Receiver Function Inversion

Advanced Studies Institute on Seismological Research

Kuwait City, Kuwait - January 19-22, 2013

Jordi Julià

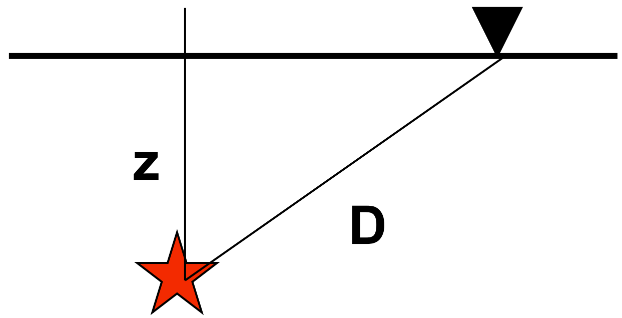
Universidade Federal do Rio Grande do Norte, Brasil

Outline

- Introduction to Inverse Theory:
  - Forward and inverse problems
  - Iterative solution: LSQ and damped LSQ
  - Generalized inverse
- Inversion of Receiver Functions:
  - Method of Ammon *et al.* (1990).
  - The non-uniqueness problem.
- Case Studies in Spain:
  - Ebre basin (Julià et al., 1998)
  - Neogene Volcanic Zone (Julià et al., 2005)

Forward Problem / Inverse Problem

- Seismic location:
  - Data: travel times
  - Unknowns: hypocentral coordinates and origin time.
  - *A priori* information: station locations and propagating medium velocities.



$$t_i = t_0 + D_i/V$$

- Forward problem:
  - Predict travel times from known hypocentral location and origin time.
- Inverse problem:
  - Obtain hypocentral location and origin time from observed travel times.

Setting up the (forward) problem

We define a vector of observations **d** and a vector of parameters **m** as:

$$\mathbf{d} = (t_1, t_2, \ldots, t_N)^T \qquad \mathbf{m} = (t_0, x_0, y_0, z_0)^T$$

so that

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

where $F_i(\mathbf{m})$ is

$$t_i = t_0 + (1/v)\,[(x_i-x_0)^2+(y_i-y_0)^2+(z_i-z_0)^2]^{1/2}$$

Inverse theory provides means for finding an operator $\mathbf{F}^{-1}(\mathbf{d})$, so that

$$\mathbf{m} = \mathbf{F}^{-1}(\mathbf{d})$$

Iterative solution

The forward problem for seismic location is non-linear. An approach is to turn it linear by doing a Taylor expansion around a trial solution $\mathbf{m}_0$

$$\mathbf{d} \approx \mathbf{F}(\mathbf{m}_0) + \nabla\mathbf{F}|_{\mathbf{m}0} \cdot(\mathbf{m}-\mathbf{m}_0)$$

and drop 2nd and higher order terms, so that

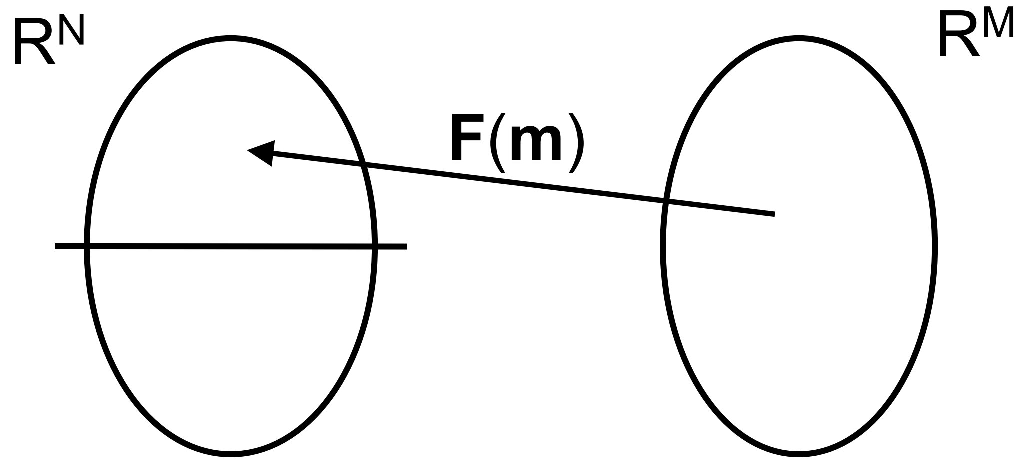
$$\Delta\mathbf{d} = G\cdot\Delta\mathbf{m}$$

Where $\Delta\mathbf{d}=\mathbf{d}-\mathbf{F}(\mathbf{m}_0)$, $\Delta\mathbf{m}=\mathbf{m}-\mathbf{m}_0$, and

$$G = \nabla\mathbf{F}|_{\mathbf{m}0} = \begin{pmatrix} \partial t_1/\partial t_0 & \partial t_1/\partial x_0 & \partial t_1/\partial y_0 & \partial t_1/\partial z_0 \\ : & : & : & : \\ \partial t_N/\partial t_0 & \partial t_N/\partial x_0 & \partial t_N/\partial y_0 & \partial t_N/\partial z_0 \end{pmatrix}$$

If we can determine G^{-1}, then $\mathbf{m}_{i+1} = \mathbf{m}_i+\Delta\mathbf{m}_i$

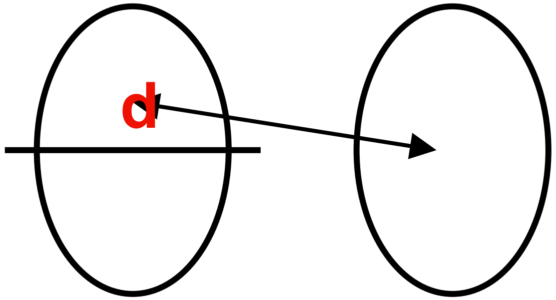
Classifying Inverse Problems



The (linear) vector function $\mathbf{d}=\mathbf{F}(\mathbf{m})$ maps the parameter space into a subspace of the data space.

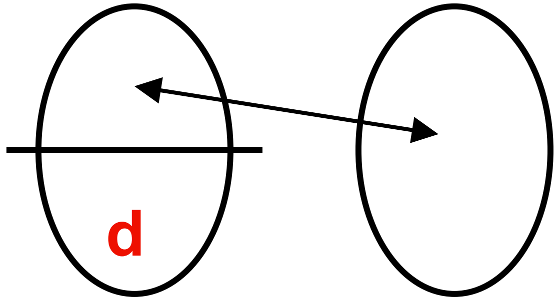
The ability of establishing an inverse mapping $\mathbf{m}=\mathbf{F}^{-1}(\mathbf{d})$ depends on the details of the forward mapping.

Ideal case



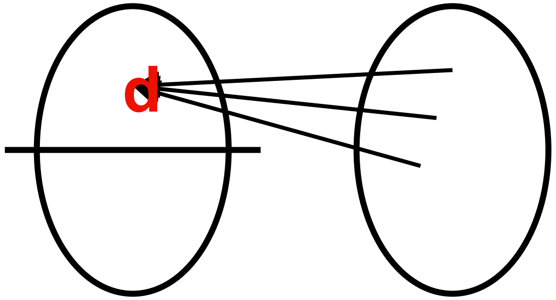
Each vector **d** relates to one and only one vector **m**.

Overdetermined case



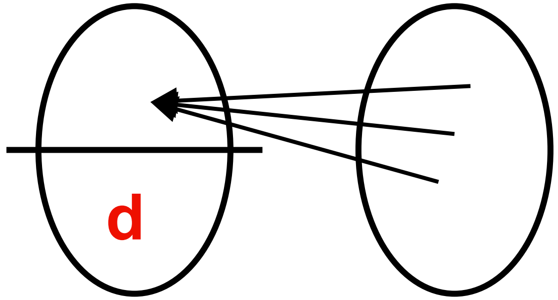
There is no exact solution, so we must choose one that is close enough.

Underdetermined case



There are multiple solutions. We must pick one.

Mixed-determined case



There are no exact solutions and many that are equally close.

