Geophysical Prospecting, 2008, **56**, 505–526 doi:10.1111/j.1365-2478.2008.00710.x

First-arrival traveltime tomography using second generation wavelets

Matthieu Delost, Jean Virieux* and Stéphane Operto

[1]*UMR Géosciences Azur, Université de Nice, Sophia Antipolis 250, Rue A. Einstein, 06560 Valbonne, France*

Received February 2007, revision accepted February 2008

ABSTRACT

Wavelet decomposition of the slowness model allows a multiscale description of the seismic first-arrival time tomography. We propose the introduction of the so-called second generation wavelets that could be used for any mesh structure and do not require a number of samples, such as the power of two in each direction for fast wavelet transform. A linearized procedure for inverting delayed travel-times considering either slowness coefficients or wavelet coefficients has been set up with an efficient ray tracing at each iteration of the inversion procedure. Wavelet decomposition over constant patches (Haar wavelet) or over linear patches (Battle-Lemarie wavelet) of coefficients at different scales are inverted as unknowns of the tomographic linearized system. Reconstruction of these coefficients depends dynamically on the local resolution when considering dense ray coverage. On simple synthetic examples, it has been found necessary to perform a local resolution analysis for specifying wavelet coefficients to be inverted. This resolution analysis could be performed for an initial smooth reconstructed medium and by designing a bit mask operator it allows fine scales to be inverted in specific areas of the model where the resolution is high while not being inverted in other areas where the resolution is poor: the wavelet decomposition will ease the multiscale reconstruction. A few synthetic examples, such as crosshole tomography or surface-surface tomography illustrate the multiscale feature of wavelet tomography. The second generation wavelet approach seems to be a flexible and rather promising tool for controlling the resolution variation of seismic first-arrival tomography.

INTRODUCTION

First-arrival delayed traveltime tomography applied to wide aperture seismic data has proved to be an interesting tool for the investigation of the Earth's structures (Aki and Lee 1976; Spakman and Nolet 1988; Hole 1992; Zelt and Smith 1992; Benz *et al.* 1996; Zhang, ten Brink and Toksöz 1998). The spatial resolution is controlled both by the theoretical resolution imposed by the relation between velocity structure and first-arrival traveltimes and by the data acquisition configuration as well as the structure itself, which may lead to uneven ray coverage. The theoretical resolution power is limited by the size of the first Fresnel zone for each ray. There is an antinomy between the limited resolution power and the numerical sensitivity kernel corresponding to rays. Rays have no thickness and, therefore, have an intrinsic infinite sensitivity although in practise discretization of the model space always smoothes out this feature (Dahlen 2004). Indeed, one raypath induces high frequency information along the ray, even at locations where resolution is expected to be low. Inversion can build up models that have a high degree of data fitting but which are far from the true model containing artefacts. The limited resolution of the traveltime tomography is accounted for by going through regularized tomographic inversion (Farra and Madariaga 1988; Delprat-Jannaud and Lailly 1993; Zhang *et al.* 1998) often based on the Tikhonov regularization approach (Tikhonov and Arsenin 1977). The covariance matrices for the model as well as for the data (Tarantola and

*E-mail: jean.virieux@geoazur.unice.fr

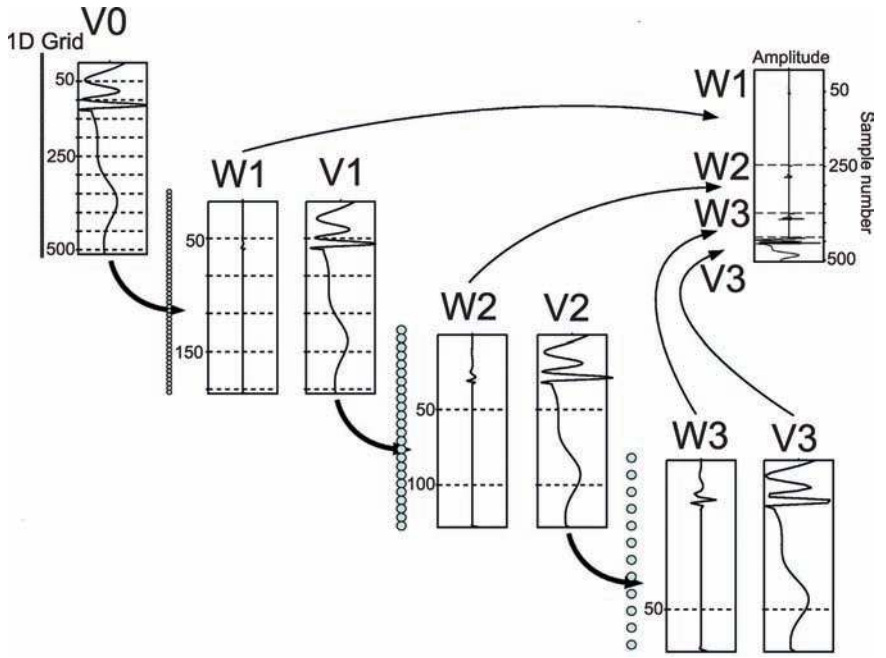

Figure 1 Projection of a 1D signal over 3 levels of wavelets: the wavelet coefficients are kept at each scale while only the final V_3 coefficients are kept for 1D signal reconstruction without loss of information.

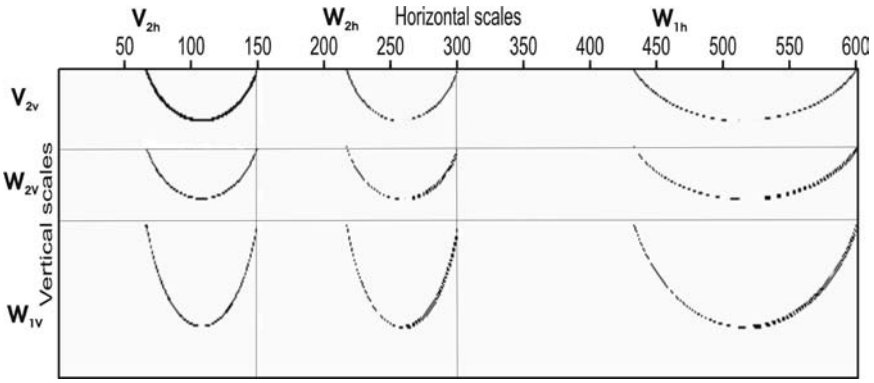

Figure 2 2D Mallat representation and storage of a 2D signal after wavelet transforms. Only the coarsest level of space V has been kept while wavelet coefficients and scale coefficients are deployed inside the same matrix as proposed by Mallat (1989a,b).

Valette 1982; Tarantola 1987) introduce a precise theoretical background although estimation of these matrices is still a difficult task (Monteiller *et al.* 2005). Another alternative regularization is based on smoothing constraints that are easier to implement than the Laplacian (Menke 1984; Zhang *et al.* 1998). We choose this penalty regularization through a Gaussian smoothing operator over slowness coefficients as the standard tomographic procedure that we shall use here for comparison. Whatever we choose, regularization strategies lack spatial adaptivity. Frequently, when the smoothing is tuned to remove the footprints of raypaths in poorly illuminated areas, short-scale features that could be resolved in other areas of the model according to the expected resolution are smoothed simultaneously, leading to a loss of information in the well-resolved areas.

The resolution varies locally and one could also locally adapt the Gaussian operator in the penalty procedure. For example, considering surface acquisition geometries, the shallow structures will be sampled both by short- and long-offset rays running across the heterogeneities with different azimuths, while the deeper structures will be sampled by long-offset rays, only with a much narrower range of azimuths. The resolution

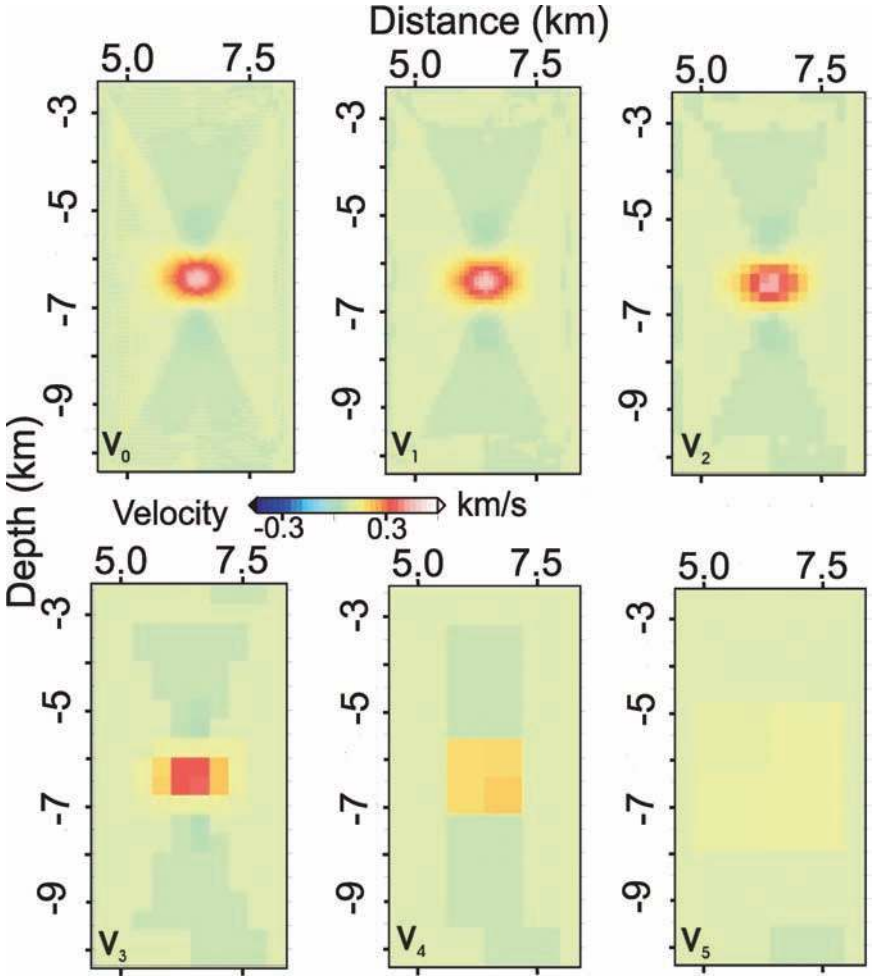

Figure 3 Decomposition of a V_0 slowness model for a crosswell experiment onto the V_i spaces using 5 levels of wavelets in order to highlight the smoothing related to scale changes.

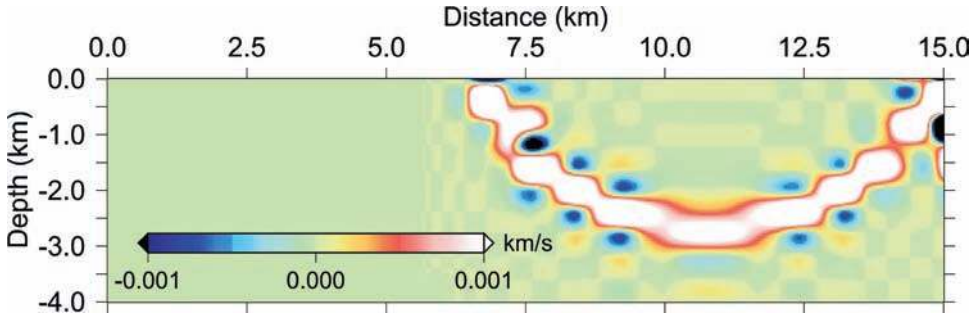

Figure 4 Non-zero Fréchet coefficients of a selected ray for the V_4 grid corresponding to the fourth level of wavelet decomposition. Please note the thickness of the ray related to the discretization at that scale.

decreases quite significantly with respect to depth for this geometry. Two other factors creating non-uniform ray coverage are the presence of low velocity anomalies that induces shadow zones in the ray coverage and non-uniform source and/or receiver arrays, especially in the case of passive tomography.

