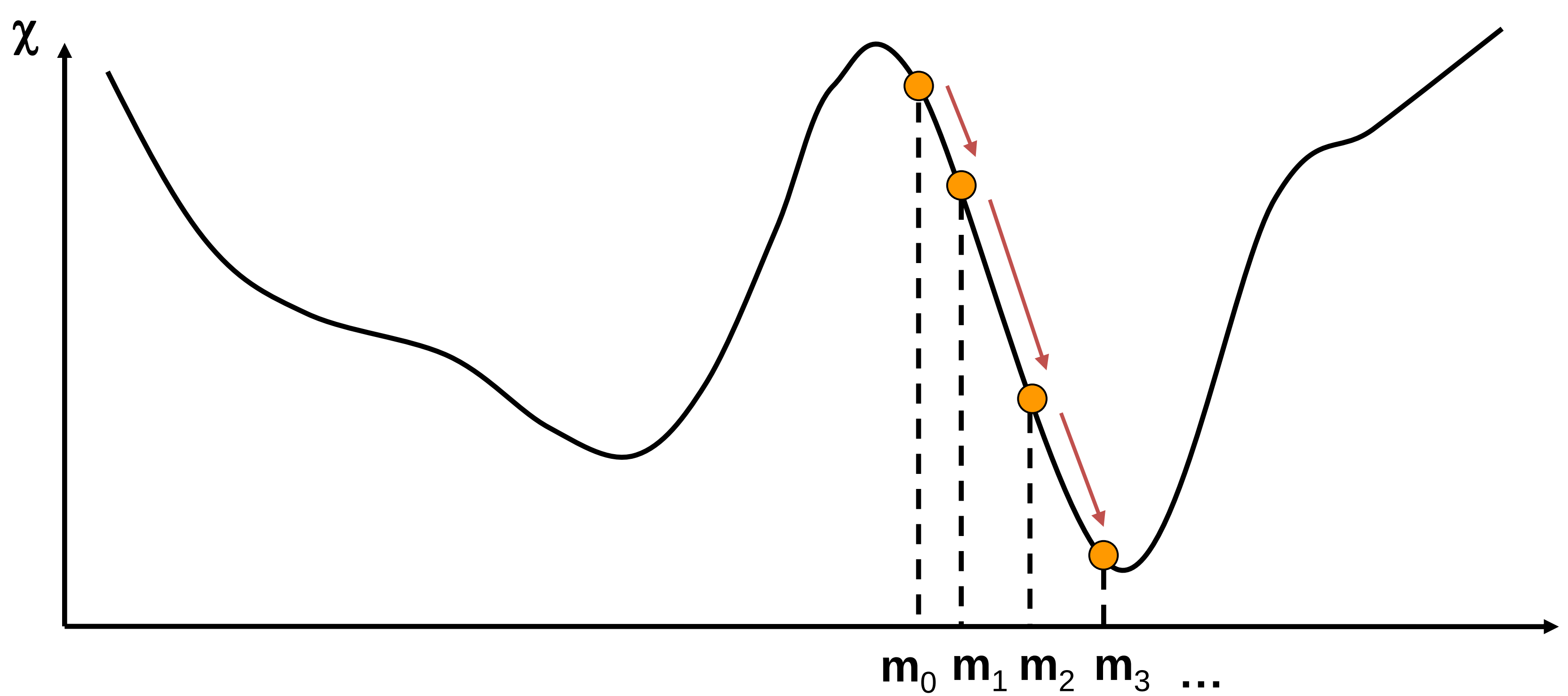


Non-linearity and multiple minima

1. Start from initial Earth model $\mathbf{m}_0$

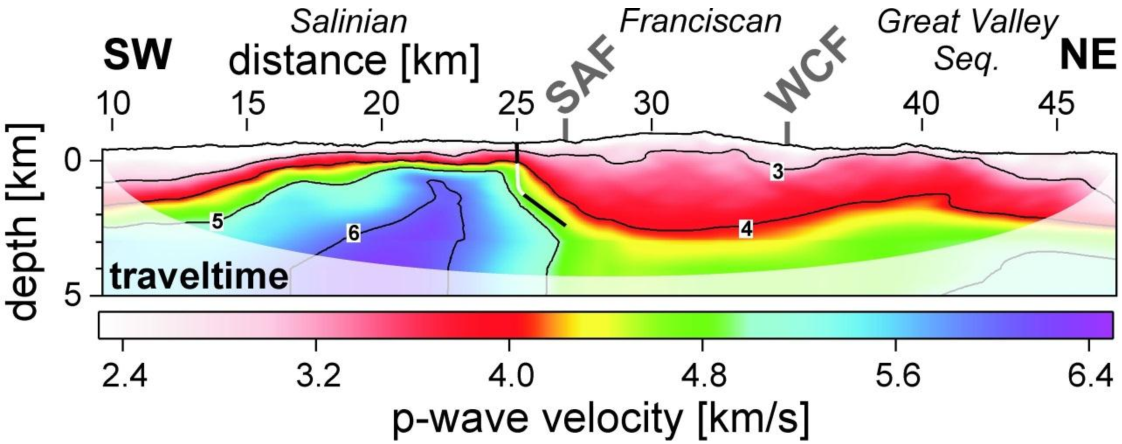
2. Update according to $\mathbf{m}_{i+1} = \mathbf{m}_i + \gamma_i \mathbf{h}_i\,, \quad \text{with} \quad \chi(\mathbf{m}_{i+1}) < \chi(\mathbf{m}_i)$

step length (γ_i) — descent direction ($\mathbf{h}_i$)



Non-linearity and multiple minima

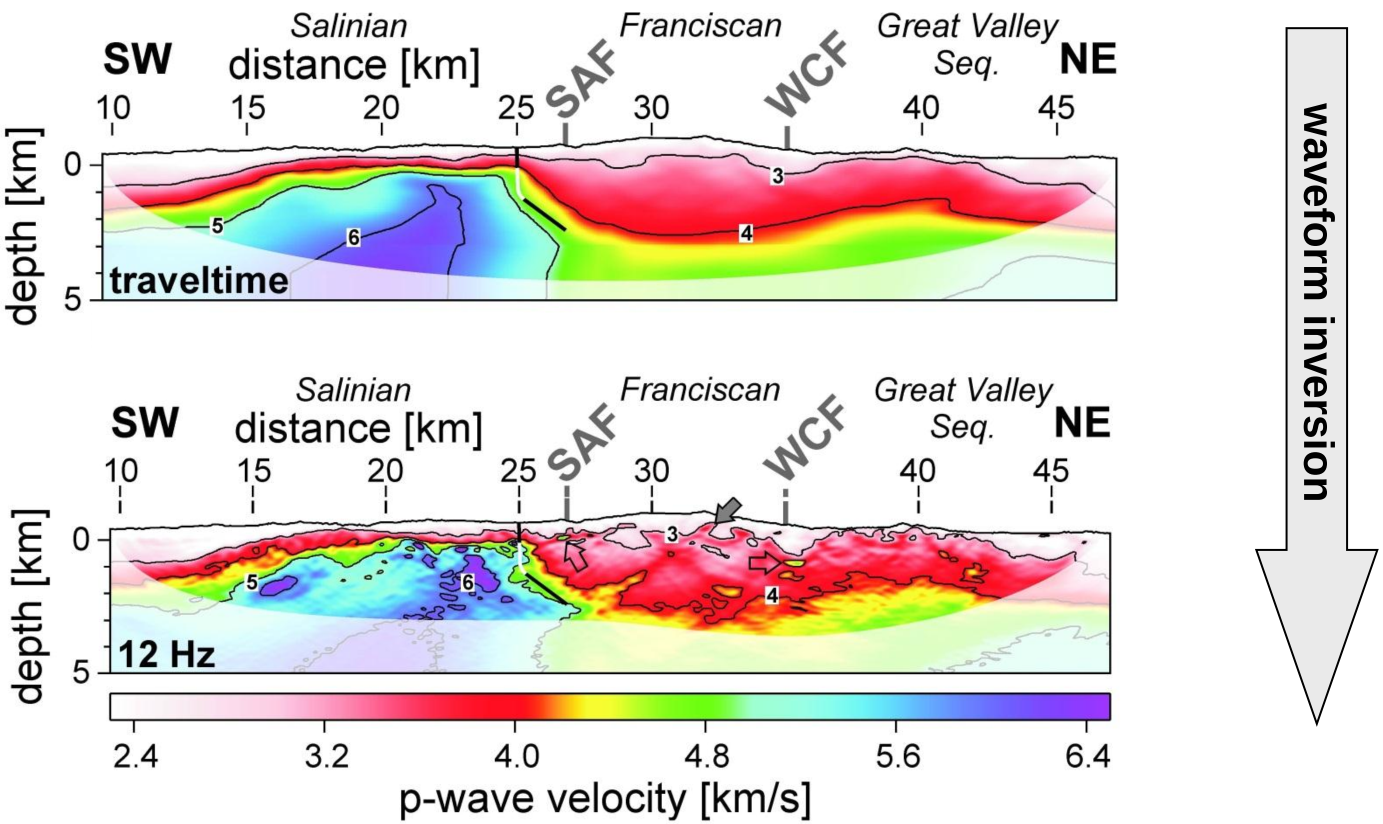
- Gradient methods are local.
- Convergence to the global minimum relies on a good initial model.
- Good initial model: e.g. long-wavelength model from ray tomography.



Bleibinhaus et al., 2007

Non-linearity and multiple minima

- Gradient methods are local.
- Convergence to the global minimum relies on a good initial model.
- Good initial model: e.g. long-wavelength model from ray tomography.



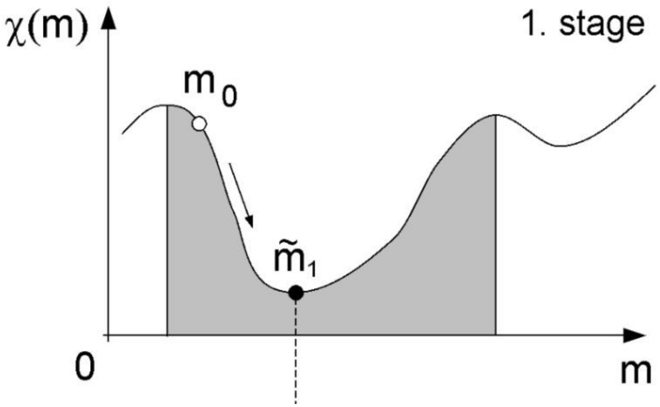
Bleibinhaus et al., 2007

The multi-scale approach

- Sufficiently good initial models are often not available.