Least squares solutions

In order to define "close" in the data space we need to introduce a metric. A popular choice is the L_2 norm, where the "distance" E between vectors is

$$E = (\mathbf{d}-\mathbf{F}(\mathbf{m}))^T(\mathbf{d}-\mathbf{F}(\mathbf{m}))$$

The "closest" solution is obtained by minimizing E and is given by

$$G^{-1} = [G^TG]^{-1}G^T$$

To choose among the multiple solutions that are equally "close" we pick the one that is minimum length

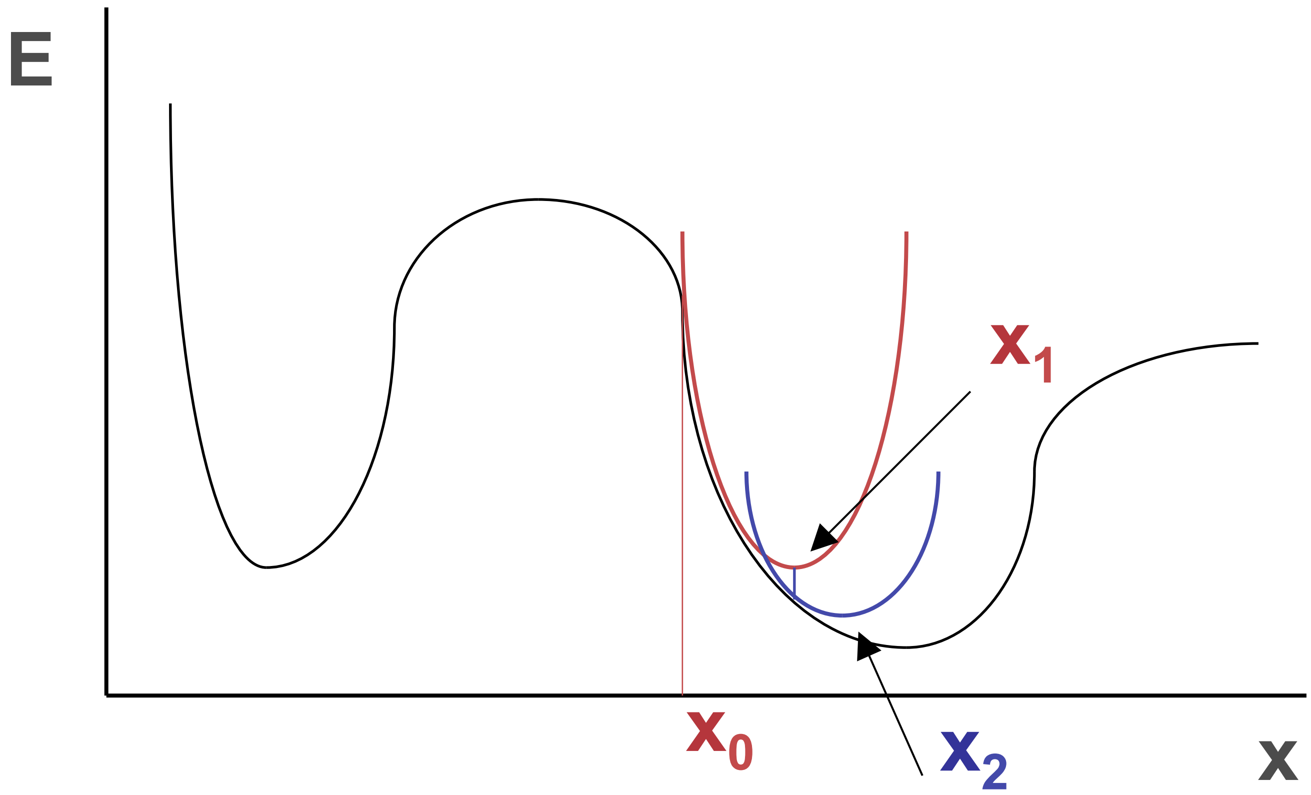
$$E = (\mathbf{d}-\mathbf{F}(\mathbf{m}))^T(\mathbf{d}-\mathbf{F}(\mathbf{m})) + \vartheta^2 (\mathbf{m}^T\mathbf{m})$$

This is called the "damped least squares" solution and is given by

$$G^{-1} = [G^TG+\vartheta^2 I]^{-1}G^T$$

Iterative least squares solution

The figure below gives a graphical illustration of how iterative least squares works:



Generalized Inverse Solution (I)

Another way of obtaining G^{-1} is based on the singular value decomposition (SVD) of matrix G.

It can be shown that, in general, any matrix G can be decomposed according to

$$G = U \Lambda V^T$$

Where $U = [\mathbf{u}_1,\ldots,\mathbf{u}_N]$ is a base in the data space, $V=[\mathbf{v}_1, \ldots,\mathbf{v}_M]$ is a base in the parameter space, and Λ is a N x M matrix given by

$$\Lambda = \begin{bmatrix} \Lambda_P & 0 \\ 0 & 0 \end{bmatrix}$$

where Λ_P is a pxp diagonal matrix, with $p \leq M$. The diagonal values λ_i are called the singular values.

Generalized Inverse Solution (II)

If we define $V=[V_P,V_0]$ and $U=[U_P,U_0]$, we can write that

$$G = U_P \Lambda_P V_P^T$$

so that

$$G^{-1} = V_P \Lambda_P^{-1} U_P^T$$

The difficult part is to choose a value for p, as singular values can be small but NOT necessarily zero. Options are:

1) We choose $\lambda^{-1} = \lambda/(\lambda^2+\vartheta^2)^{-1}$. Then the SVD inverse is the damped least squares solution.

2) We choose $\lambda^{-1} = 0$, for λ small. Then the SVD inverse is called generalized inverse or natural solution.

Inversion of Ammon *et al.* (1990)

The inversion scheme developed by Ammon et al. (1990) is based on the "jumping" version of the iterative LSQ solution:

- Creeping

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

  $\mathbf{d} = \mathbf{F}(\mathbf{m}_0) + \nabla F|_{\mathbf{m}0} \; (\mathbf{m}\text{-}\mathbf{m}_0)$

  $\delta\mathbf{y} = \nabla F|_{\mathbf{m}0} \, \delta\mathbf{m}$

- Jumping

  $\mathbf{d} + \nabla F|_{\mathbf{m}0} \, \mathbf{m}_0 = \mathbf{F}(\mathbf{m}_0) + \nabla F|_{\mathbf{m}0} \, \mathbf{m}$

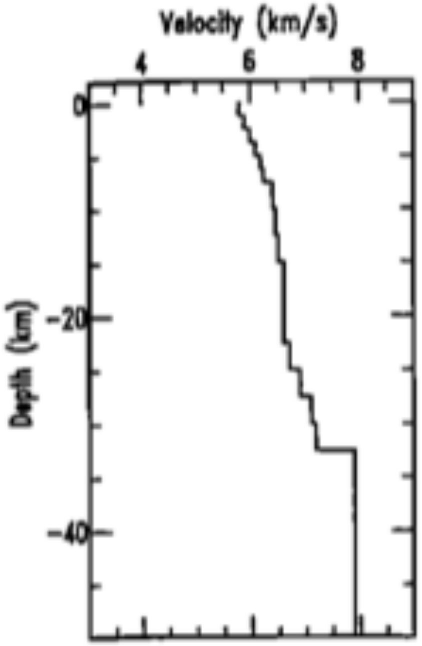
  $\Delta\mathbf{d} + \nabla F|_{\mathbf{m}0} \, \mathbf{m}_0 = \nabla F|_{\mathbf{m}0} \, \mathbf{m}$

- LSQ Norm

  $E = ||\Delta\mathbf{d} - \nabla F|_{\mathbf{m}0} \, (\mathbf{m} - \mathbf{m}_0)||^2$

Over-parameterization & regularization

Velocity models are over-parameterized through a stack of many thin layers of constant thickness and unknown S-velocity. A smoothness constrain is needed to stabilize the inversion.



$$\begin{cases} \Delta \mathbf{d} + \nabla F\, \mathbf{m}_0 = \nabla F|_{\mathbf{m}0}\, \mathbf{m} \\ \mathbf{0} = \sigma\, D\, \mathbf{m} \end{cases}$$

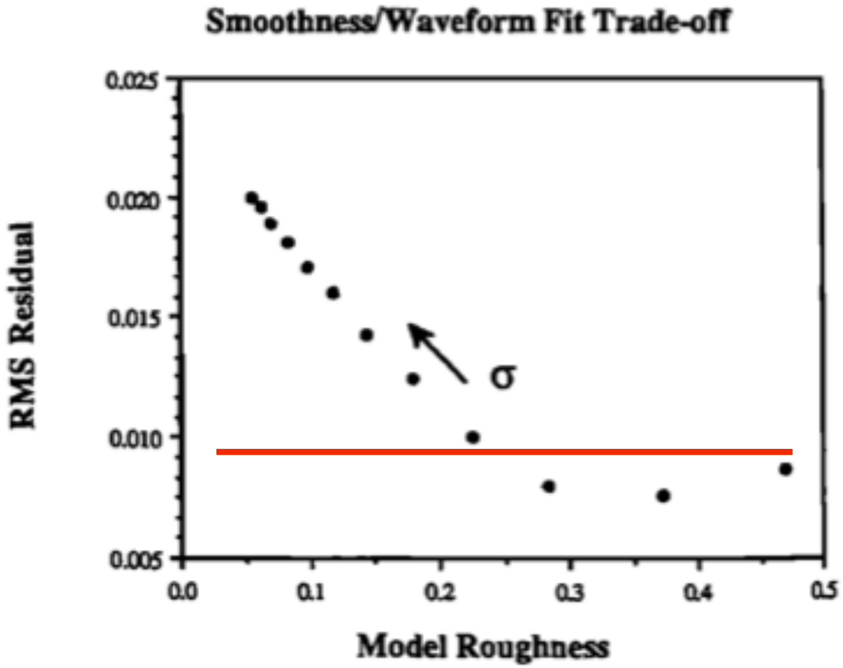
$$D\,\mathbf{m} = \begin{bmatrix} 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ & & \vdots & & \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ \vdots \end{bmatrix}$$

$$E = \|\, \Delta \mathbf{d} - \nabla F\, (\mathbf{m}-\mathbf{m}_0)\, \|^2 + \sigma^2\, \|\, D\mathbf{m}\, \|^2$$

Choosing the smoothness parameter

To determine the smoothness parameter σ a "preliminary" inversion is performed and a trade-off curve is built from the RMS error and the model roughness.