In order to consider this spatially varying resolution power, one could introduce adaptive parameterizations based on non-structured grids (Vesnaver 1994; Michelini 1995; Vesnaver 1996; Spakman and Bijwaard 1998; Spakman, Bijwaard and Engdahl 1998; Böhm and Vesnaver 1999; Böhm, Galuppo and

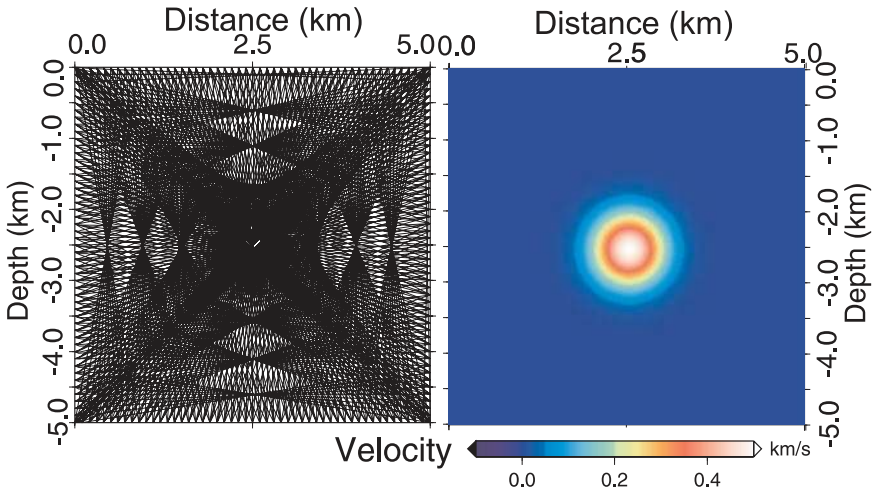

Figure 5 Synthetic example: a circular positive anomaly on top of an homogeneous model as shown on the right-hand panel. Please, note the dense ray coverage on the left-hand panel with a local undersampling centre of the circular anomaly.

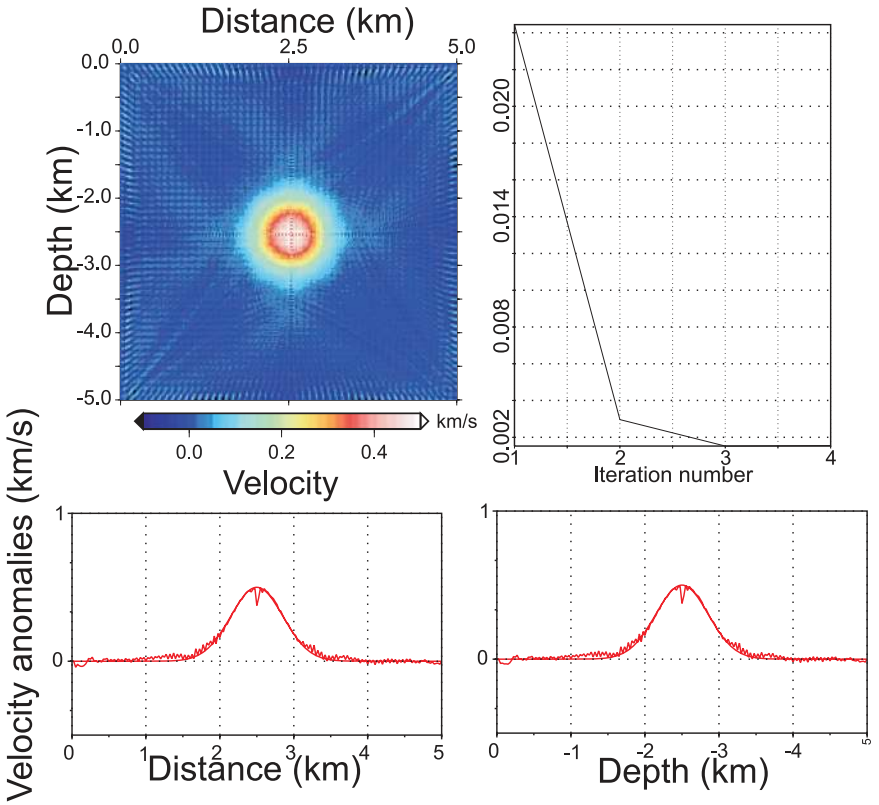

Figure 6 Tomographic inversion without any regularization or damping. On the top left-hand panel, the reconstruction anomaly and the ray sampling. On the top right-hand panel, the misfit behaviour leading to a minimal RMS error of 1.52 *ms*. Four iterations are enough for convergence. Please note high frequency noise and the glitch along geometrical axes related to poor ray coverage as shown by the horizontal section (bottom left panel) and by the vertical section (bottom right panel).

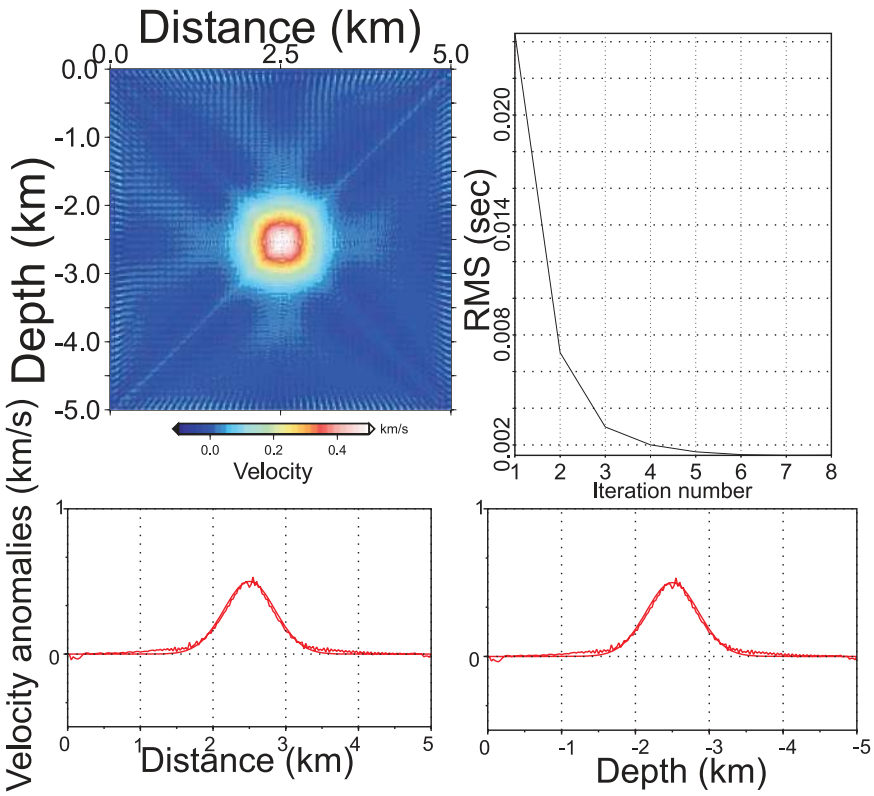

Figure 7 Tomographic inversion using wavelet with two vanishing moments and two levels. The convergence is obtained with an RMS error of 1.43 *ms* after eight iterations. See Fig. 6 for the description of the four panels.

Vesnaver 2000; Trinks *et al.* 2005; Zhang and Thurber 2005). The size of the elementary cells is adapted to the local ray coverage such that the amount of control for each model parameter tends to be uniform. More promising is the multigrid approach, where the inversion procedure recovers values on different grids depending on the resolution power (Madych 1999; Chiao and Kuo 2001; Chiao and Liang 2003; Zhou 2003), based on wavelet tools (Mallat 1999). In this article we shall propose a further investigation of this multigrid approach using a new wavelet transform named as second-generation wavelets (Sweldens 1997).

WAVELET TRANSFORMATION

Over the last few years, many constructions of wavelets have been introduced both in mathematical analysis and in the signal processing literature. In mathematical analysis, wavelets were originally constructed for the analysis and description of among other signals, geophysical signals by using translations and dilations of one specified function called the mother wavelet. A mathematical framework has been set up by the "French school" (Daubechies, Grossmann and Meyer 1986; Grossmann and Morlet 1984). In signal processing, wavelets found their way in the context of quadrature mirror filters (Mintzer 1985; Vetterli 1986; Nguyen and Vaidyannathan 1989; Vaidyannathan 1993; Woods and O'Neil 1986). The connection between the two approaches has been performed through the introduction of multiresolution analysis and the fast wavelet transform by Mallat (1989a,b) and Meyer (1999). An important step has been achieved with the construction of orthogonal, compactly supported wavelets by Daubechies (1988). Since then, several generalizations to the biorthogonal or semi-orthogonal cases were presented (Vetterli and Herley 1989; Cohen, Daubechies and Feauveau 1992). Biorthogonality allows the construction of symmetrical wavelets and thus the use of linear phase filters.

For seismic tomography, wavelets can be used for the description of the model at various scales (Chiao and Kuo 2001; Chiao and Liang 2003; Zhou 2003). The slowness model is expanded on a wavelet basis and the tomographic system is solved for wavelet coefficients. The wavelet transform of the model provides a compact multigrid representation of the

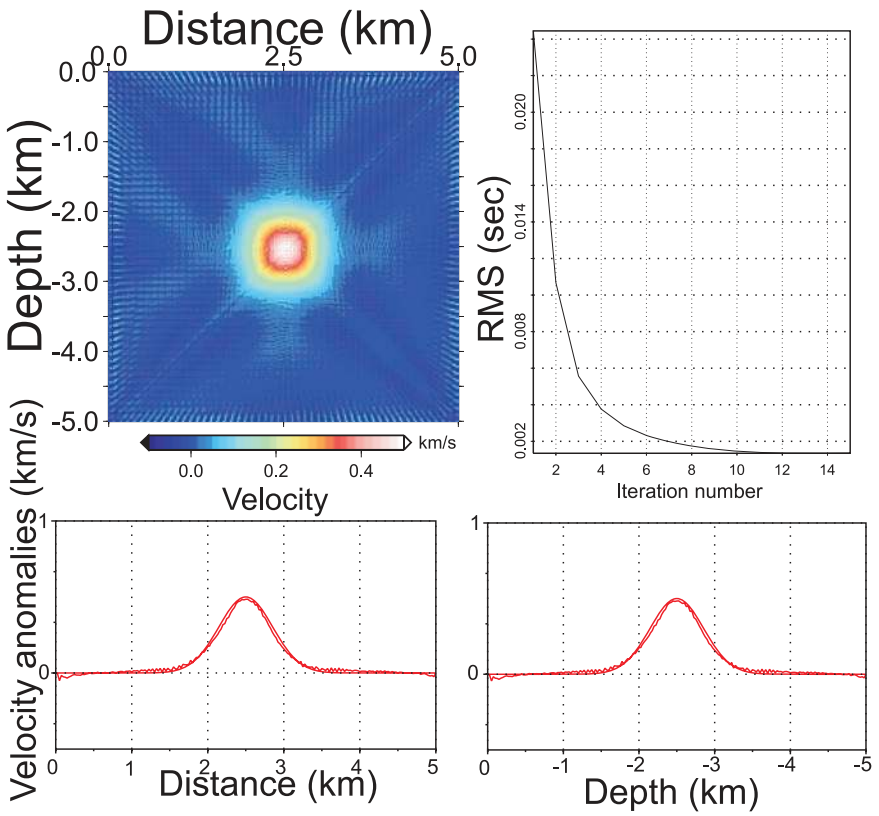

Figure 8 Tomographic inversion using wavelet with two vanishing moments and four levels. The convergence is obtained with an RMS error of 1.35 *ms* after thirteen iterations. See Fig. 6 for the description of the four panels.

