Joint inversion of P-wave velocity and density, application to La Soufrière of Guadeloupe hydrothermal system

Olivier Coutant, M.L. Bernard, François Beauducel, Florence Nicollin, M.-P. Bouin, S. Roussel

To cite this version:

Joint inversion of $P$-wave velocity and density, application to La Soufri`ere of Guadeloupe hydrothermal system

O. Coutant,1 M. L. Bernard,2 F. Beauducel,3 F. Nicollin,4 M. P. Bouin,3 and S. Roussel1

1ISTerre, CNRS, Universit´e Joseph Fourier, Grenoble, France. E-mail: olivier.coutant@ujf-grenoble.fr
2Universit´e Antilles-Guyanne, Pointe-`a-Pitre, Guadeloupe, France
3CNRS, Institut de Physique du Globe, Paris, France
4G´eosciences Rennes, CNRS, Universit´e de Rennes, France

SUMMARY
We present the result of a 3-D gravity and $P$-wave traveltime joint inversion applied to the hydrothermal system of La Soufri`ere of Guadeloupe. The joint inversion process is used here to overcome the different resolution limitations attached to the two data sets. $P$-wave traveltimes were obtained from three active seismic surveys that were conducted from 2001 to 2007. Gravity data collected during a microgravity campaign is described in a companion paper. We use a joint inversion process based on a Bayesian formulation and a deterministic iterative approach. The coupling between slowness and density is introduced through a supplementary constraint in the misfit function that tries to minimize the distance between parameter values and a theoretical relationship. This relationship is derived from measurements on samples representative of Mt Pel ´ee of Martinique and La Soufri`ere volcanoes. We chose a grid discretization that leads to an under-determined problem that we regularize using spatial exponential covariance between the nodes parameters. Our results are compared to geophysical electromagnetic results obtained using resistivity and VLF surveys. They confirm the presence of highly contrasted dense/fast and light/slow zones in La Soufri`ere dome and crater basement. Our images suggest however that some non-conductive zones may be massive andesite bodies rather than argilized zones, and that these bodies may have deeper roots than hypothesized.

Key words: Inverse theory; Tomography; Volcanic hazards and risks.

1 INTRODUCTION
La Soufri`ere of Guadeloupe is one of the active volcanoes of Lesser Antilles, located 60 km south of Soufriere Hills (Montserrat) and north of Dominica and Martinique islands. La Soufri`ere is the most recent part of the Grande D´ecouverte-Soufri`ere volcanic complex which has been the only active volcano in Guadeloupe in the last 10 000 yr. Its last magmatic eruption occurred 1530 A.D. (Boudon et al. 2008) and culminated, after a complex eruption, by the formation of the current lava dome. During historical period, all hydrothermal activity and several phreatic explosive eruptions have taken place from fractures and vents on and close to this dome (Komorowski et al. 2005). The last phreatic eruption (1976–1977) was one of the most important and led to preventive evacuation of 72 000 inhabitants for several months. In fact, this significant seismo-volcanic unrest was interpreted as a still-born magmatic eruption (Feuillard et al. 1983; Villemant et al. 2005; Boichu et al. 2011). Furthermore, numerous studies (see e.g. Boudon et al. 2007) on land and offshore have demonstrated that one of the major risk on the Lesser Antilles volcanoes which are under the influence of heavy tropical rain and hydrothermal activity that mechanically weakens edifice rocks, is partial or complete flank collapse.

Since 1992, La Soufri`ere volcano exhibits a moderate but progressively increasing activity (shallow seismicity with long period events, acidic fumaroles with high velocity flux, hot springs temperature anomaly) continuously monitored by the Volcanological and Seismological Observatory (OVSG) operated by Institut de Physique du Globe de Paris. This long-term and misunderstood unrest has motivated, about 10 yr ago, a large community to concentrate their research efforts on the superficial hydrothermal system characterization. Various geophysical studies have been carried out to improve our understanding of the hydrothermal dynamics and inner geological structures (see for instance Jacob et al. (2005), Nicollin et al. (2006), Bernard et al. (2006)). But seismic methods that are classically used on volcanoes have not been applied at La Soufri`ere since Dorel et al. (1979), mostly because of the low rate of natural seismicity and the difficulty to operate dense seismic networks under tropical environment.
The combination of different geophysical imaging methods and its implementation as joint or combined inversions is a field of growing methodological interest. From Lines et al. (1988) who proposed a ‘cooperative inversion’, to Moorkamp et al. (2011) who proposed a global strategy for the joint inversion of MT, gravity and refraction data, different studies show implementation of either (i) simultaneous inversions (e.g. Onizawa et al. 2002, for seismic and gravity), or (ii) sequential inversions (e.g. Tondi et al. 2000; Parsons et al. 2001), for seismic and gravity. These inversions use a direct relationship between parameters (e.g. Maceira & Ammon 2009 for surface wave and gravity) or a structural constraint using the cross-gradient of the parameter fields (e.g. Gallardo & Meju 2004). Geometrical constraints can also be introduced directly with discontinuity interfaces like in Tikhotsky & Achauer (2008) or Basuyau & Tiberi (2011). In this study, we perform a joint inversion, using a direct relationship derived from laboratory measurements on samples and a discrete node grid.

The main goal of this tomography is to obtain better constraints on the geological structure of La Soufrière hydrothermal system, that is the dome inner structure but also its basement that has not yet been imaged, while it may play an essential role in potential flank destabilization. The paper is organized in four sections: first we present the seismic and traveltime independent inversions, secondly, we briefly present the gravimetric data and inversion detailed in a companion paper, thirdly we introduce the joint inversion and show different tests to assess the validity of the method and results. Finally, we discuss the results in view of other geophysical studies and results.

1.1 Seismic data and traveltime modelling

The seismic data were acquired during three successive active surveys that took place between 2001 and 2007. Two different configurations where used: (i) a limited number (10) of standalone seismic stations recorded many small dynamite shots (100–300 g) fired along the ‘Chemin des dames’ trail around and on top of the dome, (ii) due to the weak signal noise ratio, we then deployed a large number of sensor using the High Resolution Imaging (HRI) facilities (Coutant et al. 2008) along different profiles to record few large dynamite shots (1–4 kg) fired in 2 m deep holes. A total of 262 recordings sites and 106 shooting site (Fig. 1) were used. Two hertz vertical geophones signals were recorded with a 250 Hz-sampling rate. The field conditions at La Soufrière are quite extreme due to the weather conditions (strong wind and rains) and the dense vegetation that strongly limits the area where instruments can be deployed. The limited azimuthal coverage and the poor signal/noise ratio for sensors deployed on the windward side are the reasons for the lack of seismic data on the northern side of the dome. We finally
could use 1003 thousand P-wave arrival times with a picking accuracy of about 5–10 ms. Data with larger error estimates were rejected. The natural seismicity is rather low on La Soufrière and has a limited spatial extent. We decided not to include it in the data set.

We analysed some of the 2-D seismic lines used for the survey to infer an average 1-D velocity model from 0 to 100 m depths. This model, given in Table 1, was derived using the plunplus-minus method (Hagedoorn 1959) for three profiles: one across the dome, and two along the dome basement. For depths below 100 m, this model is extended downward with the same velocity gradient until the P-wave velocity reached a value of 2700 m s\(^{-1}\), the average P-wave velocity obtained by Dorel et al. (1979) for the dome.