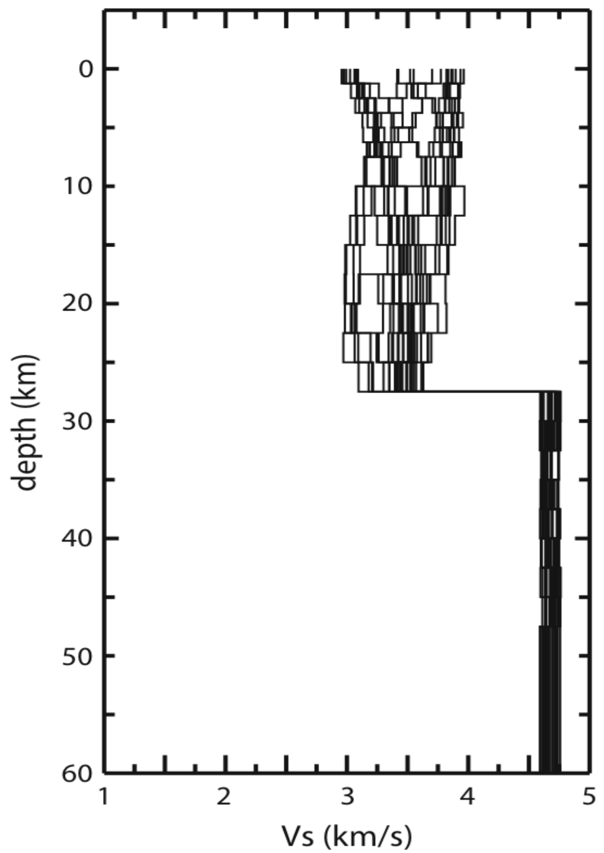
$$\text{roughness} = \sum_{i=1}^{n-2} \frac{|\alpha_i - 2\alpha_{i+1} + \alpha_{i+2}|}{(n-2)}$$

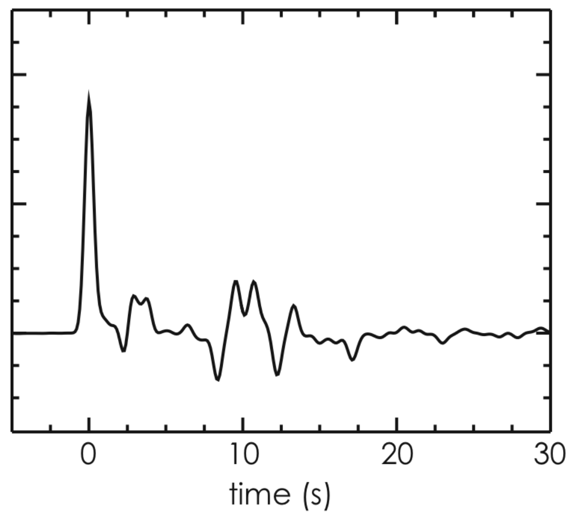


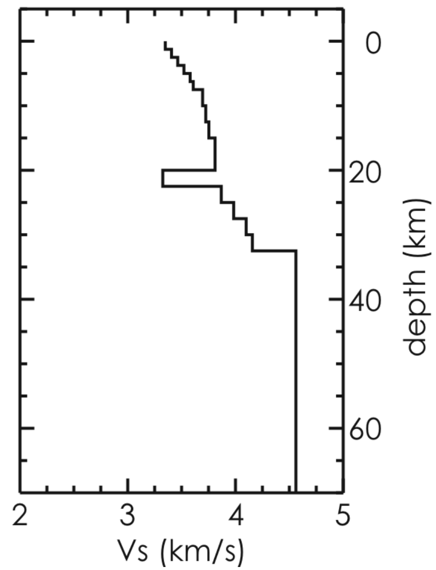
A value for σ is chose, for instance, from the noise level from the transverse RF.

