Geophys. J. Int. (2005) **162,** 685–695 doi: 10.1111/j.1365-246X.2005.02702.x

Two-dimensional inversion of direct current resistivity data using a parallel, multi-objective genetic algorithm

Christoph Schwarzbach, Ralph-Uwe Börner and Klaus Spitzer

Institut für Geophysik, TU Bergakademie Freiberg, Gustav-Zeuner-Straße 12, 09596 *Freiberg, Germany. E-mail: schwarzb@geophysik.tu-freiberg.de*

Accepted 2005 May 27. Received 2005 May 20; in original form 2004 October 8

SUMMARY

We introduce the concept of multi-objective optimization to cast the regularized inverse direct current resistivity problem into a general formulation. This formulation is suitable for the efficient application of a genetic algorithm, which is known as a global and non-linear optimization tool. The genetic inverse algorithm generates a set of solutions reflecting the trade-off between data misfit and some measure of model features. Examination of such an ensemble is highly preferable to classical approaches where just one 'optimal' solution is examined since a better overview over the range of possible inverse models is gained. However, the computational cost to obtain this ensemble is enormous. We demonstrate that at the current state of computer performance inversion of 2-D direct current resistivity data using genetic algorithms is possible if state-of-the-art computational techniques such as parallelization and efficient 2-D forward operators are applied.

Key words: electrical resistivity, finite-difference methods, numerical techniques, stochastic inversion.

1 INTRODUCTION

The geophysical inverse problem can be described as the task to deduce the subsurface distribution of physical properties from physical measurements which are made at the earth's surface or within the subsurface. This deduction process may be divided into two parts (Scales & Snieder 2000)—estimating inverse models from a number of given data and appraising the inverse models with respect to the unknown true earth model. Solutions of the estimation problem in the direct current resistivity case are non-unique due to a finite number of measurements (a finite number of both current injection and potential sampling points), restricted access to the boundary of the volume under investigation as well as finite precision of data aquisition (Friedel 2003). In short, there exists a set of inverse models all of which explain a given data set within error limits. The estimation of this set of inverse models can be achieved either statistically (Bayesian approach) or deterministically. The deterministic approach can be characterized by a search for those inverse models that aim at minimizing data misfit and—possibly—model constraints taking into account *a priori* information or assumptions about smoothness, least entropy, least structure, etc. We will only consider the deterministic view of the estimation problem here.

The minimization task of the inverse problem is performed using a genetic algorithm. Genetic algorithms can be classified as non-linear, global and stochastic optimization schemes which are inspired by processes of biological evolution and genetics. For a general description see, for example, the monographs of Holland (1975), Davis (1996) and Deb (2001). Application of a genetic algorithm has several advantages over classical optimization methods. Whereas classical methods such as quasi-Newton or conjugate gradients require the calculation of derivatives—the Jacobian of the model-data mapping—as well as an appropriate initial guess, genetic algorithms do not require these. Using derivative information, classical minimization techniques are very efficient to locate a minimum with high accuracy but at the same time are prone to be trapped in suboptimal points, that is, to fail the global minimum of a multi-modal function. Genetic algorithms, on the other hand, are a more global and robust but much slower approach. Furthermore, they can be adjusted to find a set of solutions to multi-modal and even multi-objective optimization tasks. An optimization problem is called multi-objective if a number of possibly conflicting scalar objective functions are to be optimized simultaneously. In contrast to the genetic algorithm, the classical approach only yields one solution corresponding to one local minimum and a particular linear combination of the objective function's vector components. Genetic algorithms, therefore, open the possibility to gain a better overview over the range of inverse models and to take a further step on the way from estimating a single inverse model to appraising the ensemble of inverse models.

Genetic algorithms have been applied in a number of geophysical disciplines—first of all in seismics (Stoffa & Sen 1991; Gallagher *et al.* 1991; Wilson & Vasudevan 1991; Sen & Stoffa 1992; Kennett & Sambridge 1992), later for resistivity sounding inversion (Sen *et al.* 1993) and the inversion of magnetotelluric (Everett & Schultz 1993), gravity and magnetic data (Boschetti *et al.* 1997). For further references see, for example, Gallagher & Sambridge (1994) and

Sambridge & Mosegaard (2002). In resistivity and magnetotelluric methods, application of genetic algorithms to the inverse problem was restricted to models with a very limited number of parameters and models which could be evaluated using fast, semi-analytical algorithms. More complex models require numerical techniques, that is, finite difference, finite element or integral equation methods. Using finite difference modelling, Chunduru *et al.* (1996, 1997) compared a genetic algorithm with simulated annealing and found the simulated annealing to be more efficient. Only recently, the combination of a genetic algorithm with a 2-D magnetotelluric finite difference modelling code has been reported (Pérez-Flores & Schultz 2002).

In this article we present a fully non-linear inverse algorithm to find an ensemble of inverse models. We introduce the concept of multi-objective optimization to the direct current resistivity problem so as to treat the simultaneous minimization of data misfit and model restriction in an as much general as possible way. The inverse algorithm combines a genetic algorithm and a 2-D finite difference modelling code. Whereas Chunduru *et al.* (1996, 1997) used a spline parameterization to reduce the number of parameters to be determined during inversion, we face the challenging task to search for the block parameters corresponding to the underlying finite difference grid directly.

The multi-objective formulation of the inverse problem will be the subject of the following section. Next, the genetic algorithm will be outlined followed by the description of refinements we applied to our finite difference forward modelling code in order to enhance speed and accuracy. The last section will present two synthetic data examples the genetic inversion algorithm was applied to and conclude with inversion results of a resistivity field data set.