- The multi-scale approach is an empirical strategy that helps to overcome this problem

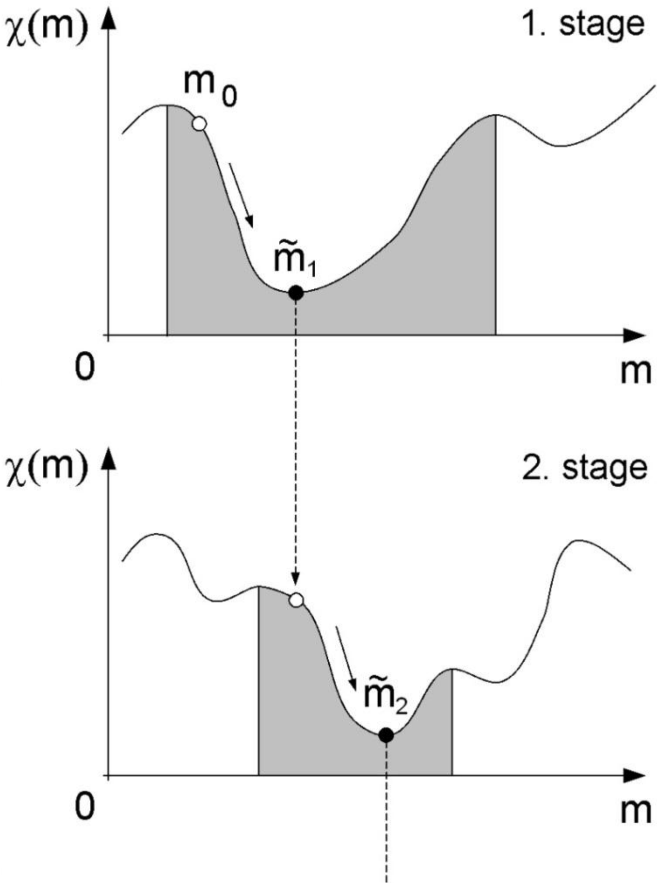
The multi-scale approach



long-period data

$\longrightarrow$ long-wavelength structure

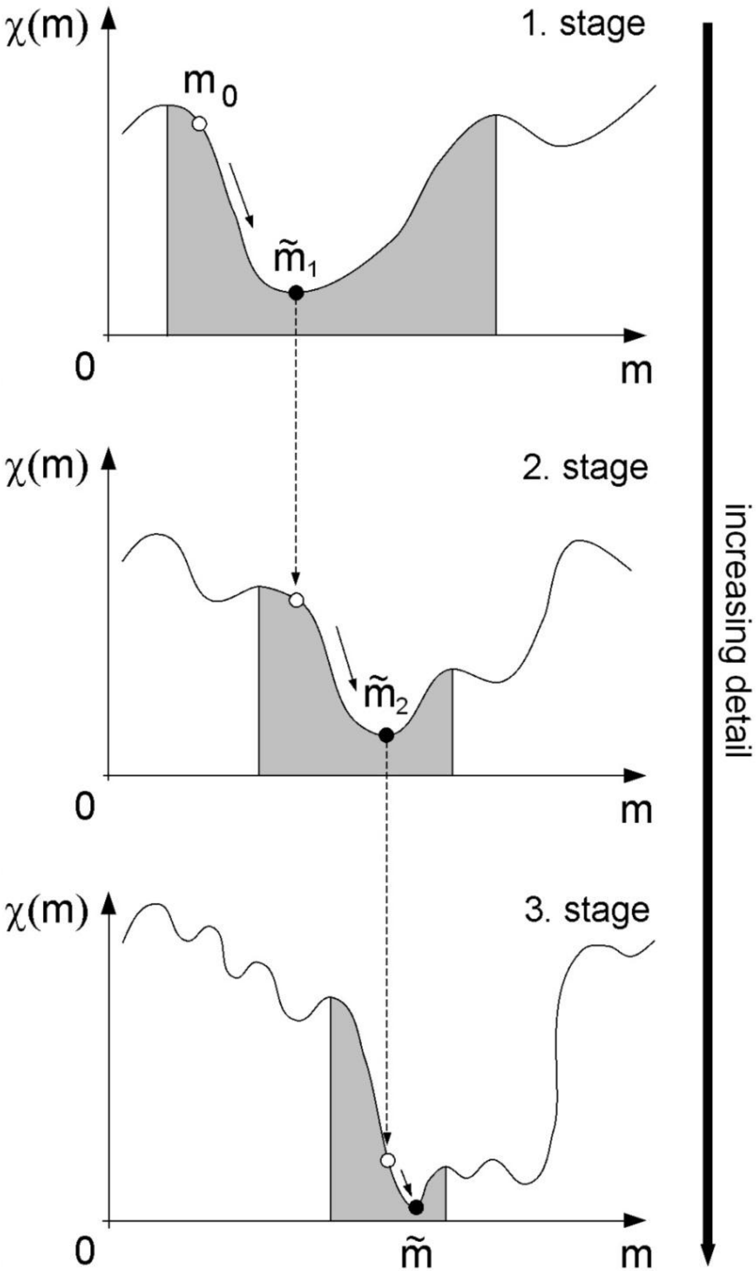
The multi-scale approach



shorter-period data

⟶ shorter-wavelength structure

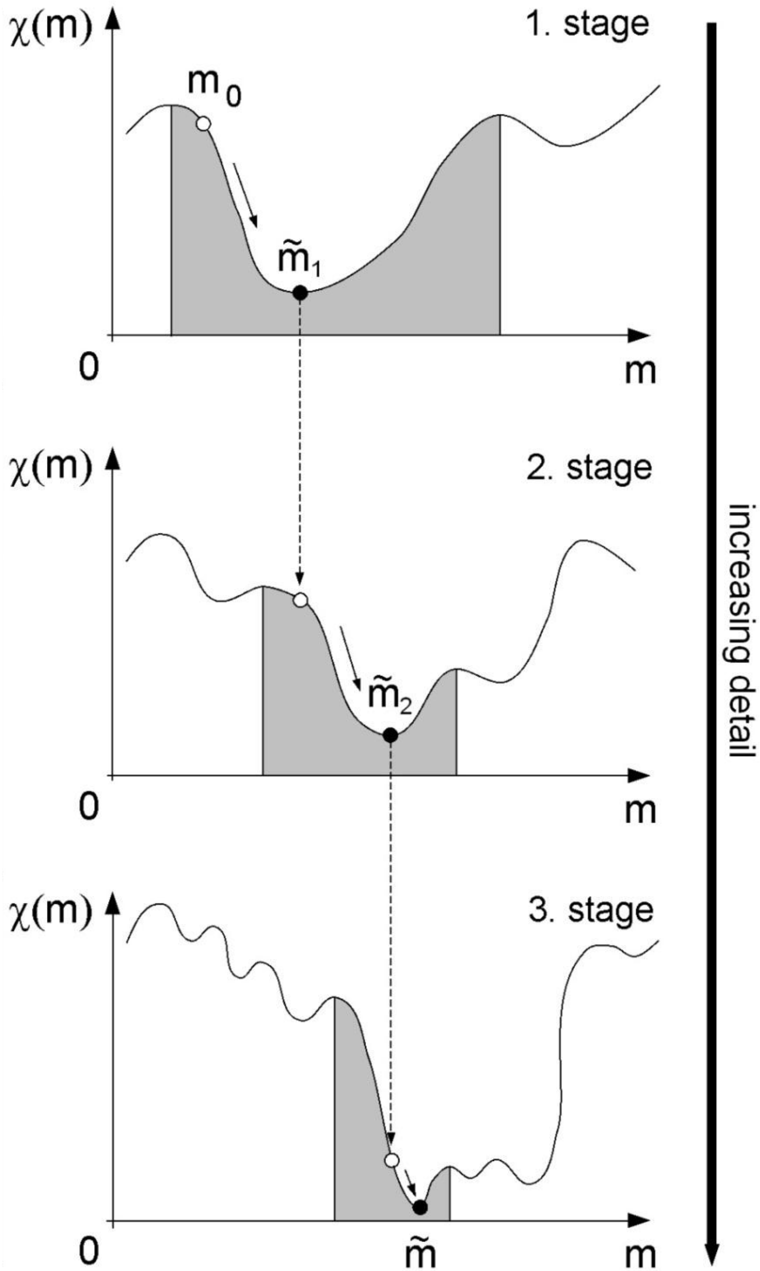
The multi-scale approach

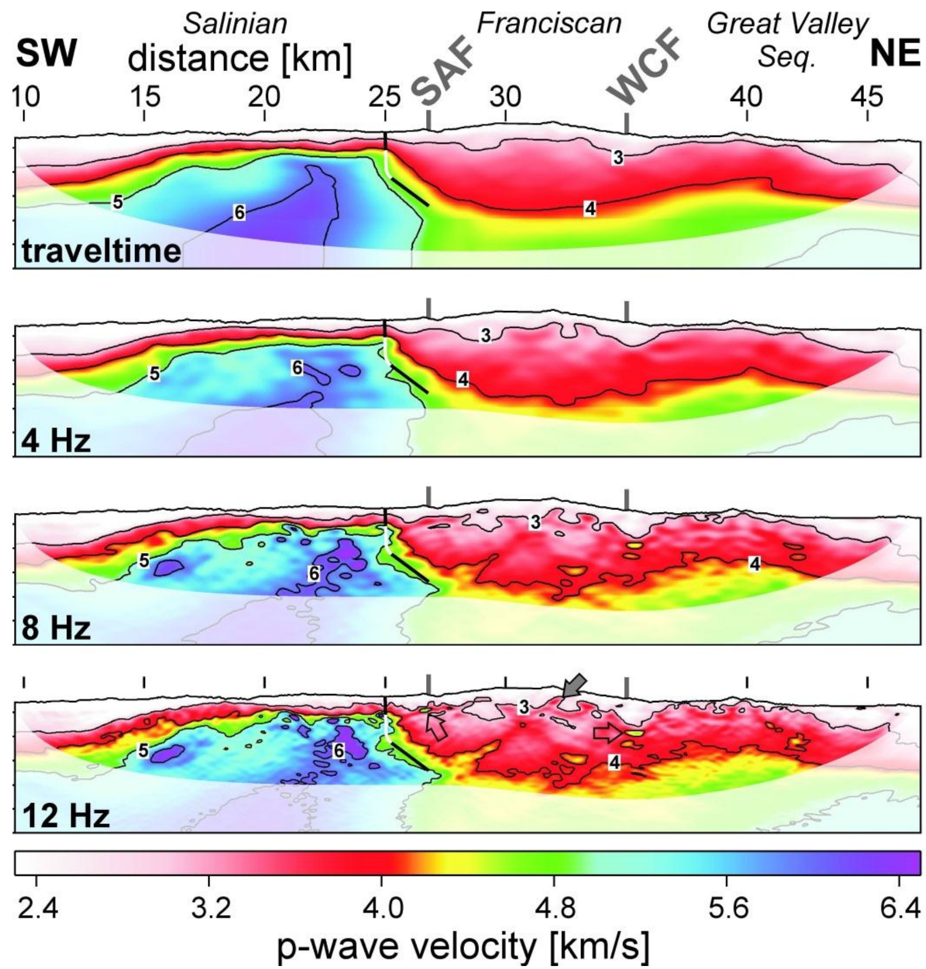


short-period data

$\longrightarrow$ short-wavelength structure

The multi-scale approach





Bleibinhaus et al., 2007

Misfit functionals

Misfit functionals: Principal Desiderata

A misfit functional suitable for waveform tomography should:

1) extract as much information as possible from seismic waveforms

2) be quasi-linearly related to Earth structure (tame non-linearity)

Misfit functionals: Principal Desiderata

A misfit functional suitable for waveform tomography should:

1) extract as much information as possible from seismic waveforms

2) be quasi-linearly related to Earth structure (tame non-linearity)

„P.S.: Do you use amplitude information? If so, be really careful with the temporary data. I am prepared to bet that you can have amplitude errors of a factor of up to 10 …