P-wave traveltime computation is based on Podvin & Lecomte (1991) finite difference method applied to a regular grid with 141 × 141 × 101 nodes and a grid spacing of 10 m. This grid delimits a volume of 1400 × 1400 × 1000 cubic metres centred laterally on the dome summit (coordinate origin (0,0) on Fig. 1). The topography is given by a DEM whose vertical accuracy is about 2 m on the dome itself (Beauducel 2001) and 5 m at the foothill. For an average P-wave velocity of 1000 m s\(^{-1}\), the traveltime error due to the DEM imprecision is about 2 ms on the dome, and 5 ms at the foothill. The grid step used for the traveltime and the derivative computations is smaller than the grid step used for inversion because the method assumes a constant slowness in each cell. This grid is rebuilt at each iteration with a tri-linear interpolation of the coarser inversion grid. For the inversion, we used a 29 × 29 × 21 nodes grid with a regular spacing \(\Delta = 50\) m.

Volcanic media are strongly heterogeneous and one must consider the traveltime computation step with a particular care. For La Soufrière, one must also take into account a very strong topography, since the elevation varies from 900 to 1460 m inside the 1.4 × 1.4 km model. The method of Podvin & Lecomte (1991) with the improvement proposed by Latorre et al. (2004) to compute the traveltime by integrating the slowness \(S_p\): 
\[T_{\text{calc}} = \int S_p \, dl\]  
along the ray is able to handle strong velocity heterogeneities. It guarantees in our configuration a traveltime accuracy better than 1 ms.

1.2 Traveltime inversion

We perform a classical traveltime inversion using Latorre et al. (2004) implementation of traveltime and derivative computation. We invert for slowness values defined at the nodes of a 29 × 29 × 21 nodes grid with a regular spacing of \(\Delta = 50\) m and delimiting the same volume as the traveltime computation grid. About 9000 nodes are actually inverted, the remaining being located above free surface. Although the model extends down to an elevation of 400 m, 1000 m below the dome summit (1460 m), we must note that the P-wave rays do not propagate below an elevation of 850 m. We use the Tarantola & Valette (1982) approach to obtain the parameter vector \(S\) that minimizes the cost function \(C(S)\) 
\[C(S) = (g_i(S) - T_{\text{obs}})^T C_{\text{S}}^{-1} (g_i(S) - T_{\text{obs}}) + (S - S_{\text{prior}})^T C_{\text{S}}^{-1} (S - S_{\text{prior}}).\]  
where \(C_{\text{S}}^{-1}\) and \(C_{\text{S}}^{-1}\) are the traveltime and slowness covariance matrices, \(T_{\text{obs}}\) and \(g_i(S)\) the observed and calculated traveltimes, \(S_{\text{prior}}\) the a priori slowness model. The function \(g_i(S)\) is non-linear, therefore we use an iterative quasi-Newton algorithm and compute the \(n+1\)th iteration as 
\[S^{n+1} = S^n + \mu_n \left(\Gamma_i^n C_{\text{S}}^{-1} \Gamma_i^n + C_{\text{S}}^{-1}\right)^{-1} \left(\Gamma_i^n C_{\text{S}}^{-1} (T_{\text{obs}} - g_i(S^n)) + C_{\text{S}}^{-1} (S^n - S_{\text{prior}})\right),\]  
where \(\Gamma_i^n\) is the matrix of the \(g_i(S)\) partial derivatives at iteration \(n\). \(\mu_n \leq 1\) is the step length, or scaling factor, that we keep constant and adjust by trial-error as high as possible to prevent the iterative process to diverge. The data covariance matrix \(C_{\text{T}}^{-1}\) is assumed to be diagonal (uncorrelated data), with a standard error \(\sigma_T = 5\) ms, our lower bound error estimate. We assume for the parameter covariance, a spatial, isotropic and exponential correlation between slowness values at points \((x,y)\) 
\[
C_S(x,y) = \sigma_S^2 e^{-\frac{|x-y|}{\Delta}},
\]  
where \(\Delta\) is the coarse grid step (50 m), \(D_S\) the dimensionless correlation distance and \(\sigma_S^2\) the slowness parameter variance. The motivation for using a non-zero correlation between parameters is to use a smaller grid mesh than the resolution brought by the data, and to control the smoothing directly inside the inversion process. This is an alternative to using large grid element, and to present smoothed results. Several authors proposed to use exponential correlation (e.g. Montagner & Jobert 1988; Tarantola 2005). Recently, Monteiller et al. (2005) showed that this covariance matrix has interesting properties, since its inverse is sparse and has a banded structure. We use the numerical expressions derived in appendix A that allow a direct computation of \(C_S^{-1}\) elements.

The starting model could not be defined from a preliminary 1-D inversion because of the topography. We then build a 3-D model by considering the average velocity profiles given in Table 1. This 1-D vertical seismic profile \(V(|z|)\) is applied to each node \((x,y,z)\) of the model by setting the P-wave velocity to values equal to \(V(|z-\text{topo}(x,y)|)\).

To determine the optimal values for \((D_S, \sigma_S)\), we first set the dimensionless correlation distance \(D_S\) to adjust the number of independent parameters to the number of data. Assuming that a correlation distance of \(D_S\) reduces the degree of freedom by \(D_S\) in one direction, a first estimation is obtained by setting \(D_S\) such that 9000/\(D_S^3\) ~ 1000, that is, \(D_S \sim 2\).

<table>
<thead>
<tr>
<th>Depth below surface (m)</th>
<th>0</th>
<th>15</th>
<th>60</th>
<th>100</th>
<th>118</th>
</tr>
</thead>
<tbody>
<tr>
<td>(V) (m s(^{-1}))</td>
<td>456</td>
<td>1158</td>
<td>1421</td>
<td>2500</td>
<td>2700</td>
</tr>
</tbody>
</table>

© 2012 The Authors, *GJI*, 191, 723–742

Geophysical Journal International © 2012 RAS
We then follow the strategy of Vergely et al. (2010), using the L-shaped misfit curve (Hansen 1992). We start from a ‘reasonable’ choice for \((D_3, \sigma_b)\), for example, \(\sigma_b = 0.1 \text{ s km}^{-1}\), \(D_3 = 2\), then we try different \((D_3, \sigma_b)\) values where \(\sigma_b\) and \(D_3\) vary inversely. An increase of \(\sigma_b\) reduces the influence of the initial model and a decrease of \(D_3\) reduces the dependence between the parameters. This would be the choice for a well-constrained inverse problem. On the contrary, a \(\sigma_b\) decrease and \(D_3\) increase will regularize a poorly constrained problem. An ‘optimal’ choice is obtained for the maximum curvature of the L-shaped misfit curve, \(\| g_{\text{fit}}(S) - T_{\text{obs}} \|^2 \) versus \(\| S - S_{\text{true}} \|^2 \) (Fig. 2a). We select a set of ‘optimal’ values for \((D_3, \sigma_b) = (1.5 \pm 0.5, 0.5 \pm 0.2 \text{ s km}^{-1})\). The \(\pm 0.5\) values indicate some variations around these values that yield similar misfit values.

