Geophys. J. Int. (2004) **157,** 890–900 doi: 10.1111/j.1365-246X.2004.02237.x

Mantle viscosity inference: a comparison between simulated annealing and neighbourhood algorithm inversion methods

N. Piana Agostinetti,[1] G. Spada[2] and S. Cianetti[1]

[1] *Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy. E-mail: spina@ingv.it*
[2] *Istituto di Fisica, Università di Urbino, Urbino, Italy*

Accepted 2003 December 23. Received 2003 December 23; in original form 2003 April 24

SUMMARY

Two direct search methods, simulated annealing and neighbourhood algorithm, are applied to the inversion of the viscosity profile of the mantle using relative sea level time-histories for the Hudson Bay region. In problems characterized by a low-dimensional model space ($N_d = 2$ in this study), the two inversion methods show comparable performances. When a larger number of dimensions is involved (specifically $N_d = 6$), we directly show that simulated annealing is less effective than neighbourhood algorithm in overcoming the obstacles that are found in the model space when our specific data set is employed. This study confirms that modifications of the conventional Monte Carlo inversion method, such as simulated annealing and neighbourhood algorithm, are viable tools to determine the viscosity profile of the mantle, which, until recently, has been mainly tackled by means of linearized techniques.

Key words: mantle viscosity, neighbourhood algorithm, postglacial rebound, simulated annealing.

1 INTRODUCTION

In the last decade, the problem of constraining the viscosity profile of the mantle has been investigated using either linear inverse theory or Monte Carlo (MC) methods (e.g. Ricard *et al.* 1989; King 1995; Mitrovica & Peltier 1995; Peltier 1998; Spada 2001). While in linearized approaches an initial plausible model is iteratively improved using local derivative information, the MC methods require a large number of forward computations whose predictions are evaluated according to their fit to the data.

The decrease of CPU time needed to solve the forward problem and the non-uniqueness and non-linearity of the postglacial rebound inversion problem (Spada *et al.* 1992; Milne *et al.* 1998) have recently stimulated studies aimed to apply the MC search methods to constrain mantle viscosity (Spada 2001). However, the classical undirected MC search methods, based on a uniform pseudo-random sampling, are known to be inefficient when the number of parameters is large (Mosegaard & Tarantola 2002). The improvement of search efficiency is the main purpose of alternative optimization methods, such as Genetic Algorithm (GA) and Simulated Annealing (SA) (a comprehensive review on this subject is given by Sambridge & Mosegaard (2002)). In importance sampling algorithms, the exploration of the models space occurs at a rate dictated by the posterior probability density. Pure MC and SA inversion techniques have another advantage with respect to the classical linearized methods, in that no explicit expression for the probability density is needed (e.g. no Gaussian constraint on the *a priori* probability density distribution). However, in SA methods, which are primarily designed for global optimization problems, the search in the parameters space is only partly dictated by information coming from previously sampled models. Moreover, both SA and GA need a non-obvious initial tuning to converge to the optimum model. This may cause serious disadvantages if the direct search methods are employed, as we will explicitly show in the following for the specific case of SA.

A new inversion method, the Neighbourhood Algorithm method (NA hereafter), has been recently introduced by Sambridge (1999a) for teleseismic receiver functions inversion. The NA method is a two-stage technique. The first is the search stage, in which an ensemble of models is generated with a sampling density inversely proportional to the value of a given objective function. In the second stage, called the appraisal stage, useful statistical information are extracted from all the models of the ensemble previously generated. In both stages the NA method employs a simple geometrical construction to enhance the sampling efficiency and to decrease the CPU time cost.

In this paper, we present a benchmark between the results of the SA inversion, coupled to the Metropolis decision rules, and that of the first stage of the NA inversion. The performances of these two inversion methods are compared in the specific case of a postglacial rebound data set. In a previous study (Cianetti *et al.* 2002, hereafter referred to as CI02) only the SA method was employed and no attention was devoted to other possible stochastic inversion procedures. In this work, we mainly emphasize the optimization part of the inverse problem and we limit our attention to qualitative resolution analyses. The performances comparison done here is, in our opinion, relevant in view of large-scale inversions of postglacial rebound signatures, in which the algorithm efficiency is a crucial factor. Here we consider postglacial relative sea level (RSL) data sets

from eight sites across the Hudson Bay, which are known to provide reliable inferences of the mantle viscosity structure particularly in the region extending between the transition zone and the top of the lower mantle (Mitrovica & Peltier 1995). This same data set was employed by CI02 in their SA inversion.

In the first section, we briefly review the direct search strategies and focus on the NA methods and tuning of the search parameters. The second section is devoted to the presentation of the results for two distinct parametrizations of the rheological profile of the mantle. Finally, we discuss the two techniques employed here and draw our conclusions.

2 METHODS

The relative sea level (RSL) optimization problem can be stated in the following way: given a set of RSL time histories from field measurements and a parametrization of the Earth viscosity profile, find the sets of parameters that provide an acceptable fit with the RSL data. The inverse methods that can be employed to address this problem can be classified into three main categories: exhaustive undirected searches, gradient methods and direct derivatives-free methods (Sambridge & Mosegaard 2002).

MC (e.g. Spada 2001) and classical grid (e.g. Velicogna & Wahr 2002) searches belong to the first category. These inversion methods, which can even be applied to strongly non-linear problems, are suited for a complete exploration of model space, but can be computationally very expensive when a large number of parameters is involved or when the fine details of the misfit function are to be resolved.

Gradient search methods are by far the most known and widely employed techniques for geophysical data inversion. Because they usually imply a linearization in the models space, these search methods are prone to be trapped in local minima of the misfit function and are not well suited to indicate trade-offs in the parameters space (Sambridge & Mosegaard 2002).

The last category includes global direct search methods. The two most widely known methods, GA and SA, based on evolutionary biology and thermodynamics natural analogues, are less prone to converge toward local minima, but require a careful tuning of the search parameters, like crossover probabilities in GA (King 1995) or cooling factors in SA (see CI02). Unlike GA and SA, the recently introduced NA method (Sambridge 1999a) is based only on a geometrical construction in the model space. In this section, the basic features of the NA and SA will be briefly illustrated and compared. For an exhaustive presentation of NA the reader is referred to Sambridge (1999a) and related articles (Sambridge 1998, 1999b, 2001).

In our implementation of NA, the Earth's mantle is modelled by a set of parameters $\mathbf{m} = (m_1, \ldots, m_{N_d})$, where N_d is the number of dimensions of the models space over which we intend to minimize the misfit between predictions and observations [in our investigation, the misfit is explicitly computed using eq. (1) below]. The *a priori* conditions on the model parameters are given in the form of bounds as $m_i^{\min} < m_i < m_i^{\max}$, $(i = 1, \ldots, N_d)$. Both the NA and SA methods can, however, deal with more complex constraints (Sambridge 1999a). Once the bounds are given, NA generates a random sample of n_s models and divides the parameters space into neighbourhoods defined as Voronoi cells. Each of them contains a sampled model and represents the region closer to that model than to the others in the L_2 norm sense. The ensemble of all of the Voronoi cells (Voronoi diagram) forms a coverage of the parameters space, which is employed to direct subsequent samplings. The misfit is computed for each of the n_s models sampled in the parameter space to identify the n_r best performing ones. The search is then continued within the Voronoi cells of the n_r models so identified. This self-adaptative scheme is iterated to concentrate the search in the most promising region of the models space. According to our experience with the postglacial rebound inversion problem, the CPU time required for the construction of the Voronoi cells in the NA may be two orders of magnitude smaller than that needed to solve the forward problem. Hence, the geometrical construction that characterizes the NA does not significantly affect its efficiency. We also notice that NA is suited for parallelization and its MPI implementation is now available (http://rses.anu.edu.au/∼malcolm/na/na.html).