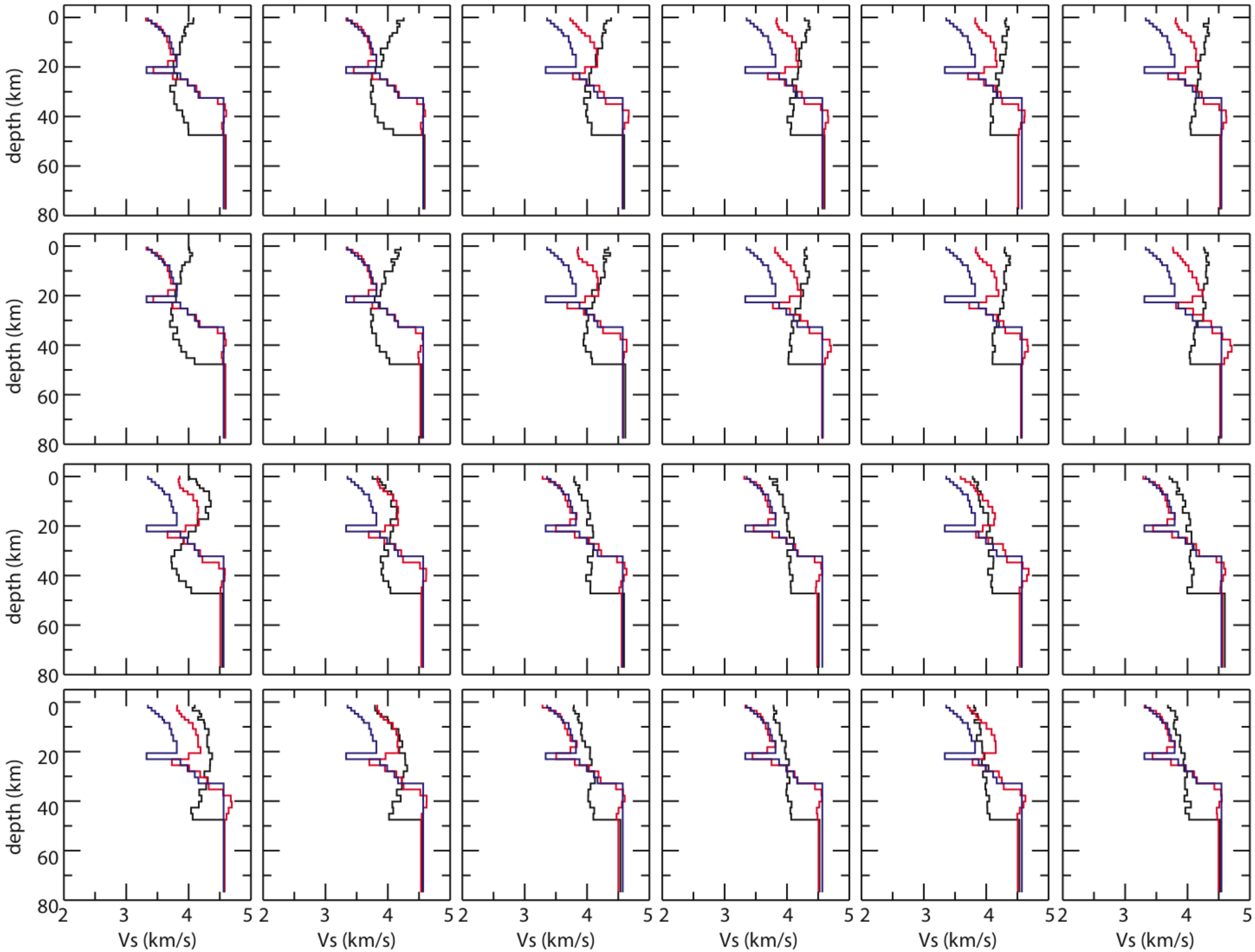
The non-uniqueness problem

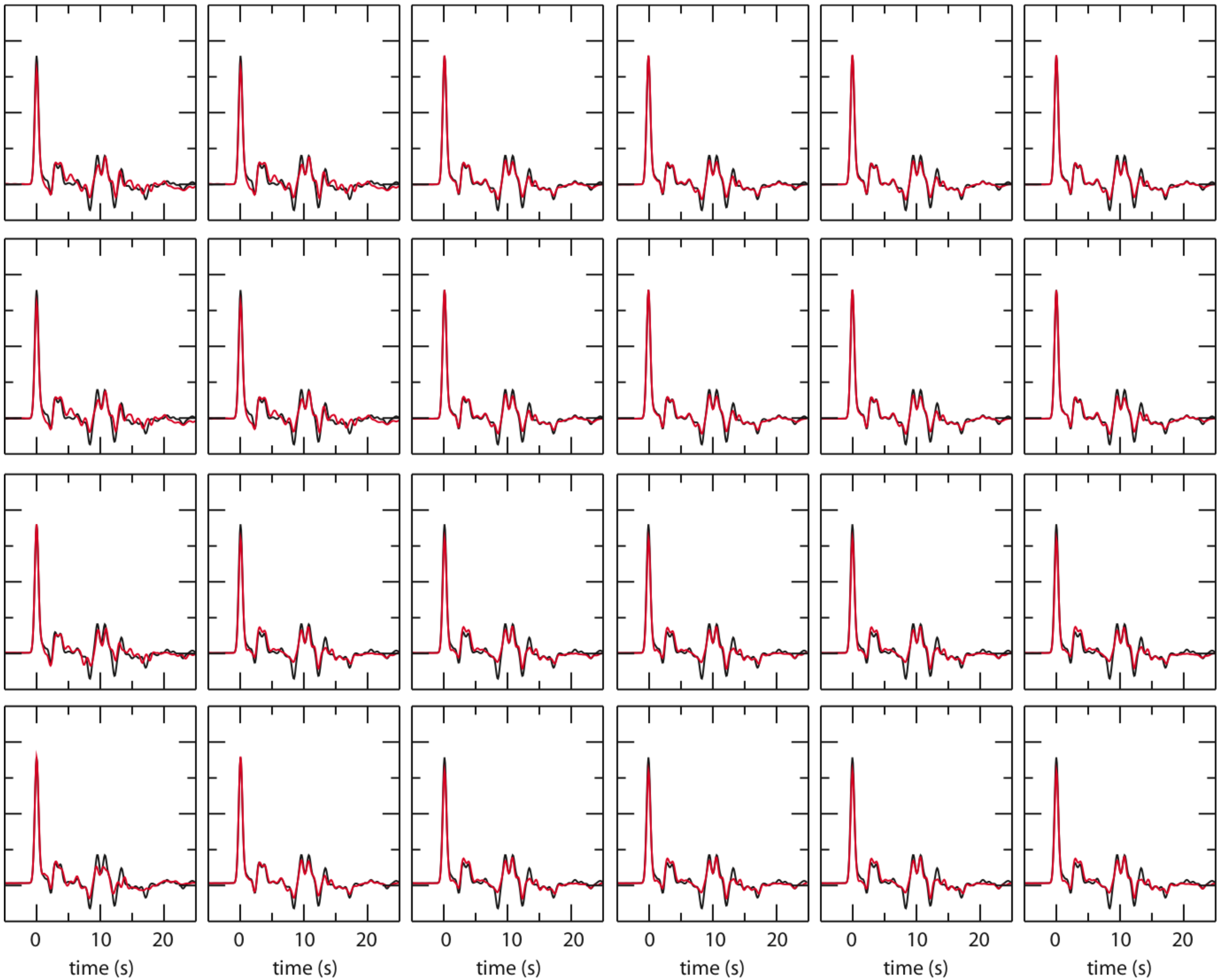
Ammon et al. (1990) showed that the modeling of receiver function waveforms is non-unique.



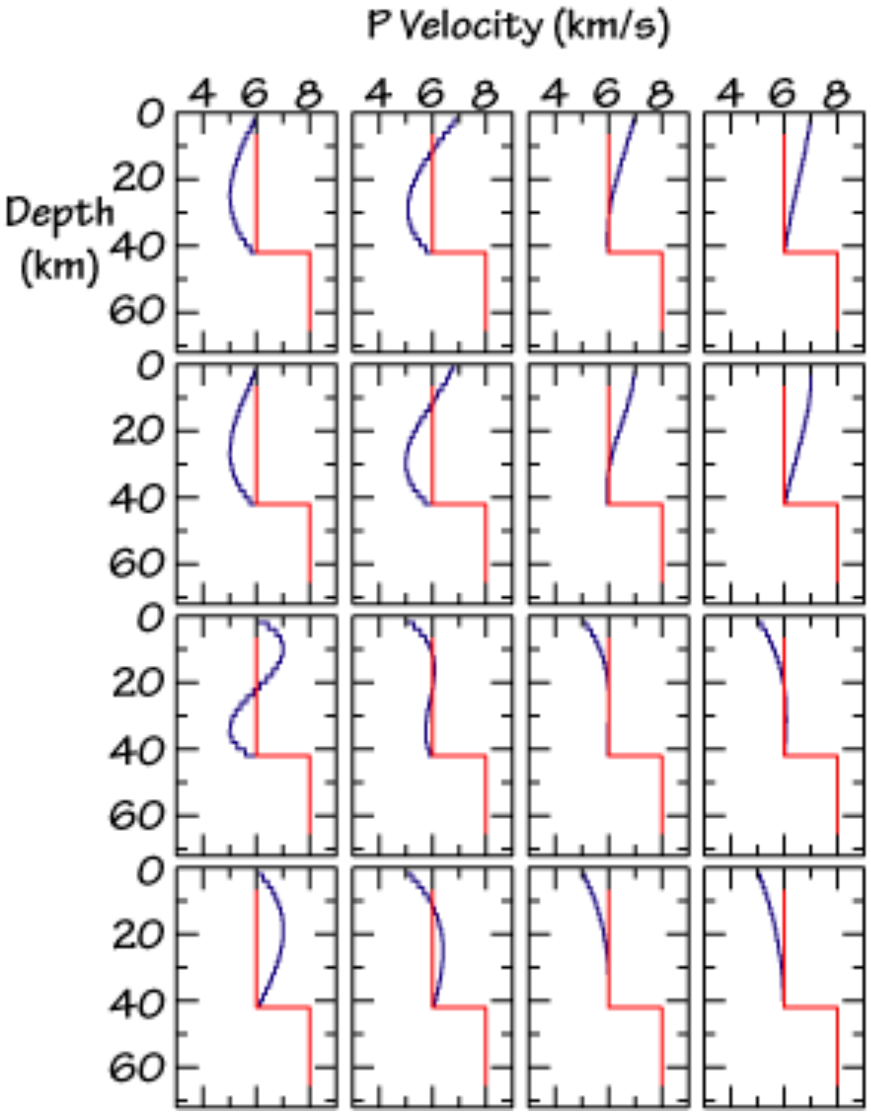




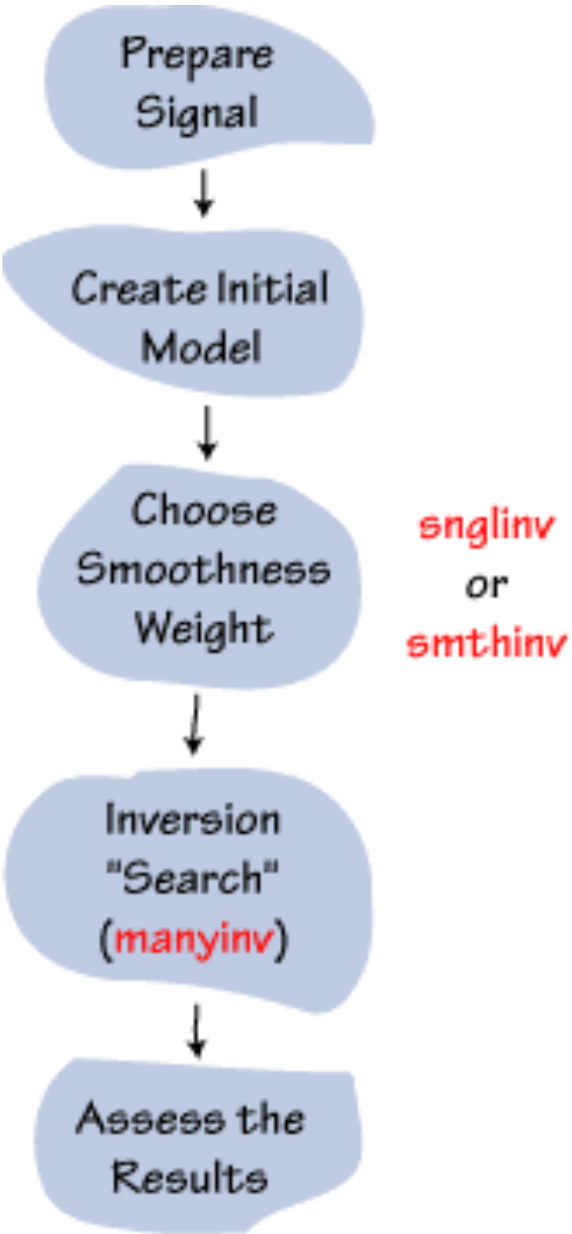

depth (km)
Vs (km/s)
0
20
40
60
80
2
3
4
5


0
10
20
time (s)

The perturbation scheme



- Many ‘starting’ models are obtained by perturbing an initial model.
- The perturbation scheme includes:
  - A cubic perturbation (up to a max value)
  - A random perturbation (up to a max %)
- Velocities above a cut-off value are not cubically perturbed.