The results obtained for the slowing inversion are shown in Fig. 3 and are presented with an estimation of the resolution. Since the size of this 3-D inverse problem is reasonable, we can compute the resolution operator defined by \(R = I - C_\Sigma C_\Sigma^{-1} \) where the posterior covariance matrix is computed as \(C_\Sigma = (\Gamma_2^\frac{1}{2} C_\rho^\frac{1}{2} \Gamma_1 + C_T^\frac{1}{2})^{-1}\). Resolution links the inversion solution \(S\) to the “true” solution \(S_{\text{true}}\) and the prior model \(S_{\text{prior}}\) by: \(S = R S_{\text{true}} + (I - R) S_{\text{prior}}\) (e.g. Tarantola 2005). We use for each node \(i\) the restitution ‘r-factor’ defined as \(r^i = \sum_j R_{ij}\) that represents the sum of the contributions of the ‘true’ values. This factor identifies the nodes that are poorly resolved, with no contributions from the ‘true’ values: \(r^i \ll 1\). An estimate of the spatial resolution can also be computed from the resolution spatial spreading defined by

\[
\delta' = \sqrt{\sum_{j} \text{dist}(i,j)^2 R_{ij}/r^i}
\]

analogue to the resolution spreading as defined in Menke (1989). \(\delta'\) represents the average (rms) distance of all the contributions from nodes \(j\) to node \(i\). When a parameter is well resolved by the inversion (i.e. \(r^i \sim 1\)), \(s\) discriminates between a parameter that is spatially well resolved with minimum contribution from neighbour nodes, \(s \sim 0\) (i.e. \(R_{ij} = \delta_{ij}\)), and a parameter which is resolved by the prior-covariance smoothing, \(s > 0\).

We plot on top of cross-section the contour values (red curve) for an \(r\)-factor value of 0.75. A larger value would unfortunately not be visible in Fig. 3(a), demonstrating clearly that the resolution of the traveltime tomography is limited especially toward the north of the model. The inversion yields an rms reduction of 83 per cent with a final traveltime \(\text{rms}_{\text{final}} = 13\) ms.

2 GRAVITY DATA AND INVERSION

We use data collected by Gunawan (2005) for the Basse-Terre Guadeloupe Island. A total of 359 measurements are available for the South of Basse-Terre, and from those points, we keep the 145 points located on and around the summit (Fig. 4). For the inversion, we use the same methodology that Coutant et al. (2012) used to invert the gravity field of South Basse-Terre Island. The method is as follows: we invert a Bouger anomaly (Fig. 4) where the correction due to the topography of the area \(G_{\text{topo}}\) is divided into two contributions: (i) constant density layers, limited on top by the topography; (ii) a finite parallelepiped volume, limited on top by the topography, where density is prescribed at a grid nodes. For (i) and (ii) the topography can have arbitrary precision. For (ii) the grid is identical to the grid used for traveltime inversion (29×29×21, \(\Delta = 50\) m), and uses also a tri-linear interpolation between the node values.

The different terms used to compute the Bouger anomaly include a bathymetry correction \(G_{\text{bathy}}\) that was estimated to be around \(-10\) mgal on and around the dome (Gunawan 2005), and a regional factor \(G_{\text{reg}}\) estimated to be of the order of 92 mgal.

The accuracy of the topography \(G_{\text{topo}}\) and free air \(G_{\text{fai}}\) corrections are strongly dependent on the altitude precision. We used here the altitude given by a composite DEM build from three sources: (i) a DEM from 1:25000 toponographic map with a 50 m horizontal sampling; (ii) a DEM built from satellite picture with a 10 m horizontal sampling and a 5 m vertical standard error valid for the whole inverted zone; (iii) a DEM built from aerial photography, with a 5 m horizontal sampling, a 2 m vertical standard error, valid on the dome itself. These DEMs vertical accuracies were estimated from comparison with GPS measurements. We estimate that the Bouger anomaly error due to altitude determination is lower than 0.5 mgal on and around the dome.

The topography correction \(G_{\text{topo}}\) was obtained by assuming a constant density of 2.71 on the island above sea level and 2.3 for the material located above the elevation of 1150 m (Gunawan 2005). The numerical accuracy of the correction was found to be better than 0.001 mgal (Coutant et al. 2012). It can be expressed as a linear function of the node density values \(\rho_1\) as

\[
G_{\text{topo}} = G_1 + \sum G_2 \rho_1.
\]

The linear inversion was performed using a formulation similar to the traveltime inversion

\[
\hat{\rho} = \rho_o + (\Gamma_2^\frac{1}{2} C_\rho^\frac{1}{2} \Gamma_2 + C_T^\frac{1}{2})^{-1} \Gamma_2 C_T^\frac{1}{2} (G_{\text{calc}} - G_{\text{obs}}),
\]

where \(G_{\text{obs}} = G_{\text{measure}} - G_{\text{bathy}} - G_{\text{fai}} - G_{\text{reg}} - G_1\) and \(G_{\text{calc}} = \sum G_2 \rho_1 - G_2 \rho_0\). The data covariance matrix is diagonal, \(C_\rho = \sigma^2_\rho I\) and does not contain any correction to enhance the density contribution at depth. We choose for the standard error \(\sigma_\rho\) a value of 0.2 mgal.

The density covariance matrix uses the same exponential correlation as above: \(C_\rho = \sigma_\rho^2 e^{-\frac{d}{2\Delta}}\). A first estimation of the correlation distance is deduced as above from the number of data (9000/\(D_{\rho}^2 = 145\)) and yields \(D_{\rho} = 4\). We determine a set of optimal values for \(D_{\rho}\) and \(\sigma_\rho\) using the same procedure as the traveltime inversion and obtain a maximum curvature of the L-shaped misfit function for \((D_{\rho}, \sigma_\rho) = (1.5 \pm 0.5, 0.3 \pm 0.1)\). We obtain with these parameters a rms reduction of 53 per cent with a final \(\text{rms}_{\text{final}} = 0.79\) mgal. The result of the inversion is shown in Fig. 3(b), the red curve delimit the zone where the \(r\)-factor is larger than 0.9. The shape and size of the \(r = 0.9\) curve show clearly that density is more constrained than slowness, especially toward the north of the model.
Figure 2. L-shaped misfit curves. (a) L-curve for gravity and slowness inversions. Data residuals are plotted versus distance between final and initial model for different values of $\sigma$. For slowness inversion, parameter error $\sigma$ larger than 0.3 s km$^{-1}$ lead to instable inversion as denoted by dotted line. Large values (>3.5) of correlation distance $\Delta$ yield final solutions that depart from initial model. (b) L-curve for joint inversion. The sum of data residuals is plotted versus the coupling residual. The best coupling is obtained for a value between 0.2 and 0.3 s km$^{-1}$. 

© 2012 The Authors, *GJI*, 191, 723–742
Geophysical Journal International © 2012 RAS
Figure 3. Results of the two independent inversions. Thick black line delimits free surface, elevation levels every 50 m. (a) $P$-wave velocity; thick red line delimits the volume where $r$-factor is higher than 0.75; (b) Density; thick red line delimits the volume where $r$-factor is higher than 0.9.
Traveltime and gravity joint inversion

729

3 JOINT INVERSION

The traveltime tomography shown in Fig. 3(a) is limited by a lack of resolution due to the source/receiver configuration especially toward the north of the model. In parallel, the gravity inversion performed on the same zone (Fig. 3b) is under-determined, smooth, with a poor depth resolution. The objective of the joint inversion then is to combine the two methods to improve the overall resolution and to obtain a global image of the full volume that we investigate. A simple strategy would be to present, side by side, the results of the two separate inversions and to build a global model. Here, we proposed to build this global model from a joint inversion of the two data sets.