The search and exploration strategies of SA differ significantly from that of NA. In our implementation of SA, the path towards the optimum models is in fact determined by the Metropolis rule (Metropolis *et al.* 1953) in conjunction with an annealing schedule. Accordingly, given a model with misfit M, a new one with misfit M' is accepted if $M' \leq M$, and accepted with a probability $\exp[-(M' - M)/T]$ otherwise, where T is a control parameter that plays the role of temperature. In agreement with CI02, we employ an annealing with geometric cooling. As the temperature decreases, the probability of going uphill in the models space diminishes and the SA walk is progressively driven in the direction of decreasing misfit. To improve the exploration performances, CI02 have modified the SA random walk so that if it crosses the *a priori* boundaries on the model parameters, a restart is performed by a new random model. The outcomes of the restarts strategy, which is also adopted here, will be further discussed in the following section.

In NA, the construction of the boundaries of the Voronoi cells allows employment of the information from previously sampled models and a global search in the parameters space to be performed. On the other hand, SA exploits previously sampled models through the cooling schedule, which generally relies only on the number of previously accepted models, and Metropolis decision rule, which only involves a test on the last model sampled (see CI02). Our experience shows that, even for $N_d = 2$, NA explores the misfit landscape more exhaustively in both low- and high-misfit regions.

The tuning of the two search parameters n_s and n_r in NA is crucial to our work. It should be observed that the choice of n_r and n_s is not dictated by the forward model computation, but only by the dimensions of the models space and total number of samplings planned. It is known that large values of n_r and n_s might be more suited for exploration than exploitation (Sambridge 1999a). In spaces of small dimensions the risk of being trapped in local minima is significant. This is results from space saturation, which occurs when the Voronoi cells are no longer neighbours to each other. As shown in Sambridge (2001), saturation is achieved with only 200 models in a 5-D space. We choose $n_r = 50$ and $n_s = 100$, which might provide a slow convergence but limits the possibility of being trapped in local minima of the misfit function. Concerning SA, the tuning of search parameters, such as the temperature schedule and length of the steps in the models space, is constrained by the data set used, by the forward model employed for the predictions and form of the objective function. To be consistent with CI02, here we use a geometric cooling schedule, with $T = T_0\alpha^n$, where $T_0 = 1$ is the initial temperature, $\alpha = 0.99$ is the cooling factor and n is the number of previously accepted models, but we are aware that other more efficient schedules could be employed. The *a priori* constraints on the model parameters will be in the form of bounds, as done for the NA inversions.

The misfit between observed ($\mathbf{d}^{\rm obs}$) and predicted ($\mathbf{d}^{\rm pre}$) data is evaluated using the following objective function:

$$M_{\mathbf{m}} = \frac{1}{N}\sum_{k=1}^{N}\left[\frac{d_k^{\rm pre}(\mathbf{m}) - d_k^{\rm obs}}{\sigma_k}\right]^2, \quad (1)$$

where N is the number of data and σ_k ($k = 1, \ldots, N$) their uncertainties. For the data set used here, $N = 34$. The data employed in this study are given in Walcott (1972) for six of the eight sites of the Hudson Bay considered (C. Henrietta Maria, Churchill, Keewatin, Southampton Is., Ungava Pen. and Ottawa Is.). For Richmond Gulf and James Bay, we use the revised data sets from Mitrovica *et al.* (2000). We refer the reader to CI02 and Mitrovica *et al.* (2000) for a discussion of the differences between the original compilation of Walcott and more recent revisited compilations (Mitrovica & Peltier 1995; Tushingham & Peltier 1991). To weaken the dependence of the results on the details of the late Pleistocene ice load history, we employ the RSL parametrization introduced by Mitrovica & Peltier (1995), which imply a normalization by the oldest datum of each time series. Mitrovica & Peltier (1995) showed that inference of mantle viscosity based upon such parametrization is particularly robust. The forward computations are based on the model by Spada *et al.* (1992), which includes an incompressible Maxwell viscoelastic mantle, a fluid inviscid core and an elastic lithosphere. The synthetic RSL at time t_{BP} is computed as $S(\mathbf{x}, t_{BP}) - S(\mathbf{x}, t_p)$ where S is the sea level variation at site $\mathbf{x}$ and t_p is present time. To gain better insight into the NA and SA performances, we will employ in the following the two different parametrizations illustrated in Table 1. In the first, there are two unknowns (the dimension of the parameters space is therefore $N_d = 2$). They are the base-10 logarithms of the upper- and lower-mantle viscosities (in Pa s units), denoted by ν_{UM} and ν_{LM}, respectively. The boundary between upper and lower mantle is at 670 km depth and the elastic lithosphere has a thickness of 120 km. In the second parametrization, characterized by $N_d = 6$, three unknowns are the base-10 logarithms of the upper-mantle (ν_{UM}), transition zone (ν_{TZ}) and lower-mantle (ν_{LM}) viscosities, and the remaining three correspond to the lithosphere (H_L), upper-mantle (H_{UM}) and transition zone (H_{TZ}) thicknesses, expressed in units of km. The maximum length of the random step, for SA, is 0.25 units along the ν axes and 10 km along the H axes of the parameters space.

Other details of the forward model employed here are described and justified by CI02. They include a uniform ocean load, a maximum harmonic degree $l_{\rm max} = 36$ in the spectral representation of the displacement and geoid height, and the use of the ICE-3G deglaciation time history of Tushingham & Peltier (1991).

Table 1. Earth mantle parametrization and parameters bounds employed in this study. N_d represents the dimension of the parameters space, the symbol ν is the base-10 logarithm viscosity of a given layer (expressed in unit of Pa s) and H is the layer thickness (units are km). Labels L, UM, TZ, and LM refer to the (elastic) lithosphere, upper mantle, transition zone and lower mantle, respectively.

N_d	Parameter	Allowed range	
2			
	ν_{UM}	20.0	22.5
	ν_{LM}	20.5	22.5
6			
	ν_{UM}	20.0	22.0
	ν_{TZ}	20.0	22.0
	ν_{LM}	20.5	22.5
	H_L	80.0	140.0
	H_{UM}	230.0	390.0
	H_{TZ}	100.0	400.0

3 RESULTS

As a first step in our study, we have run synthetic tests using ICE-3G predictions as fictitious observed data. The synthetic data uncertainties are those of the real RSL observations. This simulated inversion is done in order to ascertain how accurately the NA recovers a known viscosity profile and to explore the shape of the misfit landscape.

Figs 1 and 2 show the results from the synthetic tests for $N_d = 2$ and $N_d = 6$, respectively. It can be recognized that the misfit landscape for $N_d = 2$ has the form of a narrow valley and exhibits a clear trade-off between the two viscosity parameters. The optimum solution (triangle), with a misfit of $\sim 10^{-4}$, does not differ significantly from the known solution (circle) corresponding to the viscosity profile $\nu_{UM} = 21.0$ and $\nu_{LM} = 21.3$, which is that implied in model ICE-3G (Tushingham & Peltier 1991). The synthetic inversion for $N_d = 6$ (Fig. 2) fully recovers the viscosity profile of the mantle (the minimum misfit is again $\sim 10^{-4}$), but fails to correctly reproduce the H parameters (see Table 1). The uncertainty on the H_{TZ} parameter (Fig. 2d) is of ~ 100 km, a significant fraction of its nominal value (this is also true for the remaining H parameters, not shown in Fig. 2). The average misfit in Figs 1 and 2 is ~ 2.5 and ~ 3.5, respectively. From eq. (1), this means that, on the average, the synthetic data are reproduced within the 2σ level.