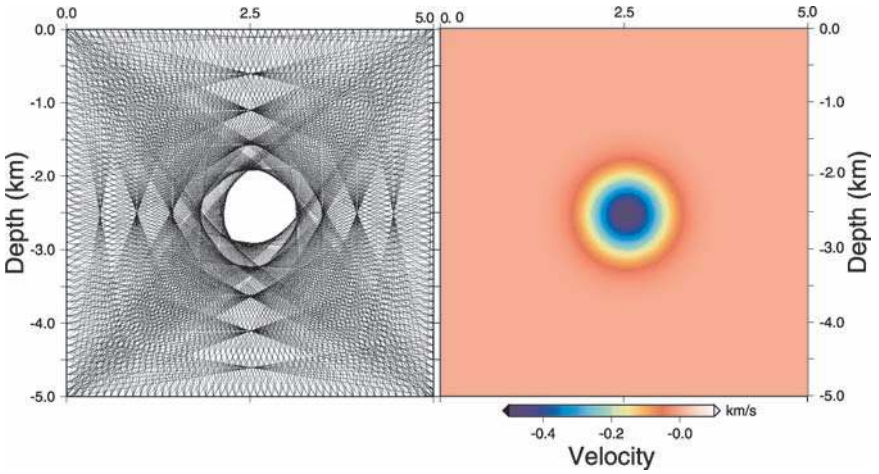

Figure 9 Synthetic example: a circular negative anomaly on top of an homogeneous model as shown on the right-hand panel. Please, note the different ray coverage on the left-hand panel compared to the positive anomaly case, essentially at the center of the box. The central value could not be recovered in the centre and is entirely controlled by the parameterization we shall use.

model, thanks to orthogonal or bi-orthogonal basis functions built by translating (localization property) and dilating (scaling property) a mother wavelet and a scaling function. These functions are the basis of functional spaces generally denoted V_i for the scaling functions and W_i for the wavelet functions. The spaces verify certain properties and constitute what is called a multiresolution analysis (Mallat 1999). Each level i corresponds to a grid size. Generally, the resolution of each

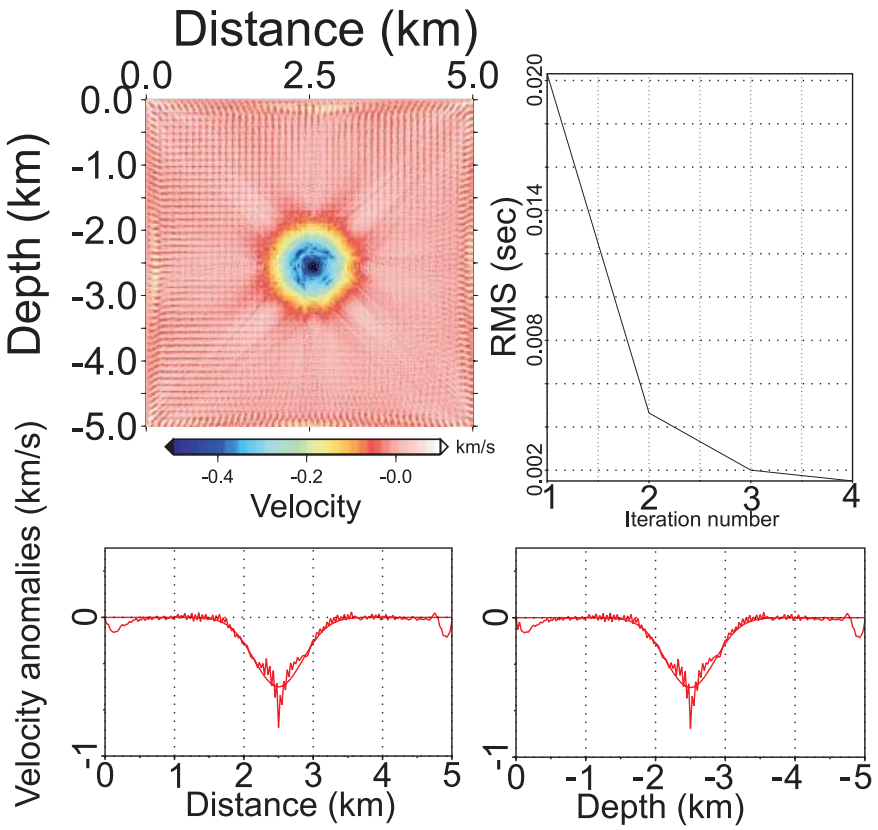

Figure 10 Tomographic inversion without any regularization or damping. On the top-left panel, the reconstruction anomaly and the ray sampling. On the top-right panel, the misfit behaviour leading to a minimal RMS error of 1.50 *ms*. Four iterations are enough for convergence. Please note high-frequency noise and the glitch along geometrical axes related to poor ray coverage as well as edge effects for similar reasons as shown by the horizontal section (bottom-left panel) and by the vertical section (bottom-right panel).

grid decreases by a factor of 2 and the mesh spacing is set accordingly when the level i increases. The starting point is the slowness model on the finest grid of the multiresolution representation (generally V_0). The wavelet transform applies a cascade of orthogonal restrictions of the slowness model over different approximation grids. The restrictions of the slowness model on a coarse grid are encoded by the scaling coefficients. These coefficients are the components along the scaling functions (generally denoted ϕ_i), basis functions of space V_i. At each step, the increment of information lost during the restriction on the coarser grid is encoded by the wavelet coefficients before proceeding to the next grid. Those coefficients are associated to wavelet functions (generally denoted ψ_i), the basis functions of the spaces W_i. At the last iteration of the orthogonal transform, we end up with the wavelet coefficients on each grid plus the scaling coefficients on the coarser grid (see Fig. 1 for a simple example with a 1D signal). The inverse wavelet transform proceeds in the other direction from the coarse grids to the finer ones.

A multidimensional wavelet basis can be built by a tensor product of 1D basis. The decomposition of the 2D slowness model of a wavelet basis can be written as:

$$\begin{aligned} u(x,z) = &\sum_{j=1}^{J}\sum_{n=-\infty}^{+\infty}\sum_{i=1}^{I}\sum_{m=-\infty}^{+\infty} c^{ww}_{j,n,i,m}\psi_{j,n}(x)\psi_{i,m}(z) \\ &+ \sum_{j=1}^{J}\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} c^{ws}_{j,n,I,m}\psi_{j,n}(x)\phi_{I,m}(z) \\ &+ \sum_{n=-\infty}^{+\infty}\sum_{i=1}^{I}\sum_{m=-\infty}^{+\infty} c^{sw}_{J,n,i,m}\phi_{J,n}(x)\psi_{i,m}(z) \\ &+ \sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} c^{ss}_{n,I,m,J}\phi_{J,n}(x)\phi_{I,m}(z), \end{aligned} \tag{1}$$

where c^{ww}, c^{ws}, c^{sw} and c^{ss} are the wavelet and scaling coefficients. The indices i, j and n, m are the scale and localization indexes, respectively. The indices I and J denote the coarsest levels of the multiresolution approximation. A compact representation of the wavelet coefficients of a 2D signal has been proposed by Mallat (1989a,b) (Fig. 2).

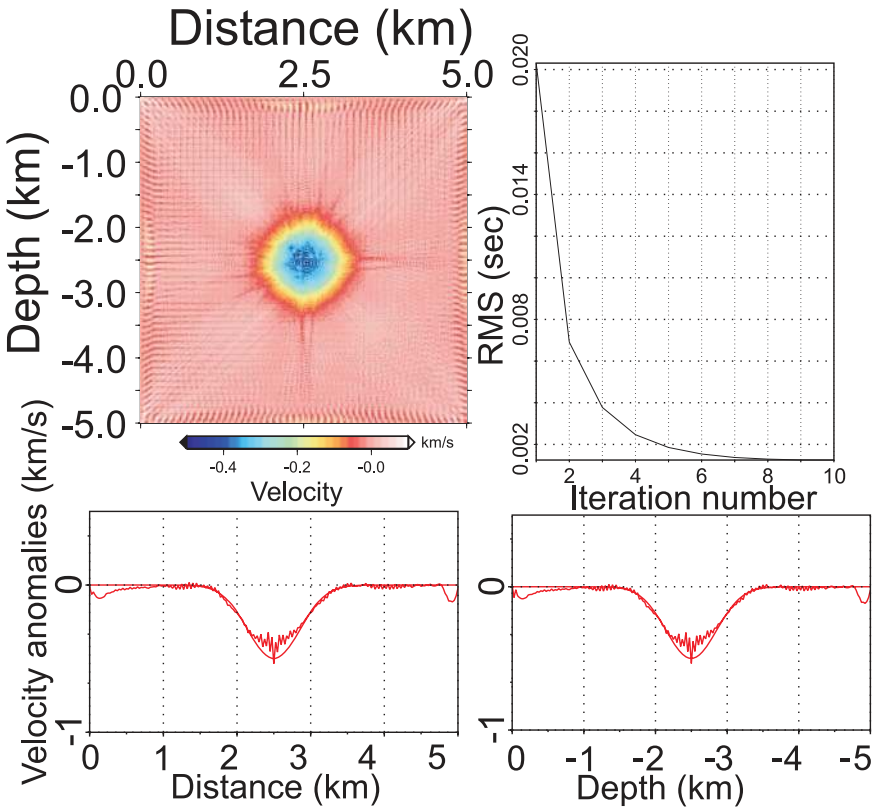

Figure 11 Tomographic inversion using wavelet with two vanishing moments and two levels. The convergence is obtained with an RMS error of 1.24 *ms* after ten iterations. See Fig. 10 for the description of the four panels.

As a Fourier driven transformation, wavelet transformations encounter some relevant drawbacks: the number of samples has to be a power of two, samples have to be regularly spaced and one must deal with periodicity which complicates boundary conditions when working on edges or surfaces. To overcome these difficulties, the concept of biorthogonal wavelets constructed by the lifting scheme has been introduced (Sweldens 1994; Sweldens and Schroder 1995; Sweldens 1997), leading to the so-called second generation wavelets. Wavelets are not necessarily translations and dilations of a defined function. Such wavelets can be adapted to curved edges and surfaces delimiting quite complex domains. They are concerned by weighted and irregular samples. The lifting scheme leads to a faster, in-place calculation of the wavelet transform.

In terms of signal processing, projections on spaces V and W can be seen as high- and low-pass filtering operators, respectively. A procedure of direct and inverse transforms can then be interpreted as a perfect reconstruction filter bank (h_1, g_1, h_2, g_2), where (h_1, g_1) are the filters for the direct transform and (h_2, g_2) are the filters for the inverse transform. In the orthonormal transform, $h_1 = h_2 = h$ and $g_1 = g_2 = g$ and (h, g) are called conjugate mirror filters. In the biorthogonal case, relaxing the orthogonality condition implies working with larger spaces such as $\tilde{V}$ and $\tilde{W}$ spaces, respectively the dual spaces of V and W. Dual spaces are used for the direct transform and normal spaces are used for the inverse transform (or *vice versa*). The associated perfect reconstruction filter bank is denoted $(h, g, \tilde{h}, \tilde{g})$. The lifting scheme is based on the very simple property demonstrated by Sweldens (1994): given a biorthogonal filter, another one can be built as follows:

$$(h, g, \tilde{h}, \tilde{g}) \Rightarrow (h, g_1, \tilde{h}_1, \tilde{g})$$

$$\begin{cases} \tilde{H}_1(w) = \tilde{H}(w) + S(2w)\tilde{G}(w), \\ G_1(w) = G(w) - S(2w)H(w), \end{cases}$$

where capital letters stand for the trigonometric polynomials associated with script letter filters and the function S is a simple trigonometric polynomial. This polynomial is chosen so that the wavelet and dual scaling functions satisfy the desired properties (as the number $\tilde{N}$ of vanishing moments of

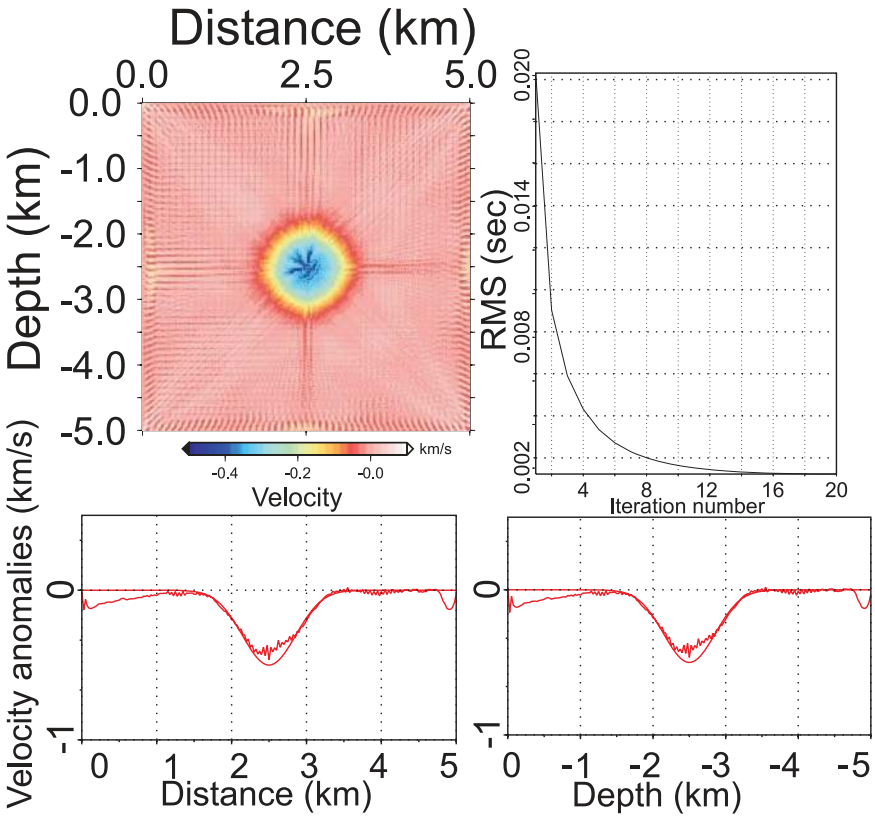

Figure 12 Tomographic inversion using wavelet with two vanishing moments and four levels. The convergence is obtained with an RMS error of 1.23 *ms* after twenty iterations. See Fig. 10 for the description of the four panels.

the dual wavelet, for example). Of course, a dual lifting can also be defined in a similar way:

$$(h, g, \tilde{h}, \tilde{g}) \Rightarrow (h_1, g, \tilde{h}, \tilde{g}_1)$$

$$\begin{cases} H_1(w) = H(w) + \tilde{S}(2w)G(w), \\ \tilde{G}_1(w) = \tilde{G}(w) - \tilde{S}(2w)\tilde{H}(w). \end{cases}$$

With this operation, the real scaling function and the dual wavelet functions can then also be attuned with a specific function $\tilde{S}$ so that they have a requested number N of vanishing moments. These properties assert that, given a biorthogonal filter, one can construct another one satisfying chosen properties. Starting from a simple biorthogonal filter, it is then possible to bootstrap one's way up to a multiresolution analysis with the desired properties on primal and dual wavelets.