3.1 Slowness/velocity versus density relationship

Different approaches are used to express the density–velocity relationship in joint gravity seismic velocity inversions. The well-known Birch’s law (Birch 1961) is a linear $\rho(V_p)$ relation valid for crustal and mantle rocks derived from sample measurements. During the inversion, it may be kept constant for the whole or part of the model (Basuyau & Tiberi 2011), or depends on node locations (Tondi et al. 2003). Gardner et al. (1974) provide relations for less dense material, and Nafe & Drake (1963) for sedimentary, possibly porous material. Onizawa et al. (2002) use Nafe & Drake (1963) data to derive their own analytical quadratic relation. Maceira & Ammon (2009) interpolate between the latter and Birch’s law for different range of density. Onizawa et al. (2002) were concerned with volcanic materials with high porosity. They choose Nafe-and-Drake’s law for sediment in absence of further data available from Isu-Oshima volcano. We benefit at La Soufrière from laboratory measurements on density and $P$-wave velocity performed by Bernard & Zamora (2012) on 43 volcanic samples coming from the neighbour Mt Pelée volcano at Martinique. Their data set that covers a large range of porosity (4 to 73 per cent) is representative of Mt Pelée products (dome, lava, scoria, pumices, ...). The two volcanoes are close (150 km) with similar kind of activity and lava composition (mainly andesitic). Ongoing similar measurements (Bernard 2012, personal communication) conducted on 15 samples from La Soufrière deposits and particularly the dome (massive and scoriaceous parts) confirm the data obtained so far at Mt Pelée with $V_p$ velocity as high as 5 km s$^{-1}$ (ultrasonic measurement at 1 MHz) in the massive part of the dome (porosity: 2–4 per cent). The porosity ranges between 1 and 35 per cent in the samples under study. Velocity and density measurements for water-saturated samples are displayed in Figs 5(a) and (b) in both ($\rho$, $S_p$) and ($\rho$, $V_p$) diagrams. We report also Birch’s law, and Onizawa et al. (2002) quadratic relation ($\rho$, $S_p$). Values measured on saturated massive samples display a clear linear trend on Fig. 5(a) highlighted by the red dot regression line. This line however is unable to represent materials that are highly porous, partially or fully empty or filled by water. We then decide to keep the ($\rho$, $S_p$) linear trend for high density ($\rho > 1.5$), and to use an asymptotic trend for low-density materials. This asymptote goes through the two particular points: air (0, 0 km s$^{-1}$) and water.
Figure 5. $P$-wave slowness (a) or velocity (b) versus density relations. Three relations are displayed: linear ($\rho, V_p$), Birch (1961), green; quadratic ($\rho, V_p$), Onizawa et al. (2002) using Nafe & Drake (1963), black; linear ($\rho, S_p$) fit to Mt. Pelée and La Soufrière samples, red dotted line; our relation based on ($\rho, S_p$) fit plus asymptotic trend for weak density, red. (c): ($\rho, S_p$) relation after joint inversion for all nodes (black) and 25 per cent best resolved nodes (grey).
Our relation is represented in red in both \( (\rho, V_p) \) and \( (\rho, S_p) \) axes and is expressed as
\[
S = g_3(\rho) = \left( 1 + \frac{\rho_0}{\rho} \right) (S_0 + \alpha \rho),
\]
where \( \rho_0 = 0.3, S_0 = 0.71 \text{ ks}^{-1} \) and \( \alpha = -0.2 \text{ ks}^{-1} \). We plot on the same Figs 5(a) and (b) the values of P-wave velocity and density that were obtained from the two independent traveltime and gravity inversions for the best 25 per cent resolved nodes. Velocity spans almost the whole range of values but densities are limited to the [2-2.7] range. We recall here (see Section 2) that density average values will depend on the regional correction term applied during the gravity inversion. The correction term obtained from Coutant et al. (2012) (92 mgal) yields this range, but a higher correction would result in values shifted toward lower density values.

3.2 Formulation

We follow the method proposed by Onizawa et al. (2002) where the coupling between velocity and density is introduced through a supplementary least square constraint in the cost function. This method adapted to our problem yields the new constraint to be minimized:
\[
(S - g_3(\rho))^T C^{-1}_3 (S - g_3(\rho)) \quad \text{where} \quad S = g_3(\rho) \quad \text{is the ‘theoretical’ relation between} \ S \ \text{and} \ \rho \ 
\]
and \( C_3 \) is the covariance diagonal matrix \( C_3 = \sigma_3^2 I \) that specifies the degree of coupling.

The complete cost function \( C(S, \rho) \) that we now minimize writes
\[
C(S, \rho) = (T_{\text{obs}} - g_1(S))^T C^{-1}_T (T_{\text{obs}} - g_1(S)) + (G_{\text{obs}} - g_2(\rho))^T C^{-1}_G (G_{\text{obs}} - g_2(\rho))
\]
\[
+ (S - S_{\text{prior}})^T C^{-1}_3 (S - S_{\text{prior}}) + (\rho - \rho_{\text{prior}})^T C^{-1}_\rho (\rho - \rho_{\text{prior}}) + (S - g_3(\rho))^T C^{-1}_3 (S - g_3(\rho)).
\]

The first two terms correspond to constraints brought by the data, the two next correspond to the \textit{a priori} information, and the last is the new constraint between the parameters. This last constraint can be viewed as a supplementary data set \( \delta \), where \( \delta = S - g_3(\rho) \) represents the variation from the theoretical relationship between density and slowness. We then express the joint inversion procedure as a function of the two independent inversions. We first define two new parameter and observation vectors \( m \) and \( d_{\text{obs}} \), and the related \( g(m) \) relationship
\[
m = \left( \begin{array}{c} S \\ \rho \end{array} \right); \quad d_{\text{obs}} = \left( \begin{array}{c} T_{\text{obs}} \\ G_{\text{obs}} \\ \delta \end{array} \right); \quad g(m) = \left( \begin{array}{c} g_1(S) \\ \Gamma_2 \rho \\ S - g_3(\rho) \end{array} \right).
\]

The inverse matrices for parameter \( C^{-1}_M \) and data \( C^{-1}_D \) covariance, and the derivative matrix \( G \) write
\[
C^{-1}_M = \begin{bmatrix} C^{-1}_S & 0 \\ 0 & C^{-1}_\rho \end{bmatrix}; \quad C^{-1}_D = \begin{bmatrix} C^{-1}_T & 0 & 0 \\ 0 & C^{-1}_G & 0 \\ 0 & 0 & C^{-1}_\rho \end{bmatrix}; \quad G = \begin{bmatrix} \Gamma_1 = \frac{\partial g_1}{\partial S} & 0 \\ 0 & \Gamma_2 \\ I & -\Gamma_3 = -\frac{\partial g_3}{\partial \rho} \end{bmatrix}.
\]

Using matrix and vector blocks eq. (8) now writes
\[
C(m) = (g(m) - d_{\text{obs}})^T C^{-1}_D (g(m) - d_{\text{obs}}) + (m - m_{\text{prior}})^T C^{-1}_M (m - m_{\text{prior}}).
\]