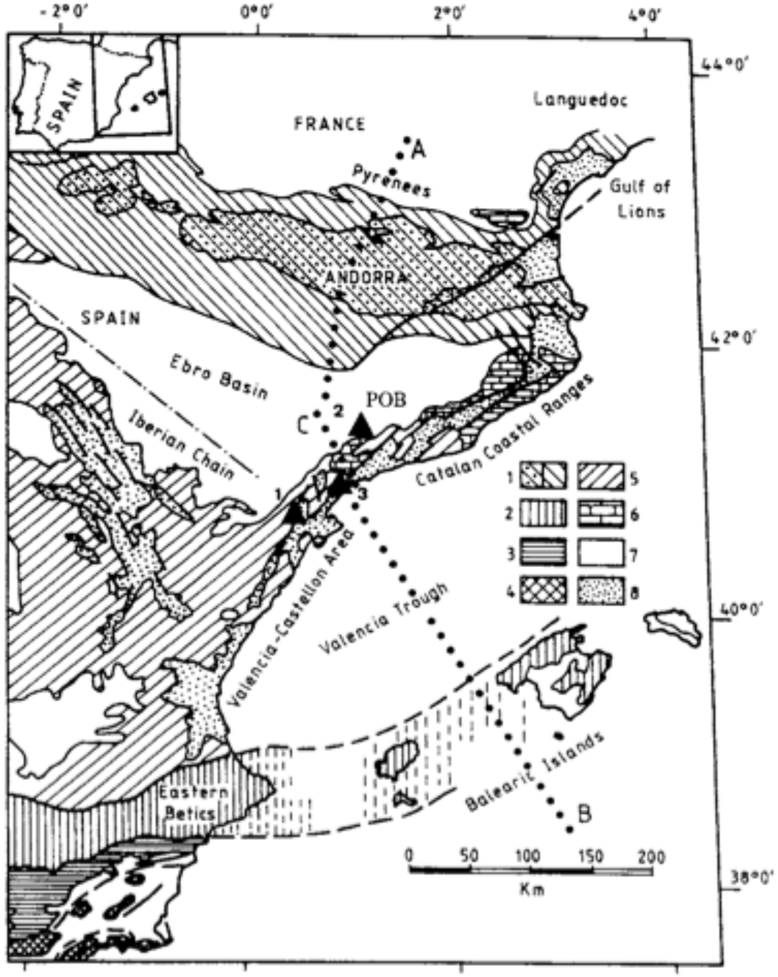
SUMMARIZING:

The inversion scheme proposed by Ammon et al. (1990) for the modeling of receiver functions is:

1) Construct an initial model with a stack of many thin layers.
2) Determine the smoothness parameter through a "preliminary" inversion.
3) Investigate the multiplicity of solutions by perturbing the initial model into many starting models.
4) Choose a model from *a priori* and independent information.

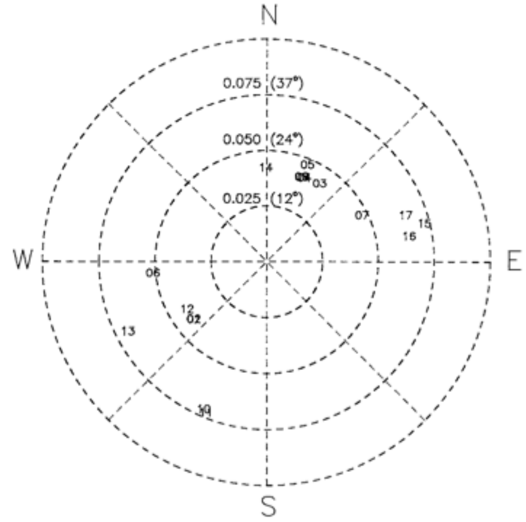
The receiver structure of the Ebre Basin

(Julià et al., BSSA, 1998)

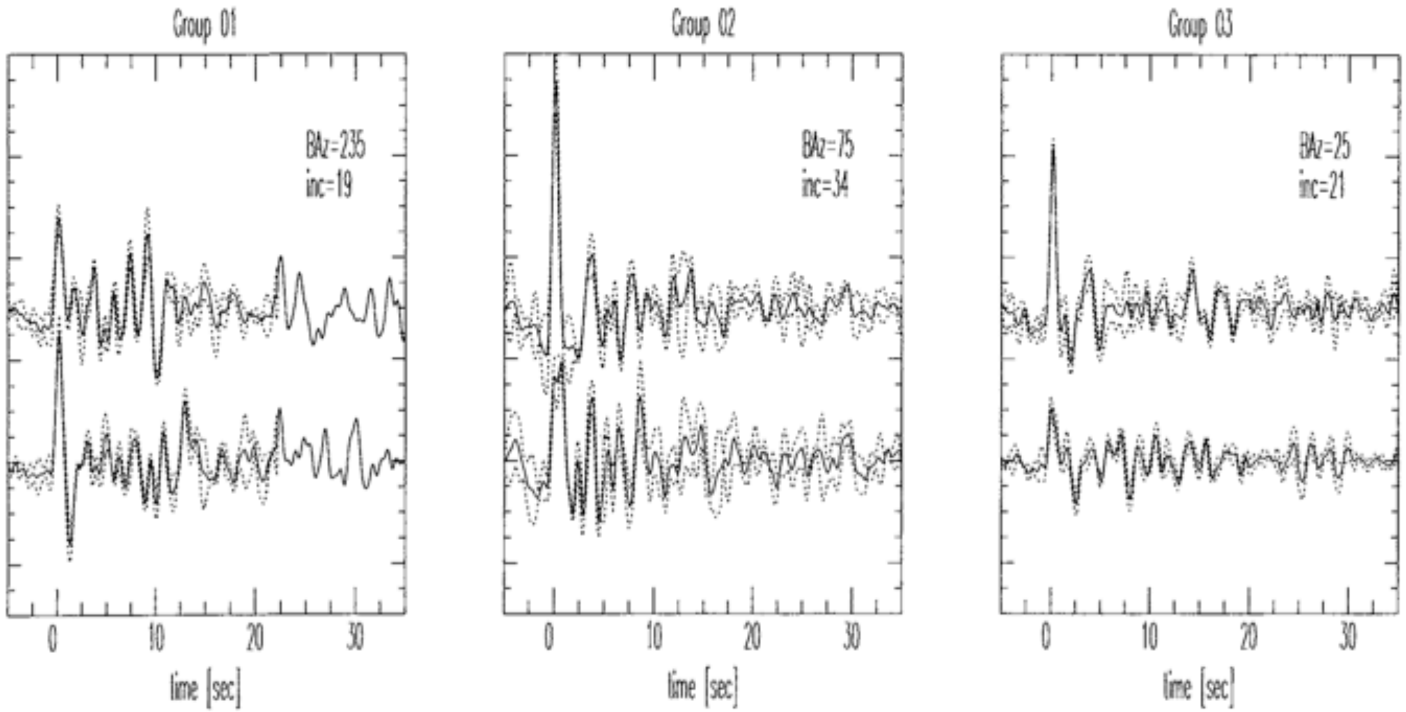


- It's an foreland basin that formed during the Alpine orogeny.
- Filled with deposits from the adjacent mountain ranges.
- Highly non-uniform on the edges.
- Highly uniform along the central axis.

Computing receiver functions at POB

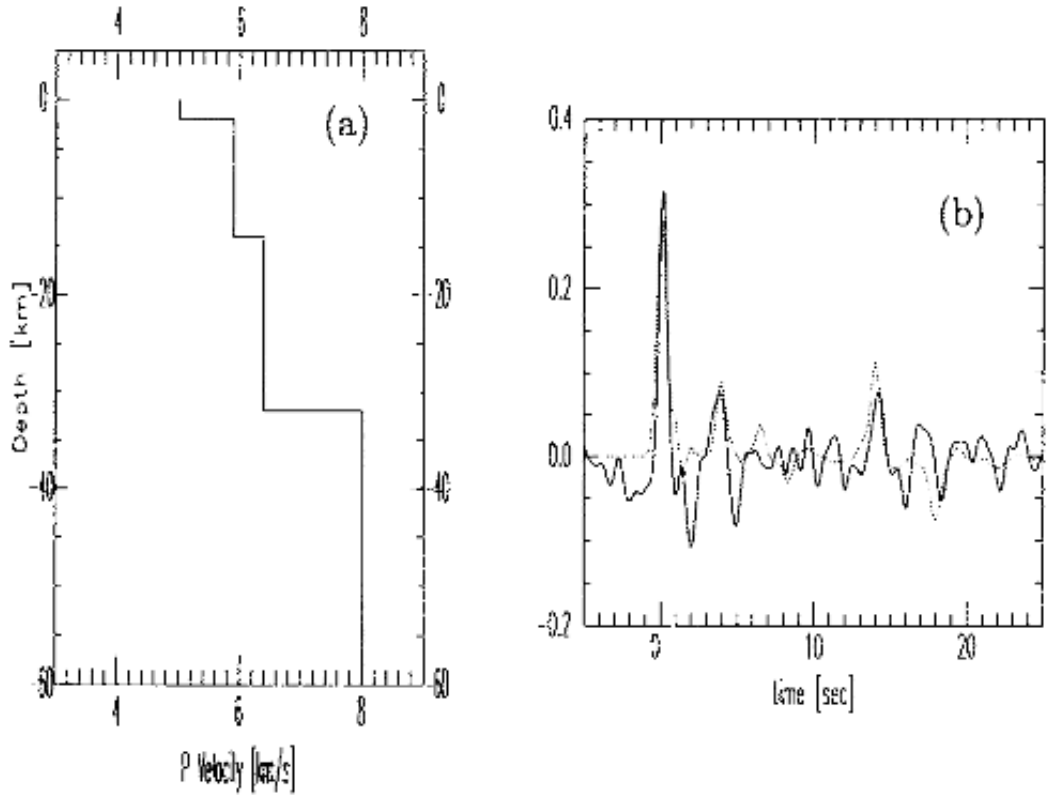


Receiver functions were computed from short-period recordings using the “water-level” method.