This type of caution even applies for many permanent observatory stations. I could show you an interesting figure on that subject, which shows that approximately 40% of permanent broad-band stations have amplitude errors >10% …“

Misfit functionals: Principal Desiderata

A misfit functional suitable for waveform tomography should:

1) extract as much information as possible from seismic waveforms

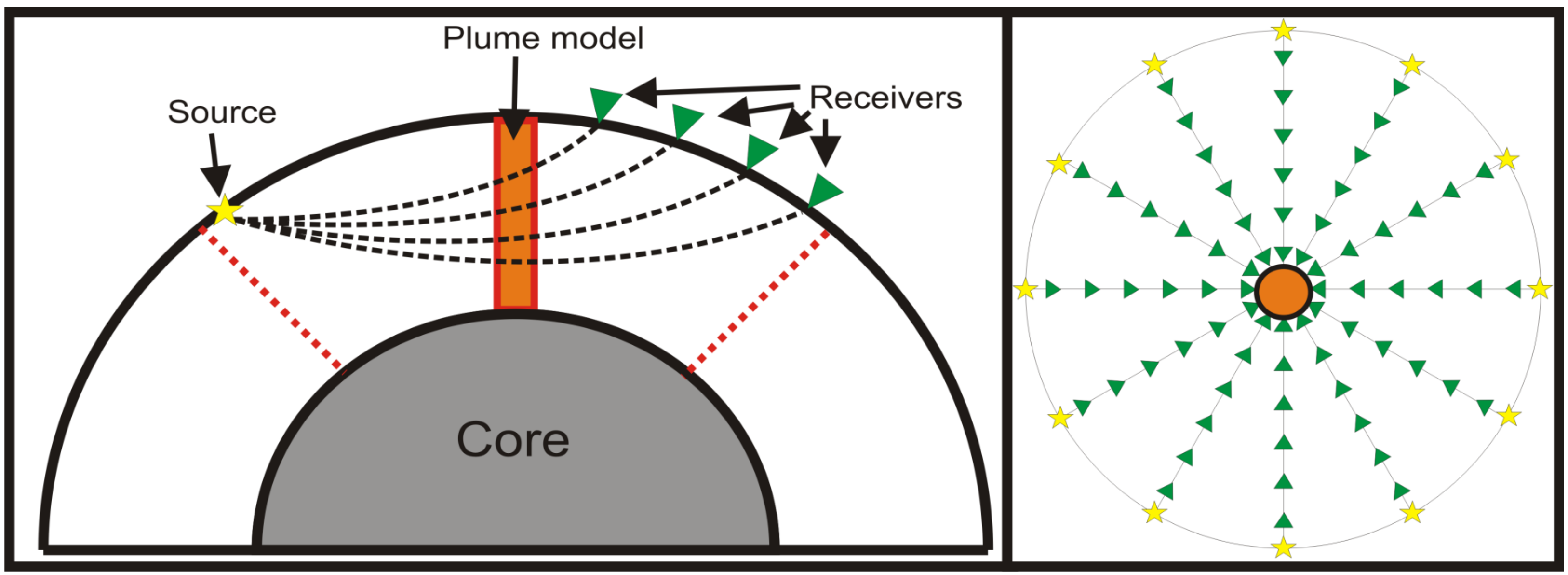
2) be quasi-linearly related to Earth structure (tame non-linearity)

3) be independent of absolute amplitudes

Design of misfit functionals

Example 1: Hunting for plumes in the mantle (Florian Rickers)

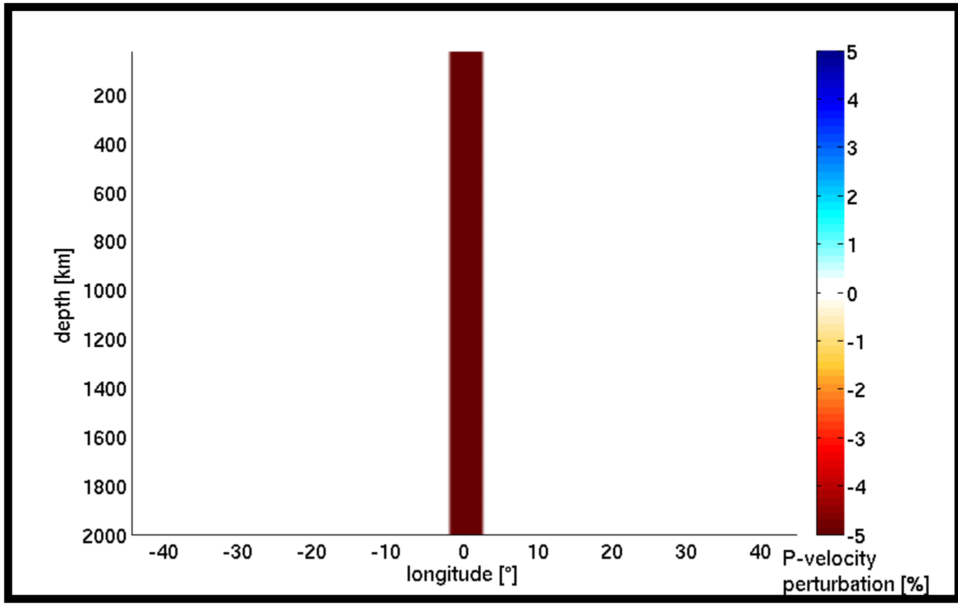
- Synthetic experiment to study the detectability of mantle plumes.
- Plume model (d=500km, -5% P wave speed, -10% S wav speed).
- Idealistic source-receiver geometry.
- Measurement of cross-correlation time shifts in 30 s P waves



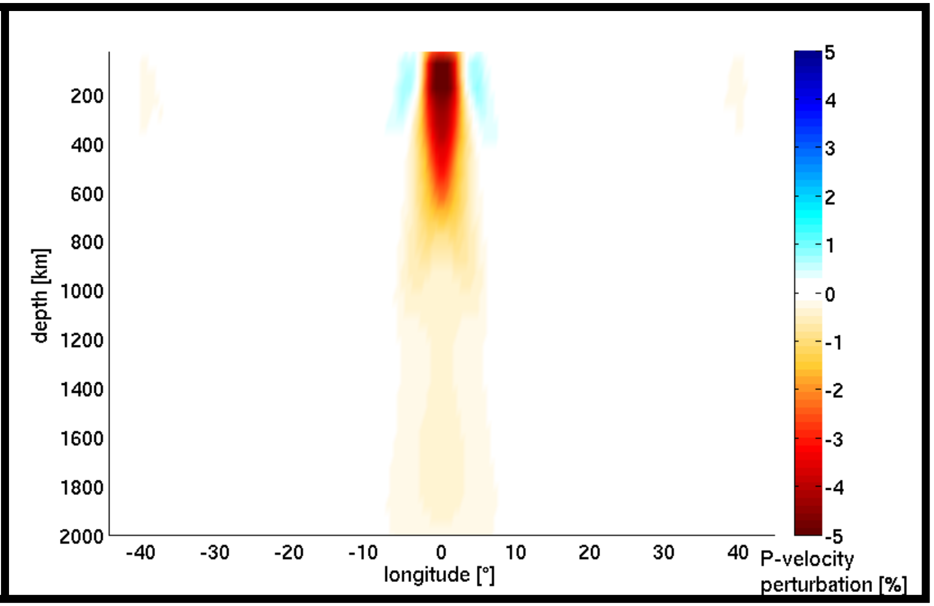
Design of misfit functionals

Example 1: Hunting for plumes in the mantle (Florian Rickers)

- Misfit reduction after 7 conjugate-gradient iterations: **95 %**
- Traveltimes are essentially explained.
- The plume remains diffuse and is not visible below 600 km.



