Supplementary material

Assessing the reliability of local earthquake tomography for crustal imaging: 30 years of records in the Western Alps as a case study

J. Virieux[1], A. Paul[1], M. Langlais[1], G. Janex[1], P. Guéguen[1], A. Helmstetter[1], and L. Stehly[1]
[1] Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, UGE, ISTerre, *38000 Grenoble, France.*

1 Local earthquake tomography as an ill-posed inverse problem

1.1 Text S1: The penalty approach as an augmented linear system

As mentioned in the section 3.3, the model increment Δm is defined in a compact way by

$$\arg\min_{\Delta m} ||\mathcal{J}\Delta m - \Delta t||^2 + \lambda||\mathcal{L}\Delta m||^2 + \epsilon||\Delta m||^2. \tag{1}$$

The smoothing operator $\mathcal{L}$ and the minimal-norm requirement depend on two hyperparameters: the scalar hyperparameter λ and the damping hyperparameter ϵ.

The first step is the choice of an initial model made up of initial V_p and V_s models (in fact slowness models) and initial hypocenter locations, described by two sub-models $m = (s, h)$ and increments $\Delta m = (\Delta s, \Delta h)$, where the hypocenter (resp. slowness) parameter vector is denoted by the symbol h (resp. s). Initial parameter values play an important role in the model search. The inversion aims at computing a better model (with a better fit to observed arrival times). At iteration k, model m_k is perturbed by $\Delta m_k = (\Delta s_k, \Delta h_k)$, such that the predicted arrival time $t(m_k) + \partial t/\partial m(m_k)\ \Delta m_k$ is closer to the observed arrival time t_{obs}. The linear system for the model increment Δm is:

$$W_D\mathcal{J}_k\Delta m_k = W_D\frac{\partial t}{\partial m}(m_k)\Delta m_k = W_D(t_{obs} - t(m_k)), \tag{2}$$

where a weighting matrix W_D is added for removing outliers from the arrival time data set.

Because different physical quantities are considered in this linear system, we scale each parameter of the four classes $(V_p, V_s, (x, y, z), t_0)$ by the maximum norm of columns of the sensitivity matrix for the corresponding class.

By solving this adimensional scaled system in the least-squares sense, we get a new model m_{k+1} with better fit to observed data and a data misfit function given by:

$$\mathcal{C}_D(\mathbf{s}, \mathbf{h}) = \frac{1}{2}\sum_{nobs} W_D(t_{obs} - t(m))^2, \tag{3}$$

Considering constraints in the velocity model is needed for this mixed-determined structure of seismic tomography. Instead of requiring exact constraints through Lagrangian multipliers, the simpler penalty approach is often preferred. This approach consists in adding new linear relations to the previous linear system both for the smoothing and damping components. We therefore consider the first-order seven-points finite-difference Laplace operator $\mathcal{L}$ over the grid of velocity parameters. The spatial-derivation operator D of this operator for each node of the inversion gridwhich leads to the general system:

$$\begin{bmatrix} W_D\frac{\partial t}{\partial s}_k & W_D\frac{\partial t}{\partial h}_k \\ \lambda D & 0 \end{bmatrix} \begin{bmatrix} \Delta s_k \\ \Delta h_k \end{bmatrix} = \begin{bmatrix} W_D(t_{obs} - t(m_k)) \\ 0 \end{bmatrix}. \tag{4}$$

The hyperparameter λ will limit variations of the spatial second derivative of the slowness model, leading to smooth velocity models while still trying to fit the observations. Such smoothing operator could be split into Cartesian components. Often, differences are made between horizontal and vertical smoothing conditions complementing possible horizontal and vertical grid discretization. Using the same penalty strategy, the LSQR algorithm adds hyperparameter ϵ to this adimensional extended linear system for making it solvable, leading to a new extended linear system:

$$\begin{bmatrix} W_D \frac{\partial t}{\partial s}_k & W_D \frac{\partial t}{\partial h}_k \\ \lambda D & 0 \\ \epsilon I & 0 \\ 0 & \epsilon I \end{bmatrix} \begin{bmatrix} \Delta s_k \\ \Delta h_k \end{bmatrix} = \begin{bmatrix} W_D(t_{obs} - t(m_k)) \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{5}$$

This system highlights the difference between slowness/velocity parameters for continuous model description and hypocenter parameters which are discrete by definition, at least in the ray formulation. Ingredients of the inverse problem are the inversion grid discretization, the initial grid-related velocity and hypocenter values, and the two hyperparameters constraining the ill-posed problem of local earthquake tomography.

1.2 Text S2: How to introduce prior model information?

Target-oriented seismic reconstruction of upper-crust reservoirs could be constrained by information from well data (Asnaashari et al., 2012) or expected geological features (Guitton, 2012). At the lithospheric scale, geodynamic interpretation may help such a design of a prior model: such a model may attract somehow the running model of the inversion scheme through iterations. Another new hyperparameter will control the influence of this prior model in the global misfit estimation. Such a design is beyond the scope of this current work. However, two simpler inversion strategies can be considered by applying an operator once the slowness perturbation Δs has been estimated. The slowness perturbation could be transformed either by a smoothing kernel (Fomel and Claerbout, 2003) or by a Total-Variation (TV) kernel (Dahl et al., 2009) before updating the slowness model. These operators essentially filter out high-wavenumber components in the velocity model, essentially artificially filled by the ray geometry assumption. They are commonly used for full-waveform inversion.
The first strategy promotes specific features of the slowness perturbation. Because of its numerical efficiency, a tensorial Gaussian filter is often considered: it has another smoothing influence than the Laplacian-smoothing operator. Other alternative more elaborated filtering operators with specific boosting of expected geological features can be considered as well (Guitton, 2012; Wellington et al., 2017). Again, such an operator depends on prior information regarding the expected model to be specified in the case study.
The second strategy promotes more or less equal-value zones: a reasonable assumption of more or less uniform geological structures. It is derived from image denoising techniques.
Such operations are not consistent with the minimization problem. They are only valid because the reduction of unwanted features does not affect the decrease of the data misfit function.

Limited illustrations of these smoothing strategies are given for making the reader aware of the numerous ways of updating the velocity model. For example, the Gaussian smoothing operator with a characteristic length of 20 km in all three directions (Fig. S1 left) gives a velocity model with nearly same data misfit than the model obtained with the seven-points Laplacian operator. Of course, the strong vertical smoothing illustrates the impact of such filtering strategy. Similarly, the TV operator can be applied to the slowness perturbation at each iteration. Such operator depend on various parameters not described in this work. They control the expected seismic velocity variations over characteristic zones. The reconstructed model (Fig. S1 right) displays less pronounced low-velocity structures at shallow depths. These low-velocity zones are very sensitive to the requested smoothness of the velocity reconstruction, mitigating possible interpretation of these quantitative values.

