# Potential and limits of double-difference tomographic methods

Jean-Luc Got[1]*, Vadim Monteiller[1], Jean Virieux[2] and Stéphane Operto[3]

[1]*Laboratoire de Géophysique Interne et Tectonophysique, CNRS, Université de Savoie, 73376, le Bourget-du-Lac, France,* [2]*Laboratoire de Géophysique Interne et Tectonophysique, CNRS, Université Joseph Fourier, 38406, Saint-Martin d'Hères, France, and* [3]*Géosciences Azur, CNRS, Université de Nice – Sophia Antipolis, 06560, La Valbonne, France*

## ABSTRACT

Double-difference tomographic methods use directly accurate time delays computed between similar signals. Such methods are designed to image very heterogeneous media, such as volcanoes or fault zones. In seismological applications, similar signals are recorded at a given station from earthquakes sharing similar and close-by sources. In seismic exploration experiments, similar signals are often recorded at neighbouring receivers. After a brief presentation of the tomographic algorithm used, a seismological application is summarized. The potential and limits of double-difference tomographic methods are explored using various numerical experiments. They show that two effects are competing in double-difference tomography: (i) the degradation of the stability of the inversion due to the geometrical proximity of the rays used in the differentiation and (ii) the decrease in modelling error, which allows improving the stability of the inversion and using smaller quantities of *a priori* information when data are sufficiently accurate. The best resolution is obtained for an optimal value of the inter-source or inter-receiver distance. For optimal values of these distances and *a priori* information, tomography using traveltime differences provides significantly better resolved results than using traveltimes.

## INTRODUCTION

Improving our knowledge of the earth's interior often requires improving tomographic methods. Improvements may arise from instrumental, methodological or theoretical advances. In the last ten years, there has been a growing interest in Earth sciences for differential or interferometric methods, mostly because variations of some quantities are easier to estimate than their absolute values. This is especially true in seismology and signal processing where such methods are being more and more widely used. Differential studies have been especially applied to the location of earthquakes, knowing the relative locations of earthquakes being easier than knowing their absolute position. Computing the position of earthquakes directly from differential times is now termed 'double-difference loca-

tion'. This computation is efficient when time delays are computed from cross-correlation of similar signals, which can lead to a sub-sample accuracy. Waldhauser and Ellsworth (2000) proposed to linearize this location problem and set up a system of 'double-difference equations'. This is a linear system in which the Jacobian matrix is merely the difference of the matrices used for each individual location. One can show that the absolute position of the earthquakes may be computed from time differences in an exact velocity model, as long as the conditioning (estimated as the ratio between the major and the minor eigenvalue) of the system is sufficiently low. This is the case when at least one hypocenter is sufficiently far from the others (Monteiller *et al.* 2005): the ray path geometry allows not only the computation of the relative positions but also the computation of the barycenter of the earthquake set.

It was tempting to take advantage of the accuracy of cross-spectral time delays to infer not only the earthquake positions

---

*E-mail: jlgot@univ-savoie.fr

but also the seismic velocities in the surrounding propagation medium (Zhang and Thurber 2003; Monteiller *et al.* 2003). However, using such accurate measurements of the time delays in a tomographic inversion implied designing a coherent numerical scheme, whose accuracy is able to keep the information brought by the data. It required designing (i) an accurate solution for the travel-time computation, (ii) a stable and optimally regularized inversion, (iii) a way to optimize this regularization and (iv) an adapted processing of outliers. Monteiller *et al.* (2005) designed a double-difference tomographic algorithm fulfilling the above four requirements. In the following, (1) we will summarize the major features of this method, (2) we will present a seismological application and (3) we will explore the method using numerical experiments.

## METHOD

### Computation of the travel-times in a heterogeneous medium

Podvin and Lecomte (1991) provided a robust solution of the eikonal equation in the case of heterogeneous media. Its accuracy is however directly dependent on the cell size used to numerically represent the medium. To avoid extreme computational requirements, Monteiller *et al.* (2005) proposed to interpolate the traveltime field and to perform an *a posteriori* ray-tracing by computation of the traveltime gradient and path integration. Such a procedure allows improvement by more than one order of magnitude, the accuracy of the traveltime computation: it reaches the O $(10^{-5})$ relative accuracy needed to use time delays in double-difference tomography, at a low computational cost (Fig. 1).

### Regularizing efficiently the inverse problem

Double-difference tomography is a nonlinear (ill-posed) problem. It may therefore be posed as:

$$g(\mathbf{m}) = \mathbf{d} \tag{1}$$

where $g(\mathbf{m})$ represents the solution of the direct problem for the model parameter vector $\mathbf{m}$ and $\mathbf{d}$ is the traveltime data vector. In the case of an active seismic experiment, model parameters are seismic velocities, whereas in the seismological case model parameters comprise both hypocentral and velocity parameters.

In the hypothesis that both the data and model parameters follow a Gaussian distribution, the maximum of the *a posteriori* probability density function is given by the minimum of
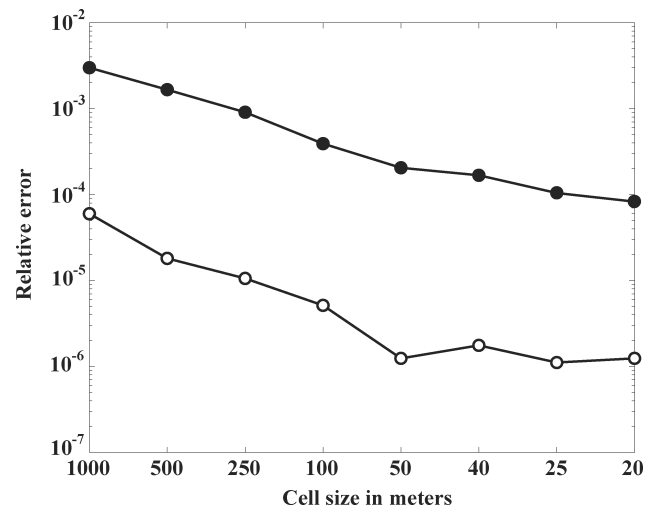


**Figure 1** Accuracy of *a posteriori* finite-difference traveltime recomputation as a function of the discretization of the finite-difference grid. Bold circles: error after Podvin-Lecomte finite-difference computation of travel-times; open circles: error after *a posteriori* ray-tracing in the traveltime field and recomputation of traveltimes.

the (cost) function (Tarantola and Valette 1982):

$$(g(\mathbf{m}) - \mathbf{d})^T \mathbf{C}_d^{-1}(g(\mathbf{m}) - \mathbf{d}) + (\mathbf{m} - \mathbf{m}_0)^T \mathbf{C}_m^{-1}(\mathbf{m} - \mathbf{m}_0) \tag{2}$$

where $\mathbf{C}_d$, $\mathbf{C}_m$ and $\mathbf{m}_0$ are respectively the data covariance matrix, the *a priori* model covariance matrix and the *a priori* model vector. Therefore, minimizing the cost function implies fitting the data (data misfit, left-hand term) with a minimum-norm model (penalty function, right-hand term). Minimizing equation (2) following a Gauss-Newton scheme leads to iteratively solve the system (Monteiller *et al.* 2005):

$$\begin{pmatrix} \mathbf{C}_d^{-1/2} \mathbf{G}_k \\ \mathbf{C}_m^{-1/2} \end{pmatrix} \delta \mathbf{m}_k = \begin{pmatrix} \mathbf{C}_d^{-1/2}(\mathbf{d} - g(\mathbf{m}_k)) \\ \mathbf{C}_m^{-1/2}(\mathbf{m}_0 - \mathbf{m}_k) \end{pmatrix} \tag{3}$$

where $\mathbf{G}_k$ is the traveltime derivative matrix at the iteration $k^{th}$.

Regularization of the solution of equation (3) is achieved through the use of an adapted quantity of *a priori* information, brought by the *a priori* covariance matrix. It leads to the computation of $\mathbf{C}_m^{-1/2}$. With the direct computation of this inverse on a large scale not being possible, coefficients of $\mathbf{C}_m^{-1/2}$ are first computed on a small scale using a Lanczos decomposition (Monteiller *et al.* 2005):

$$\mathbf{C}_m^{-1/2} = \mathbf{U}\mathbf{S}^{-1/2}\mathbf{U}^T \tag{4}$$

where $\mathbf{U}$ is an orthogonal matrix, $\mathbf{U}^T = \mathbf{U}^{-1}$ and $\mathbf{S}$ is a diagonal matrix containing the eigenvalues of $\mathbf{C}_m$.
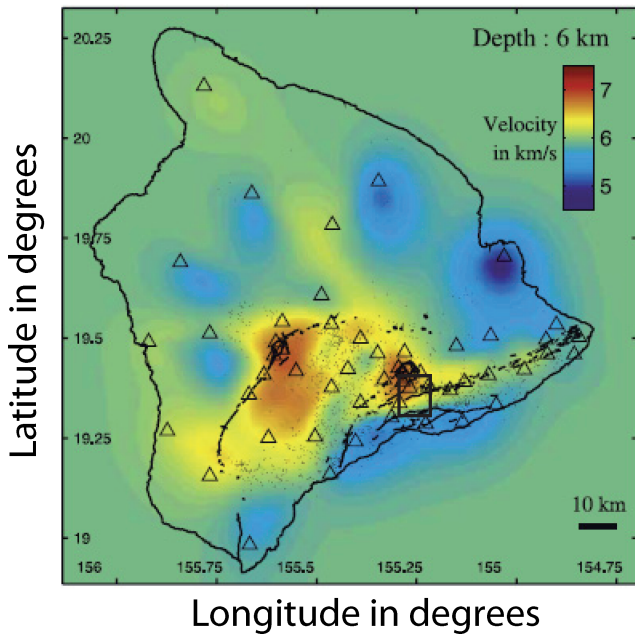
**Figure 2** Map of Hawaii Island showing the main topographic features, the epicenters (black dots) of the earthquakes used for the regional scale tomography (represented at 6 km depth) and the location of the HVO short-period seismic stations (open triangles). Colour scale represents P-wave velocity. The rectangular box limits the area in which double-difference tomography is performed (Figs 5–7a).