We seek the minimum of \( C(m) \) by using the quasi-newton iterative solution
\[
m_{\text{iter}+1} \approx m_n + M (G^T C^{-1}_D G + C^{-1}_M)^{-1} (G^T C^{-1}_D (d_{\text{obs}} - g(m_n)) + C^{-1}_M (m_n - m_{\text{prior}})),
\]

where \( M = \begin{bmatrix} \mu_s & 0 \\ 0 & \mu_\rho \end{bmatrix} \). In this formulation the coupling between \( S \) and \( \rho \) is expressed in \( g(m) \) and its derivatives \( G \). One could demonstrate that when the coupling relationship (eq. 7) is linear, the parameter covariance matrix \( C^{-1}_M \) can alternatively express this linear coupling through its off-diagonal covariance terms. This indeed goes back to the definition of the covariance between two parameters. Eq. (12) can be reformulated using slowness and gravity inversion formulas. The quasi-newton iterative solution becomes
\[
\begin{bmatrix} S_{\text{iter}+1} \\ \rho_{\text{iter}+1} \end{bmatrix} = \begin{bmatrix} S_n \\ \rho_n \end{bmatrix} + \begin{bmatrix} \Gamma_1 C^{-1}_T \Gamma_1 + C_{\delta}^{-1} \\ \Gamma_1 C^{-1}_G \Gamma_2 + \Gamma_1 C_{\rho}^{-1} \Gamma_3 \end{bmatrix} + C_{\delta}^{-1} \begin{bmatrix} \mu_s \left[ \Gamma_1 C_{\delta}^{-1} \Delta T + C_{\delta}^{-1} \Delta \delta + C_{\delta}^{-1} (S_n - S_{\text{prior}}) \right] \\ \mu_\rho \left[ \Gamma_1 C_{\rho}^{-1} \Delta G - \Gamma_2 C_{\rho}^{-1} \Delta \delta + C_{\rho}^{-1} (\rho_n - \rho_{\text{prior}}) \right] \end{bmatrix},
\]

where \( \Gamma_1 \) and \( \Gamma_3 \) iteration number dependencies are omitted. To solve eq. (13), we use different but constant step-length factor for slowness (non-linear, \( \mu_s < 1 \)) and density (linear, \( \mu_\rho = 1 \)). For each iteration, the system is decomposed and solved iteratively to keep the same
order for the conditioning of the two independent problems. Details of the solution and the expression of the resolution matrix are given in appendix B.

One difficulty is to choose the set of optimal covariance and correlation distances. The joint inversion is a competition between the constraints brought by the data set, the a priori models and the coupling. The relative weight of the data versus a priori knowledge has to be determined by the study of the two independent problems. For each independent problem, the data covariance or error estimate is somewhat arbitrary and can be compensated by the parameter covariance to give similar results. For instance, setting the traveltime error estimate to \( \sigma_T = 0.1 \) s and the parameter a priori error to \( \sigma_S = 1.e-3 \) s \( \text{km}^{-1} \) will give the same result as \( \sigma_T = 1 \) s and \( \sigma_S = 1.e-2 \) s \( \text{km}^{-1} \), except for the absolute value of the cost function.

We then proceed as follow: \((D_\rho, C_\rho, C_T)\) and \((D_S, C_S, C_T)\) are determined from the two independent inversions. The two problems are then scaled to adjust their respective level of data misfit, by dividing \( C_\rho \) and \( C_T \) by a common factor. Finally, the coupling \( C_S \) is adjusted to improve the inversion result in unresolved zone, while not degrading in the already resolved zone. To determine the optimal adjustment, we plot the L-shaped curve value of the misfit function relative to the data: \( ||T_{\text{obs}} - T_{\text{calc}}||^2 + ||G_{\text{obs}} - G_{\text{calc}}||^2 \) as a function of the coupling misfit \( ||S - g_1(\rho)||^2 \) for different values of coupling variance \( \sigma_C^2 \). Starting from weak coupling, the coupling misfit is high while the data misfit is low and is the sum of the two independent inversions misfits. When increasing the coupling, the coupling misfit decreases, and the data misfit remains stable as long as we do not change the parameters already resolved by the data. At some point, the coupling constraint becomes predominant over the data constraint and the data misfit drastically increases. We select (Fig. 2b) the maximum of curvature of the curve as the good adjustment. The optimal value for our joint inversion was found equal to \( \sigma_S = 0.25 \) s \( \text{km}^{-1} \). All these parameters are summarized on Table 2. The number of iterations needed to reach the convergence varies typically from 15 to 20 iterations. Despite the apparent complexity of the different parameters adjustment, we noted that in all the run that we tested, the results exhibit similar characteristics. The main heterogeneities remain present and the variability is always apparent on the absolute values of velocity and density. Thus, we will focus our interpretation on the contrast rather than on the absolute values.

### 3.3 Results and analysis of resolution

The results of the joint inversion are shown in Fig. 6. Velocity (Fig. 6a) and density (Fig. 6b) are plotted for six horizontal cross-sections at elevations varying from 900 to 1400 m. The distribution of slowness and density values obtained from the inversion is summarized on Fig. 5(c) in comparison with the theoretical elevations varying from 900 to 1400 m. The distribution of slowness and density values obtained from the independent inversion (Fig. 5a, blue circle, 0.5 < \( \rho \) < 0.75) increase. The resolution is increased for slowness/velocity mostly toward the northern part of the dome and for density: 36 per cent, rms final \( = 1.08 \text{ mgal} \), instead of 0.79 mgal (see Table 2).

We evaluate the contribution of the joint inversion versus independent inversions by analysing the resolution operator \( R \) and by performing a spike test. In the case of the joint inversion, we can still express the relation between the estimated, the ‘true’ and the prior solutions for density and slowness parameters with the respective \( R_\rho \) and \( R_S \) resolution (see appendix B for details)

\[
\tilde{S} = R_\rho S_{\text{true}} + [...] \rho_{\text{true}} + [...] S_0 + [...] \rho_p
\]

\[
\tilde{\rho} = R_\rho \rho_{\text{true}} + [...] S_{\text{true}} + [...] S_0 + [...] \rho_p.
\]

We use, like for the independent inversions, the \( r \)-factors \( r_\rho \) and \( r_\rho \) defined as \( r_i^2 = \sum R_i^2 \) which represent the sum of the contribution of the ‘true’ values. We plot on top of our inversion results in Figs 6–9 the contour line for \( r_\rho^2 = 0.9 \) (density plot) and \( r_\rho^2 = 0.75 \) (slowness/velocity plot). The contribution of the joint inversion can be noted in Figs 6(a) and (b), where the two volumes corresponding to higher resolution values \( (r_\rho^2 > 0.9 \) and \( r_\rho^2 > 0.75) \) increase. The resolution is increased for slowness/velocity mostly toward the northern part of the dome and at depth (Fig. 6a compared to Fig. 3a). For the density resolution, the improvement is mostly apparent for the superficial part of the dome and in a least extent for the northern part (Figs 3b and 6b). Fig. 7 displays the spatial resolution \( s \) (Eq. 4) for slowness and density for the independent (Figs 7a and c) and for the joint inversions (Figs 7b and d). Nodes that are well resolved spatially exhibit \( s \) values as low as 2\( \Delta \) (<100 m), while poorly resolved nodes show spatial resolution larger than 6\( \Delta \) (>300 m). The major changes between independent and joint inversions resolution can be seen on slowness resolution at the summit and at depth; for density the resolution changes are mostly notable for superficial layers and at depth.