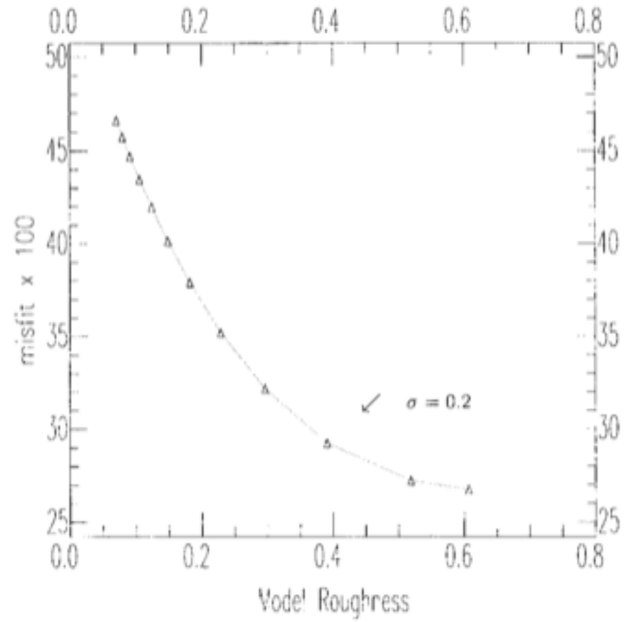
The starting model

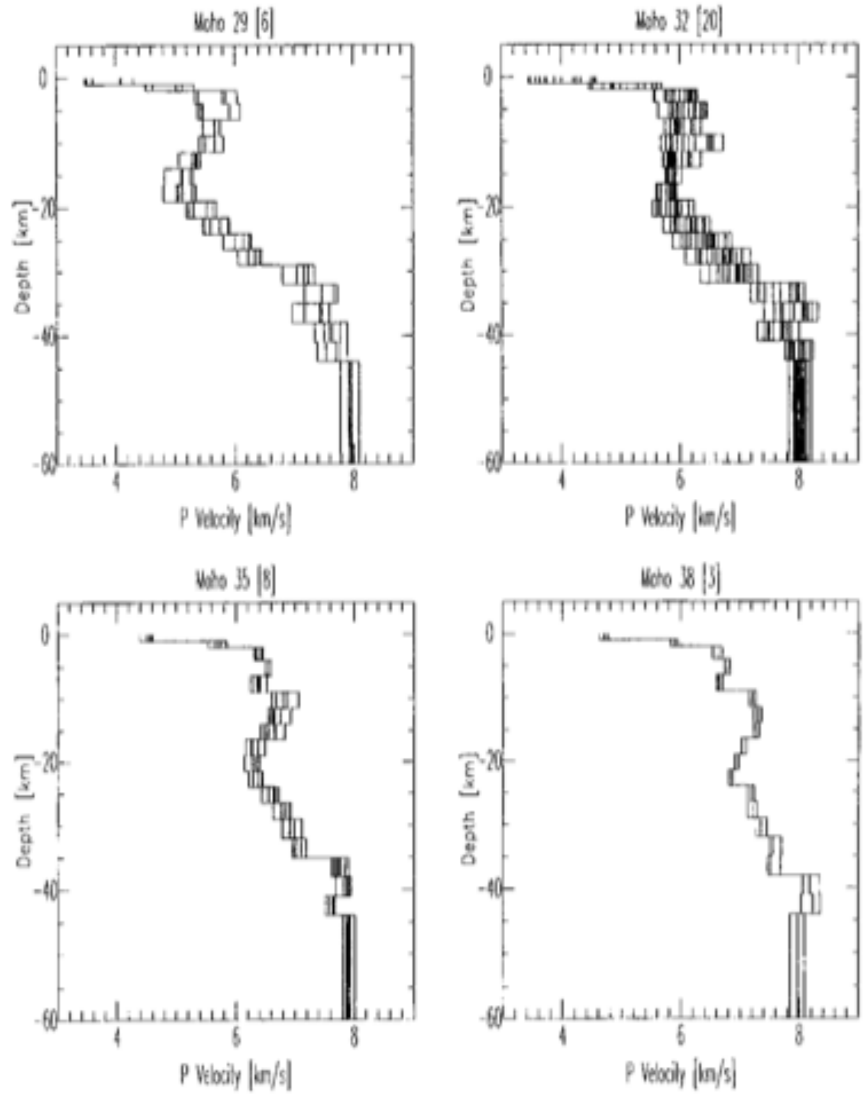
The starting model was taken from the P-wave velocity model that the Catalan Geological Survey used to locate earthquakes.



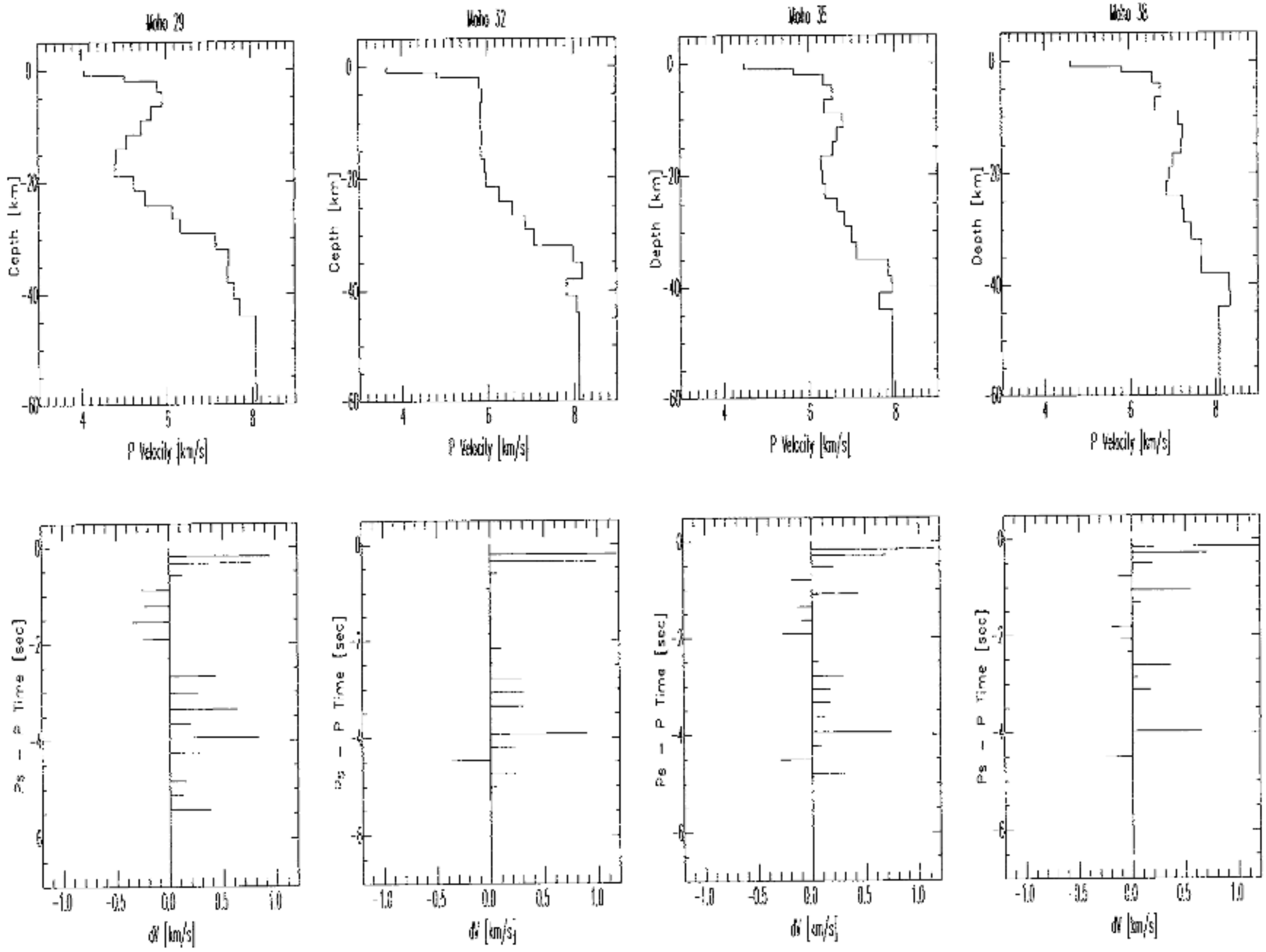
Smoothness and non-uniqueness

A smoothness parameter of 0.2 was chosen from the noise-level.
The resulting velocity models grouped into 4 families.