In the case where the correlation between nodes located at positions $\mathbf{x}$ and $\mathbf{x}'$ is represented by an exponential function:

$$\mathbf{C}_m(\mathbf{x}, \mathbf{x}') = \sigma_p^2 e^{\frac{-|\mathbf{x}-\mathbf{x}'|}{\lambda}} \tag{5}$$

where $\sigma_p^2$ is the variance of the velocity for the node located at position $\mathbf{x}$ and $\lambda$ is a correlation length-, $\mathbf{C}_m^{-1/2}$ is a band-diagonal matrix. The value of its coefficients depends only on $\sigma_p$ and $\lambda$ and does not depend on the number of model parameters, so that $\mathbf{C}_m^{-1/2}$ may be approximated accurately at a low computational expense. It is pre-computed and stored in files so that the probabilistic approach (equation (3)) may be used for large tomographic inversions. Such an approach allows the exploration of wide intervals for $\sigma_p$ and $\lambda$, and therefore the determination of the optimal quantity of *a priori* information needed to solve and regularize the solution.

We write the velocity variance into one scaled parameter

$$\sigma_\nu = \frac{\sigma_p}{\sigma_R} \tag{6}$$

where $\sigma_p$ is the *a priori* standard deviation of velocity parameters, in physical units and $\sigma_R^2 = \sigma_m^2 + \sigma_d^2$, where $\sigma_m$ is the modelling error in the theoretical differential times and $\sigma_d$

is the standard deviation of the data error distribution. $\sigma_\nu$ is therefore not expressed in physical units.

The data covariance matrix is directly related to the data probability density function (pdf). Arrival-time picks are known to exhibit long-tailed, non-Gaussian, distributions (see for example Got *et al.* 2006). Time delays, although they can be selected by using a coherency criterion, may also exhibit outliers. However, equation (2) relies upon the assumption of a Gaussian distribution of the data and the Gauss-Newton scheme for minimizing equation (2) leads to a least-square solution which is sensitive to the presence of strong outliers. Adequately using this probabilistic approach in the case of long-tailed distributions leads to the use of a more realistic pdf to represent the actual data distribution. We use the hyperbolic secant function:

$$f(x) = \frac{1}{\pi \sigma} \operatorname{sec} h\left(\frac{x - x_0}{\sigma}\right) \tag{7}$$

where $x$ is the independent variable; $f(x)$ represents a pdf with mean $x_0$ and standard deviation $\pi\sigma/2$, which behaves like the $L^1$ distribution for large values of $x$ and like a Gaussian distribution for small values. A change of variable (B. Valette, pers. comm.; see Monteiller *et al.* 2005, equations (24) to (26)) allows rewriting the inverse problem for the (long-tailed) data $x$ as a least-square inversion. This allows the robust inversion of travel-times or time delays (Monteiller *et al.* 2005; Got *et al.* 2006).

**Data processing**

The tomographic algorithm presented performs the traveltime inversion as well as the time-delay inversion. It is especially accurate when used with cross-spectral or cross-correlation time delays, given the accuracy of the traveltime delay computation and the stability of the inversion. The cross-spectral estimation of the time delay is performed by computing the slope of the phase of the cross-spectrum (see e.g., Jenkins and Watts 1968) using a weighted linear adjustment. The weight used is:

$$w(f) = \frac{C^2(f)}{1 - C^2(f)} \tag{8}$$
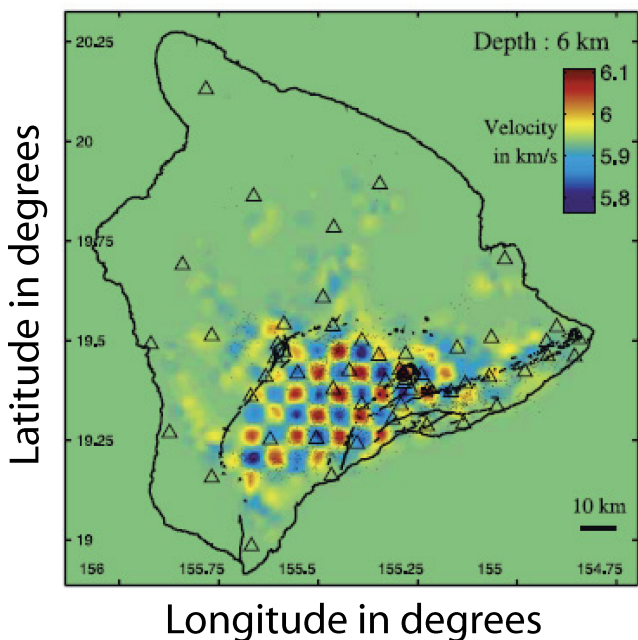
where $f$ is the frequency and $C(f)$ the coherency:

$$C(f) = \frac{\overline{\Gamma_{12}(f)}}{\sqrt{\overline{\Gamma_{11}(f)\Gamma_{22}(f)}}} \tag{9}$$

where $\Gamma_{12}(f)$, $\Gamma_{11}(f)$ and $\Gamma_{22}(f)$ are the cross-spectrum and the auto-spectra of the signals labelled 1 and 2, respectively; the bar denotes smoothing. The coherency spectrum measures
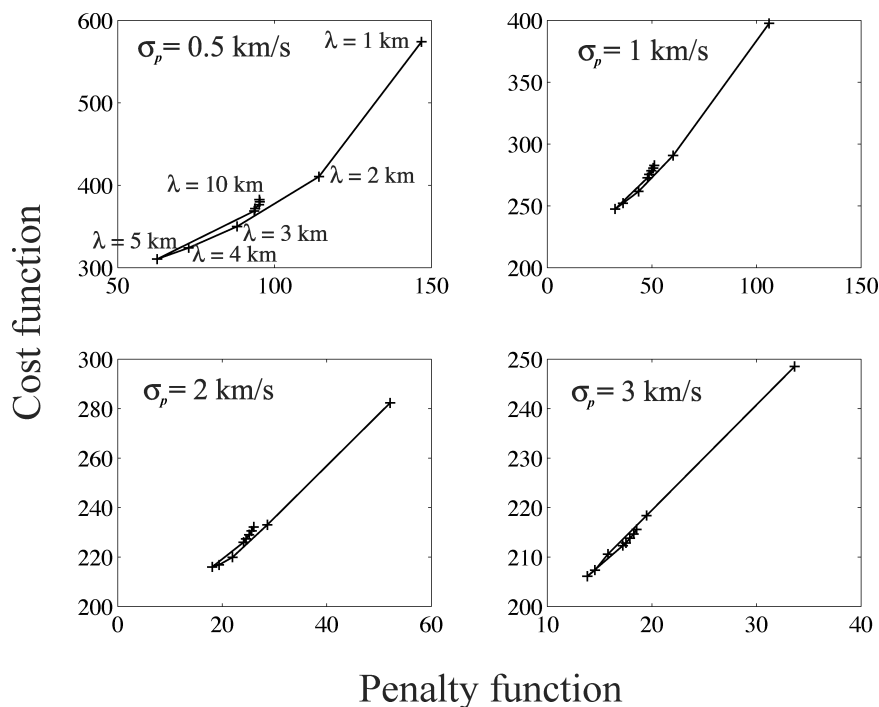
**Figure 3** Cost function as a function of the penalty function (*a posteriori* model variance) for various correlation lengths and *a priori* model variances (regional scale tomography). Cost and penalty functions (see equation (2)) are adimensional. Minimum values of these functions are used to choose the optimal correlation length and *a priori* model variance.



**Figure 4** Checkerboard test (regional scale), represented at 6 km depth.

the linearity of the filter (Wiener filter) used to relate signal 1 to signal 2.

**Application: Imaging the magmatic system of Kilauea volcano, Hawaii**

In the following we summarize an example of the application of our double-difference tomographic algorithm realized from seismological data.