2 Testing LET parameters on the high-quality *HQ-89-14* database

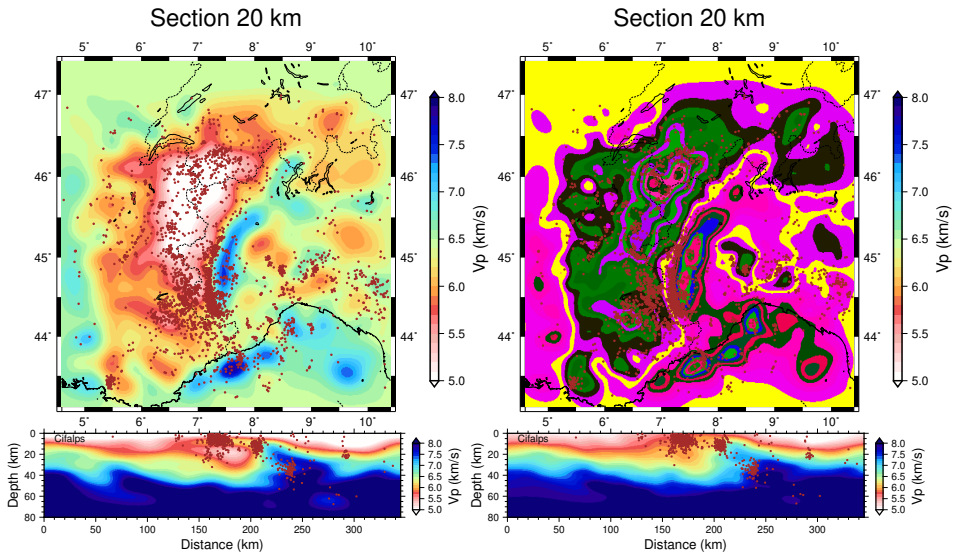

Figure S1: Influence of the smoothing (left) and the TV (right) operators applied to the slowness perturbation model. Top panels: depth slices in the P velocity model at 20-km depth; Bottom panels: P velocity depth sections along the CIFALPS profile.

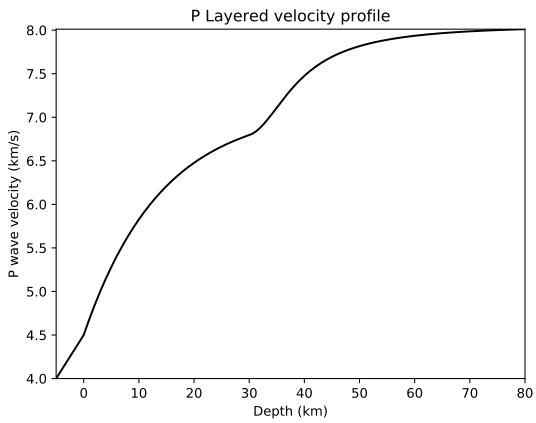

Figure S2: Initial stratified P velocity model proposed by Potin (2016). The initial S velocity is deduced by considering a constant V_p/V_s ratio of 1.6933.

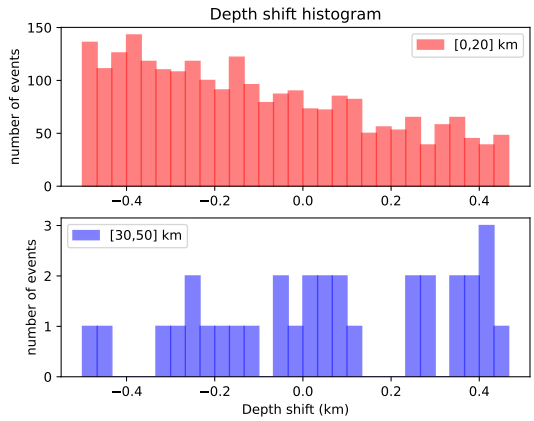

Figure S3: Histograms of hypocenter depth shifts between initial HYPO71 location and final inversion location for the 0-20 km layer (top) and the 30-50 km layer (bottom). Superficial events tend to move to shallower depths, while deeper events tend to move to greater depths.

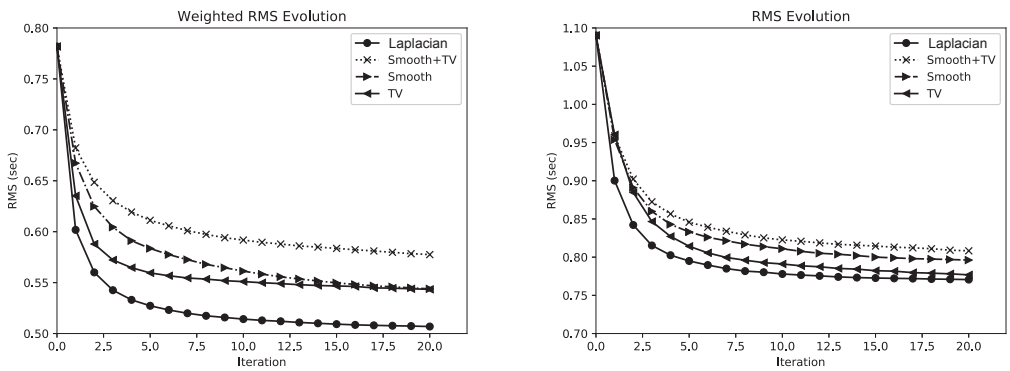

Figure S4: Weighted (left)and raw (right) data misfit reduction curves through iterations for four inversion strategies controlling the model roughness applied to *HQ-89-14* database

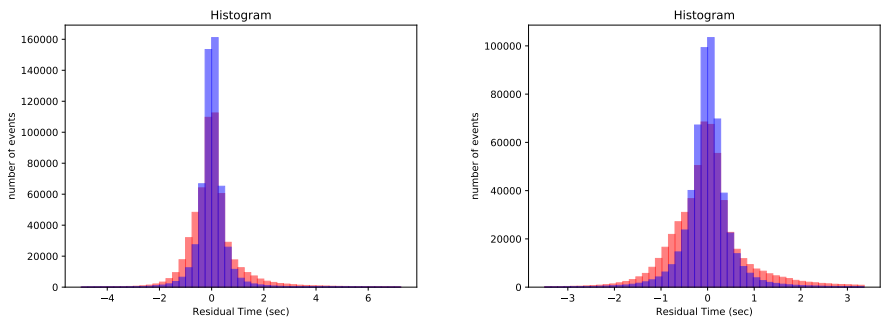

Figure S5: Arrival time residuals for database *HQ-89-14* for the initial model (red) and the recovered model (blue). Left: range [−6 s , 8 s]; Right: zoom in the range [−4 s , 4 s] that will be used from now on. The histogram for the initial model (red) has a larger tail for positive values, meaning that observed arrival times are initially higher than predicted times, while the histogram is more symmetric for the recovered model (blue).