We can expect that a contribution of the joint inversion come from the different nature of the two sensitivity kernels. Traveltime sensitivity kernel depends upon the ray coverage and should produce sharper images where coverage is dense. The gravity sensitivity kernel that derive from a potential is known to produce smoother images for increasing depths. We clearly observe that slowness resolution benefits from density

| Table 2. Covariance adjustments for independent and joint inversions. |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| \( \sigma_S \) (s \( \text{km}^{-1} \)) | \( \sigma_T \) (ms) | \( D_S \) | Reduc. | \( \sigma_\rho \) | \( \sigma_C \) (mgal) | \( D_\rho \) | Reduc. | \( \text{rms (mgal)} \) |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| ind\(^1\) | 0.5 | 5 | 1.5 | 84 per cent | 12 | 3 | 0.2 | 1.5 | 53 per cent | 0.79 |
| joint | 0.5 | 5 | 1.5 | 84 per cent | 12 | 0.75 | 0.05 | 1.5 | 36 per cent | 1.08 |
| \( \sigma_S \) (s \( \text{km}^{-1} \)) | 0.3 | 0.5 |

© 2012 The Authors, *GJI*, 191, 723–742

Geophysical Journal International © 2012 RAS
Figure 6. Results of joint inversions. (a) P-wave velocity, (b) density. Same conventions as Fig. 3.
Figure 7. Spatial resolution in grid step unit (Δ = 50 m) estimated from resolution spatial spreading at different depths and for the different inversions: (a) P-slowness independent; (b) P-slowness joint; (c) density independent; (d) density joint. Red thick lines delimit the volume where r-factor is higher than 0.75 (slowness) and 0.9 (density).
Figure 8a: slowness independent spike inversion

Figure 8b: density independent spike inversion

Figure 8. Spike test results. Black squares represent the location of six cubic +10 per cent perturbations in velocity and density. The perturbation at centre extends from 900 to 1100 m; (a) $P$-wave velocity independent inversion; (b) density independent inversion.
Figure 9. Spike test results. (a) P-wave velocity joint inversion; (b) density joint inversion.
resolution at summit, while density resolution improves at depth. We however observe that these reciprocal influences not only come from data but also from prior models. The superficial layers that exhibit excellent density spatial resolution (smaller than 3 m) in Fig. 7(d) and low density (<1.5 see Fig. 6b) are due to the velocity prior model, not from data. We are thus unable to discriminate the influence of the data versus prior model in the improvements of the joint resolutions by solely studying the resolution values.

To better quantify the joint inversion effects we also perform spike tests (Figs 8 and 9). We add six cubic velocity/slowness perturbations of +10 per cent in velocity/density to the joint inversion \( V_g \) results shown in Fig. 6(a), with size varying from \( \Delta^3 \) to \( (3\Delta)^3 \). The density model is then build using the relation \( \rho = g_0^{-1}(S) \). The spike locations correspond to the main dense heterogeneities that we wish to interpret. Following Lévêque et al. (1993), the spike test was preferred to the checkerboard. Heterogeneities that cannot be retrieved by the independent inversions (Figs 8a and b) are present in the joint inversions (Figs 9a and b). The depths of the heterogeneities are poorly resolved by the gravity inversion alone. In Fig. 8(b), spikes at 1200, 1100 and 1000 m are found above their real elevation but are more constrained in the joint inversion (Fig. 9b). The lowest elevation that can be resolved is 1000 m. The central 200 m-high perturbation is not resolved by gravity data and laterally poorly resolved by seismic data.

4 DISCUSSION AND INTERPRETATION

The volume that is investigated in our study is related to three major geological structures of the Grande Découverte-Soufrière complex. The 1530 A.D. eruption horseshoe crater is delimited at north by Morne Amic and at southwest by L’Échelle volcano (Boudon et al. 2008, see Figs 4 and 10, Komorowski et al. 2008). It is sitting in the 3100 BP eruption craters, delimited at north by a ridge leading to Nez Cassé summit (Fig. 4). Inside these two craters one can find (i) remaining of andesite spines, dome or lava flow that build up during the 3100 B.P. or 1530 A.D. eruptions, (ii) block or mega blocks from flank collapses, (iii) fractured zones with complete or partial alteration. The very high seismic velocities obtained in the inversions, confirmed by the ultrasonic measurements on Mt Pelée and La Soufrière andesite samples (Bernard & Zamora 2012) must be discussed in view of the dome growth history. They are likely signatures of andesite spines that build up during the 1530 A.D. or 3100 B.P. eruptions.

Several geophysical surveys have been conducted on La Soufrière hydrothermal system. Among others: Zlotnicki et al. (1992, 2006) hereafter denoted Z-2006 use SP (Spontaneous Potential) and VLF (Very Low Frequency) surveys to analyse the superficial conductive zone; Nicollin et al. (2006) (denoted N-2006) conducted several electrical tomography experiments across the dome with 2-D profiles similar to our seismic profiles. Many other experiments are still ongoing, including muon tomography (Lesparre et al. 2012, denoted L-2012). All these studies suggest that La Soufrière hydrothermal system is made of different massive/unaltered blocks separated by active or sealed drains. Most of the investigations conducted so far include electromagnetic methods. It is therefore difficult to discriminate between plugged clayed drain and unaltered andesite zone that both appear like highly resistive zones.

We first discuss the differences between independent and joint inversions. For density images: (i) The joint inversion adjusts the superficial layers density to the velocity that is itself close to the a priori velocity model (dark blue in Fig. 6a); (ii) the depth of heterogeneities are deeper (e.g. Figs 3b and 6b at 1300 m) in the joint inversion than in the independent inversion. The joint inversion may thus correct the lack of resolution at depth. For velocity images; (iii) the overall resolution is increased and (iv) the spatial resolution remains unchanged in already well-resolved zones.

We have reported in Fig. 10 the main heterogeneities labelled #1 to #9 for different elevations with the supposed location of the 3100 B.P. horseshoe shaped crater. We discussed these heterogeneities in regard to N-2006 resistivity results interpolated and superposed in Fig. 10, on a clockwise direction starting from summit.

At 1400 m, one observes: (i) four dense and three fast zones (#8 and #9). We cannot definitively eliminate for #9 the case of inaccurate topographic corrections in gravity data since they correspond to three higher elevations (Summit, Piton Saussure, Piton Sanner). The south #8 patches are present on both gravity and velocity at different elevations. Its position corresponds to the so-called ‘Bulge’ prominent topography; it also correlates with a high resistivity red zone (N-2006). The east #8 patch is present at 1300 m on resistivity and on muon radiography image (L-2012).

At 1300 m, the northern end of Amic crater (#1) is clearly apparent on all results with a depth that extend down to 1000 m. Below, the fast (#2) body found east of Amic crater corresponds to a shot zone (Fig. 6a) and to high-density zone (Fig. 6b) that extend the Amic dense body toward the east. These two bodies could be related to Amic Lava dome formed during the 3100 B.P. eruption (Boudon et al. 1987). This (#2) anomaly correlates with a resistive zone detected by VLF (zone A, Z-2006). Finally, the northern low resistivity values also correlate with low velocity on Fig. 6(a).