**Geodynamic setting**

The aim of the double-difference tomographic study (Monteiller *et al.* 2005) was to image a part of the Kilauea volcano, Hawaii. The Kilauea volcano is a large basaltic shield volcano located in the southern part of Hawaii Island. It is built from the almost continuous inflow of basaltic magma, provided by the activity of a mantle plume beneath the Pacific plate. The volcanic edifice and more widely the Hawaii Island, are well-instrumented by the Hawaiian Volcano Observatory (HVO) seismic network, comprising 50 telemetered seismic stations. The magmatic activity induces a strong volcano-tectonic seismicity (several thousands of local magnitude
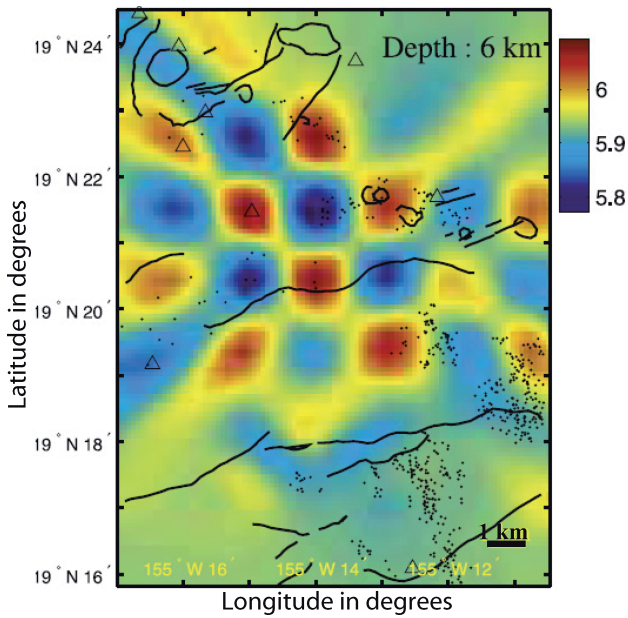
**Figure 5** Checkerboard test (double-difference tomography) realized in the south east of Kilauea caldera at 6 km depth. Colour scale indicates P-wave velocity in km/s.
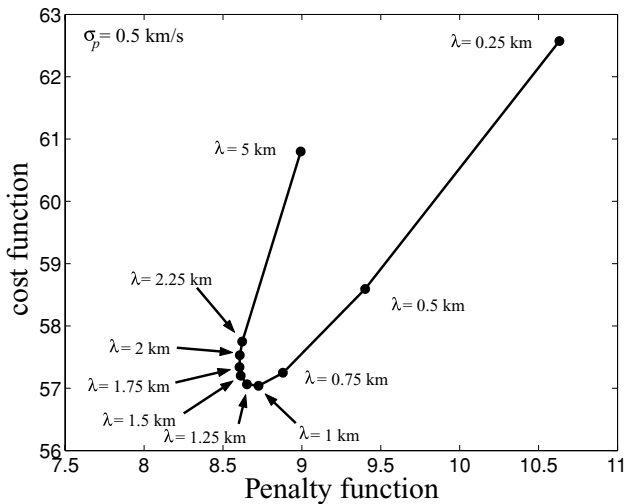


**Figure 6** Cost function represented as a function of the penalty function (*a posteriori* model variance) for various correlation lengths λ and $\sigma_p = 0.5$ km/s (double-difference tomography). Minimum values of these functions are used to choose the optimal correlation length and *a priori* model variance.

$M_l > 1$ earthquakes each year), whose most important part occurs beneath the mobile south flank of Kilauea. Numerous geodetical and seismological studies (see e.g., Swanson, Duffield and Fiske 1976; Got, Fréchet and Klein 1994; Got and Okubo 2003) have shown that most of

the south flank seismicity is due to the displacement of the south flank on the top of the oceanic crust, following a subhorizontal decollement plane. This strong microseismic activity generates numerous similar events, most of them being located along the decollement plane. These events have been well located by double-difference relocation (Got *et al.* 1994; Got and Okubo 2003). As the HVO seismic network is remarkably dense near the summit caldera (Fig. 2), the summit caldera and the East rift zone are quite well highlighted by these similar events.

**Data**

In this application, the double difference algorithm uses time delays and perturbates an initial model, chosen as a part of a regional model (scale of the island of Hawaii, characteristic dimension ∼100 km). In the first step, we computed the regional model from 41 886 high-quality P arrival times, picked from 1 358 declustered microearthquakes, recorded by at least 30 stations and selected among ∼30 000 earthquakes recorded from 1988 to 1999 by the HVO network. The model spanned a 160 km × 160 km × 36 km volume and comprised 959,077 inversion cells, each of 1 km³ volume. The use of our algorithm, applied in this first step to traveltime data only, allowed us to find a stable and accurate tomographic solution. We explored a wide range of *a priori* velocity variances and correlation lengths to find the optimal *a priori* information (Fig. 3). RMS (data misfit) drops from an initial value of ∼0.3 s for the catalog locations in the initial model to less than 0.08 s after the second iteration.