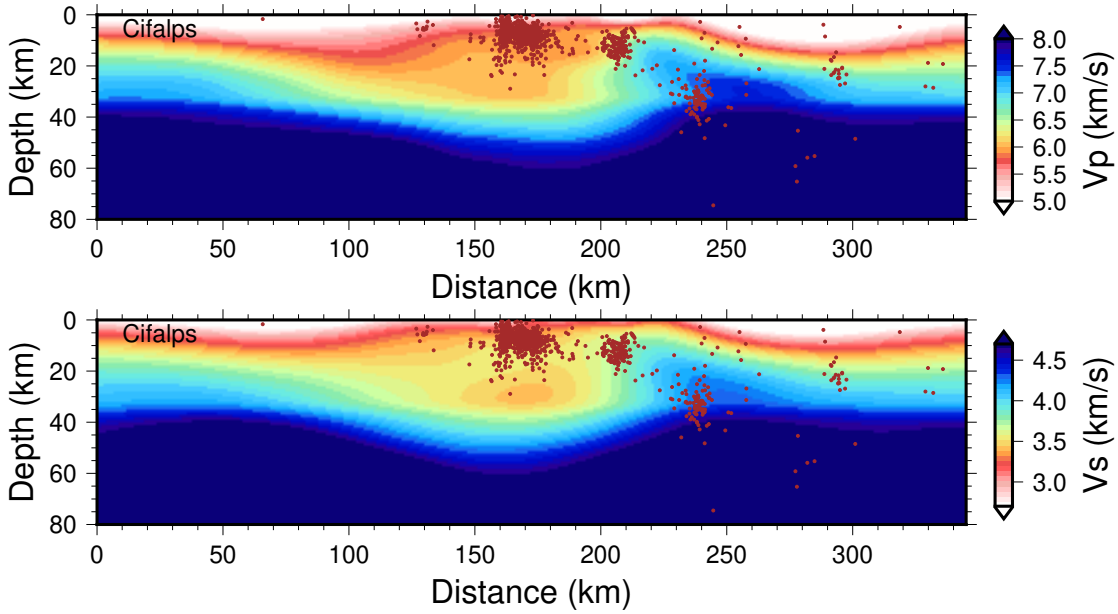

Figure S6: Vertical cross-sections along the CIFALPS profile CC' which locations in shown in Fig. 3 (top) of the main text. This reference profile, which follows the dense CIFALPS temporary experiment (Zhao et al., 2015) will be used in comparison to results of later investigations. Top: P velocity section; Bottom: S velocity section. Both models display a strong high-velocity anomaly at distances of $200 - 250$ km rising to $\sim$10-km depth. This is the so-called Ivrea body anomaly. The S velocity structure displays a deep low-velocity zone at 150 - 200 km distance and 20 - 35 km depth. This deep LVZ can be guessed in the P velocity section.

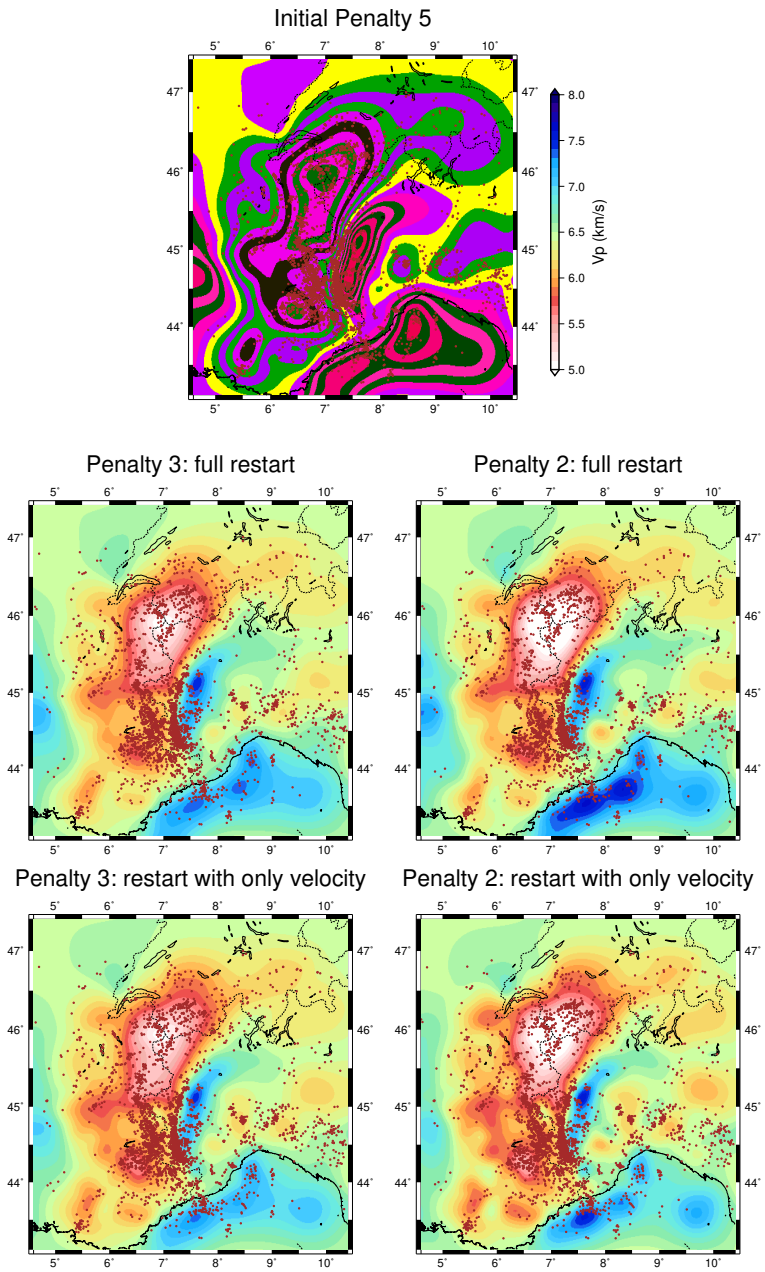

Figure S7: Results on the V_p slice at 20-km depth of a cascade of inversions with decreasing Laplacian hyperparameters. Top: result of the initial Laplacian workflow with weight $\lambda = 5$ in all directions. Middle: the resulting model (velocities and hypocenters) is then input in successive inversions with weights $\lambda_{x,y} = 3$, $\lambda_z = 2$ (left panel) and $\lambda_{x,y} = 2$, $\lambda_z = 1$) (right panel). Bottom: same as second row but with HYPO71 locations as initial model. Although velocity patterns are very similar, small differences illustrate the respective contributions of initial velocity and hypocenter parameters.