At 1200 m, the west (#3) dense and fast body likely connects to L’Échelle crater (1700 yr ago) although it is moved northward by gravity data toward Carbet river. A low density/slow velocity superficial (#4) body at 1300/1200–1100 m correlates with a low resistivity zone and Z-2006 C,D anomalies. It is interesting to correlate this #4 zone to the zone delimited by the 30 Août 1976 (b) and Gouffre-1956 (d) fractures depicted in Fig. 10 since they delimit a zone of the major hydrothermal activity along the dome flank.

Another low-density/slow velocity zone #5 starting from 1100 m is present on both velocity/density and resistivity images down to 900 m. They may be connected to the G zone of Z-2006 that was recognized as a major zone of fluid circulation.

A major high-density/velocity zone (#6) is found from 1200 m down to 900 m. It extends from Tarade piton down to the dome basement on a northeast trend. The upper part of this body only appears on the resistivity section. Piton Tarade is currently interpreted as a conglomerate.
Figure 10. Schematic of inversions results in relation with principal geological structures superposed on resistivity horizontal sections. Brown thick line: contour of horseshoe shaped 1530 AD crater after Boudon et al. 2008. Dome fracture zones plotted on elevation 900 m box only: (a) ‘Fente du nord’; (b) ‘Fracture 1956’ and (c) ‘Fracture du 30 août 1976’. Heterogeneities discussed in text: (1) to (9) dense/fast (grey) and light/slow (shaded). Resistivity horizontal sections interpolated from 1-D resistivity inversion by Nicollin et al. (2006). Interpolations are performed on and around survey profiles denoted by white dots.
of megablocks formed during 3100 B.P. collapse event. Body #6 image suggests that it may be more massive than expected, and could result from an andesite injection extending deeper than the 3100 B.P.-crater floor. It could then be considered as a more stable zone than expected from its superficial destabilization observed after the 2004 Mw 6.3 Les Saintes earthquake (Bazin et al. 2010; Feuillet et al. 2011).

Finally, minor heterogeneities are reported but not labelled on Fig. 10, except for the #7 that clearly relates to the ‘Fajau’ collapse. This is the only clear trace in the inversions of an open fracture zone, the large ‘Fente du nord’ (a) open fractured zone is not visible.

Electric and electromagnetic surveys conducted so far show evidences of basalt conductive, sealed layers and drains that are expected to follow the interfaces corresponding to the 1530 A.D. and 3100 B.P. former craters. Our results show that several dense and fast bodies (#6, #2) may cross these interfaces while bodies (#1, #3) may represent massive and stable crater rim limits. These results confirm a model of La Soufrière upper structure made of different blocks separated by conductive or sealed clayed drains. The massive blocks may have deeper roots than expected from previous investigations.

5 CONCLUSION
We have performed a joint gravity/traveltime inversion to obtain density and P-wave velocity images of La Soufrière hydrothermal system. We have adapted the method proposed by Onizawa et al. (2002) using the Bayesian approach of Tarantola & Valette (1982). Our solution of this non-linear inversion uses a deterministic quasi-newton iterative approach with direct expressions of inverse covariance matrices derived in the appendix. The joint inversion was used here to overcome the lack of resolution of the two methods taken separately. We show that the joint inversion that relies on a density-P-wave relationship derived from measurements on Mt Pelée samples fulfill this goal. Indeed, our results are valid under the condition of validity of this (ρ, Vp) relationship. The velocity results benefit from the gravity inversion coverage on the northern unresolved part. Gravity results benefit from the velocity resolution with respect to depth.

Our results show that the resistive zones that have been so far only seen by electromagnetic surveys may not be due only to argilization but may also be explained by the presence of dense massive zones, that we interpret as andesite spines resulting from 3100 B.P. or 1530 A.D eruptions. These dense bodies may have implication on the stability of the edifice. Further comparisons between available results and those from ongoing surveys (e.g. muons tomography, electromagnetic survey) will yield further information that will help to better understand the destabilization risks at La Soufrière of Guadeloupe.

ACKNOWLEDGMENTS
The study benefited from invaluable help from the staff of the ‘Observatoire Volcanologique et Sismologique’ de Guadeloupe (OVSG-IPGP). We gratefully acknowledge the ‘Parc National de la Guadeloupe’ who made these experiments inside the park limits possible. We thank Bernard Valette, Romain Brossier and Jean Virieux for fruitful discussions, Robert Guiguet and Fabrice Doré (firemen) and our colleagues from ISTerre for their help in the fieldwork. We gratefully thank the three anonymous reviewers that greatly help to improve the manuscript. Financial supports were provided through ANR contracts CATEL/RISK-VOLCAN, RISK-2008/DOMOSCAN and CNRS-INSU PNRN program.

REFERENCES
APPENDIX A

We recall here the expressions to compute directly the inverse covariance matrix $C^{-1}$ for a stationary function $\rho(m)$ evaluated at $N$ regularly spaced locations $\{m_i\}$. $C^{-1}$ is defined as

$$
C^{-1} = \text{cov}(\rho(m_i), \rho(m_j)) = \sigma^2 \text{cov}(m_i, m_j).
$$


For the 1-D case, given a grid step $\Delta$ and a covariance function $f(x)$, we define $C_{M}^{\|} = \sigma_{f}^{2} f(\Delta(i - j))$ where $C_{M}$ is a Toeplitz matrix. Toeplitz matrix are related to convolution operator and used for FIR type filtering (see e.g. R. Gray 2006, for a review). We then consider the matrix $C_{M}$ as the discrete convolution operator for an impulse response $f$, applied to the N points discrete vector $I(x)$:

$$O = C_{M} I. \tag{A1}$$

The problem of computing $C_{M}^{-1}$ is thus equivalent to finding the discrete inverse operator related to the impulse response $f_{m}$. If we consider the continuous exponential covariance function: $f(x) = e^{-\frac{\|x\|^{2}}{\delta}}$, its Fourier transform writes: $F(k) = \frac{2\pi\delta}{1 + k^{2}\delta}$ and the Fourier inverse operator $F_{m}(k) = \frac{1 + k^{2}\delta}{2\pi\delta}$ ($k =$ wavenumber). We then need to express the discrete convolution operator that will approximate $F_{m}(k)$. Different solutions are available: bilinear transform, impulse invariant transform. The derivative-analogue transform is easy to extend to 3-D and is used here: we replace the derivative operator $ik \iff \frac{1}{\Delta}$ by a centred finite differences discrete operator. The inverse convolution operator for $f(x) = e^{-\frac{\|x\|^{2}}{\delta}}$ is thus rewritten assuming a unit grid step ($\Delta = 1$) as: $I(x) = \frac{1}{2\pi\delta}(1 - D^{2}e^{i\|x\|^2}). O(x)$ and is approximated for its discrete form by $I \approx \frac{1}{2\pi\delta}[(2D^{2} + 1)O - D^{2}(O_{i+1} + O_{i-1})]$.  