**Results of traveltime tomography**

The results (Fig. 2; Monteiller *et al.* 2005) allow imaging of most of the geological features of the island. Fast cores represent calderas and rift zones. They are well correlated with dense cores identified by gravimetric studies that are probably of intrusive origin. These cores are surrounded by a low-velocity cover, which is probably of extrusive origin. It is interesting to observe that small details are stable (e.g., the tip of the rifts) and kept where data are able to constrain them, even though the velocity parameters are widely correlated in the tomographic inversion. In the volumes where data are sparser and poorly constrain the model, no spurious fluctuation appears. Model resolution can not be computed directly, due to the size of the model. It is approximated using checkerboard tests (Fig. 4), which represents the volumes where the model can be reconstructed from the data.
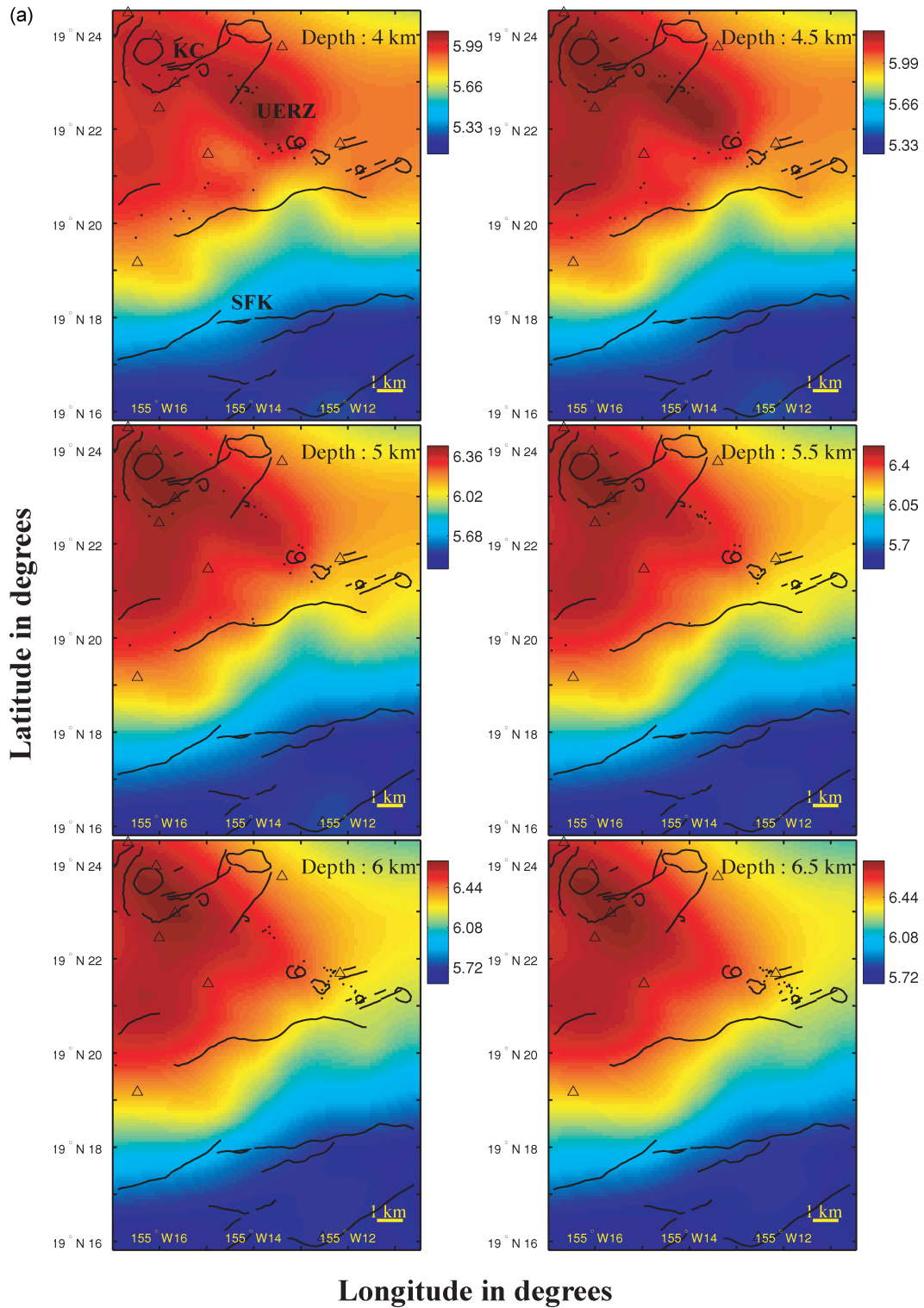
**Figure 7** (a) Maps of the P-wave seismic velocity in the south flank of Kilauea for various depths after double-difference tomography ($\lambda = 1$ km, $\sigma_p = 0.5$ km/s). Geographical labels are in the upper left subplot. KC: Kilauea caldera; UERZ: upper east rift zone; KSF: Kilauea south flank. Colour scale indicates P-wave velocity in km/s. (b) 3D representation of the high-velocity anomaly inferred from double-difference tomography in the southern caldera – upper east rift zone between 4 and 9 km depth. The volume is built from the 7% contour surfaces of the relative variation between the double-difference and the initial velocity model. Black dots represent earthquakes.
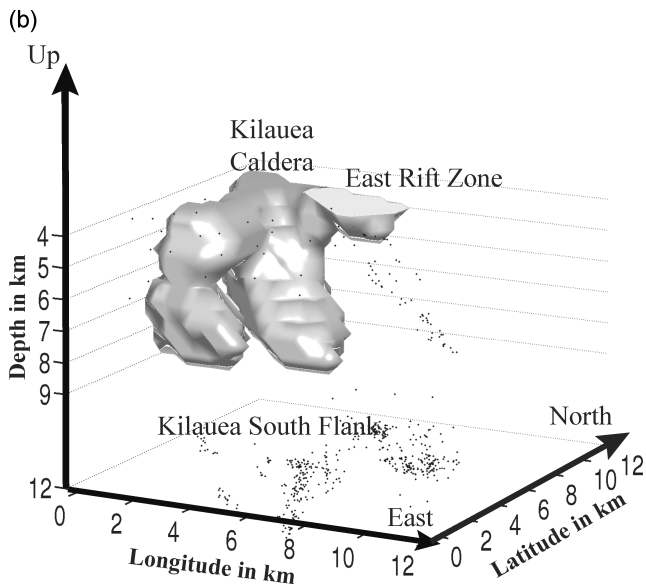
**Figure 7** Continued.
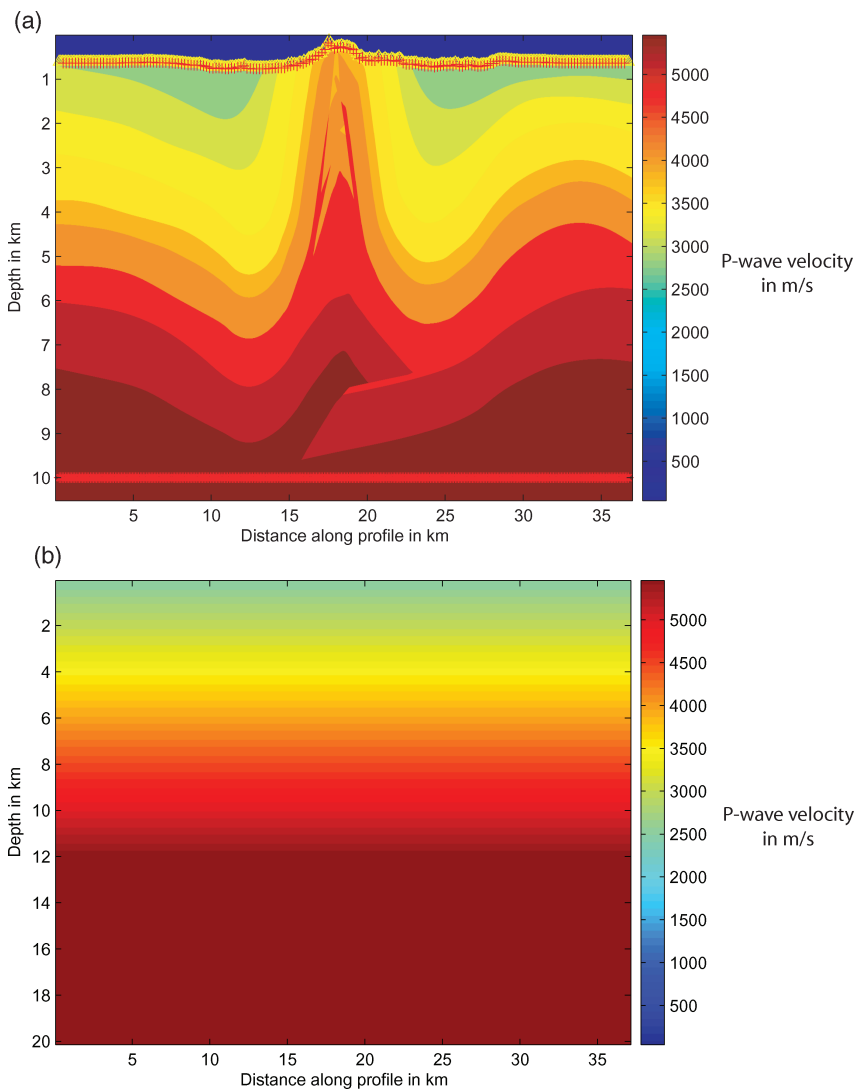
## Results of double-difference tomography

A double-difference tomographic model was built from time delays computed using a set of 614 well-recorded earthquakes in which neighbouring events are similar. Time delays were computed from a cross-spectral analysis performed on 2.56 s windows of signal (sample rate: 100 Hz). The selected event set comprised high-quality (coherency > 95%) short-distance correlations and all long-distance correlations with coherency greater than 80%. A selection of pairs on a simple coherency threshold would have mostly kept close-by events and therefore would have removed useful information from the initial data set. The double-difference tomographic model is a local model spanning a 12 km × 16 km × 12 km volume inside the regional model. This volume is sampled with a 500 m interval and comprises 20 625 inversion nodes. The initial model is the regional model in this volume. Checkerboard tests (Fig. 5) show that the model can be reconstructed from 3 km to 9 km depth below the east rift and southern caldera with a ~500 m spatial resolution. The results (Monteiller *et al.* 2005) show that time delays efficiently modified the initial velocity model in a stable and coherent manner. As in the traveltime case, a wide interval of velocity variances and correlation lengths was spanned and allowed finding the optimal *a priori* information to reach a stable solution (Fig. 6). RMS drops down to 0.01 s. The resulting image (Fig. 7) shows a complex but coherent subvertical high-velocity body located south of the caldera, which connects to the east rift zones near 4 to 5 km depth. This body may be interpreted as magma solidified at depth at the time of each intrusion/eruption around the active magma supply system. Despite the huge quantity of data such details were not revealed by using traveltimes with this event set or with larger ones (Monteiller *et al.* 2005). This application, which is the most accomplished one with this tomographic method, was allowed by the quantity and quality of available data. It illustrates the potential of this method.

## Potential and limits of double-difference tomography for active and passive geometries

Earthquakes and seismological networks do not necessarily allow imaging of areas that are of interest from the point of view of volcanic or seismic risk. Large magma storage areas, in which body wave velocities may vary by a few percent may remain undetected, especially when they are in a low-velocity zone where seismic rays tend to be sparser than in surrounding volumes. Magma storage areas may be composed of high- and low-velocity structures varying at small scales, making the detection of the low-velocity part very sensitive to the wavelength recorded. A possible solution is to investigate these areas using high-resolution active seismic experiments.

In order to investigate the potential and limits of double-difference tomography to image heterogeneous media with first-arrival data from seismological or seismic exploration experiments, we realized a series of numerical experiments. We used a synthetic velocity model (Fig. 8a) to compute theoretical arrival times using our solution (Podvin-Lecomte finite-difference computation of traveltime and *a posteriori* ray tracing and recomputation of traveltimes) of the eikonal equation. The model was discretized with 10 m wide cells, which leads to a relative accuracy in traveltimes of $10^{-5}$–$10^{-6}$ (Monteiller *et al.* 2005). This model contains spatial high-frequencies, as well as a thin (10 to 50 m thick) and very low (300 m/s) velocity zone below the surface, featuring a weathered zone. It was built from geological studies of structurally complex regions and represents a generic (eventually volcanic) orogene. Numerical experiments are performed using 2D seismic profiles. In order to investigate the effect of various source-receiver geometry, we designed two different geometries (Fig. 8a): the first one in which 181 stations spanning (200 m interval) a 37 km-long profile record at the topographic surface, the signal generated by 366 earthquakes (100 m interval) at depth (a seismological geometry) and a second one in which 366 stations (100 m interval) record 181 shots (200 m interval), both sources and stations being located at the topographic surface (a seismic exploration geometry).