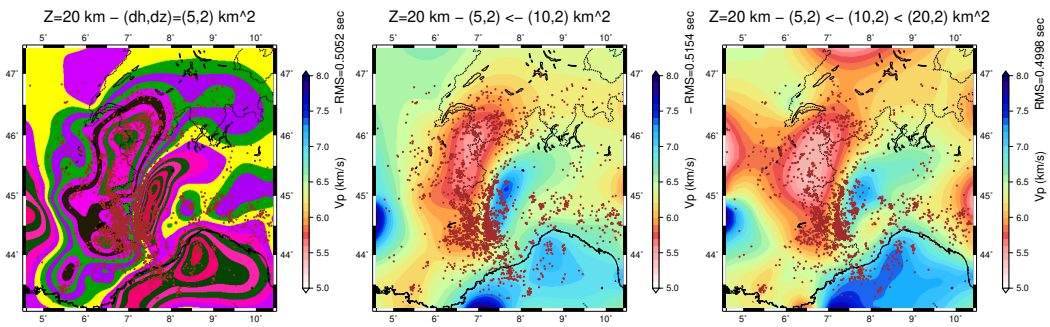

Figure S8: Influence of the grid discretisation on the 20-km depth V_p slice. Same values of hyperparameters as in the first tomography run. Left: velocity slice with a horizontal discretisation of 5 km. Center and right: velocity slices with two different grid refinement strategies (10 to 5 km, and 20 to 10 to 5 km). RMS values are almost identical for the three models. Initial coarse grid steps of 10 km or 20 km induce strong edge effects.

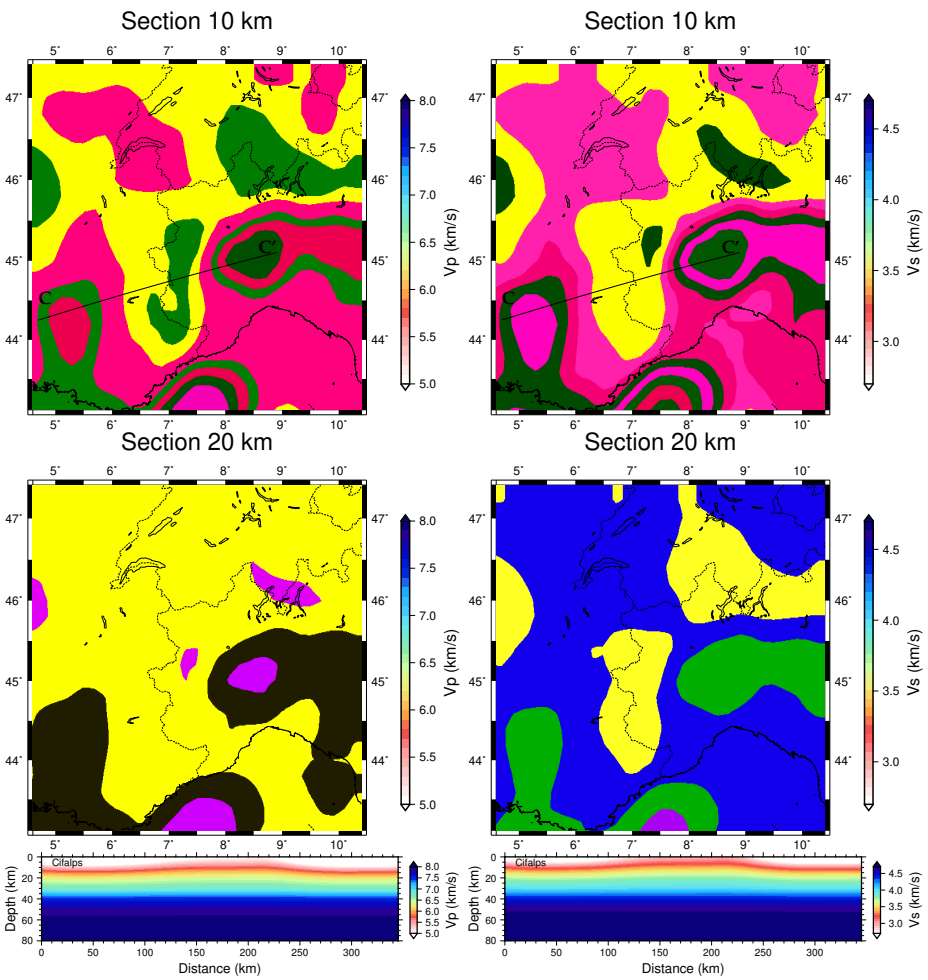

Figure S9: Initial *3D-ANT* model combining an ambient-noise-tomography model and the reference layered model: a smooth change between the 2 models is performed in the range [10 km-30 km]. V_p in the left-hand side and V_s in the right-hand side. Top: horizontal sections at 10-km depth; Middle: horizontal sections at 20 km; Bottom: vertical sections along the CIFALPS profile.

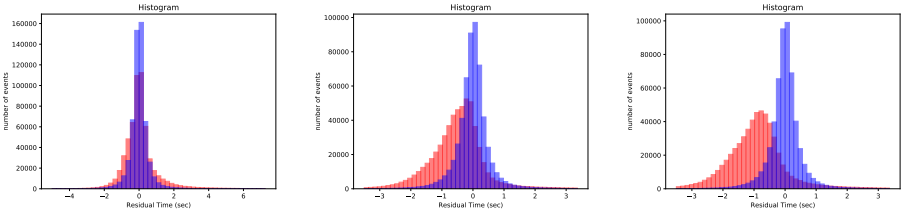

Figure S10: Arrival time residuals in the [−4 s, 4 s] range for initial (red) and recovered (blue) models. Left: initial *Z-HQ* model with HYPO71 locations; Center: initial *3D-ANT* model with HYPO71 locations; Right: initial *3D-ANT* model with NLLOC locations. More negative initial residuals appear with the NLLOC procedure while the final (blue) histogram is sharper when starting from a stratified model (right).

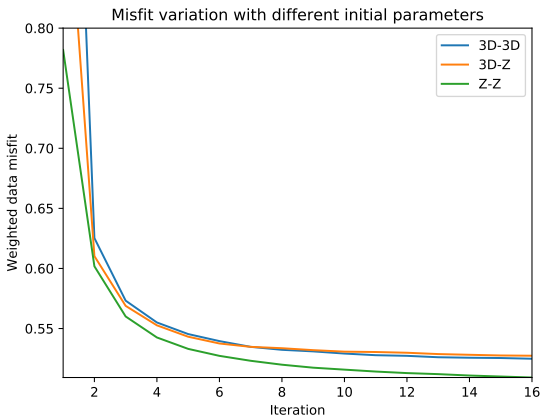

Figure S11: Date misfit evolution with iterations; Green line: initial layered model *Z-HQ* with HYPO71 locations; Red line: 3-D (it 3D-ANT) model with HYPO71 locations; Blue line: 3-D model with NLLOC locations. Again, the stratified model looks like a better initial guess than a 3-D initial model too far from the target 3-D model.