From the synthetic tests of Figs 1 and 2 we can conclude that if observed uncertainties are the only source of error, the mantle parametrization is correct and the ice load time history is known, the NA technique can retrieve the viscosity profile with a very high accuracy. However, the sensitivity of the synthetic data to the H

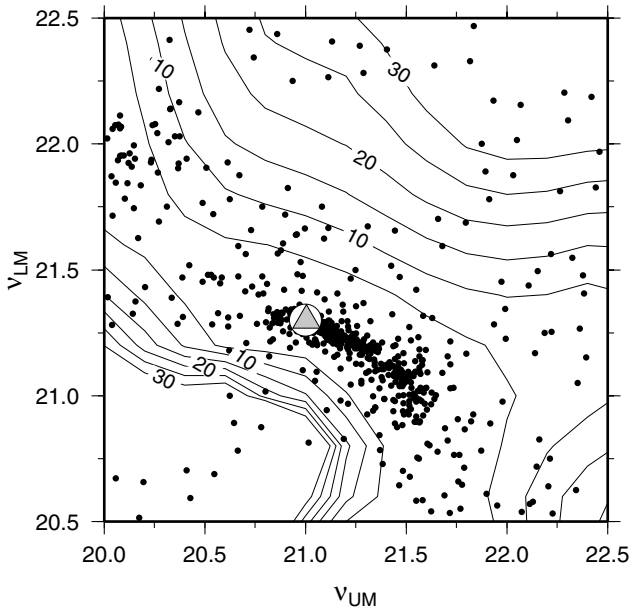

Figure 1. Synthetic test for mantle viscosity inversion performed by NA for $N_d = 2$. The only free parameters are ν_{UM} and ν_{LM} (see Table 1). The synthetic data set are obtained from a forward computation using the ICE-3G viscosity profile (white circle, with $\nu_{UM} = 21$ and $\nu_{LM} = 21.3$). Each dot denotes a sampled model. The contours of misfit $M_{\mathbf{m}}$ are also shown. The minimum misfit region is clearly the most densely sampled. The triangle shows the location of the best-fitting model, which is virtually coincident with the actual ICE-3G profile.

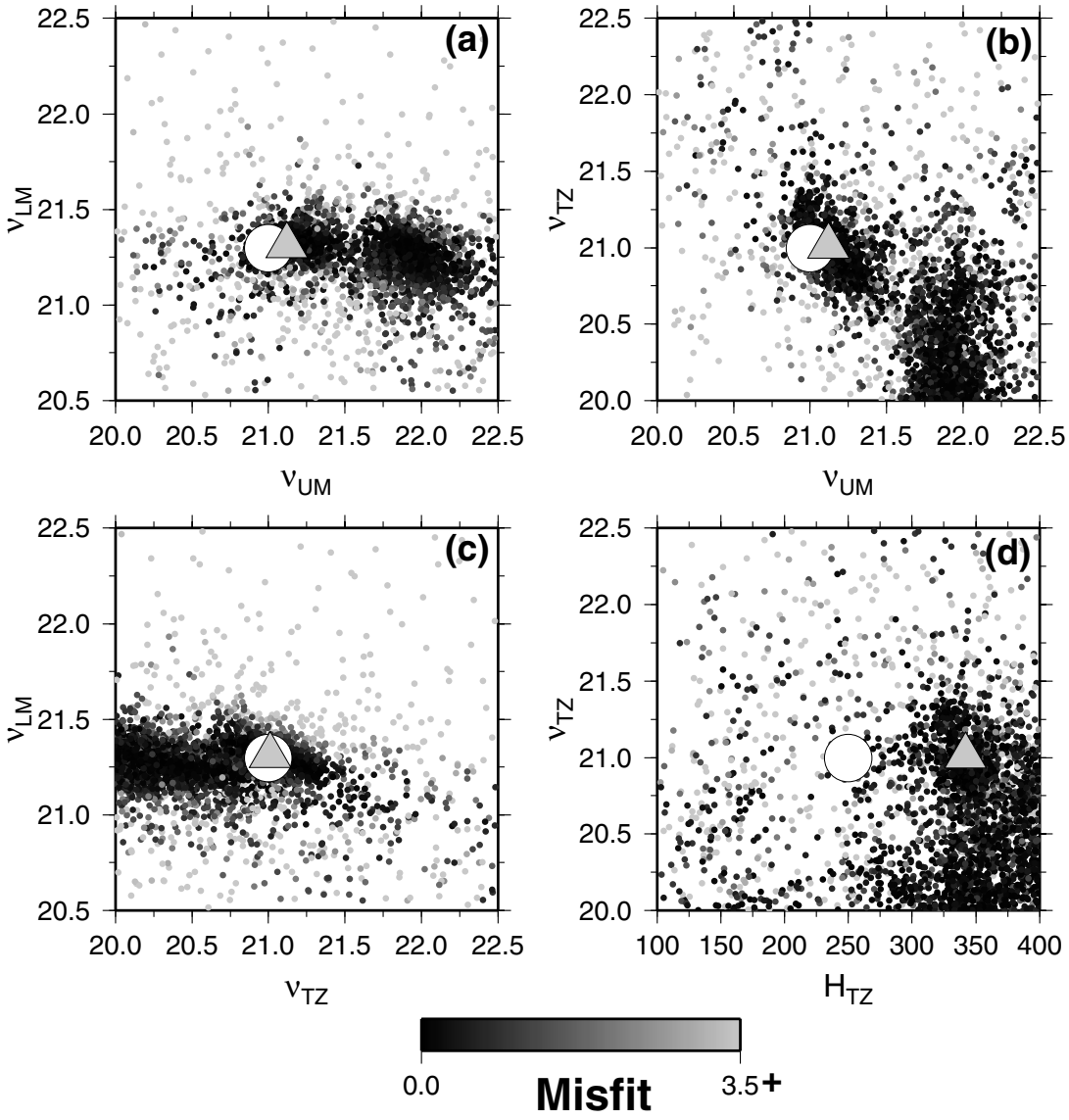

Figure 2. As for Fig. 1, but with $N_d = 6$ (see Table 1). Landscapes of viscosity parameters and of the H_{TZ} parameter are shown. Shades of gray from white to black indicate decreasing misfit values. A white circle and a gray triangle show the known viscosity model and best-fit model, respectively.

parameters is not large enough to allow for a comparably precise reconstruction of their values.

In Fig. 3 we show the misfit landscapes sampled from NA (panel a) and SA (b) for $N_d = 2$ (see Table 1). The real Hudson Bay relative sea level data are employed here and a total number of 10^3 forward models has been produced for both inversions. The arrows denote the narrow valley, which characterizes the two landscapes, and the triangles show the best-fit solutions for the two methods, which virtually coincide. The NA samples ~ 10 models even in the area of large misfit ($M \gtrsim 100$) evidenced in the bottom-left portion of each frames. This region is not visited by SA because, as the low misfit region is reached, the exploration of other portions of the parameters space is inhibited. With the tuning parameters used here ($n_s = 50$ and $n_r = 100$ for the NA and $\alpha = 0.99$ for the SA), the NA clearly appears to be more explorative than SA.

Differences in the search performances of the two methods are better evidenced in Fig. 4, which shows, for the same 2-D inversion of Fig. 3, the misfit as a function of the number of the sampled models. As in CI02, we have defined an acceptable misfit level $M_{\rm acc}$ (dotted lines) such that solutions with misfit $\leq M_{\rm acc}$ perform as the model with minimum misfit ($M_{\rm min} \simeq 36$) at the 95 per cent confidence level. With the data set considered here, by an F-test for 32 degrees of freedom we obtain $M_{\rm acc} \simeq 80$. We are aware that different criteria might be employed to define an acceptable misfit level, based on both previous experience (Ritzwoller *et al.* 2001; Shapiro & Ritzwoller 2002) and statistical treatment of models variance (Lomax & Snieder 1995), or a mixed criterion (Johnston & Lambeck 2000; Snoke & Sambridge 2002). As shown in Fig. 4, we have established another misfit level, the convergence level $M_{\rm con}$, that tells when, in the NA runs, resampled Voronoi cells become contiguous. In our inversion, $M_{\rm con} \simeq 40.0$. The arrows in Fig. 4 indicate the restart steps that are typical of the SA implementation by CI02. In the early stages of the inversion (number of sampled models smaller than ~ 100 in Fig. 4), the NA converges more slowly than SA towards low-misfit values because a significant effort is made by NA to also explore regions of large misfit (also see Fig. 3). However, it can be observed that after ~ 200 samplings, NA limits the search to models with $M < M_{\rm acc}$, while random restarts bring the SA search