(a)



(b)

We first performed first-arrival traveltime tomography and investigated large intervals for *a priori* velocity standard deviation $\sigma_v$ and correlation length $\lambda$. Sampling was made with 100 m wide cells. An initial model (Fig. 8b) was chosen to be a constant gradient model, increasing linearly from 2 000 m/s at the top of the model to 5 500 m/s at 10 km depth.

In order to quantify the quality of the results, we compute for each result:

$$R = 1 - \sqrt{\frac{\sum\limits_{n}(\hat{m} - m_t)^2}{\sum\limits_{n} m_t^2}} \qquad (10)$$

or

$$R = 1 - \frac{\sigma_r}{\sigma_t} \qquad (11)$$

where $\hat{m}$ is the estimated tomographic model, $m_t$ is the true model, $n$ is the number of cells in the model, $\sigma_r^2 = \frac{\sum_n(\hat{m}-m_t)^2}{n}$ is the model residual variance and $\sigma_t = \sqrt{\frac{\sum_n m_t^2}{n}}$. Notice that we do not remove the means in equation (10) as the mean has to be estimated during the inversion. $R$ may be thought of as an estimator of the model resolution; we will call it the pseudo-resolution. Pseudo-resolution for our initial gradient model gives $R = 85.7\%$. A $\sim$30 m/s homogeneous variation of the velocity throughout the whole model leads to a 1% variation in pseudo-resolution.

Results of the traveltime tomography (Fig. 9) show that RMS drops with increasing *a priori* standard deviation of velocity parameters, whereas $R$ tends to exhibit a maximum when standard deviation equals 350 to 500 m/s. Both quantities are almost insensitive to the variation of the correlation

**Figure 9** Pseudo-resolution *R* (in %, left axis, solid line) of the traveltime tomographic model and data misfit (RMS in s, right axis, dashed line) as a function of the *a priori* standard deviation of the velocity parameters $\sigma_p$: (a) seismological geometry; (b) seismic exploration geometry. Both results were computed for a correlation length $\lambda = 1$ km.
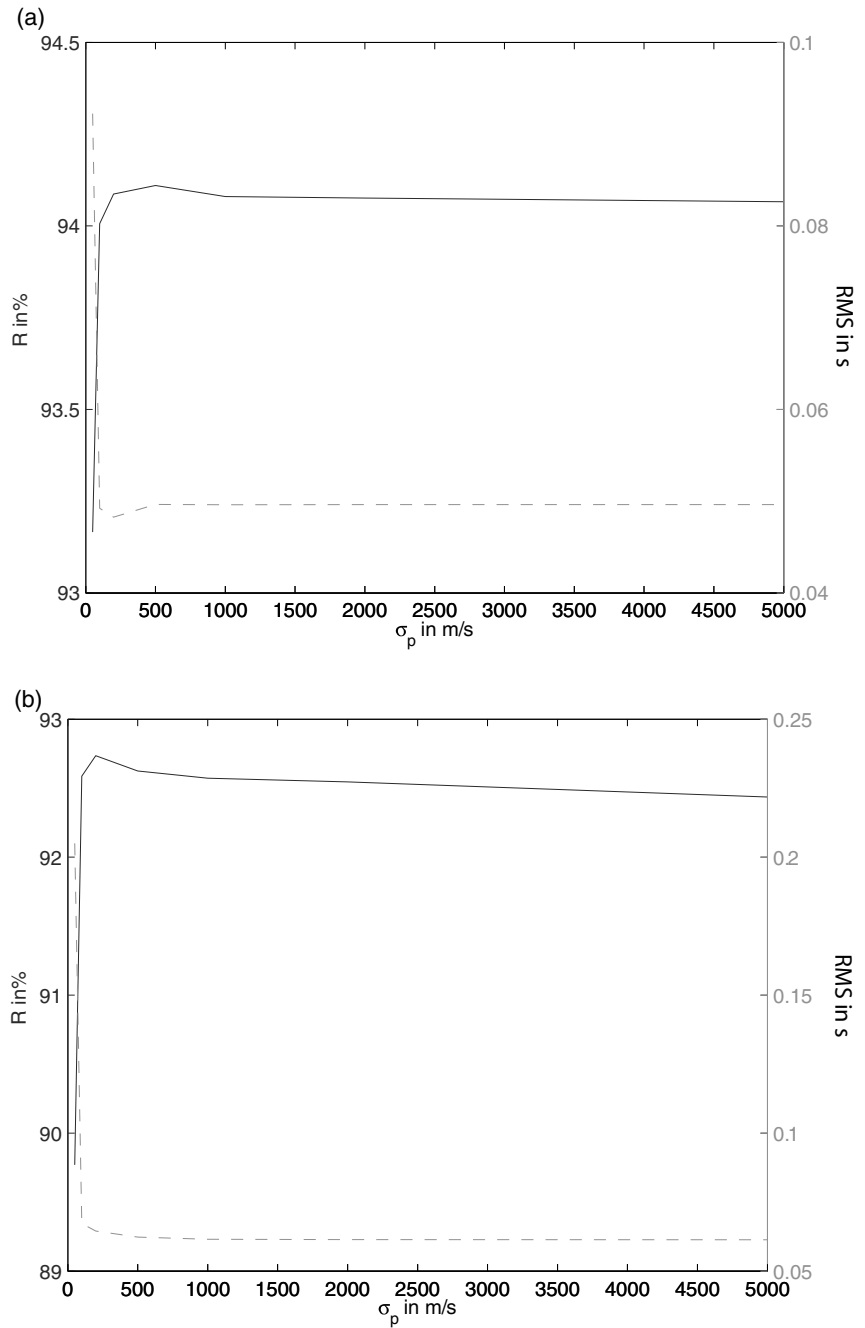


length, in the (100 m, 2000 m) interval. It shows that the model is well-constrained by the data, in most of its area. Values computed for *R* (Fig. 9) and careful inspection of the tomographic results (Fig. 10) show that the model reconstruction tends to be better in the case of the seismological geometry than in the case of superficial seismic exploration geometry, especially at intermediate depths. In the latter geometry, rays tend to be parallel at depth, which increases the condition number and diminishes the resolution.

Double-difference tomography was performed by using differences of theoretical traveltimes as data. In this case, the accuracy of the differential data is of the same order of magnitude as the accuracy of the traveltime data: in the hypothesis the traveltime error distribution is Gaussian, the standard
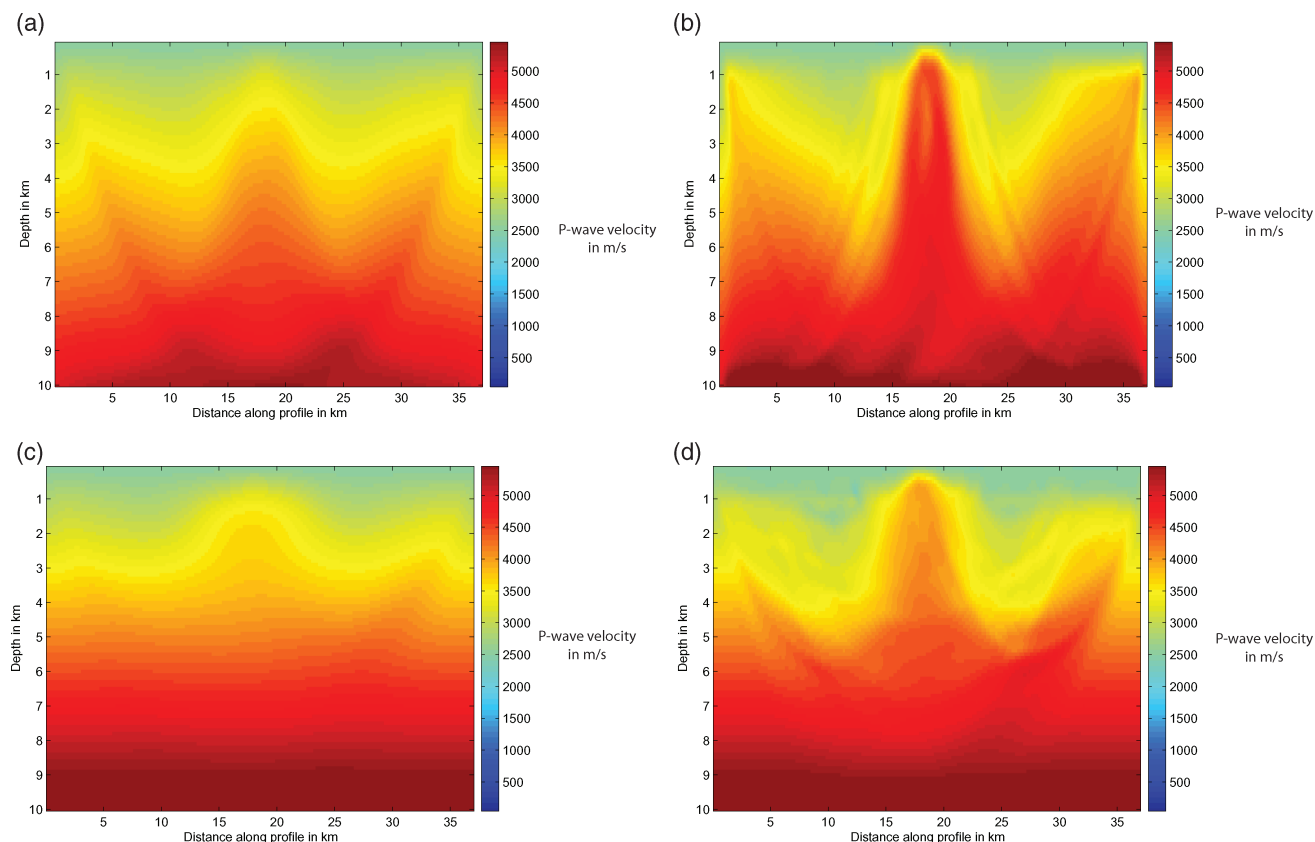
**Figure 10** P-wave velocity model inferred from traveltime tomography: (a) seismological geometry, $\sigma_p = 50$ m/s, $\lambda = 1$ km. (b) Seismological geometry, $\sigma_p = 500$ m/s, $\lambda = 1$ km. (c) Seismic exploration geometry, $\sigma_p = 50$ m/s, $\lambda = 1$ km. (d) Seismic exploration geometry, $\sigma_p = 200$ m/s, $\lambda = 1$ km. Colour scale indicates the P-wave seismic velocity in m/s.