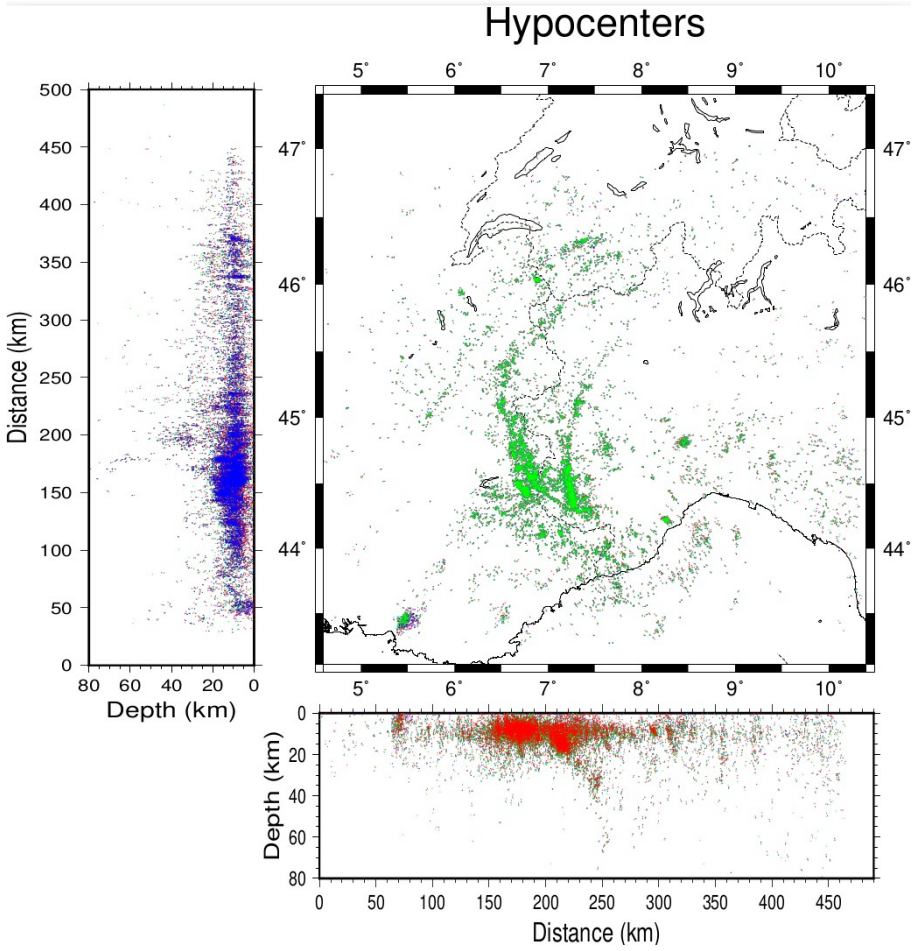

Figure S12: Final hypocenter locations after inversion with three different intial models. Blue: stratified initial model with HYPO71 locations (plotted in the foreground in the left panel); Red: 3-D initial model with HYPO71 locations (in the foreground in the bottom panel); Green: 3-D initial model with NLLOC locations (in the foreground in the central panel).

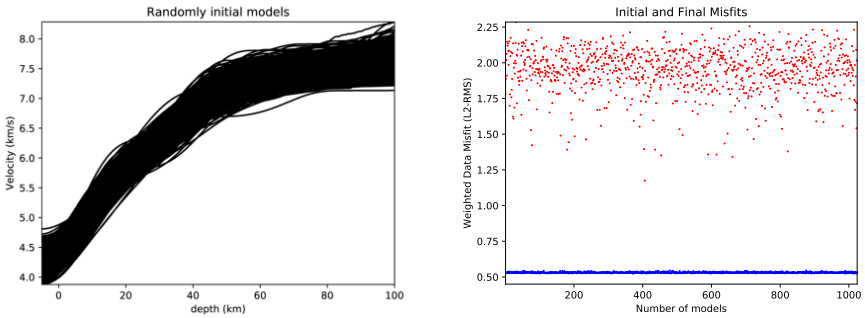

Figure S13: Left panel: Random initial stratified models computed around Potin (2016)'s initial model. Right panel: initial (red) and final (blue) data misfits for the 1024 inversions with initial random stratified models shown in the left-hand side. Final misfits concentrate between 0.52s and 0.53s, in the same range as with Potin (2016)'s initial stratified model.

2.1 Text S3 and Figs. S13-S16: Randomly-layered (stratified) initial models

To further assess the influence of the initial velocity model, we generated a set of 1024 stratified random models around Potin (2016)'s initial model of Fig. S2. They are shown in Fig. S13 (left). Then, we performed 1024 tomographies using each model as initial velocity model, the *HQ-89-14* database, and a Laplacian-smoothing inversion procedure with the same hyperparameters as in section 4 of the main text. The initial and final data misfits are shown in Fig. S13 (right). For further analyses of the inversion results, we selected final models with a data misfit below 0.525 s, that is ∼1 % of random models. The conclusions that we will draw do not strongly depend on this selection.
We now have two sets of velocity models, a set of initial (stratified) and a set of final (3-D) models. For each set, we compute an average velocity model and a standard deviation (RMS) model. At a given depth, areas sampled by seismic waves should have lower final than initial RMS values. Figure S14 shows the RMS ratio between the final and initial P velocity models. It may be considered as a proxy of the initial-model influence on the reconstructed P velocity models. A nearly similar pattern is observed for the reconstructed S velocity models. At 10-km depth, poorly sampled zones corresponding to high values of the RMS ratio are low-seismicity areas with poor station coverage. At 20-km depth, areas of low proxy values have a high density of hypocenters.

We may also evaluate the amplitude of velocity changes during the inversions. Figure S15 displays the average percentage of velocity perturbations (with respect to the initially constant velocity) at 10 and 20-km depths. At 10 km, areas of small velocity perturbations concentrate along the borders of the tomographic box. The low-velocity anomaly at the end of the CIFALPS CC' profile is well identified. At 20 km depth, a strong (positive) velocity variation is required in the Ivrea body region.

The average final velocity model shown in Fig. S16 might be considered as a possible initial 3-D model in future inversions. However, poorly sampled zones may have velocity values depending strongly on the way the random distribution of initial velocities is computed. We disregard such possible 3-D initial models because of these unwanted low-wavenumber components.