The so-called 'lifting scheme' proceeds with this strategy: one starts with the straightforward biorthogonal filter that does nothing but sets aside odd and even indexed coefficients (the so-called Lazy Wavelet). Then one does a dual lifting operation by predicting odd indexed coefficients using even indexed ones (we use the spatial correlation between neighbours). We thus ensure that polynomials of degree up to $N-1$ are well predicted. During a second time, we update the even indexed coefficients using new odd coefficients so that $\tilde{N}$ moments of the signal are conserved. The algorithm can be written as:

$$\begin{cases} (odd_{i+1}, even_{i+1}) = Lazy(\lambda_i), \\ odd_{i+1} -= P(even_i), \\ even_{i+1} += Update(odd_i). \end{cases}$$

The inverse transform can be written similarly as:

$$\begin{cases} even_{i+1} -= U(odd_i), \\ odd_{i+1} += P(even_i), \\ \lambda_i = Merge(odd_{i+1}, even_{i+1}). \end{cases}$$

This numerical flowchart can be extended because the predict stage is a simple polynomial interpolation and the update stage is all but the resolution of linear systems: boundary conditions and number of samples are no longer a limitation of the wavelet decomposition.

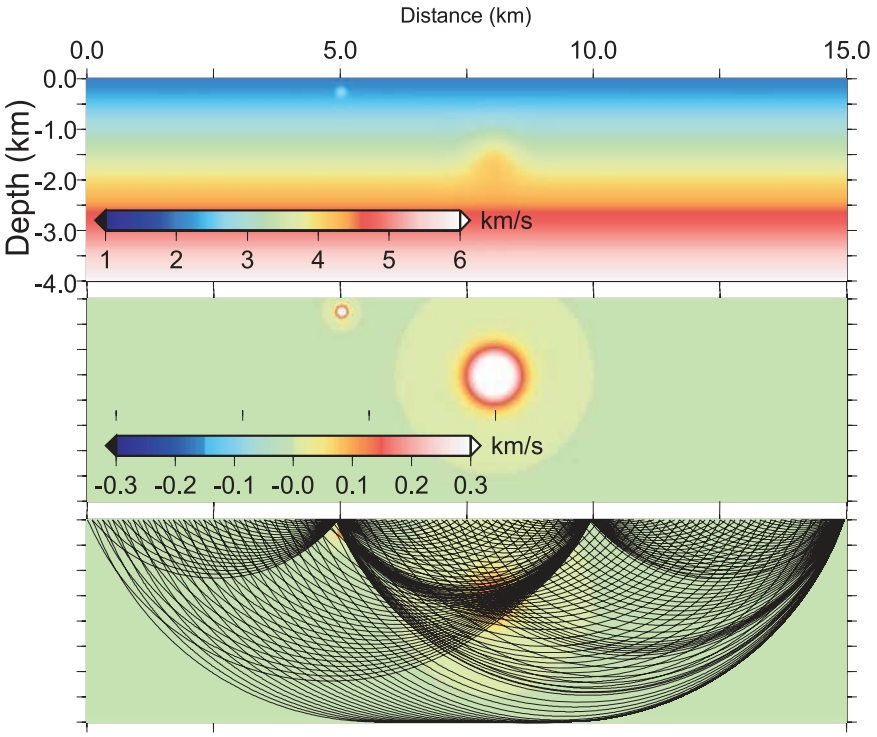

Figure 13 The synthetic model with a velocity gradient and two circular anomalies on the top panel (one is positive and the other negative). The middle panel shows only these two anomalies with a central amplitude of 0.5 km/s while the bottom panel displays the ray coverage in the true synthetic model.

TRAVELTIME TOMOGRAPHY

Once rays have been computed between each source/receiver pair through an eikonal equation solver (Podvin and Lecomte 1991), the delayed traveltime tomographic system relating slowness anomaly Δu to traveltime residuals Δt written as

$$\Delta t = A\Delta u, \tag{2}$$

can be transformed into the new system

$$\Delta t = AW^t \Delta c, \tag{3}$$

where the A matrix is the sensitivity or Fréchet matrix and the equation $\Delta c = W\Delta u$ denotes the transformed wavelet parameters through the transform matrix W. The matrix AW^t can be built by computing the 2D wavelet transform of each row of the original matrix A. A row of the sensitivity matrix has $n_1 \times n_2$ model parameters in a 2D geometry where the number of horizontal grid points is n_1 and the number of vertical grid points is n_2: it contains the contribution of one ray to the Fréchet derivatives. After transformation, we have the same number of unknowns but they are in the wavelet domain and partial derivatives are for scale and wavelet coefficients. Once the system equation (3) has been solved through an iterative procedure as the LSQR algorithm (Paige and Saunders 1982), the slowness perturbation coefficients Δu in the standard geometrical space are obtained by the inverse wavelet transform of $\Delta u = W^t \Delta c$. We shall plot velocity models from these slowness values.

Traveltime tomography in the wavelet domain has an interesting feature since coefficients are associated to levels of the wavelet transform and, consequently, related to scales and local resolutions, adaptivity strategy is much easier to handle through scale hierarchy. The decomposition of a simple circular anomaly leads to a sequence of grids for each scale level using square wavelet basis functions as shown in the Fig. 3. Controlling incremental perturbations through wavelet coefficients between these different descriptions of the true model taken as the one in the V_0 space is a hierarchical procedure. The Fréchet derivatives along a ray will extend their area of sensitivity depending on the scale where are of the wavelet coefficients, as shown in the Fig. 4. The sensitivity of a ray is related to the spatial discretization of the slowness model at the given wavelet scale.

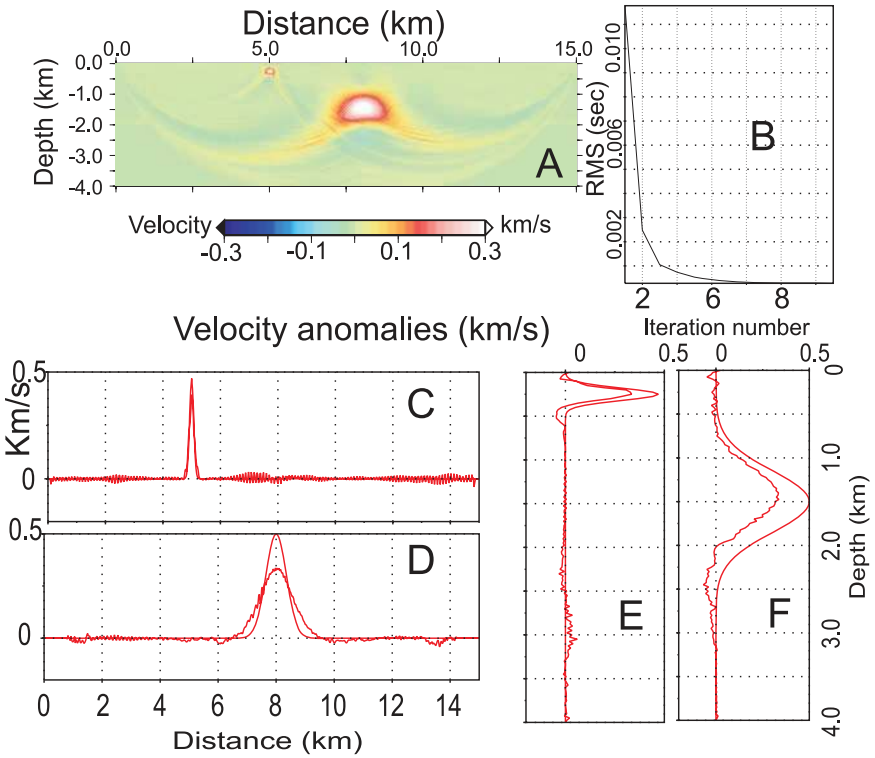

Figure 14 Reconstruction in the slowness domain (panel A) whithout any regularization except the one deriving from the iterative procedure for the matrix inversion. The RMS error goes down to 1.8 *ms* after 10 iterations (panel B). The small heterogeneity is well recovered, as shown on the horizontal profile (panel C) and the vertical profile (panel E) while the deeper heterogeneity presents oscillations as shown on the horizontal profile (panel D) and the vertical profile (panel F). Please note the ray sampling footprint in this undamped inversion: recovering the superficial anomaly requires too high a resolution for the reconstruction of other areas of the structure.

The adaptivity could arise automatically from the inversion procedure. As we shall see later, for highly sampled synthetic examples, this is not the case. Therefore, the adaptivity of the parameterization should be implemented through *a priori* constraints and we propose the application of a bit mask operator to each row of the sensitivity matrix AW^t, zeroing coefficients in areas of the model where poor resolution is expected at a given scale. In other words, this adaptivity selection can be implemented via a linear operator M applied to the Fréchet derivatives in the wavelet domain through the following extended linear system

$$\Delta t = AMW^t \Delta c, \tag{4}$$

in a quite natural way. Designing this operator will depend on local resolution which must be estimated. For surface acquisition, one could, for example, define an empirical rule relating the resolution of the model at a given depth to the minimum source-receiver offset of the rays sampling this depth. The source-receiver offset would provide an estimate of the width of the first Fresnel zone at a given depth and, hence, of the expected resolution of the tomography varying with depth. According to the sampling theorem, only grids with mesh spacing twice smaller than the width of the first Fresnel zone would be locally involved in the inversion: a simple rule for the construction of the M operator. We will discuss later how we can propose a more general strategy for building this operator. Let us remark that the wavelet transform also has the capacity to automatically adapt the parameterization to the resolution in the case where the real resolution is lower than the intrinsic one, such as in the case of a shadow zone. In areas where no rays illuminate the medium, partial derivatives of the traveltime with respect to the wavelet coefficients at the finest scales will be zero, due to the second generation property. By contrast, the partial derivative of the traveltime with respect to the wavelet coefficients of the coarsest scales will be non-zero, due to the spatially extended sensitivity of the coarse grid representations of the rays. Hence, only the coarse-scale wavelet coefficients of the slowness model will

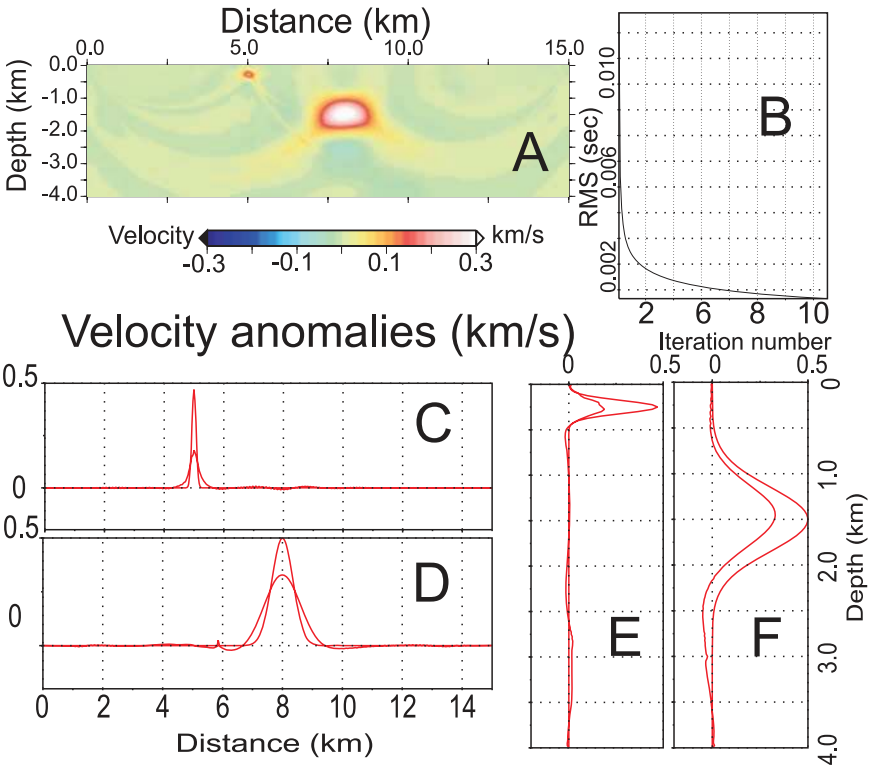

Figure 15 Reconstruction using a penalty inversion with a Gaussian regularization: the correlation length is uniform and taken as the value of 500 *m* in the entire domain. The RMS error reaches the same value of 1.8 *ms* as the undamped inversion after 12 iterations. The deepest heterogeneity has a smoother shape than in the Fig. 14, but the shallower one registers a loss of high-frequency content and thus is not well reconstructed. See the Fig. 14 for a description of the panels.

be non-zero after inversion and will automatically lead to a smooth representation of the true model.