deviation of differential data is $\sqrt{2}$ times that of the traveltime data. In the case where the locations of both the sources and the receivers are considered as perfectly known, no improvement in the knowledge of the medium may be expected from the improved knowledge of the source locations. Furthermore, in the case where differences of traveltimes are used, the double-difference differential operator introduces equations that are not linearly independent, which degrades conditioning and does not increase the quantity of information. Therefore it would be rather unexpected that, in this case, the use of double-difference tomography would improve the tomographic result. Some uncertainty in the behaviour of the algorithm arises from the differentiation of the derivatives and the introduction of an optimal *a priori* information to regularize the double-difference system.

We therefore performed a series of double-difference tomographies for various velocity standard deviations, correlation lengths and inter-event (seismological case) or inter-receiver (seismic exploration case) distances. Results

(Figs 11–12) show the variation of the RMS and pseudo-resolution $R$ as a function of $\sigma_v$ and $\lambda$ and the corresponding tomographic results. In the superficial seismic exploration case, the inversion is unstable for $\sigma_v$ greater than 250 m/s, which shows that the conditioning of the system for this source-receiver geometry may be poor and can not lead to high resolution. As a result, the pseudo-resolution $R$ for the seismic exploration geometry is found to be lower than for the seismological geometry.

The pseudo-resolution $R$ exhibits significant variations as a function of $\sigma_v$, $\lambda$ and inter-event or inter-receiver distance $\delta$. Results (Figs 11–12) show that $R$ strongly drops when $\delta$ is less than 500 m. This result clearly indicates that data from pairs having larger distances $\delta$ contribute far more to the model reconstruction than data corresponding to smaller $\delta$. This may limit the interest of the double-difference tomography in the case where the only possible (correlated) pairs have small distances $\delta$; it overall underlines the need of computing relatively accurate time delays for distant (500 m to 1 km) pairs.
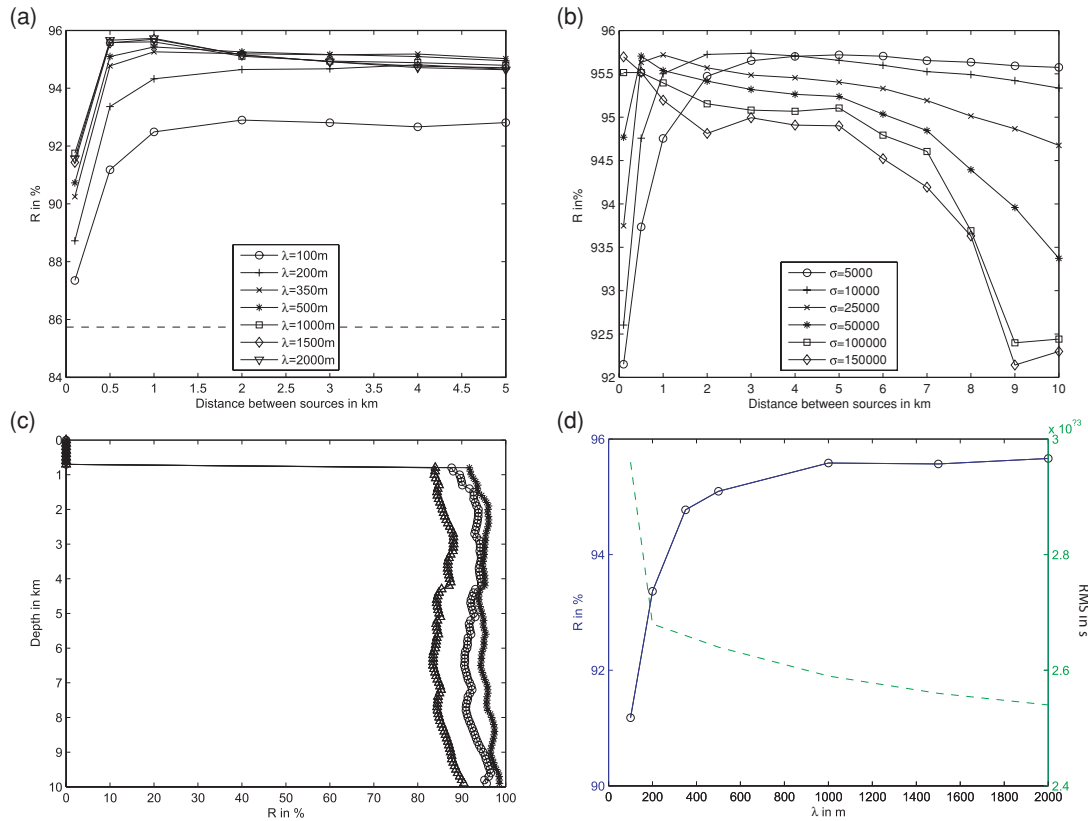
**Figure 11** Seismological geometry: (a) pseudo-resolution $R$ of double-difference tomographic models as a function of the inter-source distance $\delta$ for various correlation lengths $\lambda$ in the seismological geometry. A maximum is reached in the pseudo-resolution for a critical correlation length $\lambda \sim 1$ km and an optimal inter-source distance $\delta \sim 500$ m. The pseudo-resolution strongly drops below this distance threshold. Dashed line figurates the pseudo-resolution of the initial model. (b) Pseudo-resolution $R$ as a function of the inter-source distance $\delta$ for various *a priori* standard deviations in velocity parameters $\sigma_v$ in the seismological geometry. A maximum is reached in the pseudo-resolution for an optimal $\sigma_v$, diminishing with the inter-source distance $\delta$. At high $\sigma_v$, pseudo-resolution is very sensitive to inter-source distance $\delta$. High pseudo-resolution at short inter-source distance and high $\sigma_v$ is induced by the very accurate traveltime differences. (c) Pseudo-resolution as a function of depth for traveltime tomography (open circles; $\sigma_p \sim 500$ m/s, $\lambda = 1$ km) and double-difference tomography (stars; $\delta = 500$ m, $\lambda = 1$ km, $\sigma_p \sim 500$ m/s). Pseudo-resolution of the initial model (open triangles) is given as a reference. (d) Pseudo-resolution (in%, left axis, solid line) and RMS (in s, right axis, dashed line) as a function of $\lambda$ for double-difference tomography ($\delta = 500$ m, $\sigma_p \sim 500$ m/s). (e) Pseudo-resolution (in%, left axis, solid line) and RMS (in s, right axis, dashed line) as a function of $\sigma_p$ for double-difference tomography ($\delta = 500$ m, $\lambda = 1$ km). (f) Double-difference tomography result with optimal parameters ($\delta = 500$ m, $\lambda = 1$ km, $\sigma_p \sim 500$ m/s). Colour scale indicates the P-wave seismic velocity in m/s. (g) Double-difference tomography result with $\delta = 100$ m, $\lambda = 100$ m, $\sigma_p \sim 500$ m/s. Colour scale indicates the P-wave seismic velocity in m/s.

Figures 11–12 also show that pseudo-resolution $R$ increases with correlation length. It allows to show that the probabilistic solution brought by the algorithm given by equation (3) using the covariance kernel (equation 5) is essentially different of the one given by a simple averaging of adjacent cells. When the inverse problem is not strongly nonlinear – which is the case when the propagation medium and its model are relatively smooth – the solution may be understood as the stack of a set of models, each of them being a solution of the inversion. These models have a deterministic and a random part: their stack enhances the common deterministic part

and tends to annihilate the random part. Therefore increasing the correlation length tends, to some extent, to improve the pseudo-resolution. Figures 11(a) and 12(a) show that there is a critical correlation length above which no improvement in pseudo-resolution is brought. An interesting feature is that a statistical mode appears for the largest correlation lengths, for an optimal distance $\delta$. This mode may be related to a characteristic dimension of the model.