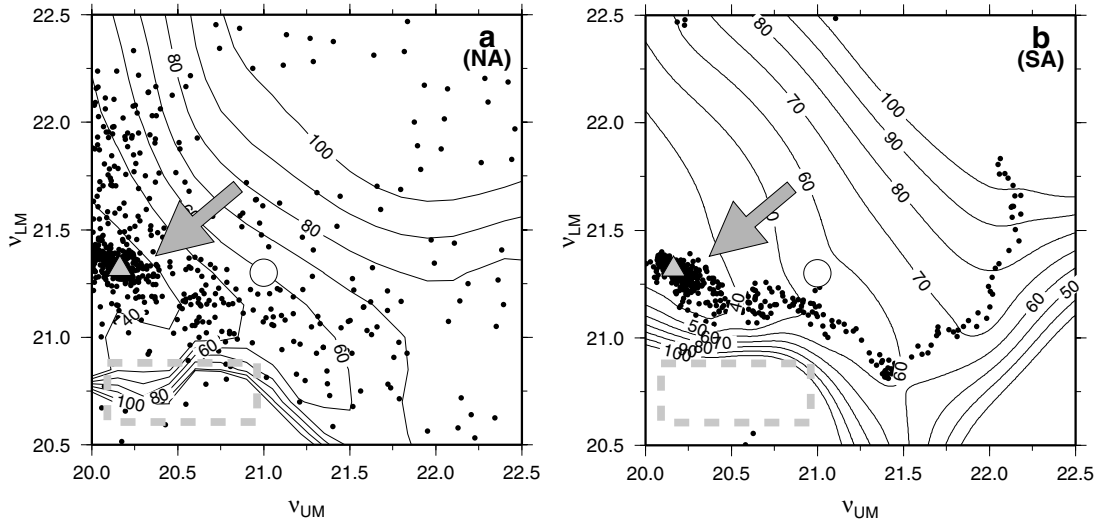

Figure 3. Mantle viscosity inversion using NA (panel a) and SA (b). Here $N_d = 2$ and the two inverted mantle parameters are ν_{UM} and ν_{LM}. Symbols are as for Fig. 1. The arrows enlighten the shape of the valleys, which characterize the misfit landscape in the 2-D parameters space. In addition to the sampled points, we have also qualitatively depicted the misfit function contour lines. The dashed rectangles show differences in the samplings of the two methods in the region of large misfit.

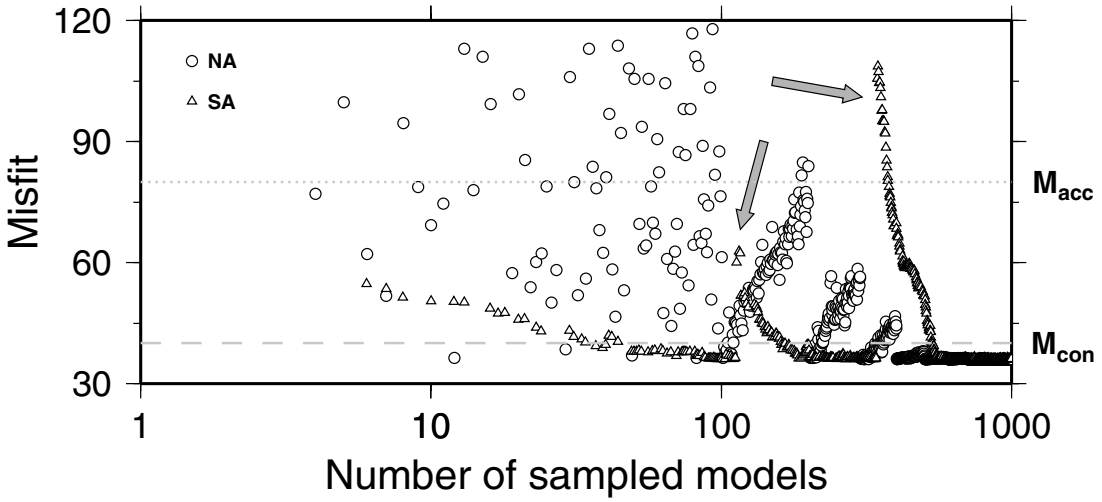

Figure 4. Misfit values as a function of the number of sampled models, for SA (triangles) and NA (circles). The arrows show the restart points in the SA method. Dotted and dashed lines indicate the acceptable ($M = M_{\rm acc}$) and convergence ($M = M_{\rm con}$) misfit level, respectively.

back to regions characterized by large M values. It is obvious that different tuning parameters would have produced different histories in Fig. 4 for both SA and NA. However, the results shown well depict the different strategies of the two inversion methods. In the NA, an initial stage during which the whole models space is sampled is followed by stages of resampling in more restricted regions, which may individually imply an increase in the misfit values. Nevertheless, as the inversion proceeds, the peak misfit values obtained during each resampling decrease systematically. In the SA, on the other hand, a decreasing misfit trend is always visible between two restarts, except perhaps during very short periods. Because the convergence is largely dictated by random restarts, in SA the course towards the valley that encompasses the misfit minima is less regular than when NA is used.

In Fig. 5 we present the results of the NA search for $N_d = 6$ (see Table 1). Here the final ensemble contains 3×10^3 models. Similar to the 2-D inversion of Fig. 3, the landscapes of the viscosity parameters are sampled more exhaustively by the NA method. For the SA inversion (panel b), the restarts which have been performed during the inversion are clearly visible in the form of localized tracks, which prevent a coherent visualization of the misfit landscape. From the convergence study of Fig. 6 we notice that, similar to the 2-D problem of Fig. 4, the branches that characterize the SA approach converge to low-misfit values faster than NA. However, after only ~ 750 trials, NA samples exclusively in the acceptable region ($M < M_{\rm acc}$) and after ~ 1000 it converges to the global minimum.

4 DISCUSSION AND CONCLUSIONS

In this section, we compare the practical outcomes of two inversion methods on the basis of the results obtained in the previous section, but we also consider further points. Our purpose is to clarify positive and negative aspects of two direct inversion methods and to illustrate