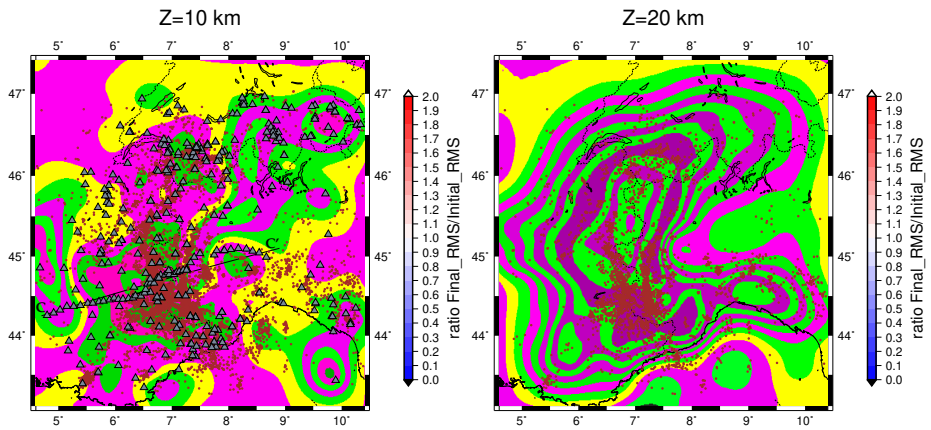

Figure S14: Maps at 10 and 20 km depth of the ratio between RMS values of final P velocity models and RMS values of initial P velocity models. Low values mean that different inversions with different initial stratified models provide similar results. Note the greater variability of final results at 10 than at 20-km depth, the influence of earthquake distribution at both depths, and the influence of station coverage at 10 km. This ratio is a good proxy of the influence of the initial model.

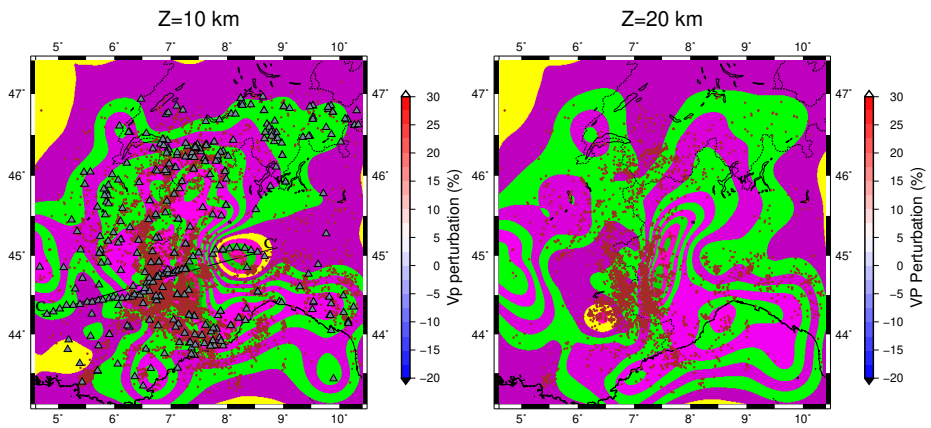

Figure S15: Maps at 10 and 20 km depth of the average velocity perturbations in the final models with respect to their initial models. The maps display areas where fitting arrival time data requires significant velocity changes.

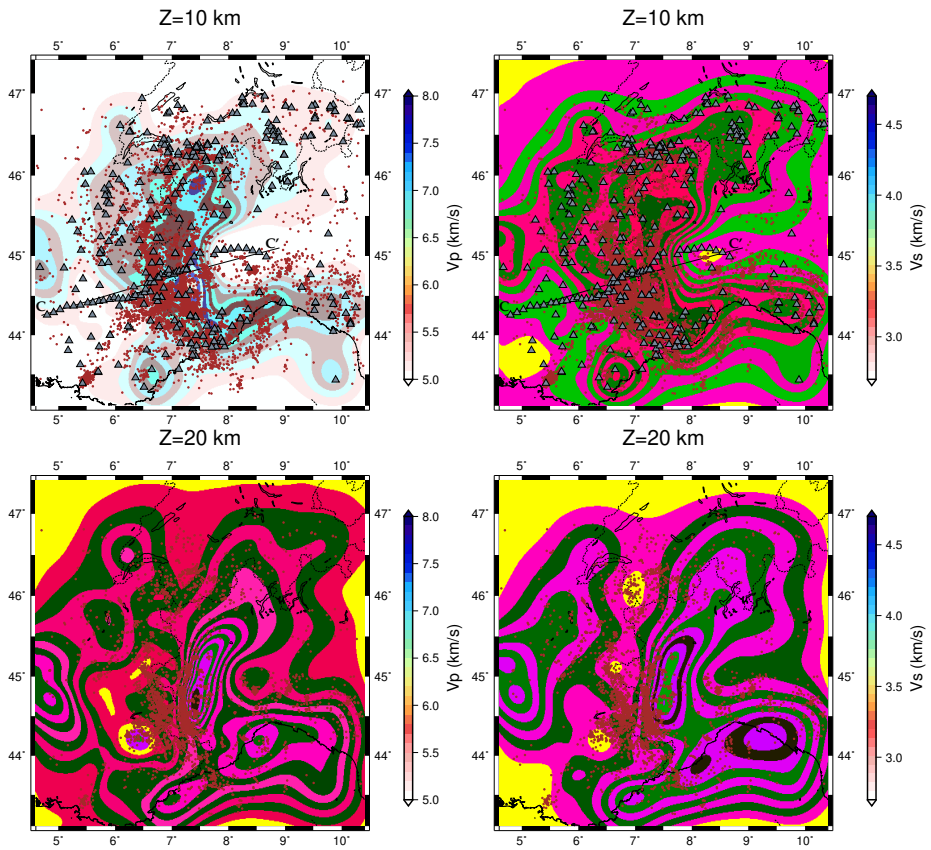

Figure S16: Depth slices at 10-km (top) and 20-km depths (bottom) in the average P (right) and S (left) velocity models reconstructed after inversions with initial random stratified models.

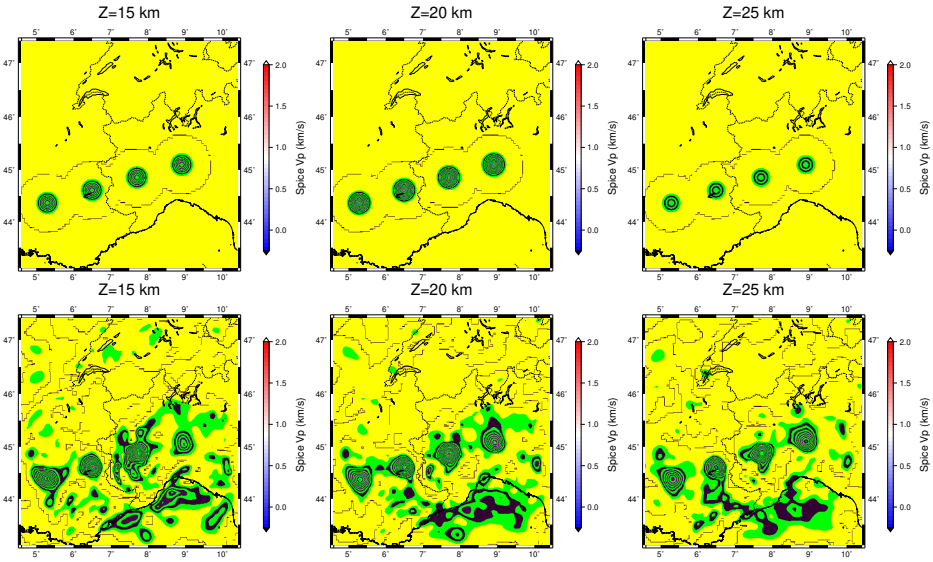

Figure S17: Horizontal sections at 15, 20 and 30-km depth for individual spike tests with a Gaussian V_p anomaly centred at 20 km-depth along the CIFALPS profile; Top: input synthetic model; Bottom: recovered model. Each spike is recovered independently, even if the four models are plotted together.