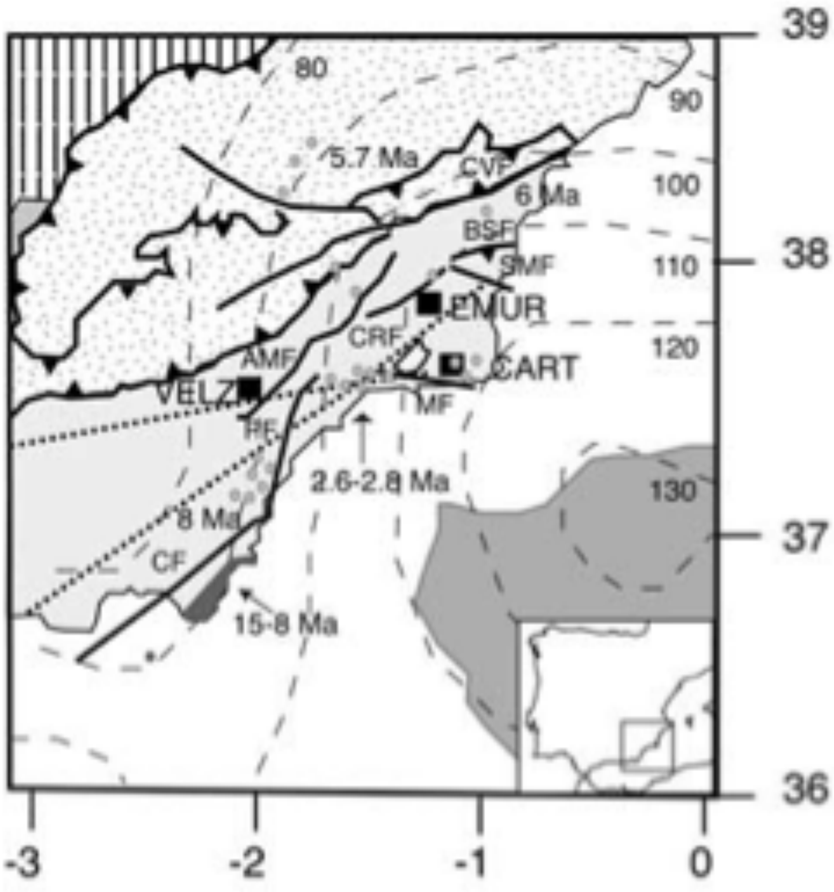


What do receiver functions constrain?

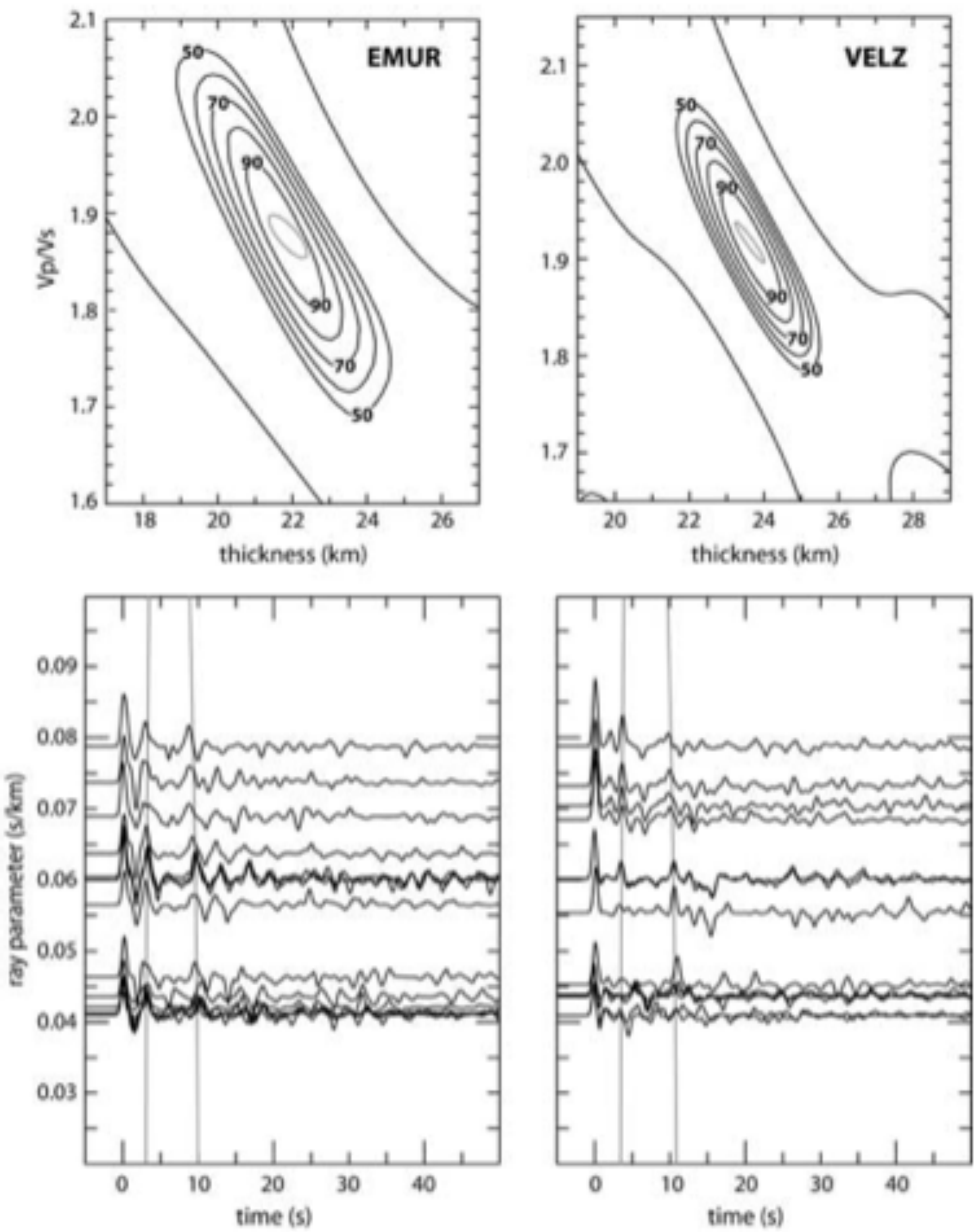


Seismic signature of intra-crustal magmatic intrusions in the Eastern Betics

(Julià et al., GRL, 2005)



- Bounded by the Palomares and Alhama de Murcia faults.
- Postulated as a structurally distinctive block.
- Characterized by high heat-flow values.
- Widespread strike-slip faulting.
- Neogene volcanism (2.6 - 2.8 Ma).