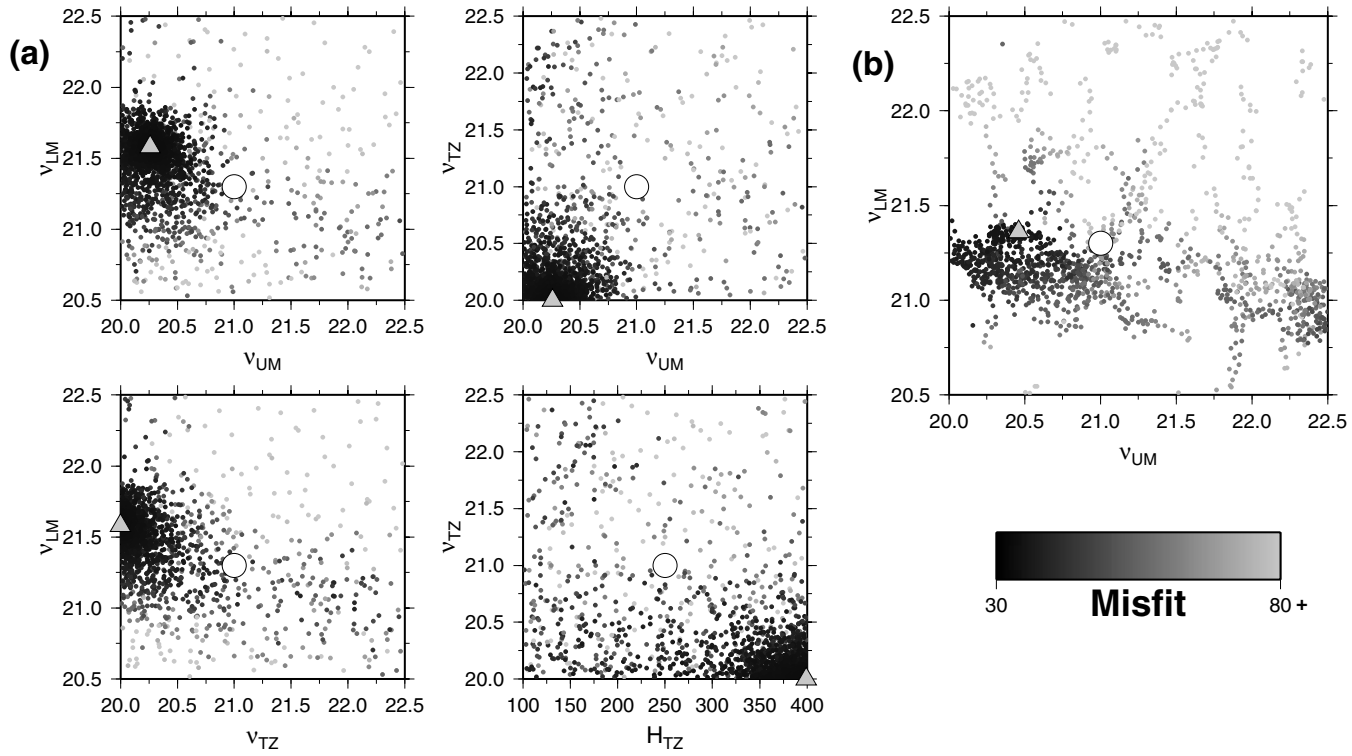

Figure 5. As for Fig. 3, but for a six-parameter search. Landscapes of viscosity parameters and of the H_{TZ} parameter are shown. Same notation as for Fig. 1. NA and SA are used in (a) and (b), respectively. For SA we only show the misfit landscape for ν_{UM} and ν_{LM}.

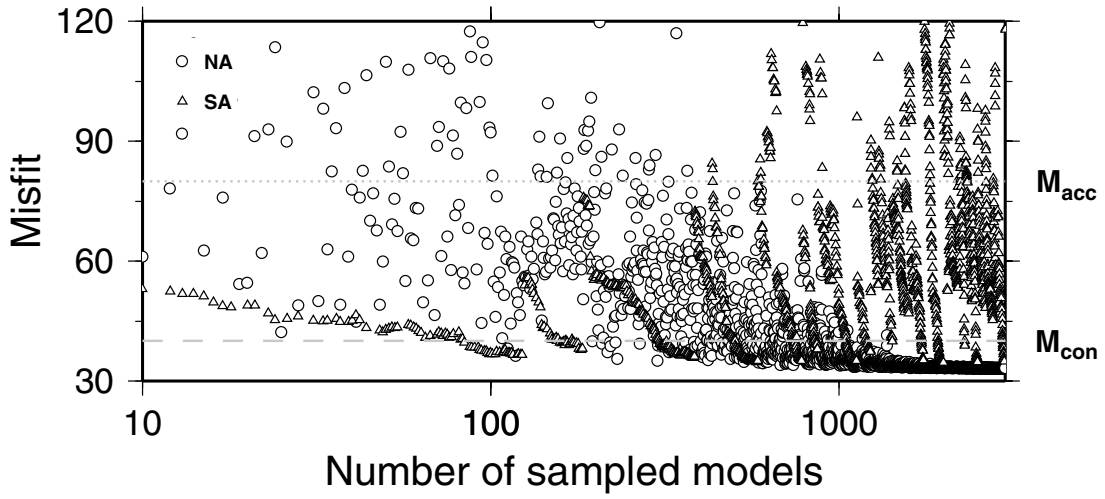

Figure 6. As for Fig. 4, but in the case $N_d = 6$. Dotted and dashed horizontal lines indicate the acceptable and convergence misfit level, respectively.

our experience with the specific problem of RSL data inversion. Our findings on the two inversion methods are restricted to the particular problem addressed in this paper and are not general (theoretically provable) results.

In order to compare the relative efficiency of the two inversion methods we consider the ratio

$$\mu(M) = \frac{n_{\rm acc}(M)}{n_{fw}}, \tag{2}$$

where $n_{\rm acc}(M)$ is the number of models with misfit $\leq M$ sampled by NA or accepted by the Metropolis rule in SA, and n_{fw} is the number of forward computations executed during the inversion process. The number n_{fw} has been fixed *a priori* and is equal to 10^3 for $N_d = 2$ and 3×10^3 for $N_d = 6$. The larger is $\mu(M)$, the greater is the efficiency of the method.

From Table 2, which summarizes the μ values for the results of Figs 4 and 6, we observe that, with the tuning parameters employed in this study, the μ coefficient for SA is always sensibly smaller than the corresponding quantity from NA, which indicates a larger efficiency of the NA. This occurs for both $N_d = 2$ and $N_d = 6$, at the two misfit levels $M_{\rm acc}$ and $M_{\rm con}$. This is also true when the μ ratio is computed as $n_{\rm acc}(M)/n_{\rm nre}$, where $n_{\rm nre}$ is the total number of models accepted by the Metropolis decision rule in SA (these modified μ ratios are given in parentheses in Table 2).

Table 2. Relative efficiency ratio $\mu(M)$ for SA and NA inversions, computed according to eq. (2). The μ values in parentheses are, however, computed differently (see text). N_d indicates the number of unknowns.

N_d	Algorithm	$n_{\rm fw}$ ($n_{\rm nre}$)	μ ($M_{\rm acc}$)	μ ($M_{\rm con}$)
2				
	SA	1000 (551)	0.53(0.96)	0.38(0.70)
	NA	1000	0.95	0.67
6				
	SA	3000 (1496)	0.43 (0.84)	0.08 (0.17)
	NA	3000	0.96	0.79

As previously observed, in 2-D inversion the misfit landscape generally appears to be better explored by NA than SA. This occurs in regions of both large and small misfit (Fig. 3). However, the region containing the global minimum of the misfit function are well resolved in both cases. This can be also appreciated building the frequency histograms for the inverted parameters, which provide information about the data resolving power of the inversion (Mosegaard & Tarantola 1995). In Fig. 7 the frequencies are computed as the number of models falling within a given parameter interval divided by the total number of models with misfit below a certain threshold. The thresholds are $M = M_{\rm acc}$ and $M = M_{\rm con}$ for the histograms drawn by grey and black lines, respectively. SA and NA histograms are depicted in upper and lower panels, respectively. The frequency histograms show only minor differences between the two algorithms. The larger number of models within the acceptable region $M < M_{\rm acc}$ in NA (see also Fig. 4) is the reason for the more bell-shaped NA histograms. The presence of a minor local minimum close to $v_{UM} = 21.4$ (see the left panel) is better evidenced by SA. From the analysis of the 2-D histograms of Fig. 7, we cannot conclude that there is a significant difference in the quality of the solutions retrieved by SA and NA. However, as we have verified in Fig. 4 for this 2-D case, the convergence of SA is faster than that of NA, albeit the restarts may significantly disturb the convergence.