2.2 Text S4 and Figs. S16-S17: Spike tests

We follow the spike test strategy promoted by Spakman (1991) and add velocity perturbations on top of the final model m_{opt}. Since we are using a standard Laplacian-smoothing/damping approach, we consider a small velocity perturbation described by a Gaussian function centred at a given point with different characteristic lengths along the three directions. The perturbation should be small enough to avoid significant ray deviation.

Synthetic arrival times are computed in model m_{opt} perturbed by a single spike for the same earthquake-station couples as in the observed dataset. An inversion is performed starting from model m_{opt} with a few iterations. The difference between the newly recovered model and the final model is then compared to the input synthetic velocity anomaly.

As proposed by Rawlinson and Spakman (2016), we proceed with individual spikes even if they are combined for plotting. With this method, the inversion may reveal velocity anomalies at specific locations independently of spike locations, but tightly connected to station and earthquake distribution.

Input and output models are shown in Figs. S17 and S18 for an example with four independent positive anomalies at 20-km depth along the CIFALPS profile. The maximum perturbation is 1200 km/s and the characteristic lengths are 15 km in the horizontal directions and 5 km in the vertical one. Horizontal slices at 15, 20 and 25-km depths and depth sections illustrate amplitude changes during the reconstruction associated with slight spatial spreading (Figs. S17 and S18). The velocity reconstruction has an overall satisfactory quality.

A few anomalous patterns are observed whatever the location of the spike anomaly, which are highlighted by our tests with individual spikes combined in a second stage for plotting. This is the case for the anomaly at 90-km distance and 10-km depth in Fig. S18, as well as the anomalies spread in the south-eastern corner of the study region in Fig. S17. As outlined above, these spurious anomalies are connected to station and event distribution. They could lead to ambiguous interpretations of the results of tests performed with multiple spikes, even if they are sparsely distributed.

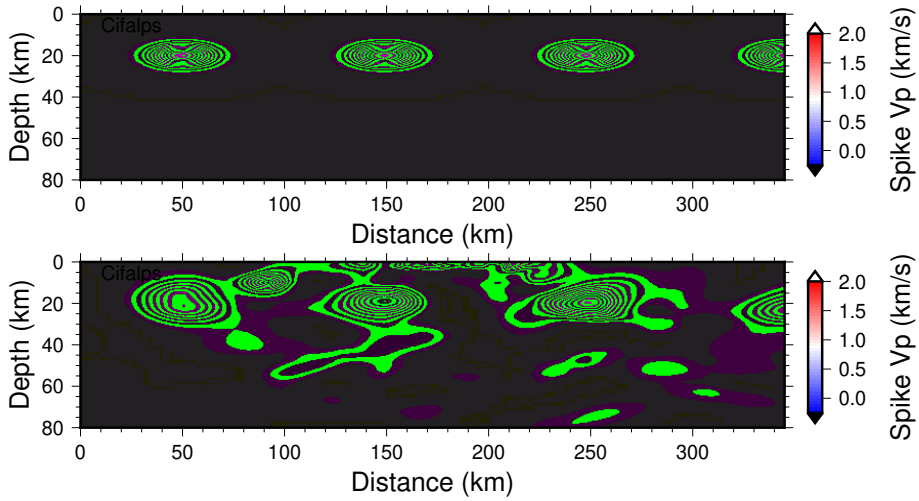

Figure S18: Depth sections along the CIFALPS profile in the same individual spike test model as in Fig. S17, with a Gaussian anomaly centred at 20 km-depth; Top: input synthetic model; Bottom: recovered model.