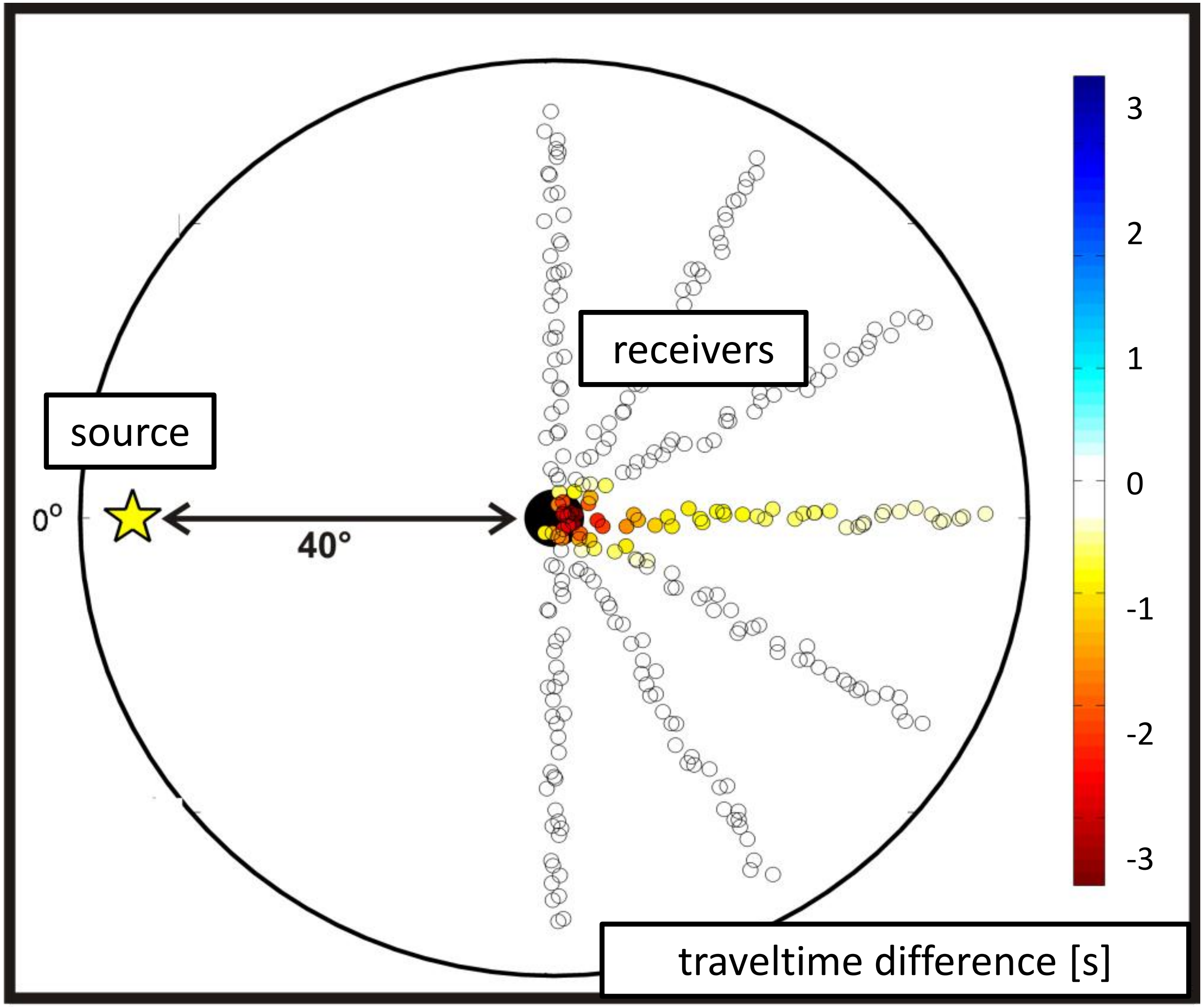
Original plume



Traveltime recovered plume

Design of misfit functionals

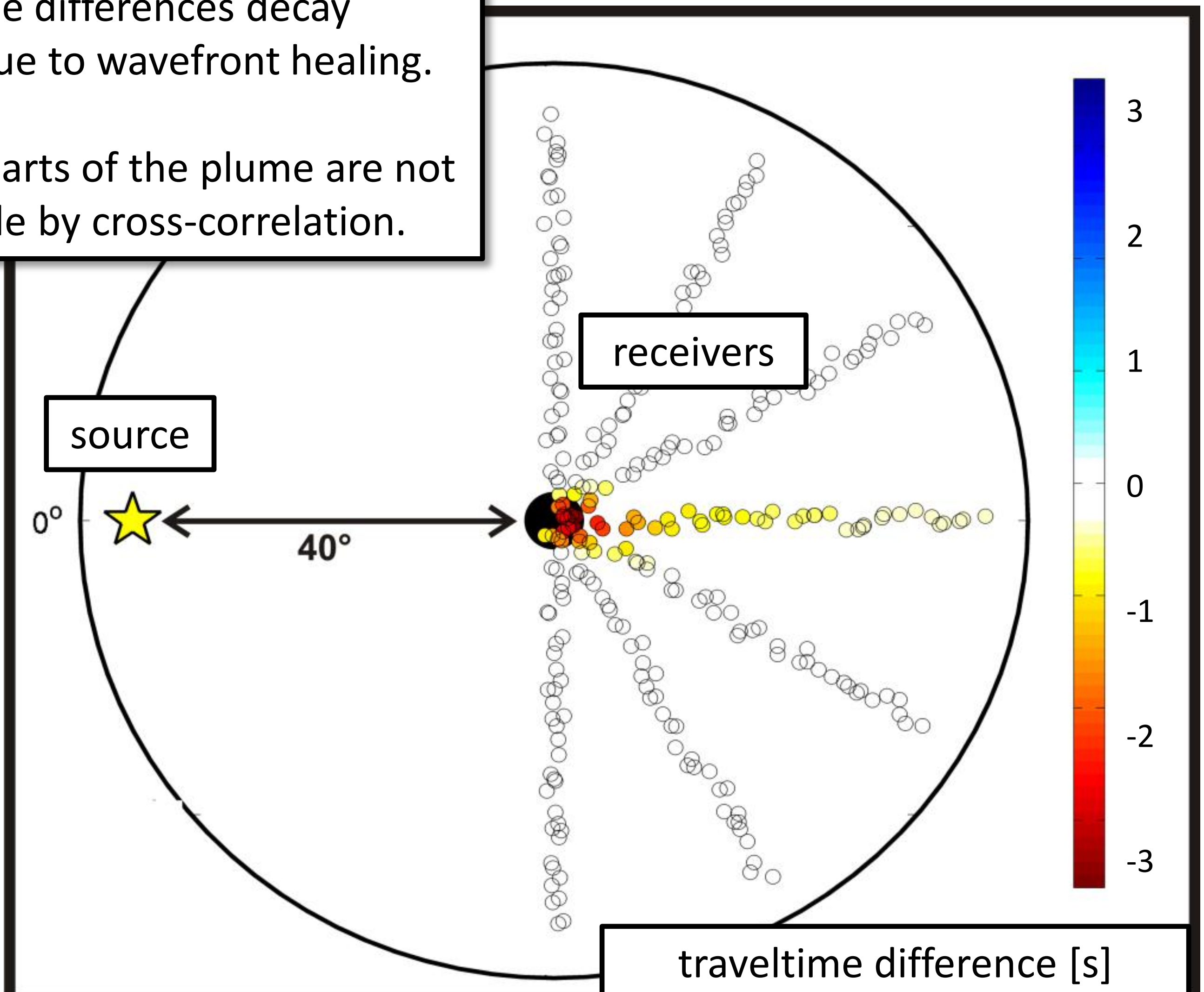
Example 1: Hunting for plumes in the mantle (Florian Rickers)



Design of misfit functionals

Example 1: Hunting for plumes in the mantle (Florian Rickers)

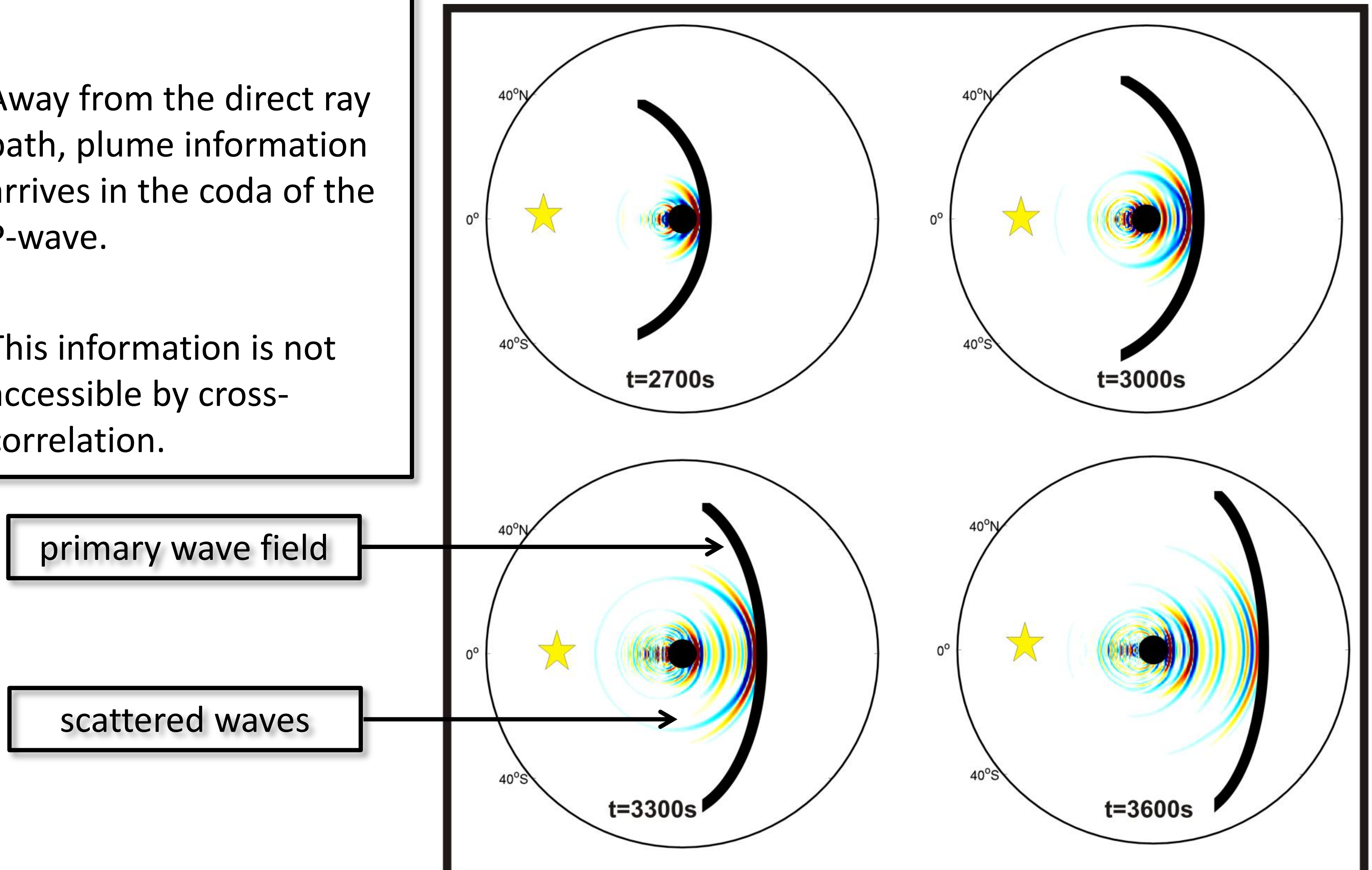
- Traveltime differences decay rapidly due to wavefront healing.
- Deeper parts of the plume are not detectable by cross-correlation.