If the dimension of the models space is enlarged, the complexity of the misfit landscape increases significantly. This can be appreciated in Fig. 8, which shows the results of the 6-D inversion illustrated in Fig. 6 from a different perspective. Frames (a) and (b) pertain to the NA and SA, respectively. An interesting feature is the presence of a bottleneck for misfit $M \approx 50$. As a result of the restart strategy of SA and the complexity of the misfit surface, a large number of samples is generated in the region $50 < M < 60$. The bottleneck causes all of these restarted paths to cross a narrow, densely populated region evidenced by the dark gray box in the v_{LM} frame of Fig. 8(b). In the corresponding panel of Fig. 8(a), such a dark cloud of points is not observed. The reason is that in the NA method the bottleneck region is visited only once during the resamplings of the parameters space and after this the attention is immediately devoted to regions of smaller misfit below that obstacle. While the NA search, driven toward a global optimization of the ensemble of models, succeeds in concentrating the analysis in promising regions, a large number of models are sampled by SA along its random walk, quite often in a previously visited region. An example is evidenced by the light

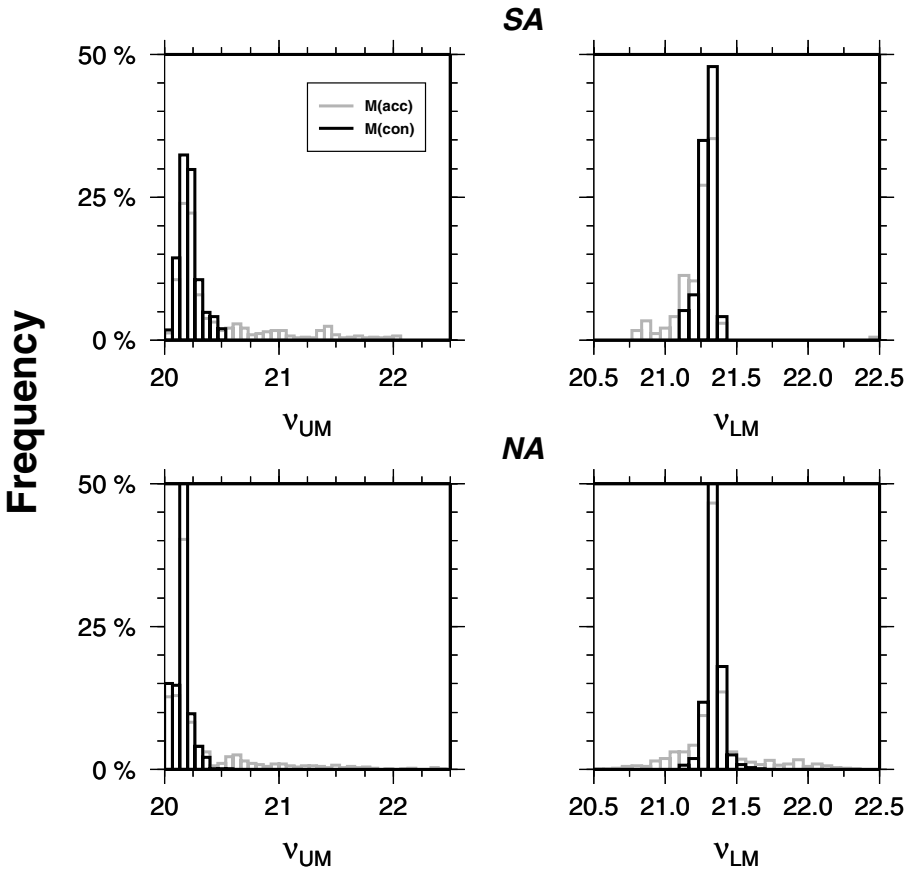

Figure 7. Frequency histograms for the parameters v_{UM} and v_{LM} in the $N_d = 2$ inversion. The frequencies are computed as the number of models falling within a given interval, divided by the total number of models with $M < M_{\rm acc}$ (gray) and $M < M_{\rm con}$ (black). Upper and lower panels are for SA and NA, respectively.

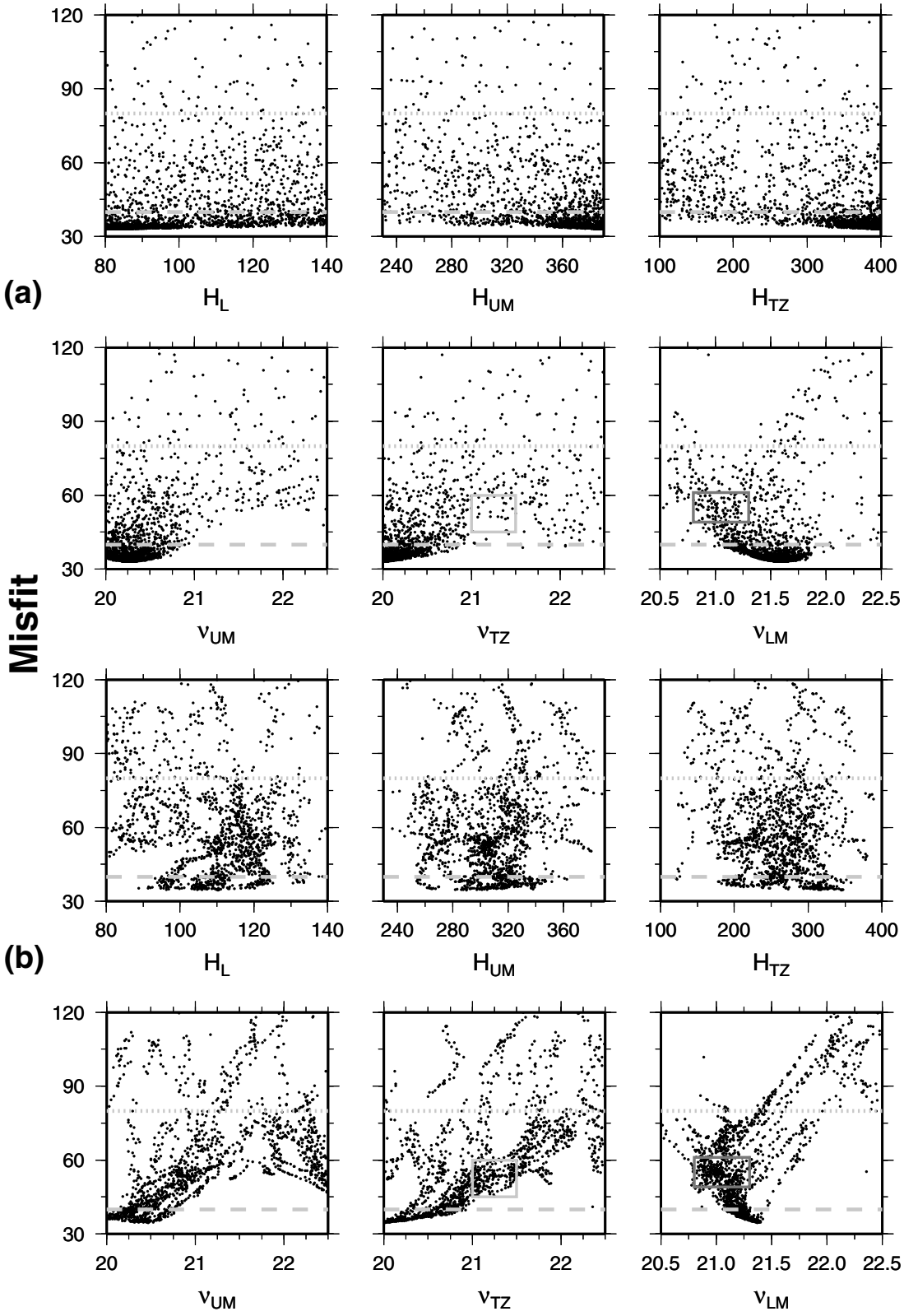

Figure 8. Misfit vs. model parameters, for the 6-D inversion of Fig. 5 using the NA (a) and SA (b). Gray dotted and dashed lines indicate $M_{\rm acc}$ and $M_{\rm con}$, respectively. See text for other details.

gray boxes in ν_{TZ} Fig. 8, where the SA sampling density is certainly excessive in such an high-misfit region. We finally observe that bottlenecks may be entry points for many minima of the misfit function and it may be wise not to leave them behind too quickly. We note from Fig. 8 that SA, after crossing the bottleneck many times, has sampled the same low-misfit regions ($M < M_{\rm con}$) as NA. Thus, we hypothesize that NA has adequately sampled the bottleneck, finding all the entry points to the region with $M < M_{\rm con}$.