Comparison of the pseudo-resolution $R$ computed from respectively the traveltime and double-difference tomography results shows that $R$ may be significantly larger in the
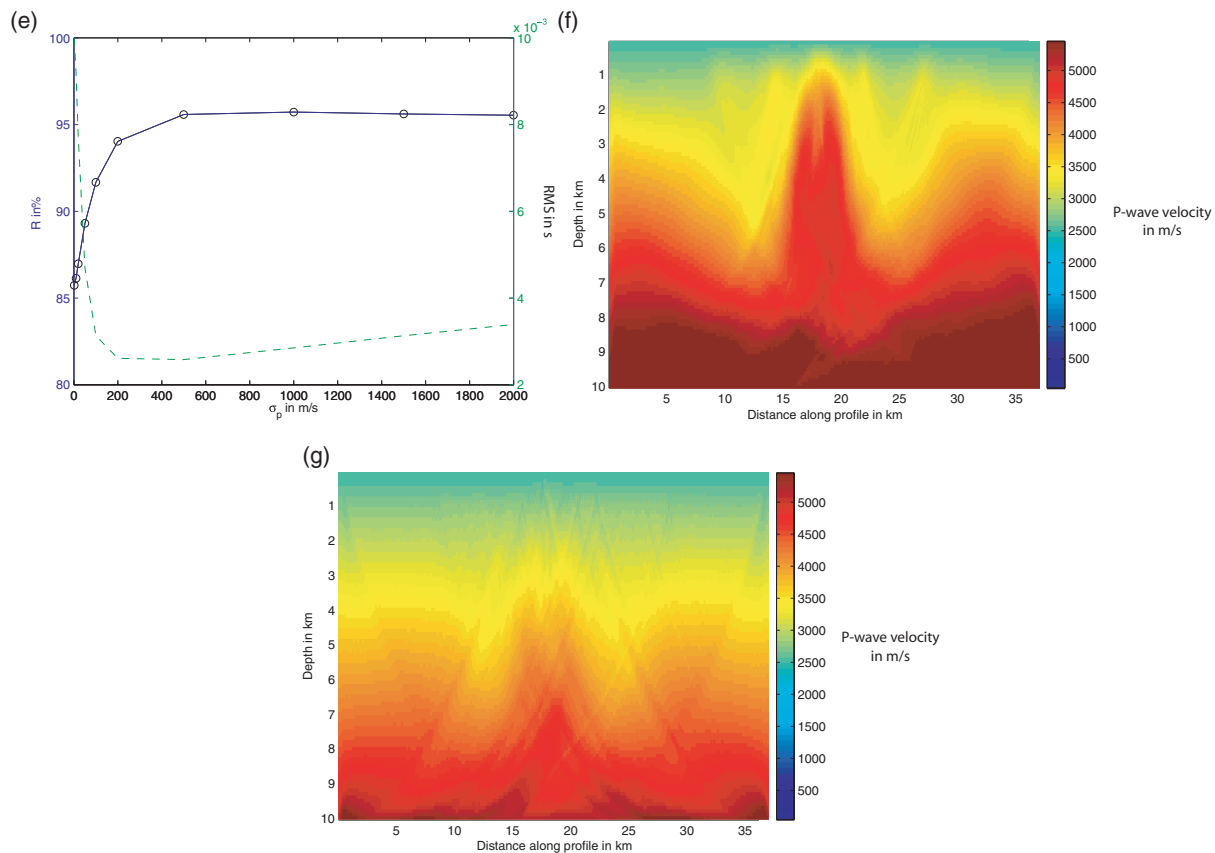
**Figure 11** Continued.

double-difference case, especially when the critical correlation length is reached at the optimal distance $\delta$. About 9% information is brought to the initial model by the inversion of differential times whereas only 7.5% is brought by the traveltime inversion. This is therefore a $\sim$15% increase in the information brought by the differential data, relative to the traveltime data. This increase in pseudo-resolution may be linked to the fact that the *a priori* $\sigma_v$ (equations 5–6) used in double-difference tomography are larger than those used in traveltime tomography. In our (synthetic) case, $\sigma_d$ (equation 6) is of the same order of magnitude for the theoretical traveltimes and their differences. Therefore the one-order of magnitude change experimentally found between the values of $\sigma_v$ used in traveltime and double-difference tomography arises from the change in modelling error $\sigma_m$: the modelling error in the difference between theoretical traveltimes is about ten times smaller than the modelling error in theoretical traveltimes.

However, differences computed from short-distance correlation can not resolve large wavelengths (Figs 11–12). Numerical instabilities are brought by the use of very close ray paths, which leads to almost singular, poorly conditioned, derivative

matrices $\mathbf{G}_k$ – whose inversion has to be carefully regularized. An optimum has therefore to be found by exploring large intervals in inter-source or inter-station distance $\delta$, correlation length $\lambda$ and *a priori* velocity standard deviation $\sigma$. Notice that the variation in $\delta$ induces a variation in modelling error and therefore (equation 6) a variation in $\sigma_p$ for a given $\sigma_v$: exploring large intervals of $\delta$ with various *a priori* velocity standard deviations implies to represent $\sigma_v$ rather than $\sigma_p$ (Figs 11b–12b). $\sigma_R$ is of the order of 50 ms in traveltime tomography and 5 ms for double-difference tomography when $\delta \sim 100$–500 m.

Computation of the pseudo-resolution is impossible as far as the true model is unknown. Comparison of the data misfit (RMS) and the pseudo-resolution as a function of the *a priori* standard deviation $\sigma_v$ or the correlation length $\lambda$ shows that both RMS and $R$ exhibit inverse L-curves (Figs 11d–e and 12d–e). Exploring the RMS or cost function as a function of $\sigma_v$ and $\lambda$ therefore allows finding the optimum *a priori* information for a given distance $\delta$. However, both RMS and cost function are found to increase with $\delta$; the low value of RMS for low $\delta$ is therefore not significant and cannot be used as a criterion to select data. As we show that the pseudo-resolution
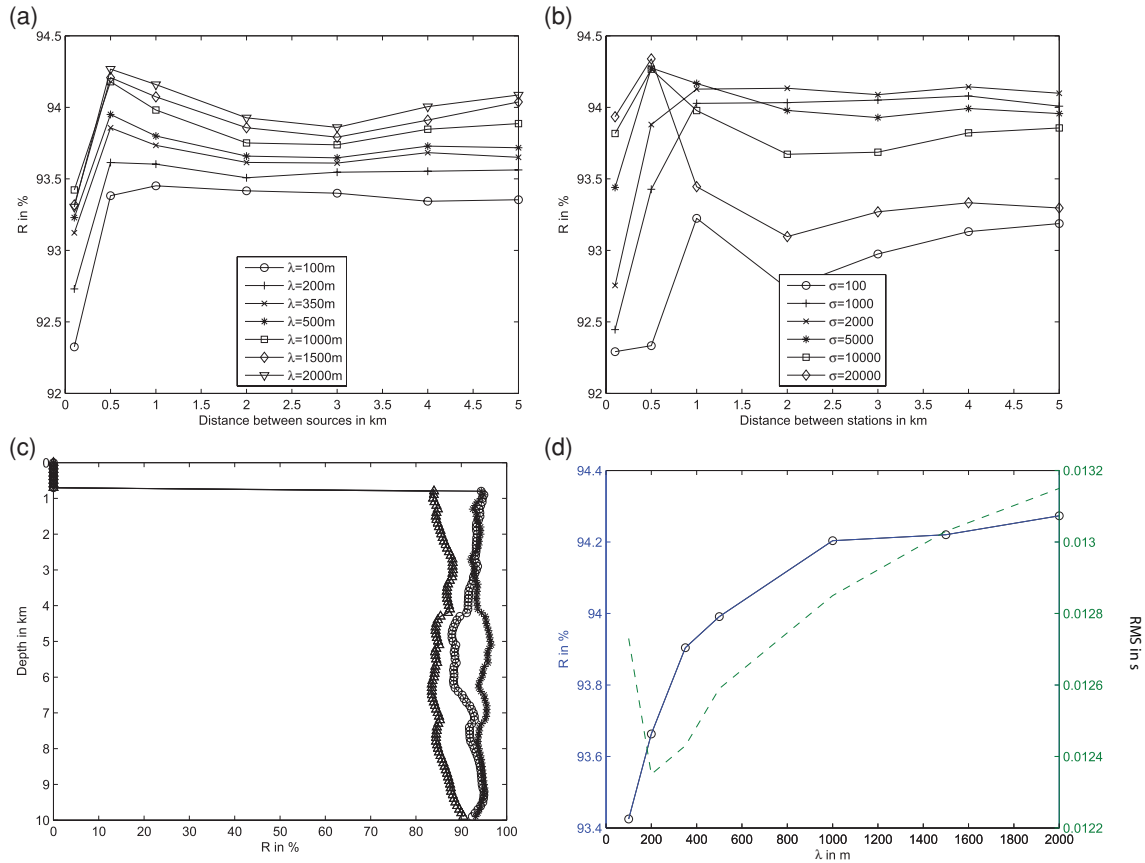
**Figure 12** Seismic exploration geometry: (a) pseudo-resolution *R* of double-difference tomographic models as a function of the inter-source distance $\delta$ for various correlation lengths $\lambda$ in the seismic exploration geometry. A maximum is reached in the pseudo-resolution from a critical correlation length $\lambda \sim 1$ km and an optimal inter-source distance $\delta \sim 500$ m. The pseudo-resolution strongly drops below this distance threshold. (b) Pseudo-resolution R as a function of the inter-source distance $\delta$ for various *a priori* standard deviations in velocity parameters $\sigma_v$ in the seismic exploration geometry. As in the seismological case (Fig. 11b), a maximum is reached in the pseudo-resolution for an optimal $\sigma_v$, diminishing with the inter-source distance $\delta$. (c) Pseudo-resolution as a function of depth for traveltime tomography (open circles; $\sigma_p \sim 500$ m/s, $\lambda = 1$ km) and double-difference tomography (stars; $\delta = 500$ m, $\lambda = 1$ km, $\sigma_p \sim 500$ m/s) for the seismic exploration geometry. Pseudo-resolution of the initial model (open triangles) is given as a reference. (d) Pseudo-resolution (in%, left axis, solid line) and RMS (in s, right axis, dashed line) as a function of $\lambda$ for double-difference tomography ($\delta = 500$ m, $\sigma_p \sim 500$ m/s). Notice the very short variation range of RMS with $\lambda$. (e) Pseudo-resolution (in%, left axis, solid line) and RMS (in s, right axis, dashed line) as a function of $\sigma_p$ for double-difference tomography ($\delta = 500$ m, $\lambda = 1$ km). Computation does not converge correctly for $\sigma_p$ greater than 200 m/s. (f) Double-difference tomography result with optimal parameters ($\delta = 500$ m, $\lambda = 1$ km, $\sigma_p \sim 200$ m/s). Colour scale indicates the P-wave seismic velocity in m/s. (g) Double-difference tomography result with $\delta = 100$ m, $\lambda = 100$ m, $\sigma_p \sim 500$ m/s. Colour scale indicates the P-wave seismic velocity in m/s.