SYNTHETIC EXAMPLES

We present synthetic examples to illustrate how the wavelet decomposition enables us to reconstruct P-wave velocity structures. First, we focus our attention on the capability of the wavelet tomography to locally adapt the resolution of the reconstruction by considering a well-controlled experience of a positive or negative anomaly. We then compare results in a more standard surface-surface configuration between a standard tomographic approach and the wavelet approach: we shall see that we need to specify *a priori* the resolution we expect locally. We propose a numerical strategy for the construction of an operator for this local resolution analysis. We illustrate how the resolution is adapted in the image of the structure for a synthetic crosswell experience. Finally, we compare performances of the standard tomography approach and the wavelet approach for a real case of a surface-surface configuration.

A densely illuminated model with positive anomalies

We investigate the performance of the wavelet tomography for automatic tuning of the resolution using a synthetic example where the ray density will be quite high. We design an acquisition system for a complete illumination of the simple circular positive velocity anomaly of 0.5 km/s that we want to reconstruct (Fig. 5) from a background model of constant velocity of 3 km/s. This background velocity model will be our starting model. The configuration is essentially a transmission configuration. Each side of the square medium has 102 sources, while 51 receivers are lying on the opposite side, leading to a total of 5 202 rays. For each inversion, we estimate the lowest RMS error from the misfit function as it flattens with linearized iterations. Moreover, we display horizontal and vertical sections for a better appreciation of the spatial frequency contents of the reconstruction.

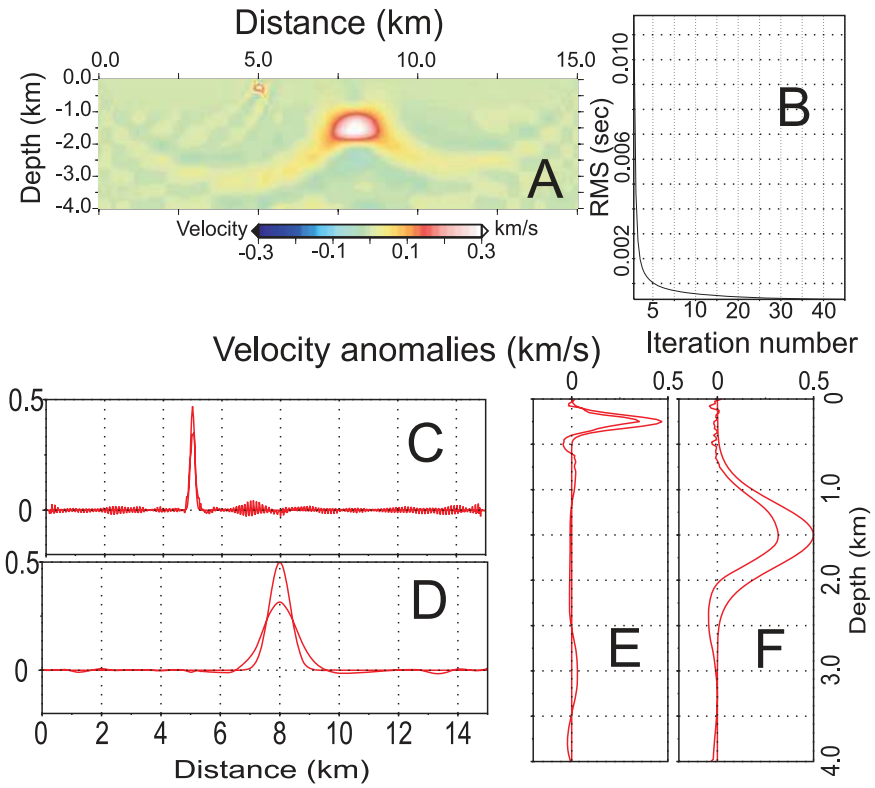

Figure 16 Reconstruction using Haar wavelets with four vanishing moments and four levels: a mask ensures that under the depth of 800 *m*, the fourth level of wavelet is the only one kept, corresponding to a resolution of 400 *m*. The RMS error reaches the same value of 1.8 *ms* as for the undamped and regularized inversions with 40 iterations. The large heterogeneity is kept smooth and the shallow heterogeneity is well reconstructed. See the Fig. 14 for a description of the panels.

The standard tomography approach using the regularization technique will work nicely for this example, as the resolution is roughly uniform in the central part of the medium. The wavelet approach should provide similar results. In order to analyse how high-frequency content will be kept under control by the multiscale feature of the wavelet approach, we have performed a standard tomography over a grid of 5 *m* step size, which gives us a 101 × 101 regular grid. For our purposes, no damping is applied to the system. The convergence is obtained after four iterations and the anomaly is accurately recovered although with high-frequency artefacts (Fig. 6): these artefacts will be nicely removed by any regularization, as a penalty, by a smooth Gaussian operator.

How these high-frequency features will be eliminated by the wavelet tomography is the question we want to examine.

When considering wavelet decomposition, we select wavelets with two vanishing moments. We consider two and four levels for the scale sampling. This means that the coarser grid size will have 10*m* resolution and 40*m* resolution, respectively. The anomaly is well recovered but the convergence to the flat level of the misfit function is slightly slower than before (Figs 7 and 8). We need eight and thirteen iterations, respectively, to reach convergence. We may underline that the RMS final error decreases as we increase the number of levels, although the decrease is quite small. For real applications with errors in the picked traveltime, this may not be detected. More importantly, the high-frequency artefacts are less noticeable, providing a better fit of the model. As we increase the number of levels, we can see the decrease of the amplitude of the high-frequency artefacts. In other words, when coarser levels are introduced, the reconstruction is driven by a better spatial coherence, essentially for broad features. We may say that the parameterization into the wavelet domain provides a better conditioning of the system to be inverted, since it is less sensitive to high-frequency perturbations. As a partial conclusion, we have found in this very particular case that the wavelet tomography adapts itself to the local resolution.

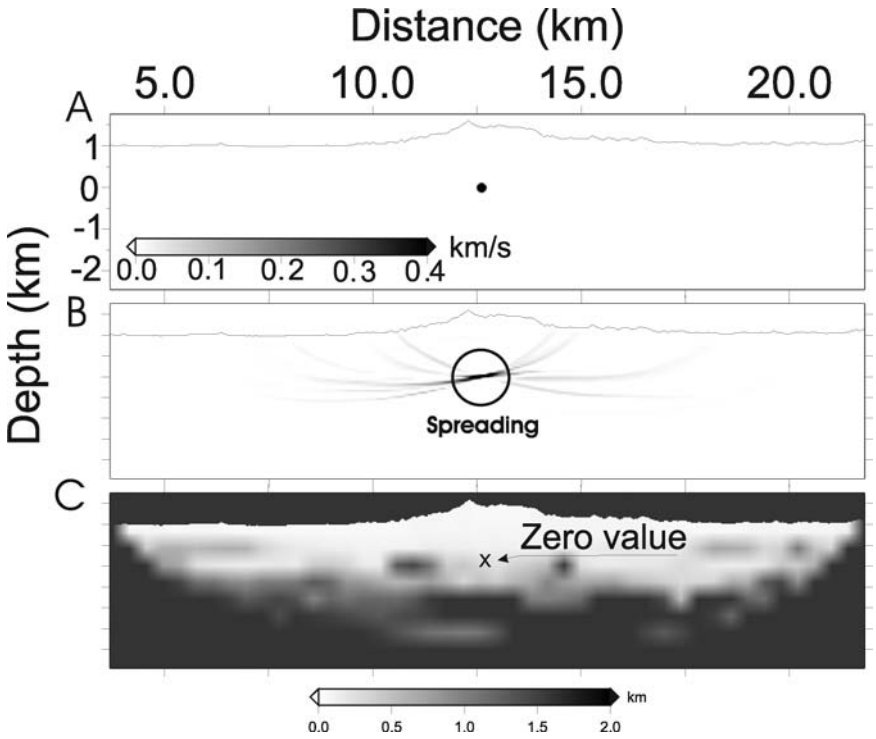

Figure 17 Resolution analysis: top panel displays the spike anomaly we wish to reconstruct, the medium panel the spreading of the amplitude with a partial amplitude reconstruction where the expected spike is and the bottom panel related to the characteristic length (circle radius where 95% of the amplitude is recovered) from which we will design a mask operator for wavelet coefficients.

A densely illuminated model with negative anomalies

We may proceed similarly when considering a negative velocity anomaly. The velocity amplitude will be −0.5 km/s and the background velocity model will be the starting model. The acquisition distribution is exactly the same as the previous example and, therefore, the number of rays is identical, allowing comparison between reconstructions of positive and negative anomalies. The reconstruction is more difficult as rays tend to avoid low-velocity zones, creating shadow zones that are not sampled by rays (Fig. 9). Velocity anomalies should not be reconstructed in the centre of the medium because rays are missing there. Therefore, the reconstructed anomaly in the centre of the medium comes from the grid discretization (and the related Fréchet derivative of these nodes if there is one) and from the starting medium without this negative anomaly inside, where the centre of the square medium is initially sampled by straight rays.

The negative velocity anomaly is retrieved as for the previous example, although fundamental differences can be observed (Fig. 10). When considering an inversion of slowness parameters, we observe high-frequency oscillations in the centre of the medium where the illumination by rays is poor in the final model: rays do not sample this zone at the end of the procedure – they do it at the beginning of the inversion since the first iteration is performed with straight rays inside an homogeneous medium. Because of this poor sampling, edge anomalies were observed and could be hardly seen for positive anomalies.

By introducing wavelet coefficients with two vanishing moments, we recover the negative anomaly with an increased number of iterations (Figs 11 and 12). The final RMS residual is smaller although the decrease is not significant. Because the anomaly we want to recover has a smooth wavelength content, the result is better when considering four levels of discretization, as we have checked. Partial reconstruction introduces numerical artefacts that spread inside the medium from edges even when we move to the four levels description. Reconstruction of the wavelet coefficients could not automatically overcome the aliasing, in spite of the dense ray coverage.

Using wavelet decomposition leads to a memory increase by a factor of three for the Haar decomposition with two vanishing moments as well as a CPU time roughly twice the one observed for standard tomography.