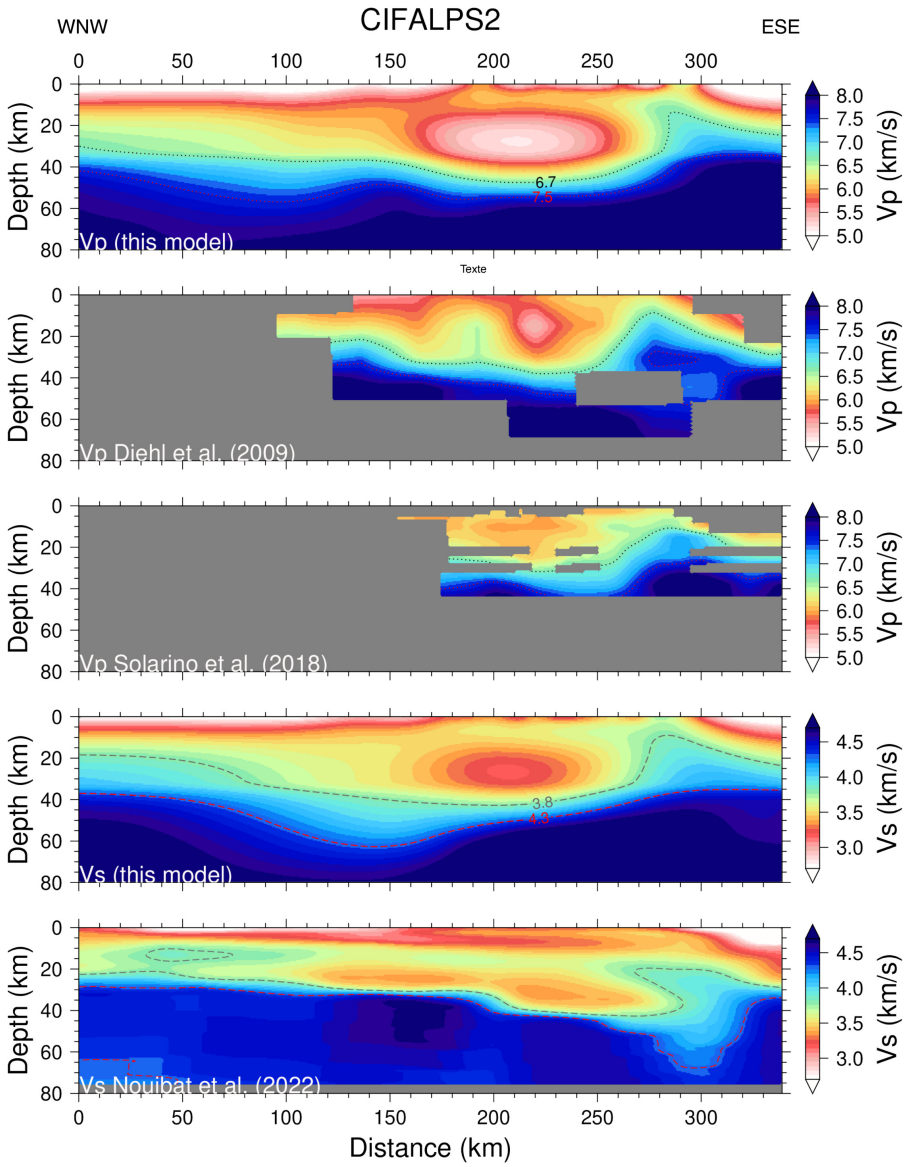

Figure S19: Depth slices in the V_p (top, second, and third rows) and V_s (fourth and bottom rows) models along the CIFALPS2 section across the northwestern Alps (location in the top left panel of Fig. 12 in the main text). The model's name is indicated in the lower left corner of each section. The 6.7 km/s (black) and 7.5 km/s (red) contours are shown as dotted lines in the three V_p sections, while the 3.8 km/s (black) and 4.3 km/s (red) contours are shown as dashed lines in the two V_s sections. The red contours may be considered as Moho proxies, while the black contours highlight the Ivrea body high-velocity shallow anomaly in the eastern side of the sections (distance range: 270 - 310 km).

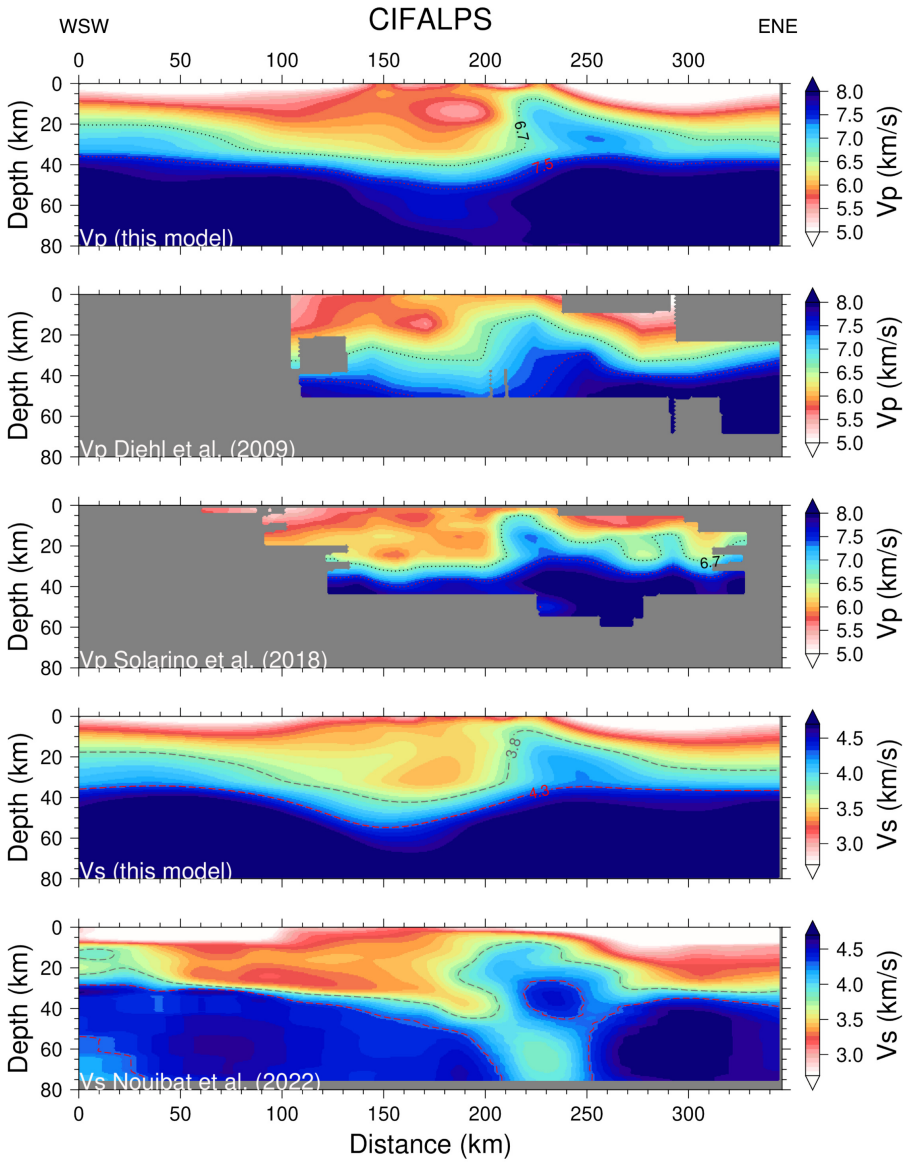

Figure S20: Same legend as Fig. S19 for the CIFALPS section across the southwestern Alps (location in the top left panel of Fig. 12 in the main text).

References

Asnaashari, A., Brossier, R., Garambois, S., Audebert, F., Thore, P., and Virieux, J. (2012). Regularized full waveform inversion including prior model information. In *EAGE, 74th Conference and Exhibition, Expanded abstracts*, page W031.

Dahl, J., Hansen, P. C., Jensen, S. H., and Jensen, T. L. (2009). Algorithms and software for total variation image reconstruction via first-order methods. *Numerical Algorithms*, 53:67–92.

Fomel, S. and Claerbout, J. F. (2003). Mutlidimensional recursive filter preconditioning in geophysical estimation problems. *Geophysics*, 68(2):1–12.

Guitton, A. (2012). Blocky regularization schemes for full waveform inversion. *Geophysical Prospecting*, 60(5):870–884.

Potin, B. (2016). *Les Alpes occidentales : tomographie, localisation de séismes et topographie du Moho*. PhD thesis, Université Grenoble Alpes.

Rawlinson, N. and Spakman, W. (2016). On the use of sensitivity tests in seismic tomography. *Geophysical Journal International*, 205:1221–1243.

Spakman, W. (1991). Delay-time tomography of the upper mantle below Europe, the Mediterranean, and Asia Minor. *Geophysical Journal International*, 107:309–332.

Wellington, P., Brossier, R., Hamitou, O., Trinh, P., and Virieux, J. (2017). Efficient anisotropic dip filtering via inverse correlation functions. *Geophysics*, 82(4):A31–A35.

Zhao, L., Paul, A., Guillot, S., Solarino, S., Malusà, M. G., Zheng, T., Aubert, C., Salimbeni, S., Dumont, T., Schwartz, S., Zhu, R., and Wang, Q. (2015). First seismic evidence for continental subduction beneath the Western Alps. *Geology*, 43(9):815–818.