Design of misfit functionals

Example 1: Hunting for plumes in the mantle (Florian Rickers)

- Plume acts as a scatterer.
- Away from the direct ray path, plume information arrives in the coda of the P-wave.
- This information is not accessible by cross-correlation.



Design of misfit functionals

Example 1: Hunting for plumes in the mantle (Florian Rickers)

Alternative measurement: **instantaneous phase differences of the complete seismogram**

- analytic signal: $\tilde{f}(t) = f(t) + i\,Hf(t)$
- instantaneous phase: $\Phi(t) = \arctan \dfrac{\operatorname{Im} \tilde{f}}{\operatorname{Re} \tilde{f}}$
- instantaneous phase misfit: $\chi = \sum_{i=1}^{N_r} \int [\Phi_i^0(t) - \Phi_i(t)]^2\, dt$

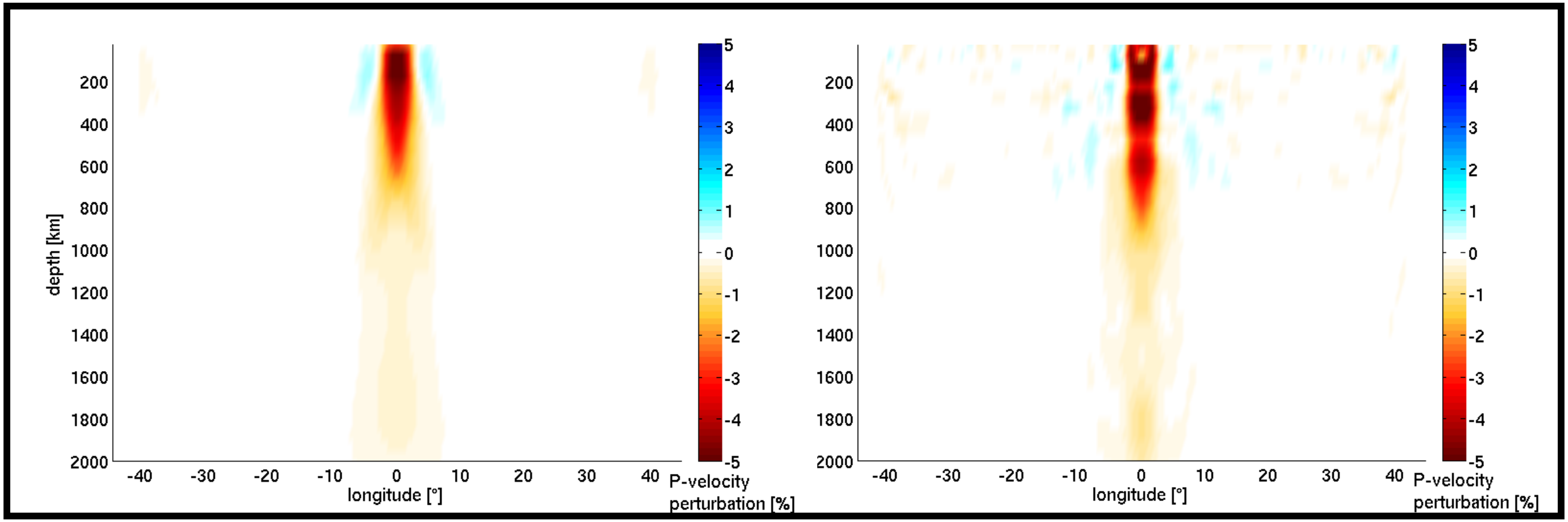
Advantages:

- Considers the complete time series, including the scattered waves.
- All the „wiggles“ are automatically weighted equally.
- Independent of absolute amplitudes.

Design of misfit functionals

Example 1: Hunting for plumes in the mantle (Florian Rickers)

- Instantaneous phase inversion for the synthetic plume.
- Use the P wave train until the arrival of the S wave.
- Misfit reduction of 90 % after 9 iterations.
- Much sharper image of the plume (… but still not to great depth).