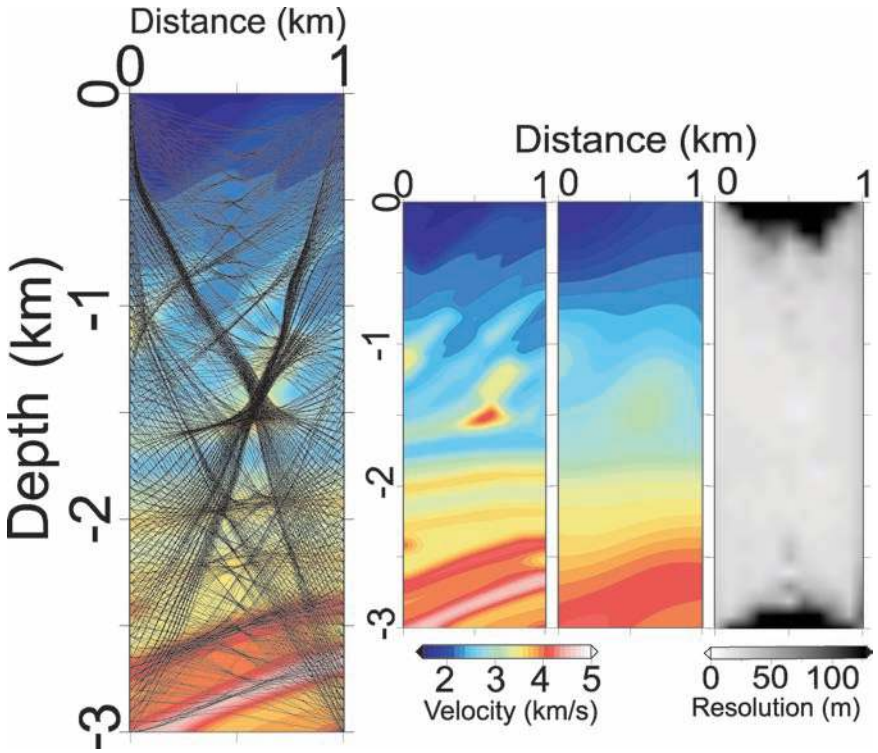

Figure 18 Resolution analysis: the left-hand panel shows the quite dense ray sampling of the initial model between two hypothetical wells while the right panel is divided into (a) the exact model on the left, (b) a smooth version of this model in the middle and (c) a resoluton analysis on the right, based on the radius of the circular extension for recovering 95% of the small velocity perturbation.

In conclusion of these two simple synthetic tests, the wavelet transformation contains, as expected, a preconditioning of the linearized system: high-frequency perturbations are less noticeable and scale decomposition increases the stability of the reconstructed system at the expense of the number of iterations. Unfortunately, this automatic tuning has been found to work tediously, especially for poorly sampled negative anomalies and may require an additional strategy for speeding up the repartition between scales, especially when considering more standard geometries of acquisition.

We shall examine how to perform this repartition in the following more realistic synthetic example where resolution is expected to vary with depth.

A typical surface-surface synthetic example with two different anomalies

Dealing with uneven ray coverage is a crucial issue in tomography and a configuration with sources and receivers on the free surface is considered: 30 sources with a maximum of 41 receivers per source provide 1030 rays that are bent by the velocity increase with depth in the structure we consider. Hence, parts of the model are expected to be known precisely, while other parts will suffer from low resolution. A spatially varying illumination related to poor ray coverage results for this configuration. In this case, the theoretical resolution is expected to be heterogeneous: in the shallow part, both long and short rays spanning across a large range of azimuths make the resolution quite high. On the contrary, the deepest parts of the model are sampled and enlightened only by longer rays with a mainly horizontal azimuth, leading to low resolution.

Wavelets deal with the multiscale nature of the information contained in the data but the inversion is not able to automatically focus anomalies on related scales. Working in the wavelet domain makes the selection of the scales of the parameter model by a predefined operator easy. Zeroing wavelet coefficients corresponding to scales too fine to be reconstructed will help in reducing the spatial aliasing of the inversion procedure. By applying a mask operator to each row of the sensitivity matrix written in the wavelet domain, we may successfully distinguish well- and poorly-resolved zones in our reconstruction. Applying a mask does not increase the computational

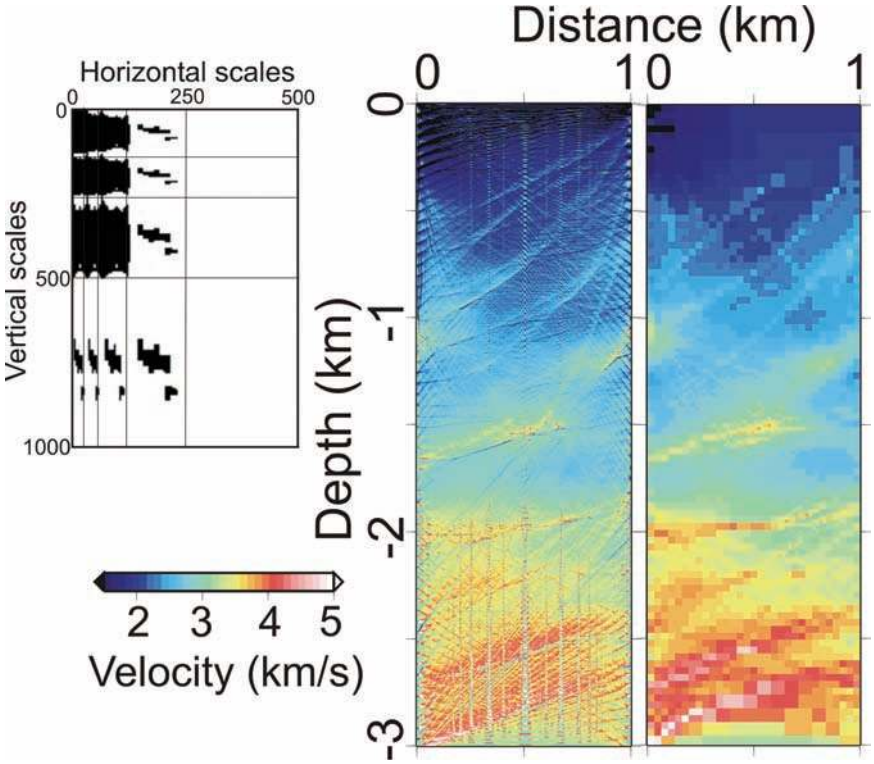

Figure 19 Tomographic inversion: the left-hand panel displays the mask operator over the four inverted scale levels from the finest to the coarsest one (the finest fifth one connected to the sampling of the MARMOUSI model has been zeroed by the mask operator), the middle panel displays the inverted model using slowness coefficients without any damping and the right-hand panel displays the inverted model using Haar wavelet coefficients over four scales. Please note that cancelling the entire finest scale level does not harm the inversion procedure and shows us that the tomography is not able to reconstruct small-scale features.

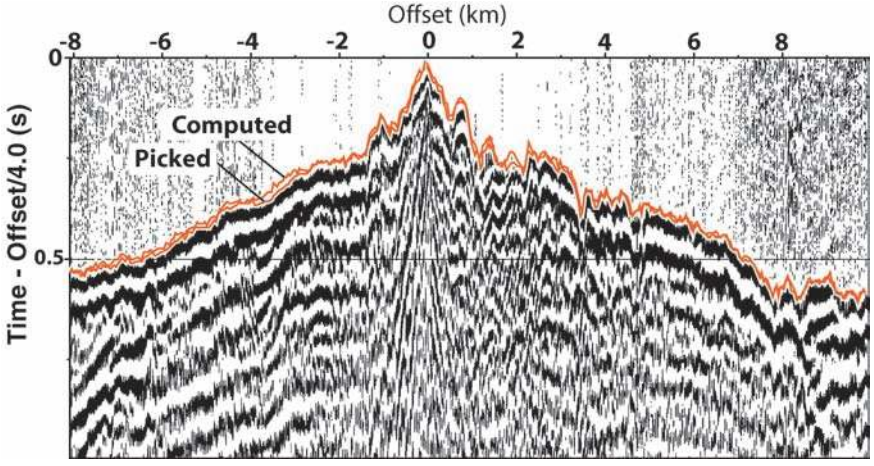

Figure 20 First-arrival picked time using industrial software for the gather of shot 311; synthetic traveltimes are displayed, illustrating that the recovered velocity model predicts traveltimes quite well, except at near offsets.

cost, as we perform a multiplication by zeroes or ones for each term of the matrix. Defining the mask may be quite demanding although pre-tabulated masks should be devised for typical data acquisition configuration.

Because of the linearized formulation, we need a method that can adapt the resolution to the theoretical one alongside the natural adaptation of the wavelet reconstruction. Wavelet parameterization makes it easy to adapt resolution locally, by

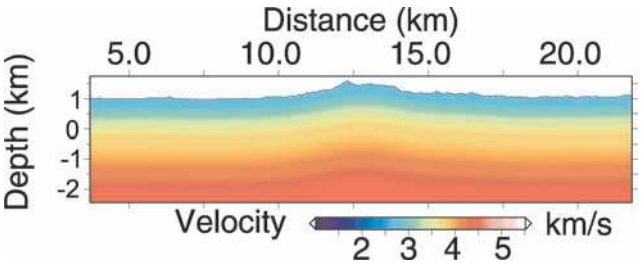

Figure 21 Initial 1D-like model for the traveltime inversion as the 1D vertical profile starts exactly from the complex topography.

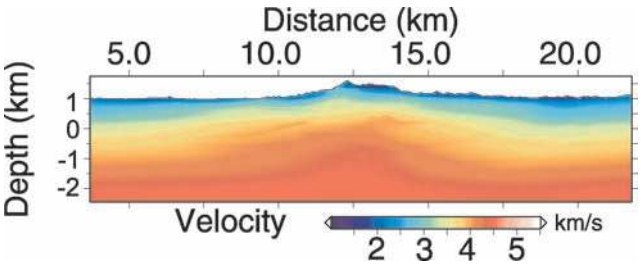

Figure 22 Reconstruction through slowness coefficients inversion using the penalty procedure based on smooth Gaussian regularization that varies with depth. In spite of the depth variation of the Gaussian operator, this constrained inversion necessary for smearing out ray footprints prevents a highly resolved reconstruction in the superficial part and illustrates the intricate structure of the different resolution scales. The RMS final error is 26 ms.

applying a simple mask. Two velocity heterogeneities, a small one in the shallow part and a wider one in the deep part, are tentatively reconstructed according to the expected local resolution. The starting model is the model with a constant velocity gradient without anomalies (Fig. 13).

Figure 14 shows the reconstruction when considering slowness coefficients without any regularization. One can easily see the smearing related with ray sampling for the broad, deepest anomaly (see sections D and F of the Fig. 14). The small heterogeneity is well reconstructed (see sections C and E of the same figure) but the model contains high-frequency perturbations. The amplitude of the widest anomaly is not well recovered due to ray enlightening, which is not as dense as in the shallow part. The results contain a reasonable amount of low-frequency information, reducing the RMS error down to 1.8 *ms*. Nevertheless, high-frequency information is quite strong and must not appear during the inversion procedure.

Because the poor ray coverage leads to the underdetermination of the linearized system, we now consider the inversion procedure with constraints. The one we have selected is the penalty approach, where a Gaussian smoothing operator between slowness parameters adds rows to the iterative linear system to be solved. The definition of the length of this operator is a difficult task, as we may over-smooth the medium by cutting down too small eigenvalues of the linearized system. The level of smoothing needed for removing high-frequency perturbation in low resolution zones leads to a loss of high frequency content in better defined parts of the model (see Fig. 15). A characteristic length of the Gaussian operator of 500 *m* has been found necessary for removing high-frequency oscillations and provides the same RMS error of 1.8 *ms* as the one observed in the undamped tomography. The broader, deep anomaly is recovered as previously because of its low-frequency content, while the shallower anomaly is not fully recovered as the smoothing operator cancels out the eigenvalues of the linear system necessary for such reconstruction. Of course, we may vary the characteristic length of the Gaussian operator with respect to depth in this simple example.

Wavelet decomposition allows different resolutions at different places in the model due to the multiscale structure. Once the Fréchet derivative is computed, we should apply an M operator designed as follows. From the surface down to the depth of 800 *m*, wavelet coefficients are retained at all scales. From the depth of 800 *m* down to the bottom, only the fourth level of wavelet coefficients is retained, corresponding to a resolution of 400 *m*, coefficients of the finer other, finer three scales are zeroed out. The inversion provides a reconstruction that varies in resolution as we move to deeper depths (Fig. 16). We have chosen the mask operator in order to separate two zones of different resolutions.

Although we cannot perform this *a priori* separation for realistic structures, this simple example illustrates the capacity of the wavelet decomposition for spatially varying resolution reconstruction: a high resolution in the shallow part and a broader resolution at depth. We may obtain the same result with standard constrained tomography by considering locally varying smoothing operators, as we have tried, or by defining covariance matrices. These alternatives do not lead to a simple implementation as the definition of the M operator.

We may conclude that wavelet decomposition adapts locally the resolution of the reconstruction by the multiplication of wavelet coefficients by a bit mask operator, when the Fréchet derivative matrix is calculated in the wavelet domain.