*R* strongly drops for small values of $\delta$, it will be preferable to select data with a criterion bearing on the relative error in traveltime difference rather than on the error itself, given that this error increases with $\delta$. Large $\delta$ for which the relative error is sufficiently low should be preferred in a selection.

Differences may be used to focus the resolution in some delimited regions, as far as the source-receiver geometry may allow it. As an example, differences between signals emitted by different sources at depth and recorded by the same receiver at the surface allow removal of the modelling error introduced

by the mis-modelled weathered zone. Consequently, superficial structures are poorly- and deep structures more correctly-resolved.

## CONCLUSION

As a conclusion, our numerical experiments show that in double-difference tomography, degradation of the conditioning due to the geometrical proximity of the rays used in the differentiation competes with the decrease in modelling error,
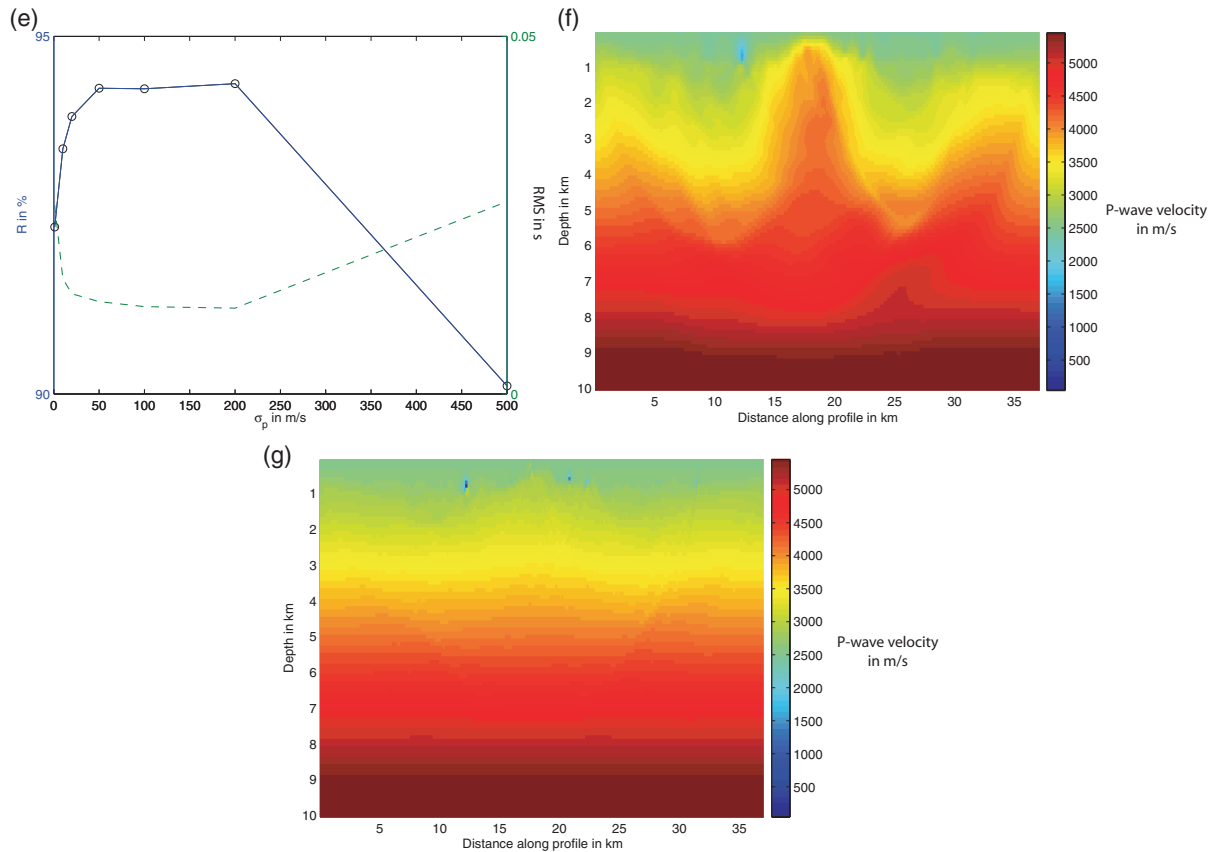
**Figure 12** Continued.

which allows, if the data are sufficiently accurate, improving the conditioning and using smaller quantities of *a priori* information. We show that the best resolution is obtained for an optimal value (or range) of the inter-source or inter-station distance, rather than for the smallest distance. This distance controls the optimal quantity of *a priori* information to be brought to regularize the inversion. In favourable geometric conditions and for optimal parameters, resolution of the double-difference tomographic method may be better than those of the conventional traveltime tomography. A well-regularized traveltime and double-difference tomographic method may be used in seismic exploration to find initial pre-stack migration velocity models in heterogeneous media.

## REFERENCES

Got J.-L., Fréchet J. and Klein F.W. 1994. Deep fault plane geometry inferred from multiplet relative relocation beneath the south flank of Kilauea. *Journal of Geophysical Research* **99**, 15375–15386.

Got J.-L., Monteiller V., Virieux J. and Okubo P. 2006. Estimating crustal heterogeneity from double-difference tomography. *Pure and Applied Geophysics* **163**, 405–430. doi:10.1007/s00024–005–0022-x.

Got J.-L. and Okubo P.G. 2003. New insights into Kilauea's volcano dynamics brought by large-scale relative relocation of microearthquakes. *Journal of Geophysical Research* **108**, 2337. doi:10.1029/2002JB002060.

Jenkins G.M. and Watts D.G. 1968. *Spectral Analysis and Its Applications*. Holden-Day, Oakland, CA.

Monteiller V., Got J.-L., Virieux J. and Okubo P.G. 2003. Double-difference tomography of Kilauea Volcano, Hawaii. *Proceedings of the EGS-AGU-EUG Joint Assembly*, Nice, France, April 2003.

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

Podvin P. and Lecomte I. 1991. Finite difference computation of traveltimes in very contrasted velocity models: A massively parallel approach and its associated tools. *Geophysical Journal International* **105**, 271–295.

Swanson D.A., Duffield W.A. and Fiske R.S. 1976. Displacement of the south flank of Kilauea volcano: The result of forceful intrusion

of magma into the rift zones. *U.S. Geological Survey Professional Paper* **963**, 1–39.

Tarantola A. 1987. *Inverse Problem Theory.* Elsevier.

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

Waldhauser F. and Ellsworth W.L. 2000. A double difference earth-quake location algorithm: Method and application to the northern Hayward fault, CA. *Bulletin of the Seismological Society of America* **90**, 1353–1368.

Zhang H. and Thurber C.H. 2003. Double difference tomography: The method and its application to the Hayward fault, California. *Bulletin of the Seismological Society of America* **93**, 1875–1889.