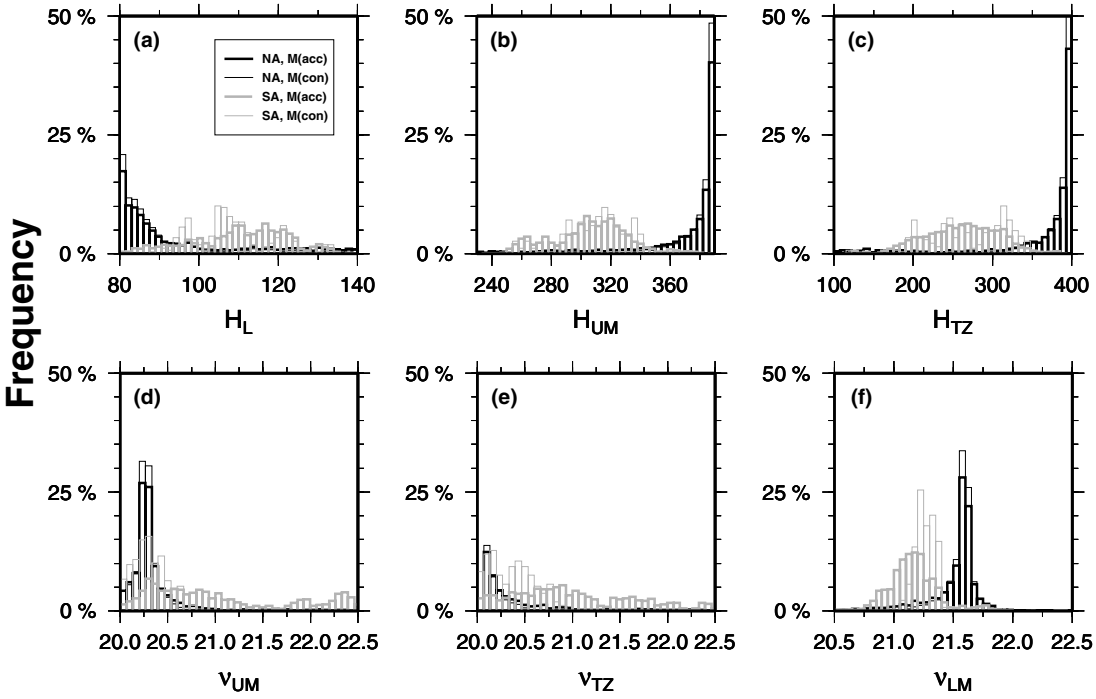

Figure 9. Frequency histograms for the six parameters in the $N_d = 6$ inversion. The frequencies are computed as the number of models falling within a given interval, divided by the total number of models with $M < M_{\rm acc}$ (thick lines) and with $M < M_{\rm con}$ (thin). Gray and black lines are for SA and NA, respectively.

The ability of NA in avoiding possible obstacles in the parameters space, which certainly constitutes an improvement with respect to SA, is the cause of the significant differences that we have found in the frequency histograms that are shown in Fig. 9. The SA histograms (gray lines) of the ν_{LM} parameter in Fig. 9 are biased towards small values by the above mentioned inefficiency of SA when obstacles are present in the misfit space. The NA solution indicates less ambiguously the specific value $\nu_{LM} = 21.6$. The upper-mantle histogram indicates that this parameter is resolved by both NA and SA, with a preferred value of ~ 20.3.

To characterize the solutions obtained from our 6-D NA inversion, we show in Fig. 10 the viscosity profiles falling below the misfit thresholds $M_{\rm acc}$ (light gray) and $M_{\rm con}$ (dark gray). According to Table 2, the number of models contained in these regions of the parameters space is 2880 and 2370, respectively. The viscosity profiles are grouped into the categories ($\nu_{UM} > \nu_{LM}$, top, and $\nu_{UM} < \nu_{LM}$, bottom) and for each of them we show the monotonous, soft transition zone and hard transition zone solutions. It is evident that at the misfit level $M_{\rm acc}$ the viscosity profile is poorly constrained because all of the six categories are densely populated. As first observed by Mitrovica & Peltier (1995) and later discussed again by CI02, the poor constraints on the rheological profile of the mantle imposed by the inversion reflect the low accuracy and inconsistency of the relative sea level available for Hudson Bay. When the models with $M < M_{\rm con}$ are considered, the solutions are characterized by a lower mantle with viscosity larger than that of the upper mantle (bottom frames). Following Shibutani *et al.* (1996), we have also shown (black) the 600 best-fitting viscosity profiles (this corresponds to 20 per cent of the number of solutions contained in the whole NA ensemble). Such a restricted family suggests an upper mantle and transition zone with viscosity logarithm well below the traditional value $\nu = 21$ (Mitrovica 1996), and a lower mantle with viscosity logarithm close to 21.3, as assumed in the construction of the ICE-3G viscosity deglaciation chronology (Tushingham & Peltier 1991). However, as a result of the data inconsistency mentioned above, our conclusions regarding the viscosity profile of the mantle have mainly a qualitative character.

The use of SA and NA inversion methods and the comparison of their performances, which was our main concern here, leads us to draw the following conclusions.

(i) In general, NA gives better performances when the details of the misfit landscape are sought. This is particularly true for explorations of multidimensional spaces ($N_d = 6$ in this study). For low-dimensional spaces ($N_d = 2$) our experience with the specific data set employed here indicate that NA and SA perform comparably well.

(ii) Unusual features related to the topology of the parameters space, such as the bottleneck in the 6-D inversion evidenced in Fig. 8, are easily circumvented by the NA as a result of its ability to concentrate the search in the most promising regions of the parameters space. SA looses much of its efficiency when obstacles are found in the models spaces.

(iii) The use of restarts in the parameters space, introduced by CI02 and also implemented in our SA searches, requires further investigations. In its present implementation, it sensibly rises the cost of the SA inversion.

(iv) According to our experience, tuning of the parameters in direct search methods is a difficult task. In SA, we have kept fixed the tuning parameters of the annealing schedule to those suggested by CI02 and we have proposed an *ad hoc* tuning for NA. It is however possible that a significant optimization could be achieved in both cases adopting more sophisticated tuning strategies.

ACKNOWLEDGMENTS

The authors thank two anonymous reviewers for their comments regarding the original manuscript and Malcolm Sambridge for