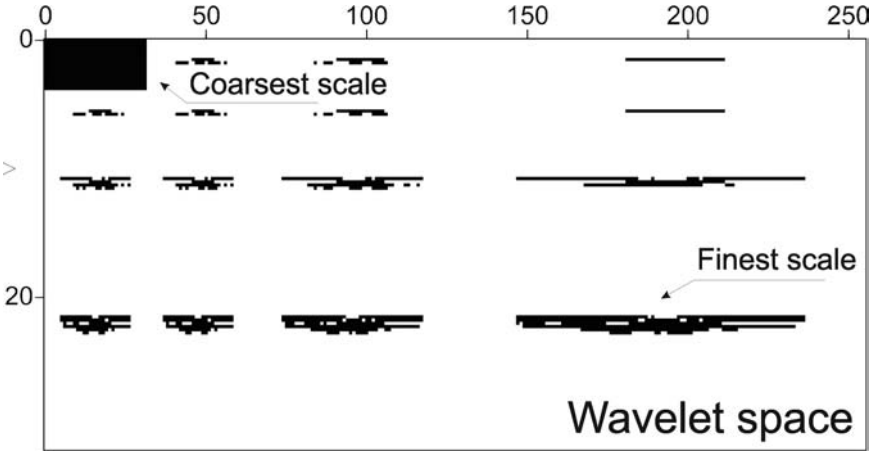

Figure 23 Bit mask construction over the already constructed smooth medium based on an analysis of the numerical estimation of the resolution matrix. The criterion for having a value of one for the mask is a local reconstruction of 95% of the velocity amplitude. Please note that the mask is set on only the nearby free surface for the finest scales.

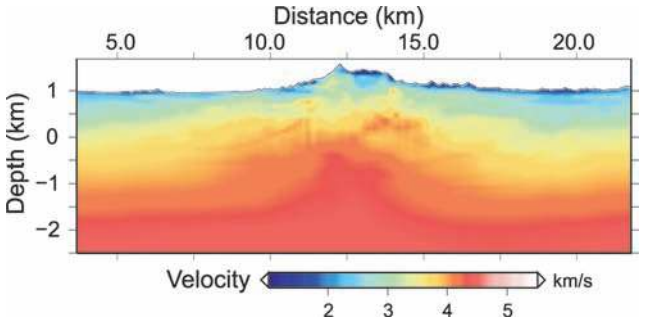

Figure 24 Reconstruction velocity model using the Battle-Lemarie wavelet decomposition: two vanishing moments have been requested and the RMS final error is 23 *ms*. Please note the finest details nearby the free surface when comparing with the model of Fig. 22.

RESOLUTION ANALYSIS FOR DEFINING THE BINARY *M* OPERATOR

For better focusing of the reconstruction through different scale levels, one has to estimate *a priori* resolution in order to overcome difficulties deriving from the linearized inversion. We assume that a model has been designed for which one can perform a resolution analysis: a standard damped traveltime tomography could be such a candidate. Starting with this model, we may want to have a rough estimation of the resolution matrix for a better performance of the inversion (Nolet, Montelli and Vireux 1999). Of course, once done, we may end up with another medium we shall use again for a new resolution analysis and so on. We cannot rule out this loop over adapted models.

For small systems, one can use a direct solver and computing the resolution matrix is quite fast. For large systems, we may rely on numerical iterative techniques as proposed by Latorre *et al.* (2004). As shown in the Fig. 17, a small local velocity perturbation is introduced with an amplitude of 0.5 km/s, making a spike with a given spatial grid extension. This amplitude is such that we have a perturbation of traveltimes above the numerical noise and rays are kept nearly undeformed. Therefore, we may keep rays and Fréchet derivatives computed in this reference model, speeding up dramatically the inversion procedure: we perform the linear inversion for different spikes in each zone of the medium. The recovered amplitude is spread out away from the spike location: a characteristic length is estimated as the radius of the circle where 95% of the amplitude of the spike has been restored and could be estimated quite efficiently for the entire medium (Fig. 17). We may zero out Fréchet derivatives for scales with the finest sampling lengths rather than the local characteristic length, leading to the bit mask operator *M*. Because we have performed a linear investigation, similar results would have been found with negative anomalies.

REALISTIC CROSSWELL EXPERIMENT: APPLICATION TO THE MARMOUSI MODEL

This procedure has been applied to a synthetic crosswell experiment performed inside the Marmousi model (Bourgeois et al. 1991; Versteeg 1994). Between two wells of 3 km depth at a distance of 1 km (Fig. 18), ray tracing between 60 sources regularly spaced in the first well and 60 receivers regularly spaced in the second well was performed for computing synthetic traveltimes. A reference model was constructed

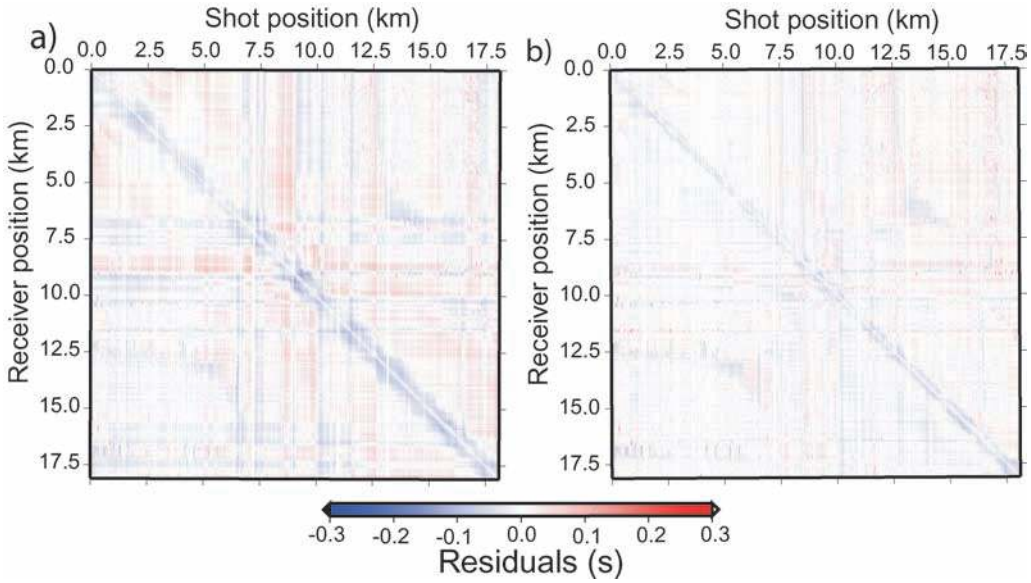

Figure 25 Final time residues after thirty iterations. Please note the too fast velocity structure at near offsets. Very local description of a low-velocity zone below the finest grid mesh should be given and could be performed through statics evaluation.

by smoothing the synthetic model and a resolution analysis performed on this model, leading to an estimation of a resolution map with a 95% threshold level (Fig. 18). This resolution analysis has led us to zero out the finest fifth level of Haar coefficients we are using in our wavelet decomposition as well as specific zones of other scale levels (Fig. 19). In other words, we may start from a very fine grid and the resolution analysis will tell us if this grid could be used for inversion. Inversion is performed using either slowness coefficients or wavelet coefficients with the mask application (Fig. 19). One can see that where the ray sampling is not fine enough especially at the bottom of the two wells, a coarser description of the reconstructed image has been designed through the resolution analysis. Wavelet procedure seems to perform better and provides better features than the slowness reconstruction as long as we are able to specify *a priori* the local resolution.

APPLICATION TO A REAL DATA SET

We have analysed a real data set with a global offset configuration using a surface-surface geometry. We have considered 702 shots with a regular spacing of 25 *m* recorded by 1440 receivers along an 18 km 2D profile with a spacing of 12.5 *m*. Let us note that the samping over shots is twice those of receivers. First-arrival traveltimes for 301 shots have been picked and a tentative macromodel of 18 km long and 4 km deep should be reconstructed using these traveltimes. As an example, the Fig. 20 displays one shot gather with an initial picked time curve and computed time curve. Our slowness inversion grid has 169 vertical nodes and 725 horizontal nodes with a spatial step of 25 *m*. We design an initial model for the tomography by fitting a mean traveltime curve with a 1D velocity model over picked times. We consider this 1D model below the given topography typical of a foothill structure (Fig. 21) using a spatial discretization of 100 *m* in both directions. The model through first traveltime tomography has to fit the data as much as possible and it also has to be smooth, especially for deep zones, as the ray sampling becomes quite poor.

We consider the standard tomography procedure with a Gaussian smoothing operator in the penalty approach that will depend on the depth, as we expect that the resolution will vary strongly with depth. A 2D velocity structure has been constructed through slowness coefficients (Fig. 22) with the same spatial discretization as the initial model. This penalty standard procedure is considered as the reference for comparison. No global damping is applied except the internal iterative procedure of the LSQR algorithm. The number of the linearized inversion is five and the RMS final error is 26 *ms*. This model seems to take into account spatial low-frequency contents while the high-frequency ones are missing.

This reconstructed model is used in the procedure (see Fig. 17) for the construction of the *M* mask operator that is displayed in the Fig. 23. We have selected three levels of the wavelet decomposition starting from 100 *m* spacing. The

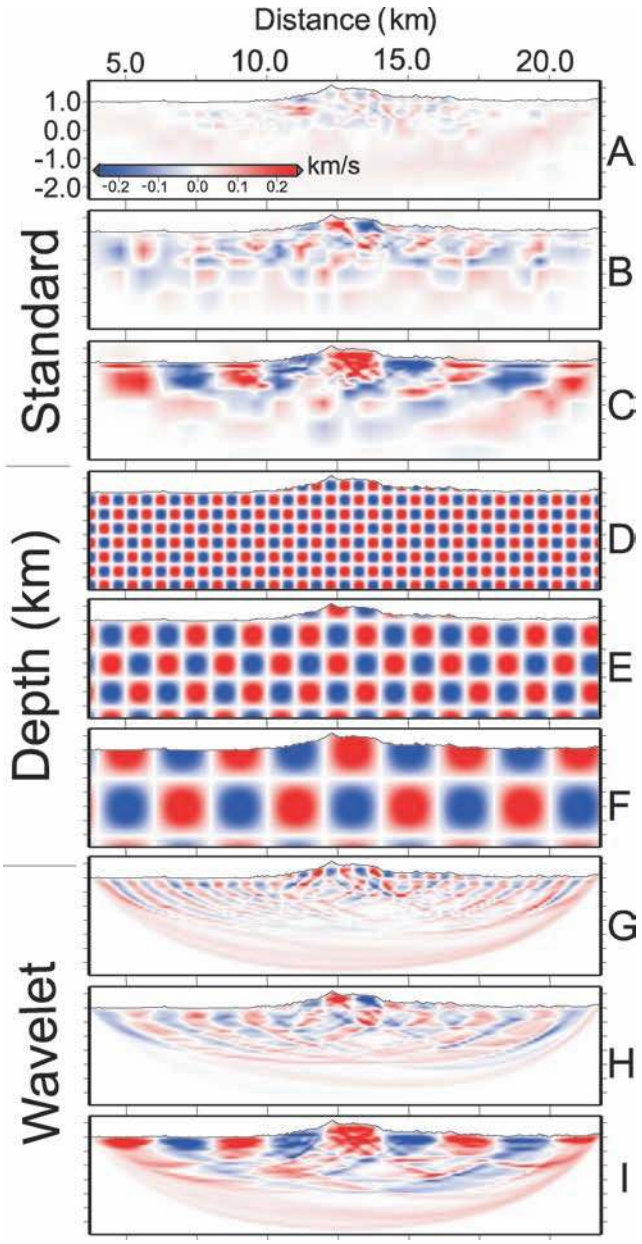

Figure 26 Checker-board test with three middle panels (D,E,F) showing different wavelengths ranging from top to bottom as 500 *m*, 1 km and 2 km. The first three panels (A,B,C) show reconstruction using the standard tomographic procedure while the last three panels (G, H, I) show reconstruction using the wavelet approach. Please note the better behaviour especially at short wavelengths when inverting wavelet coefficients.

coarsest grid is therefore with a spatial discretization of 400 *m*. Using these three characteristic lengths, we performed the resolution analysis we have described and the bit mask operator M is shown in the Fig. 23. One can see that the mask fills the entire domain at the scale of 400 *m* while the mask only fills the superficial part at the scale of 100 *m*. Of course, local variations depend on the ray sampling in the reference model (Fig. 22).