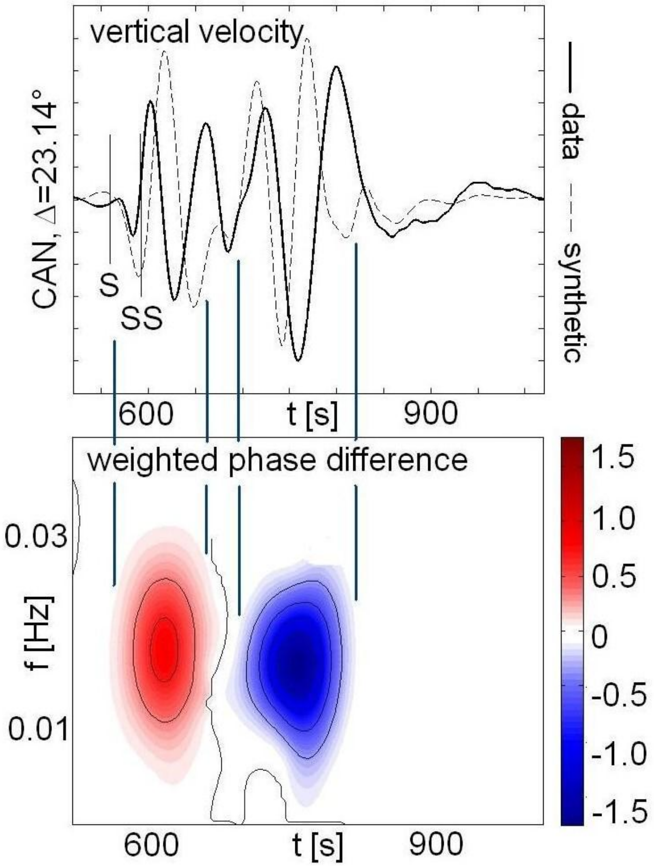


Cross correlation traveltime

Instantaneous phase

Design of misfit functionals

Example 2: Time-frequency misfits

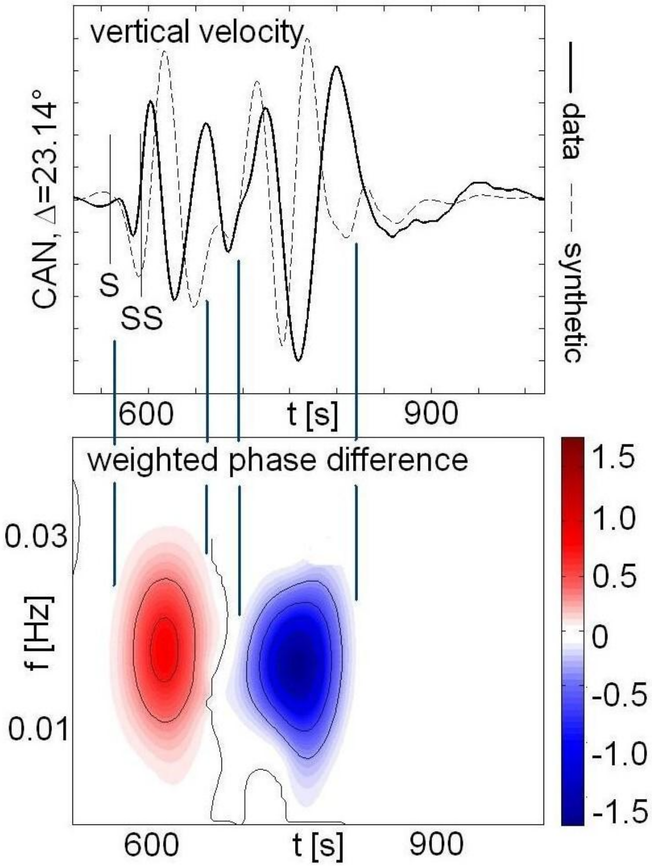


Time-frequency misfits

phase differences as functions of time and frequency

Design of misfit functionals

Example 2: Time-frequency misfits



Time-frequency misfits

phase differences as functions of time and frequency

- **quasi-linearly related to Earth structure**

  improves convergence

- **independent of amplitudes**

  reliably measurable, deep structure information

- **applicable to complex waveforms**

  interfering waves, unidentifyable waves

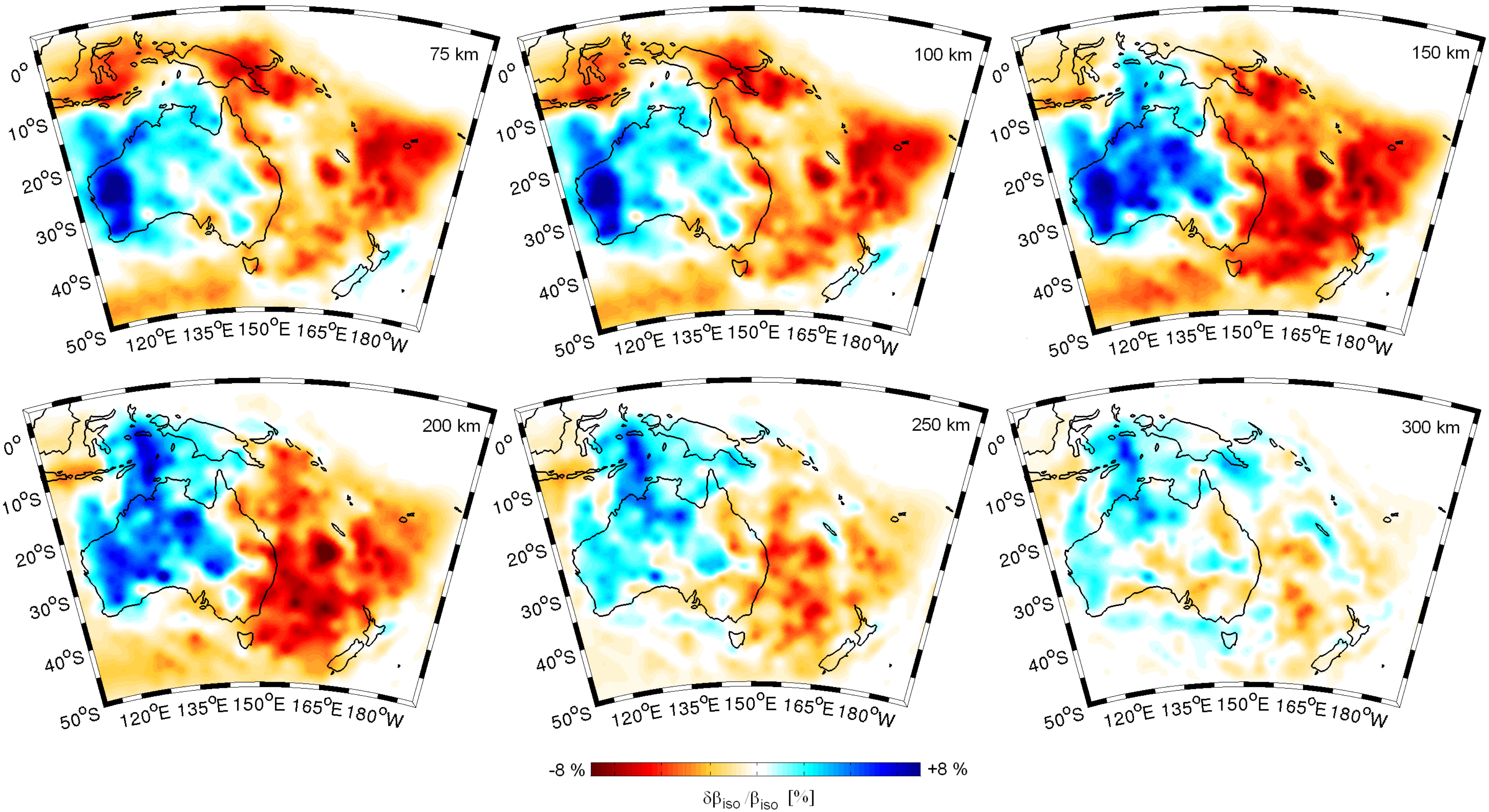
- **continuous in frequency**

  no discrete frequency bands

Design of misfit functionals

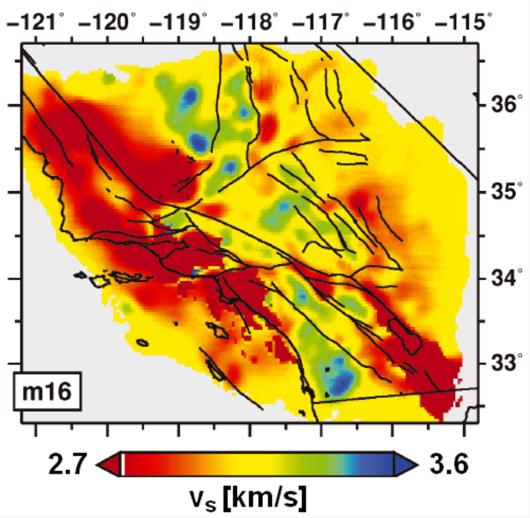
Example 2: Time-frequency misfits

isotropic S wave speed

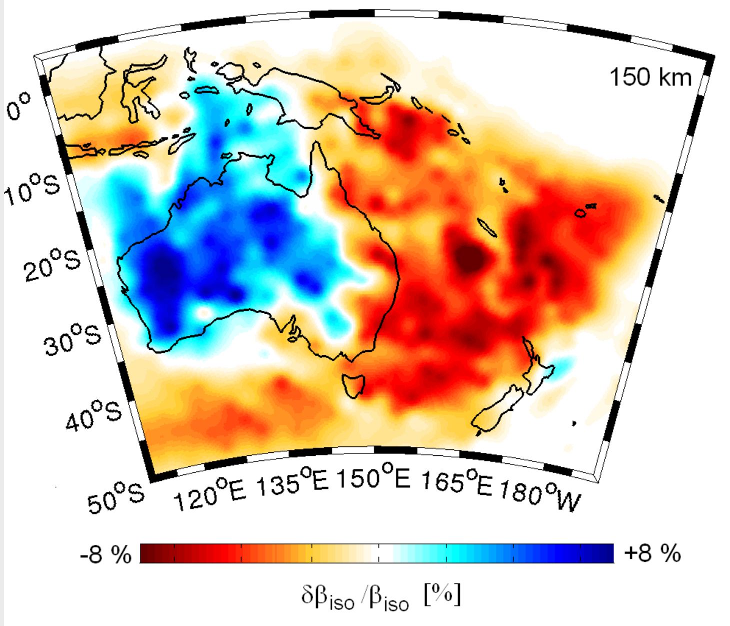


Towards quantitative resolution analysis

Resolution analysis



Southern California Tape et al. (2009,2010)



Australasia Fichtner et al. (2009,2010)

- **How can we quantify the trustworthyness of FWI images ?**
- **Arguments typically involve:**
  - **(1) visual analysis**
  - **(2) synthetic inversions**

- **Are FWI images really any better ?**
- **Is FWI worth the effort ?**

Resolution analysis

- Complications: non-linearity, multiple minima, finite computational resources, …
- Quadratic approximation of the misfit functional

$$\chi(\mathbf{m}_{\mathrm{opt}} + \delta\mathbf{m}) \approx \chi(\mathbf{m}_{\mathrm{opt}}) + \delta\mathbf{m}^{T}\mathbf{H}\,\delta\mathbf{m}$$

misfit functional

optimal Earth model

Hessian at $\mathbf{m}_{\mathrm{opt}}$

- The Hessian H:

  (1) local geometry of the misfit surface

  (2) resolution and trade-offs

  (3) H = inverse posterior covariance

Resolution analysis

- **Efficient computation of $H \cdot \delta m$ via an extension of the adjoint method**

  (1) 2 forward simulations

  (2) 2 adjoint simulations (time-reversed)

Resolution analysis

- **Efficient computation of H·δm via an extension of the adjoint method**

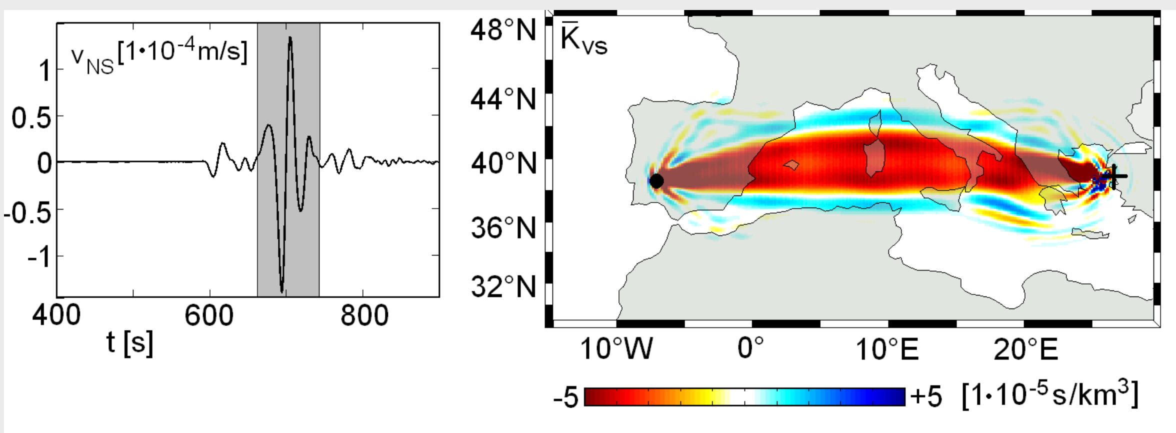
  (1) 2 forward simulations

  (2) 2 adjoint simulations (time-reversed)

- **<u>Example</u>: 25 s Love wave**

  finite-frequency traveltime

Fréchet kernel



Fichtner & Trampert, *Hessian kernels of seismic data functionals.* GJI, submitted.