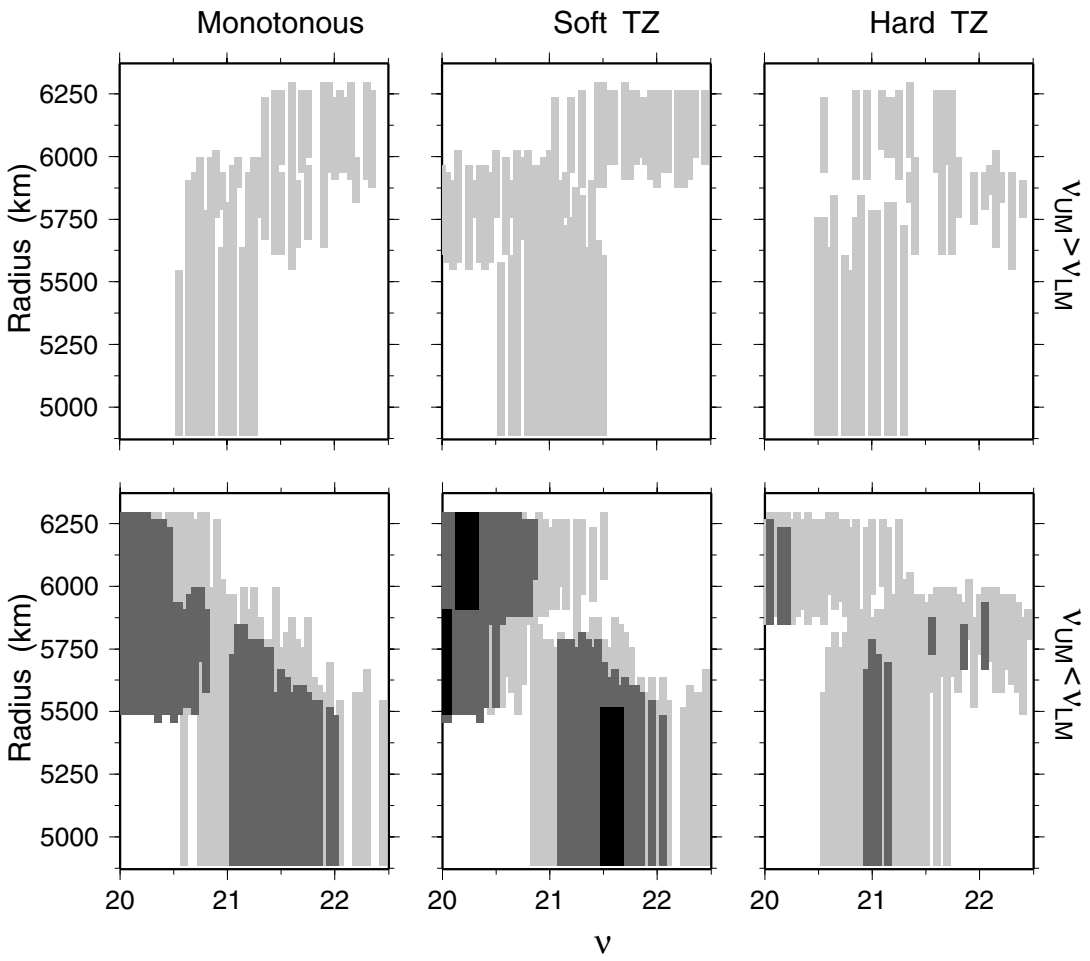

Figure 10. Mantle viscosity profiles from the 6-D NA inversion. Models with $M < M_{acc}$ and $M < M_{con}$ are shown in light and dark gray, respectively. The models portrayed in black represent the 600 best-fitting models, corresponding to 20 per cent of the entire NA ensemble.

providing us with the NA inversion code. This work has been financially supported by the Faculty of Environmental Sciences of the University of Urbino with grants 'ex-60 per cent' and by a FIRB grant of MIUR. NPA was funded by a PNRA grant. The Generic Mapping Tools software has been used for plotting the figures of this manuscript (Wessel & Smith 1998).

REFERENCES

Cianetti, S., Giunchi, C. & Spada, G., 2002. Mantle viscosity beneath the Hudson Bay: an inversion based on the Metropolis algorithm, *J. geophys. Res.,* **107,** B12, doi:10.1029/2001JB000585.

Johnston, P.J. & Lambeck, K., 2000. Automatic inference of ice models from postglacial sea level observations: theory and application to the British Isles, *J. geophys. Res.,* **105,** B6, 13 179–13 194.

King, S.D., 1995. Radial models of mantle viscosity: results from a genetic algorithm, *Geophys. J. Int.,* **122,** 725–734.

Lomax, A. & Snieder, R., 1995. Identifying sets of acceptable solutions to non-linear, geophysical inverse problems which have complicated misfit functions, *Nonlinear Processes in Geophysics,* **2,** 222–227.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, N.M., Teller, A.H. & Teller, E., 1953. Equation of state calculations by fast computing machines, *J. Chem. Phys.,* **1(6),** 1087–1092.

Milne, G.A., Mitrovica, J.X. & Forte, A.M., 1998. The sensitivity of glacial isostatic adjustment predictions to a low–viscosity layer at the base of the upper mantle, *Earth planet. Sci. Lett.,* **154,** 265–278.

Mitrovica, J.X., 1996. Haskell (1935) revisited, *J. geophys. Res.,* **101,** 555–569.

Mitrovica, J.X. & Peltier, W.R., 1995. Constraints on mantle viscosity based upon the inversion of post-glacial uplift data from the Hudson Bay region, *Geophys. J. Int.,* **122,** 353–377.

Mitrovica, J.X., Forte, A.M. & Simons, M., 2000. A reappraisal of postglacial decay times from Richmond gulf and James Bay, Canada, *Geophys. J. Int.,* **142,** 783–800.

Mosegaard, K. & Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems, *J. geophys. Res.,* **100,** B7, 12 431–12 447.

Mosegaard, K. & Tarantola, A., 2002. Probabilistic approach to inverse problems, in, *International handbook of Earthquake and engineering seismology*, pp. 237–265, eds, Lee, W.H.K. *et al.,* Academic Press, San Diego, CA.

Peltier, W.R., 1998. Postglacial variations in the level of the sea: implications for climate dynamics and solid-earth geophysics, *Rev. Geophys.,* **36(4),** 603–689.

Ricard, Y., Vigny, C. & Froidevaux, C., 1989. Mantle heterogeneities, geoid and plate motions: a Monte Carlo inversion, *J. geophys. Res.,* **94,** 13 739–13 754.

Ritzwoller, M.H., Shapiro, N.M., Levshin, A.L. & Leahy, G.M., 2001. Crustal and upper mantle structure beneath Antarctica and surrounding oceans, *J. geophys. Res.,* **106,** B12,30 645–30 670.

Sambridge, M., 1998. Exploring multidimensional landscapes without a map, *Inverse Problems,* **14,** 427–440.

Sambridge, M., 1999a. Geophysical inversion with a neighbourhood algorithm—I. Searching a parameter space, *Geophys. J. Int.,* **138,** 479–494.

Sambridge, M., 1999b. Geophysical inversion with a neighbourhood algorithm—II. Appraising the ensemble, *Geophys. J. Int.,* **138,** 727–746.

Sambridge, M., 2001. Finding acceptable models in nonlinear inverse problems using a neighbourhood algorithm, *Inverse Problems,* **17,** 387–403.

Sambridge, M. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.,* **40(3),** doi:10.1029/2000RG000089.

Shapiro, N.M. & Ritzwoller, M.H., 2002. Monte-Carlo inversion for a global shear-velocity model of the crust and upper mantle, *Geophys. J. Int.,* **151,** 88–105.

Shibutani, T., Sambridge, M. & Kennett, B., 1996. Genetic algorithm inversion for receiver functions with application to crust and uppermost mantle structure beneath eastern Australia, *Geophys. Res. Lett.,* **23(14),** 1829–1832.

Snoke, J.A. & Sambridge, M., 2002. Constraints on the *s* wave velocity structure in a continental shield from surface wave data: comparing linearized least squares inversion and the direct search neighbourhood algorithm, *J. geophys. Res.,* **107,** B5, doi:10.1029/2001JB000498.

Spada, G., 2001. Mantle viscosity from Monte Carlo inversion of very long baseline interferometry data, *J. geophys. Res.,* **106,** B8, 16 375–16 385.

Spada, G., Sabadini, R., Yuen, D.A. & Ricard, Y., 1992. Effects on postglacial rebound from the hard rheology of the transition zone, *Geophys. J. Int.,* **109,** 683–700.

Tushingham, A.M. & Peltier, W.R., 1991. ICE-3G: a new global model of late Pleistocene deglaciation based upon geophysical predictions of postglacial relative sea level change, *J. geophys. Res.,* **96,** B3, 4497–4523.

Velicogna, I. & Wahr, J., 2002. Postglacial rebound and Earth's viscosity structure from GRACE, *J. geophys. Res.,* **107,** B12, doi:10.1029/2001JB001735.

Walcott, R.I., 1972. Late quaternary vertical movements in eastern north America: quantitative evidence of glacio-isostatic rebound, *Rev. Geophys. Space Phys.,* **10,** 849–884.

Wessel, P. & Smith, W. H. F, 1998. New, improved version of the Generic Mapping Tools released, *EOS, Trans. Am. geophys. Un.,* **79,** 579.