The inverse covariance matrix $C_{M}^{-1}$ is then a three bands sparse matrix built with the stencil. $C_{M}^{-1}|_{ij} = [-\frac{\delta}{\pi}|_{i+1}|_{i} - \frac{\delta}{\pi}|_{i+1}|_{i}]$. This constraint is written for node 1 using Eq. A1 with $\mathbf{I} = \sigma_{f}^{2}[1, 0, 0, \ldots]$ and $\mathbf{O} = \sigma_{f}^{2}[1, e^{-\frac{\|x\|^{2}}{\delta}}, e^{-\frac{2\|x\|^{2}}{\delta}}, \ldots]$. This yields the relation $1 = C_{M}^{-1}|_{11} + C_{M}^{-1}|_{12} e^{-\frac{\|x\|^{2}}{\delta}}$ and hence the stencil for the first node: $[1 + \frac{\delta}{\pi} e^{-\frac{\|x\|^{2}}{\delta}}]$. The $C_{M}^{-1}|_{NN}$ term is deduced in a similar way.

For the 3-D case, the inverse covariance matrix can be build from the tensor product of the previous solution applied successively to the three dimensions. This gives however a non-isotropic smoothing operator. We prefer to build an isotropic operator using the 3-D Fourier transform of the exponential covariance function $f(x) = e^{-\frac{\|x\|^{2}}{\delta}}$: assuming again $\Delta = 1$:

$$F(k) = \frac{1}{(1 + k^{2}\delta)^{2}} \text{ where } k \text{ is the radial wavenumber.}$$

The discrete inverse operator must then approximate

$$F_{m}(k) = \frac{1}{8\pi D^{2}} (1 + D^{2}k^{2})(1 + D^{2}k^{2}).$$

We first evaluate $C$, the matrix associated to the $(1 + D^{2}k^{2})$ operator where the Laplacian $k^{2}$ is evaluated again using centred finite differences. $C$ is a seven bands sparse matrix build with a 3-D stencil (see below) that involves seven spatial nodes.

\[
\begin{array}{cccccccc}
0 & 0 & 0 & 0 & D^{2} & 0 & 0 & 0 \\
0 & -D^{2} & 0 & -D^{2} & 1 + 6D^{2} & D^{2} & 0 & -D^{2} \\
0 & 0 & 0 & -D^{2} & 0 & 0 & 0 & 0 \\
\end{array}
\]

The stencils for face, edge and vertices nodes need again an additional constraint to define their diagonal value $C^{i}$. We choose here to keep the sum of the stencil values and equal to one. This yields the diagonal values $(1 + 5D^{2}, 1 + 4D^{2}, 1 + 3D^{2})$ for respectively the face, edge and vertices nodes. Using Eq. (A1), we see that this constraint is equivalent to consider that the sum of each covariance line element (the integral of the impulse response) is constant. This induces an effect on border nodes where the variance is larger than the variance inside the model. A constraint equivalent to the 1-D case is difficult to apply since the inverse Fourier transform of $F(k) = \frac{1}{(1 + k^{2}\delta)^{2}}$: that is,

$$f(r) = e^{-\frac{\|x\|^{2}}{\delta}} \text{ cannot be evaluated simply at } r = 0.$$

Finally, we build the inverse matrix with the product: $C_{M}^{-1} = \frac{1}{8\pi D^{2}} C \ast C$. 

**APPENDIX B**

The quasi-newton iterative solution is obtain as

\[
\begin{pmatrix}
S_{e} \\
\rho_{e}
\end{pmatrix} = \begin{pmatrix}
S_{e-1} \\
\rho_{e-1}
\end{pmatrix} + \begin{pmatrix}
\Gamma_{T}^{T} C_{T}^{-1} \Gamma_{T} + C_{s}^{-1} & -\Gamma_{T}^{T} C_{s}^{-1} \\
-\Gamma_{s}^{T} C_{s}^{-1} & \Gamma_{s}^{T} C_{s}^{-1} \Gamma_{s} + C_{T}^{-1} \Gamma_{T}^{T} \Gamma_{T}
\end{pmatrix}^{-1} \begin{pmatrix}
C_{s}^{-1} & 0 \\
0 & C_{e}^{-1}
\end{pmatrix}^{-1} \begin{pmatrix}
\mu_{S} \left[ \Gamma_{T}^{T} C_{T}^{-1} \Delta T + \Gamma_{s}^{T} C_{s}^{-1} \Delta S + C_{s}^{-1} (S_{e-1} - S_{e}) \right] \\
\mu_{e} \left[ \Gamma_{s}^{T} C_{s}^{-1} \Delta G - \Gamma_{s}^{T} C_{s}^{-1} \Delta G + C_{e}^{-1} (\rho_{e-1} - \rho_{e}) \right]
\end{pmatrix},
\]

where $\Gamma_{T}$ and $\Gamma_{s}$ iteration number dependencies are omitted.

Rather than solving the complete linear system, we prefer to rewrite for each iteration $n$, the system in the form

\[
\begin{pmatrix}
A & B \\
B & C
\end{pmatrix} \begin{pmatrix}
S_{n} - S_{n-1} \\
\rho_{n} - \rho_{n-1}
\end{pmatrix} = \begin{pmatrix}
A & B \\
B & C
\end{pmatrix} \begin{pmatrix}
\Delta S_{n} \\
\Delta \rho_{n}
\end{pmatrix} = \begin{pmatrix}
E \\
F
\end{pmatrix}.
\]

© 2012 The Authors, *GJI*, 191, 723–742

Geophysical Journal International © 2012 RAS
where

\[ A = \Gamma_1^T C_1^{-1} \Gamma_1 + C_3^{-1} + C_5^{-1} \]
\[ B = -\Gamma_1^T C_3^{-1} = B^T \]
\[ C = \Gamma_2^T C_2^{-1} \Gamma_2 + \Gamma_3^T C_3^{-1} \Gamma_3 + C_4^{-1} \]
\[ E = \Gamma_2^T C_2^{-1} \Delta T + C_4^{-1} \Delta \delta + C_5^{-1} (S_n - S_p) \]
\[ F = \Gamma_2^T C_2^{-1} \Delta G - \Gamma_3^T C_3^{-1} \Delta \delta + C_4^{-1} (\rho_n - \rho_p) \]

where \( B \) is diagonal, \( A \) is sparse, \( C \) is full. This system is again solved iteratively for \( \Delta S_n \) and \( \Delta \rho_n \) as

\[ A \Delta S_n + 1 = E - B \Delta \rho_n \]
\[ C \Delta \rho_n + 1 = F - B \Delta S_n \]

then

\[ S_n = \mu_S \Delta S_n + S_{n-1} \]
\[ \rho_n = \mu_\rho \Delta \rho_n + \rho_{n-1} \]

where the step-length factors are kept constant and respectively set equal to \( \mu_S \leq 1 \) and \( \mu_\rho = 1 \).

The posterior covariance matrix can be deduced from the above expression and the expressions for the inverse of a block matrix

\[ \tilde{C}_M = \begin{bmatrix} \Gamma_1^T C_1^{-1} \Gamma_1 + C_3^{-1} + C_5^{-1} & -\Gamma_1^T C_3^{-1} \\ -C_3^{-1} \Gamma_3 & \Gamma_2^T C_2^{-1} \Gamma_2 + \Gamma_3^T C_3^{-1} \Gamma_3 + C_4^{-1} \end{bmatrix}^{-1} = \begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix} \]

We finally get the expression for the resolution operator

\[ R = I - \tilde{C}_M \begin{bmatrix} I - A'C_3^{-1} & B'C_4^{-1} \\ C'C_3^{-1} & I - D'C_4^{-1} \end{bmatrix} = \begin{bmatrix} R_S & B'C_4^{-1} \\ C'C_3^{-1} & R_{\rho} \end{bmatrix} \]