We shall consider linear patches, known as Battle-Lemarie wavelets, with two vanishing moments. After thirty iterations, a model with a higher spatial frequency content nearby the free surface was obtained (Fig. 24). Patterns related to the ray sampling have been kept under control and do not overprint the reconstructed image at depth. The RMS final error of 23 *ms* does not show a significant decrease for assessing the quality of the recovered image through this simple quality criterion. By displaying residual times between sources and receivers (Fig. 25), one may notice the better behaviour at far offsets compared to short offsets. This may lead us to either a picking difficulty at near offsets or a an insufficient dense grid sampling at the free surface. Decreasing the grid step will increase dramatically the number of unknowns, especially for the resolution estimation. Therefore by handling time shifts related to fine details nearby receivers could be captured through statics as is usually done (Kissling *et al.* 1994).

Another way to assess the resolution variation could be through a checker-board test applied to the final best resolved model. As shown by the Fig. 26, different scales are better reconstructed by wavelet decomposition than by slowness decomposition. The checker-board over the smoothest patch with a characteristic length of 2 km provides similar results when considering slowness or wavelet decompositions. Other, smaller, scales are significantly better recovered using the wavelet decomposition, showing us the interest of this decomposition.

CONCLUSION

Seismic tomography of first-arrival delayed times could be performed using a wavelet decomposition of the model space. We have shown that the second generation wavelet decomposition leads to a simple decomposition scheme that is well-suited for traveltime tomography.

Simple synthetic tests have shown that the tomographic inversion does not automatically project traveltime anomalies into the wavelet coefficients at the right scale and, therefore, the need for to insert an *a priori* resolution estimation has been confirmed. For a surface-surface configuration, the reconstruction of two circular anomalies of different radius has been performed quite successfully with a simple mask operator depending on the depth.

From the many ways of designing this binary operator, we have chosen a simple numerical technique and we have shown

that a mask operator could be designed over the wavelet decomposition in order to cancel out coefficients at too small scales for possible reconstruction. Applications to a synthetic crosshole experiment and to a surface-surface real experiment illustrate possibilities for better reconstructions of the medium.

Data misfit reduction is not significant and may be well below the noise level, which leaves us without any clear quality control criterion. The final resolution related to the *a posteriori* covariance matrix will be the quantity for assessing a better reconstruction.

The wavelet decomposition has been unsuccessful for automatically adapting the resolution during the inversion but the insertion of an *a priori* resolution is as simple as a binary operator over wavelet coefficients. Of course, penalty procedures as we have performed or covariance formulation may lead to the same final result, as these operators act as continuous functions. The wavelet decomposition might be seen as a simple way to introduce this spatially varying *a priori* resolution into the first-arrival delayed traveltime tomography.

Extension to 3D configuration should be performed and combining this wavelet decomposition and the finite frequency tomography (Montelli et al. 2004) will certainly improve local resolution of seismic imaging in the future, before tackling the amplitude analysis of seismic traces (Pratt, Shin and Hicks 1998; Operto *et al.* 2004).

ACKNOWLEDGEMENTS

This work was performed under the industry French ministry contract FSH N° CEP&M RE.1006/05. We thank TOTAL for providing us the complete data set as well as the industrial picked time. Special thanks to Aldo Vesnaver for his careful review and to Frank Adler of TOTAL for many interesting discussions that have allowed the improvement of this work.

REFERENCES

Aki K. and Lee W. 1976. Determination of three-dimentional velocity anomalies under a seismic array using first P-arrival times from local earthquakes. *Journal of Geophysical Research* **81**, 4,381–4,399.

Benz H.M., Chouet S.A., Dawson P.B., Lahr J.C., Page R.A. and Hole J.A. 1996. Three-dimentional P and S wave velocity structure of Redoubt Volcano, Alaska. *Journal of Geophysical Research* **101**, 8111–8128.

Böhm G. and Vesnaver A. 1999. In quest of the grid. *Geophysics* **64**, 1116–1125.

Böhm G., Galuppo P. and Vesnaver A. 2000. 3D adaptative tomography using delauney triangles and voronoı polygons. *Geophysical Prospecting* **48**, 723–744.

Bourgeois A., Bourget M., Lailly P., Poulet M., Ricarte P. and Versteeg R. 1991a. Marmousi, model and data. *Proceedings of the EAEG Workshop on Practical Aspects of Seismic Data Inversion*, pp. 5–16. EAEG.

Bourgeois A., Bourget M., Lailly P., Poulet M., Ricarte P. and Versteeg R. 1991b. Marmousi, model and data marmousi, model and data, the marmousi experience. *Proceedings of the EAEG Workshop on Practical Aspects of Seismic Data Inversion*. EAEG.

Chiao L.-Y. and Kuo B.-Y. 2001a. Multiscale seismic tomography. *Geophysical Journal International* **145**, 517–527.

Chiao L.-Y. and Kuo B.-Y. 2001b. Multiscale seismic tomography. *Geophysical Journal International* **145**, 517–527.

Chiao L.-Y. and Liang W.-T. 2003. Multiresolution parametrization for geophysical inverse problems. *Geophysics* **68**, 199–209.

Cohen A., Daubechies I. and Feauveau J. 1992. Biorthogonal basis of compactly supported wavelets. *Communications on Pure and Applied Mathematics* **45**, 485–560.

Dahlen F.A. 2004. Resolution limit of traveltime tomography. *Geophysical Journal International* **157**, 315–331.

Daubechies I. 1988. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics* **41**, 909–996.

Daubechies I., Grossmann A. and Meyer Y. 1986. Painless nonorthogonal expansions. *Journal of Mathematical Physics* **27**, 1271–1283.

Delprat-Jannaud F. and Lailly P. 1993. Ill posed and well posed formulation of the reflection traveltime tomography problem. *Journal of Geophysical Research* **98**, 6589–6605.

Farra V. and Madariaga R. 1988. Non-linear reflection tomography. *Geophysical Journal* **95**, 135–147.

Grossmann A. and Morlet J. 1984. Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM Journal on Mathematical Analysis* **15**, 723–736.

Hole J. A. 1992. Nonlinear high-resolution three-dimentional seismic travel time tomography. *Journal of Geophysical Research* **97**, 6553–6562.

Kissling E., Ellsworth W.L., Eberhart-Phillips D. and Kradolfer U. 1994. Initial reference models in local earthquake tomography. *Journal of Geophysical Research* **99**, 19,635–19,646.

Latorre D., Virieux J., Monfret T., Monteiller V., Vanorio T., Got J.L., and Lyon-Caen H. 2004. A new seismic tomography of Aigion area (Gulf of Corinth, Greece) from the 1991 data set. *Geophysical Journal International* **159**, 1013–1031.

Madych W.R. 1999. Tomography, approximate reconstruction, and continuous wavelet transforms. *Applied and Computational Harmonic Analysis* **7**, 54–100.

Mallat S. 1989a. Multifrequency approximations and wavelet orthonormal bases of $l_2(\mathbf{r})$. *Transactions of the American Mathematical Society* **315**, 69–87.

Mallat S. 1989b. Multifrequency channel decompositions of images and wavelet models. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **37**, 2091–2110.

Mallat S. 1999. *A Wavelet Tour of Signal Processing*. Academic Press.

Menke W. 1984. *Geophysical Data Analysis: Discrete Inverse Theory*. Academic Press.

Meyer Y. 1999. *Ondelettes et opérateurs, i:ondelettes, ii:opérateurs de calderon-zygmund, iii:(with r. coifman), opérateurs multilinéaires*. Academic Press.

Michelini A. 1995. An adaptative-grid formalism for traveltime tomography. *Geophysical Journal International* **121**, 489–510.

Mintzer F. 1985. Filter for distorsion-free multirate filter banks. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **33**, 626–630.

Monteiller V., Got J.-L., Virieux J. and Okubo P. 2005. An efficient algorithm for double-difference tomography and location in heterogeneous media, with an application to kilauea volcano. *Journal of Geophysical Research* **110**, B12306. doi:10.1029/2004JB003466.

Montelli R., Nolet G., Dahlen F.A., Masters G., Engdahl E.R. and Hung S. H. 2004. Finite-frequency tomography reveals a variety of plumes in the mantle. *Science* **303**, 338–343.

Nguyen T.Q. and Vaidyannathan P.P. 1989. Two-channnel perfect reconstruction fir qmf structures which yield linear phase analysis ans synthesis filters. *IEEE Transcations on Acoustics, Speech, and Signal Processing* **37**, 676–690.

Nolet G., Montelli R. and Vireux J. 1999. Explicit, approximate expressions for the resolution and *a posteriori* covariance of massive tomographic systems. *Geophysical Journal International* **138**, 36–44.

Operto S., Ravaut C., Improta L., Virieux J., Herrero A. and Dell'Aversana P. 2004. Quantitative imaging of complex structures from multi-fold wide aperture seismic data. *Geophysical Prospecting* **52**, 625–651.

Paige C.C. and Saunders M.A. 1982. ALGORITHM 583 LSQR: Sparse linear equations and least squares problems. *ACM Transactions on Mathematical Software* **8**, 195–209.

Podvin P. and Lecomte I. 1991. Finite difference computation of traveltimes in very contrasted velocity model: a massively parallel approach and its associated tools. *Geophysical Journal International* **105**, 271–284.

Pratt R.G., Shin C. and Hicks G.J. 1998. Gauss-newton and full newton methods in frequency-space seismic waveform inversion. *Geophysical Journal International* **133**, 341–362.

Spakman W. and Bijwaard H. 1998. Irregular cell parametrization of tomographic problems. *Annales Geophysicae* **16**, 28.

Spakman W. and Nolet G. 1988. Imaging algorithms, accuracy and resolution. In: *Mathematical Geophysics* (ed. N. Vlaar), pp. 155–187. D Reidel, Dordrecht.

Spakman W., Bijwaard H. and Engdahl E.R. 1998. Closing the gap between regional and global travel time tomography. *Journal of Geophysical Research* **103**, 30055–30078.

Sweldens W. 1994. The lifting scheme, a custom design construction of biorthogonal wavelets. Technical Report 1994:7, University of South Carolina.

Sweldens W. 1997. The lifting scheme, a construction of second generation wavelets. *SIAM Journal on Mathematical Analysis* **29**, 511–546.

Sweldens W. and Schroder P. 1995. Buidling your own wavelets at home. Technical Report 1995:5, University of South Carolina.

Tarantola A. 1987. *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*. Elsevier.

Tarantola A. and Valette B. 1982. Generalized nonlinear inverse problems solved using the least square criterion. *Reviews of Geophysics and Space Physics* **20**, 219–232.

Tikhonov A. and Arsenin V. 1977. *Solution of Ill-posed Problems*. Winston, Washington DC.

Trinks I., Singh S.C., Chapman C.H., Barton P.J., Bosch M. and Cherrett A. 2005. Adaptative traveltime tomography of densely sampled seismic data. *Geophysical Journal International* **160**, 925–938.

Vaidyannathan P.P. 1993. Theory an design of m-channel maximally decimated quadrature mirror filters with arbitrary m, having perfect reconstruction property. *SIAM Journal on Mathematical Analysis* **30**, 141–162.

Versteeg R. 1994. The marmousi experience: Velocity model determination on a synthetic complex data set. *The Leading Edge* **13**, 927–936.

Vesnaver A. 1994. Towards the uniqueness of tomographic inversion solutions. *Journal of Seismic Exploration* **3**, 323–334.

Vesnaver A. 1996. Irregular grids in seismic tomography and minimum time ray tracing. *Geophysical Journal International* **125**, 147–165.

Vetterli M. 1986. Filter banks allowing prefect reconstruction. *Signal Processing* **10**, 219–244.

Vetterli M. and Herley C. 1989. Wavelet and filter banks : theory and design. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **37**, 1057–1071.

Wesseling P. 2002. *An Introduction to Multigrid Methods*. John Wiley.

Woods J.W. and O'neil S.D. 1986. Subband coding of images. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **34**, 1278–1288.

Zelt C. and Smith R. B. 1992. Seismic traveltime inversion for 2-D crustal velocity structure. *Geophysical Journal International* **108**, 16–34.

Zhang H. and Thurber C. 2005. Adaptative mesh seismic tomography based on tetrahedral and voronoi diagrams: application to parkfield, california. *Journal of Geophysical Research* **110**, B04303. doi:10.1029/2004JB003186.

Zhang J., ten Brink U.S. and Toksöz M.N. 1998. Nonlinear refraction and reflection travel time tomography. *Journal of Geophysical Research* **103**, 29743–29757.

Zhou H. 2003. Multiscale traveltime tomography. *Geophysics* **68**, 1639–1649.