Resolution analysis

- Efficient computation of $H \cdot \delta m$ via an extension of the adjoint method

  (1) 2 forward simulations

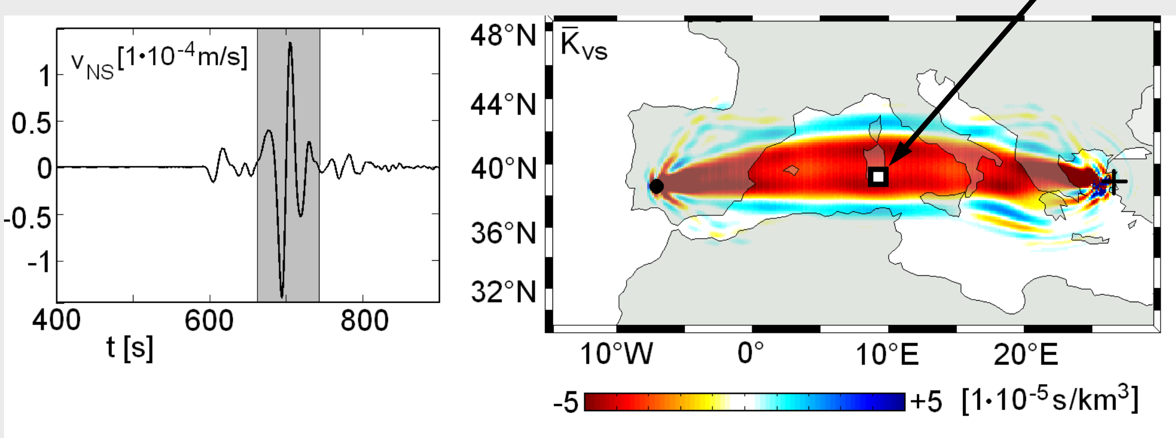
  (2) 2 adjoint simulations (time-reversed)

- Example: 25 s Love wave

  finite-frequency traveltime

$\delta m = v_s$ perturbation in pixel *k*

Fréchet kernel



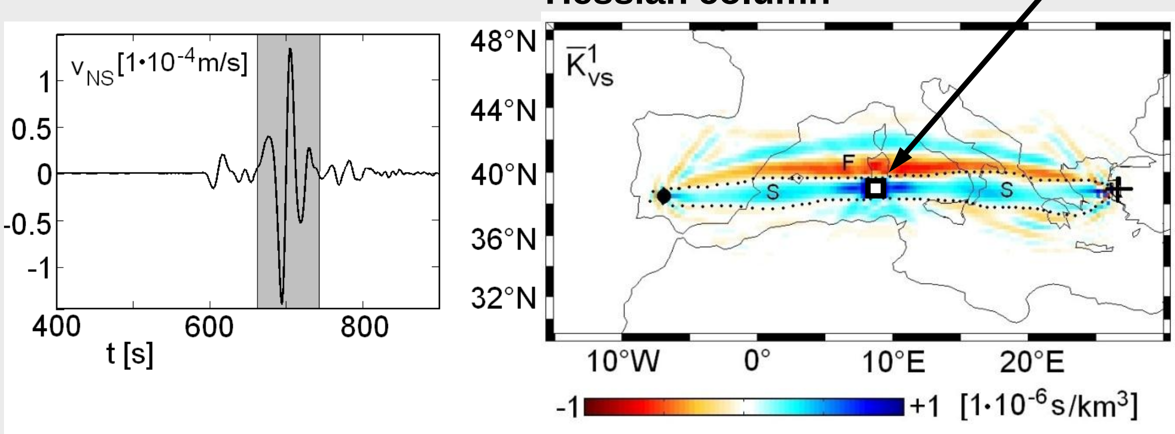
Fichtner & Trampert, *Hessian kernels of seismic data functionals.* GJI, submitted.

- **Efficient computation of $H \cdot \delta m$ via an extension of the adjoint method**
  - **(1) 2 forward simulations**
  - **(2) 2 adjoint simulations (time-reversed)**

- **<u>Example</u>: 25 s Love wave**
  finite-frequency traveltime

δm = v_s perturbation in pixel *k*

Hessian column



$$\mathbf{H} = \begin{pmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

column *k*

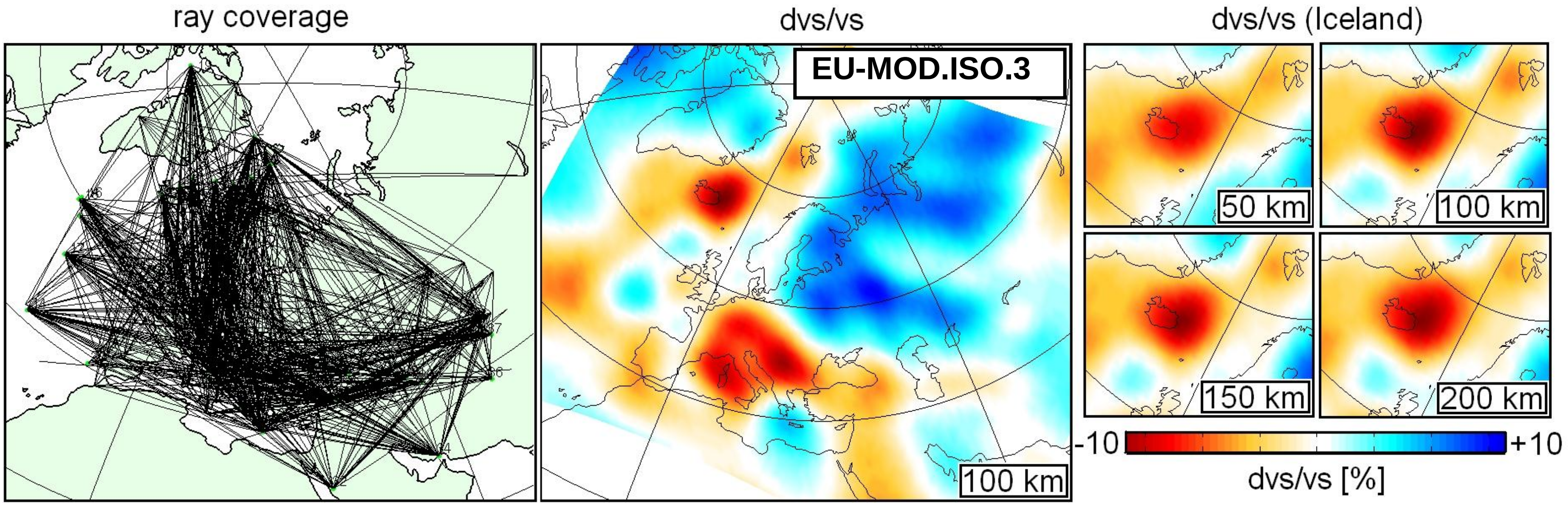
F: approximate Hessian (first-order scattering)

S: second-order scattering

Fichtner & Trampert, *Hessian kernels of seismic data functionals.* GJI, submitted.

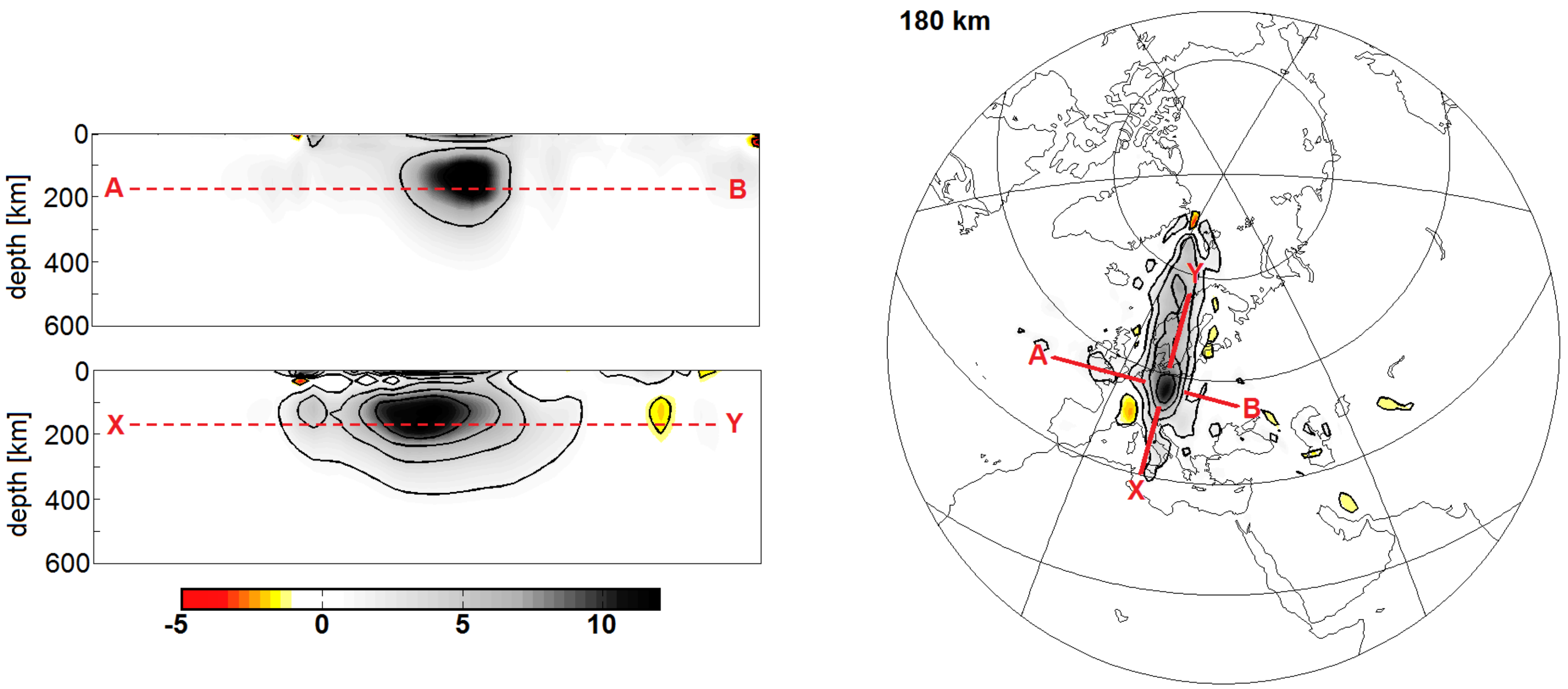
Resolution analysis

- **Long-wavelength FWI image of Europe: *EU-MOD.ISO.3***
- **Data:** **≈ 3000 waveforms**
  dominant period: 100 s
- **3 iterations**