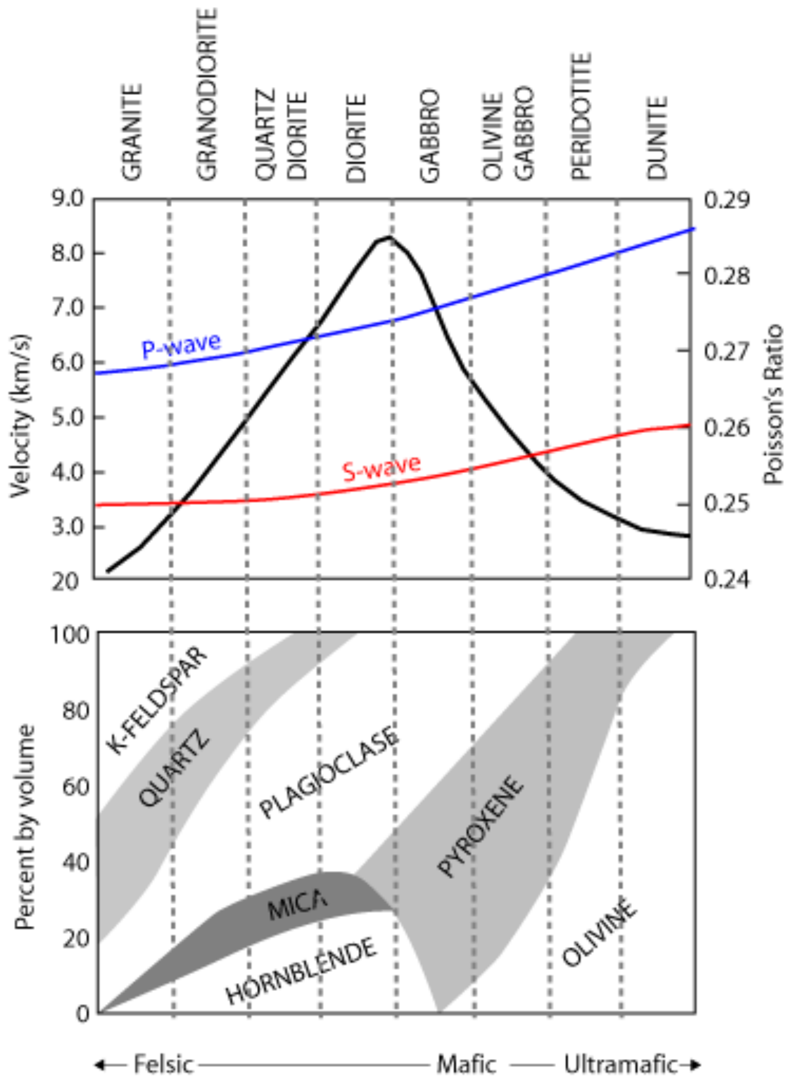


hk-stacking results

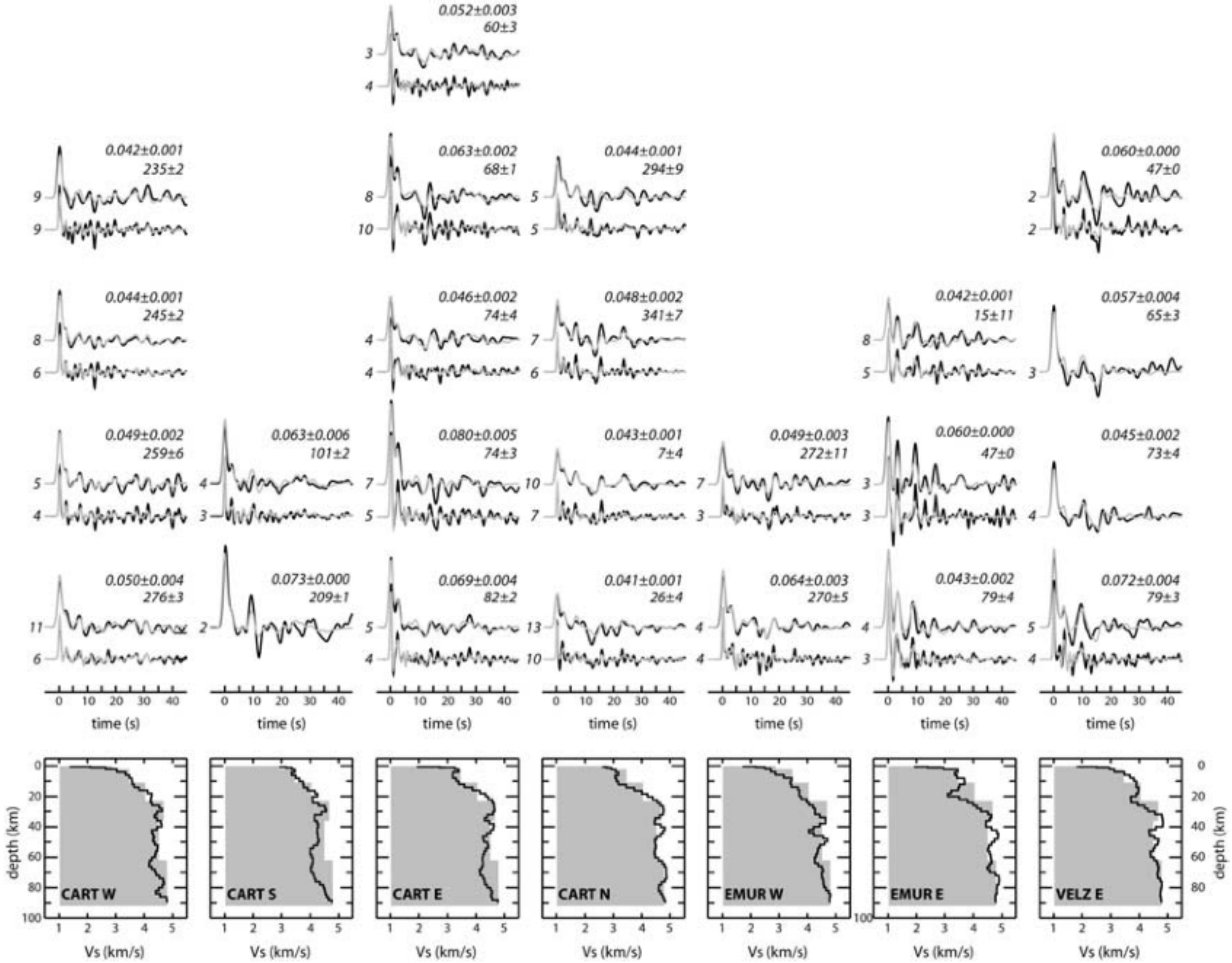
- Shallow depth for the interface, a bit over ~20 km.
- Very large Vp/Vs ratio, ~1.90 (σ ~0.31)
- Consistent with active-source profiling? (Vp ~6.3 km/s, h=~23 km)
- Or is there something else going on?

What does a large Vp/Vs ratio mean?

- The upper crust is made of granites and gneisses ($0.24 < \sigma < 0.26$).
- The lower crust is generally more mafic ($0.26 < \sigma < 0.29$).
- Large Vp/Vs (Poisson's) usually explained by
  - Mafic underplate
  - Fusió parcial

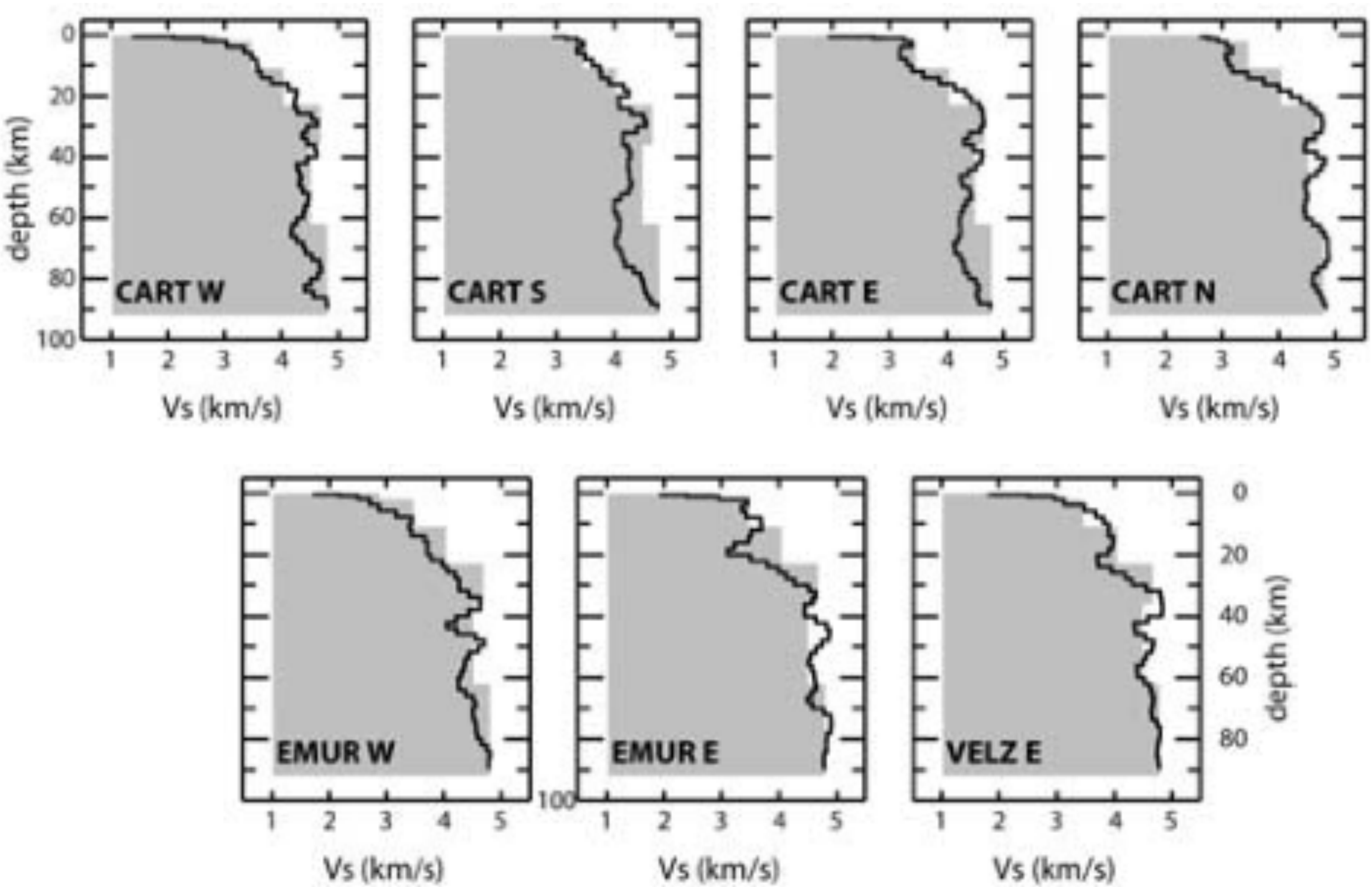


After Christensen (1996)



0.052±0.003
60±3
0.042±0.001
235±2
0.063±0.002
68±1
0.044±0.001
294±9
0.060±0.000
47±0
0.044±0.001
245±2
0.046±0.002
74±4
0.048±0.002
341±7
0.042±0.001
15±11
0.057±0.004
65±3
0.049±0.002
259±6
0.063±0.006
101±2
0.080±0.005
74±3
0.043±0.001
7±4
0.049±0.003
272±11
0.060±0.000
47±0
0.045±0.002
73±4
0.050±0.004
276±3
0.073±0.000
209±1
0.069±0.004
82±2
0.041±0.001
26±4
0.064±0.003
270±5
0.043±0.002
79±4
0.072±0.004
79±3
time (s)
depth (km)
CART W
CART S
CART E
CART N
EMUR W
EMUR E
VELZ E
Vs (km/s)

What do the inversion models reveal?



Summarizing …

- Receiver function inversions are highly non-unique.
- What receiver functions constrain are:
  - Velocity contrasts across discontinuities
  - S-P travel times between the surface and the discontinuity.
- The scheme of Ammon et al. (1990) uses a stack of thin layers and requires smoothness constraints.
- Independent *a priori* information is necessary to choose among many competing models.