Resolution analysis

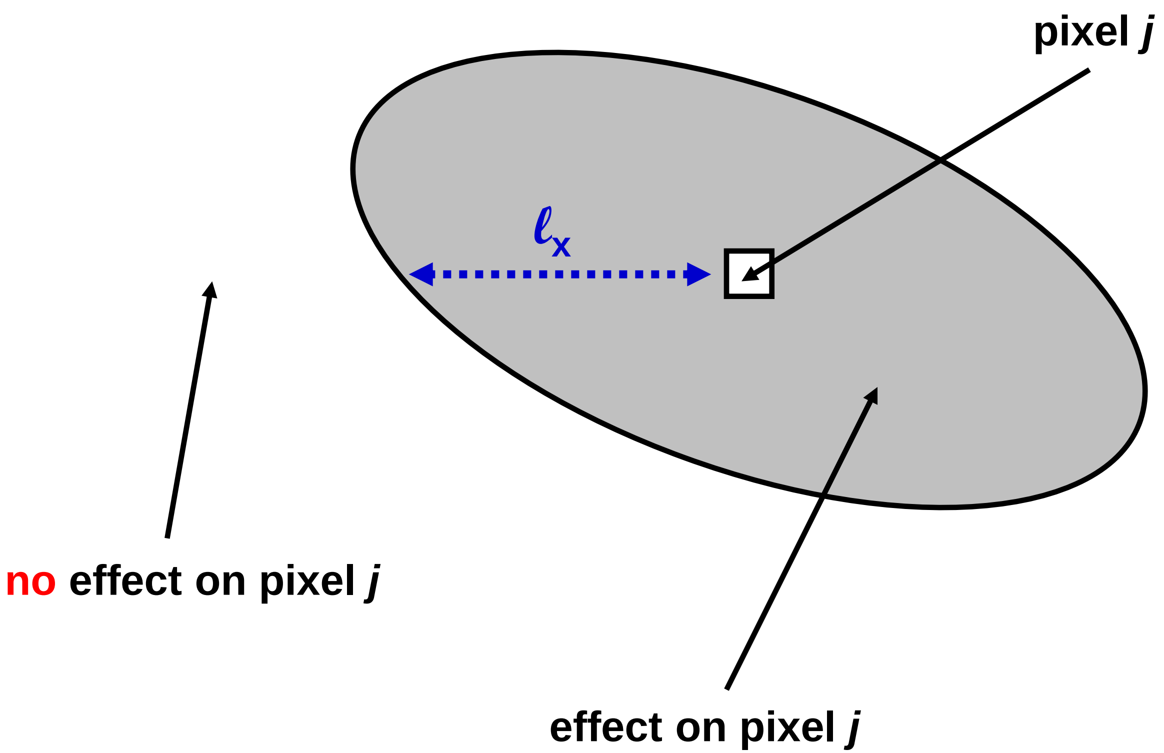
Example: A Hessian column (kernel) for northern Germany



- Ideally, we would like to have such a Hessian kernel for each point in the model.
- This would be a computationally expensive exercise, but ...

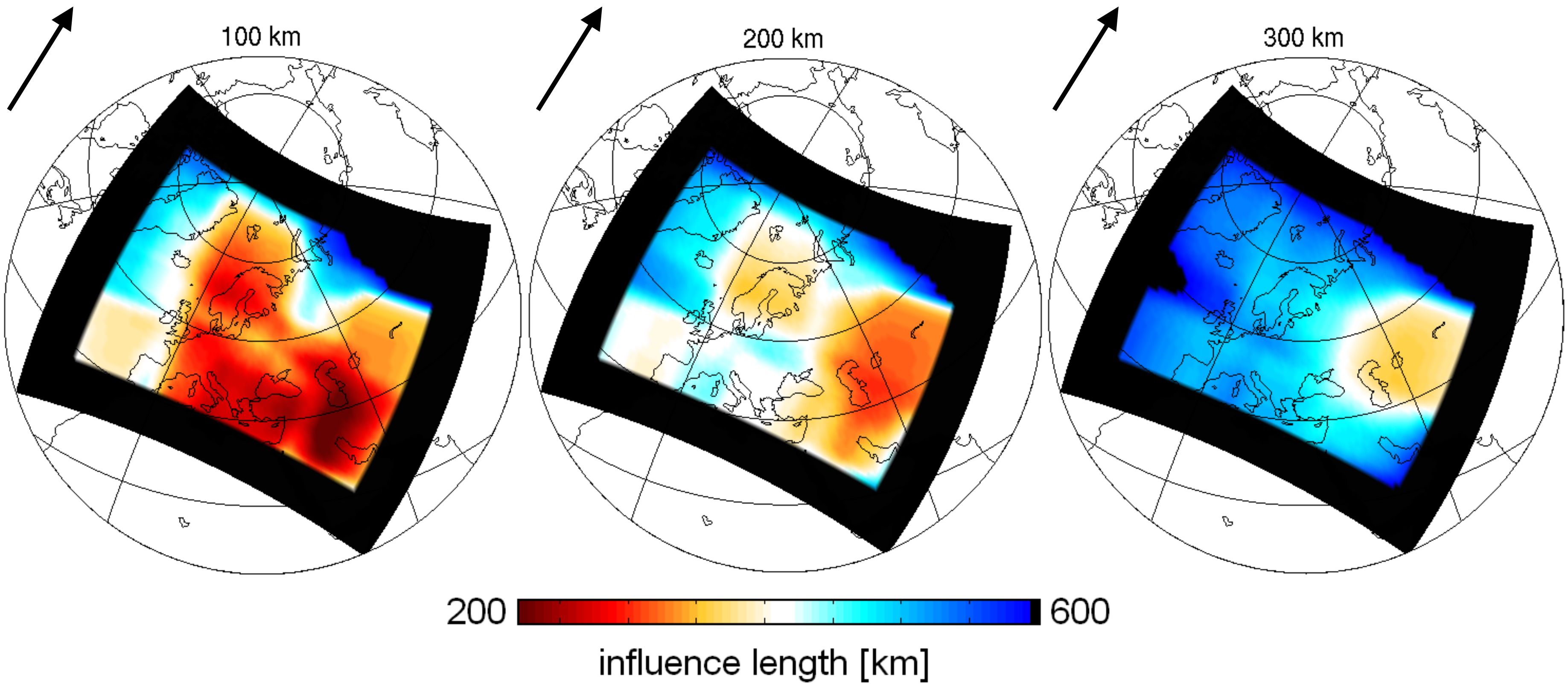
Resolution analysis

- **H·**exp(i**k·x**) = **Fourier transform of the H-columns**
- **approximate H from a small number of wave number vectors k**
- **estimate width of the <u>influence length ℓ</u> in different directions**



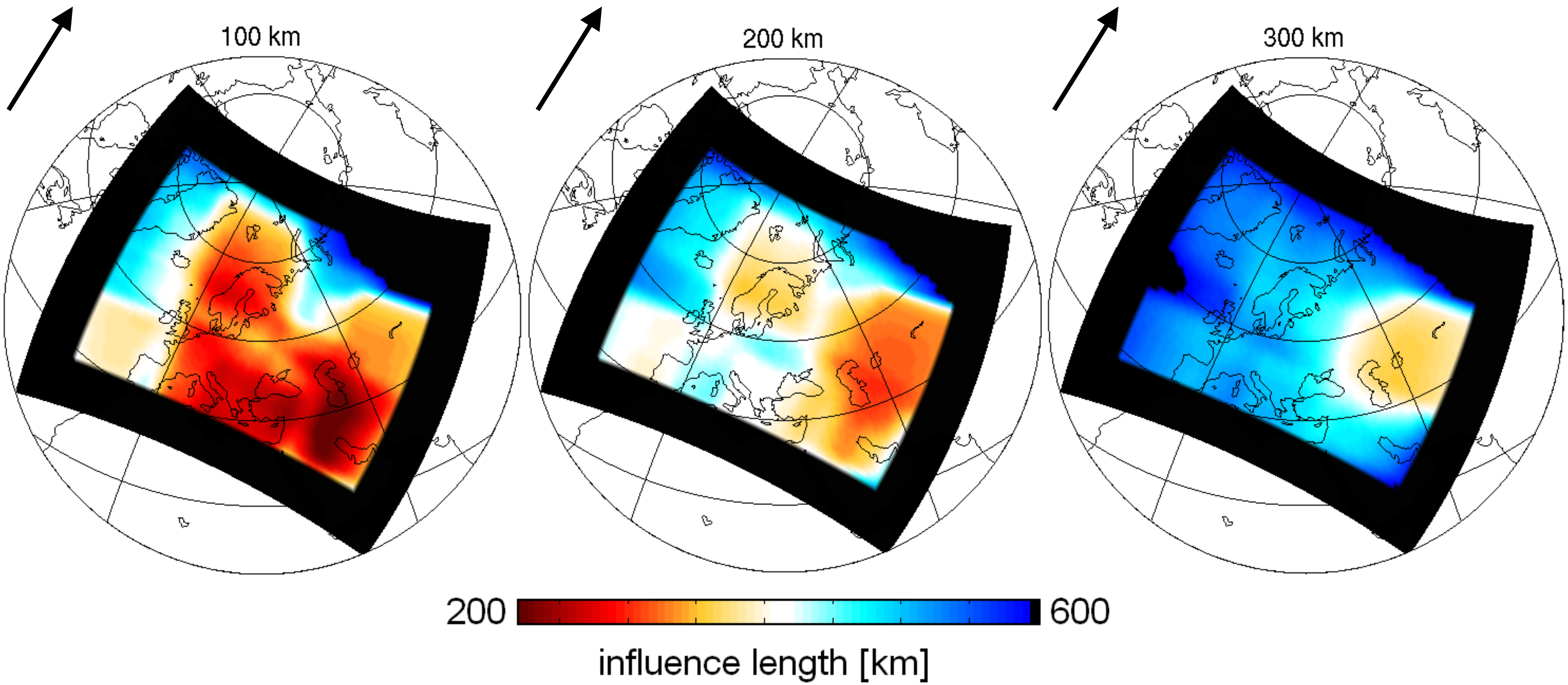
Resolution analysis

- **H·exp(i$\mathbf{k}$·$\mathbf{x}$) = Fourier transform of the H-columns**
- **approximate H from a small number of wave number vectors k**
- **estimate width of the <u>influence length</u> ℓ in different directions**



Resolution analysis

- $H \cdot \exp(ik \cdot x)$ = **Fourier transform of the H-columns**
- **approximate H from a small number of wave number vectors k**
- **estimate width of the <u>influence length</u> ℓ in different directions**



- **good estimate of H from a hand full of forward and adjoint simulations**

Resolution analysis

1. extension of the adjoint method to compute $H \cdot \delta m$

2. $H \cdot \delta m$ = superposition of approximate Hessian and second-order scattering

3. resolution and influence region of geologic structures (e.g. Iceland plume)

4. trade-offs between parameters (e.g. v_s, v_p, ρ)

5. spectral estimation of the Hessian through Fourier transforms

6. continuous distribution of the influence length

First step towards quantitative resolution analysis in full waveform inversion.

Lorentz Center Workshop on

Uncertainty Analysis in Geophysical Inverse Problems

7 – 11 November 2011
Leiden, The Netherlands

Lorentz center