2 THE INVERSE PROBLEM

Estimating inverse models from resistivity data can be described as an inverse modelling process: Take an initial guess of the subsurface conductivity distribution, calculate this model's hypothetical resistivity data set, and refine the model until synthetic data fit measured data up to a certain degree depending on the noise level of the measured data. Assuming Gaussian distributed noise, data misfit can be expressed as

$$f_1(\mathbf{m}) = \left\| \frac{\mathbf{d} - \mathbf{F}(\mathbf{m})}{\sqrt{n-1}\,\mathbf{s}} \right\|_{\ell_p} = \left(\sum_{i=1}^{n} \left| \frac{d_i - F_i(\mathbf{m})}{\sqrt{n-1}\,s_i} \right|^p \right)^{1/p}, \quad 1 \le p < \infty, \tag{1}$$

where $\mathbf{m}$ denotes the model parameters, $\mathbf{d}$ measured data, $\mathbf{s}$ their standard deviation, n the number of data points, and $\mathbf{F}(\mathbf{m})$ synthetic data as calculated by the model-data mapping. Note that the expectation of $[f_1(\mathbf{m})]^2$ is 1 if $p = 2$. We used logarithms of apparent resistivity, $\mathbf{d} = \log \boldsymbol{\rho_a}$, and of model conductivity, $\mathbf{m} = \log \boldsymbol{\sigma}$, throughout this study. Estimated standard deviation $\boldsymbol{\Delta\rho_a}$ is transformed according to $\mathbf{s} = \log(1 + \boldsymbol{\Delta\rho_a}/\boldsymbol{\rho_a})$ (Friedel 2003). The assumption of Gaussian distributed noise is no longer valid under this transformation.

Because of data noise, all inverse models which satisfy $f_1(\mathbf{m}) \le const.$ should be accepted where *const*. depends on $\mathbf{s}$. There is an infinite number of solutions to this problem. Now, either a statistical approach can be used to characterize the ensemble of inverse models or, following the deterministic view, one model with particular additional properties to be defined can be searched. The deterministic approach is applied here and achieved by

(1) replacing $f_1(\mathbf{m}) \le const.$ by minimization with respect to $\mathbf{m}$, $f_1(\mathbf{m}) \to \min!$ and

(2) restricting solutions to a certain class of models which may satisfy additional constraints.

Model space is restricted in the following ways:

(1) Since a finite difference algorithm is used to compute the synthetic data, the subsurface conductivity model is parameterized accordingly by piecewise constant conductivity blocks where the block conductivity is variable and block geometry is fixed. (Hence, the model is denoted by a vector $\mathbf{m}$.)

(2) Due to limited resolution provided by the data, it is sufficient to consider a finite and discrete set of 'representative' conductivity values only, much less than could be expressed by the standard computer floating point variables. This fact will be exploited for the use of encodings within the genetic algorithm.

(3) Minimizing data misfit only can lead to misinterpretation because noise can be mistaken as signal. Thus, a second objective function $f_2(\mathbf{m})$ is introduced and will be simultaneously minimized. f_2 can account, for example, for deviations from an initial model adding *a priori* knowledge or for a measure of roughness adding certain assumptions about a particular conductivity distribution. Following the classical approach, roughness is calculated by a first- or second-order finite difference operator applied to the discrete conductivity distribution and finally, the ℓ_2-norm is taken. This writes for the first finite difference with normalized distances and 2-D models ($\partial\sigma/\partial y \equiv 0$) of piecewise constant conductivity on a rectangular grid as

$$f_2^{(a)}(\mathbf{m}) = \left(\sum_{i=2}^{p} \sum_{j=1}^{q} (m_{i,j} - m_{i-1,j})^2 + \sum_{i=1}^{p} \sum_{j=2}^{q} (m_{i,j} - m_{i,j-1})^2 \right)^{1/2}, \tag{2}$$

where model vector $\mathbf{m}$ has been rearranged to an array and contains the logarithms of block conductivity, $m_{i,j} = \log \sigma(x, z)$ for $x_i \le x \le x_{i+1}, i = 1, \ldots, p, z_j \le z \le z_{j+1}, j = 1, \ldots, q$. Further, we propose an ℓ_1-measure which is derived by taking the first derivative of the non-discrete and logarithmized conductivity distribution $m(\mathbf{r}) = \log \sigma(\mathbf{r})$ and integrating the absolute value over the considered 2-D model domain Ω,

$$f_2^{(b)}(m) = \int_\Omega \left(\left| \frac{\partial m}{\partial x} \right| + \left| \frac{\partial m}{\partial z} \right| \right) dx\, dz. \tag{3}$$

This extends for the rectangular grid with piecewise constant σ as described above to

$$f_2^{(b)}(m) = \sum_{j=1}^{q} \Delta z_j \sum_{i=2}^{p} |m_{i,j} - m_{i-1,j}| + \sum_{i=1}^{p} \Delta x_i \sum_{j=2}^{q} |m_{i,j} - m_{i,j-1}|, \tag{4}$$

where $\Delta x_i = x_{i+1} - x_i$ and $\Delta z_j = z_{j+1} - z_j$. To see this, consider, for example, the vertical edge x_i at $z_j < z < z_{j+1}$. There, m can be written locally as

$$m = m_{i-1,j} + (m_{i,j} - m_{i-1,j}) H(x - x_i) \tag{5}$$

using the unit step function $H(x)$. Since the formal derivative of $H(x)$ is the Dirac delta function, we have

$$\partial m/\partial x = (m_{i,j} - m_{i-1,j}) \delta(x - x_i). \tag{6}$$

Taking the absolute value does not affect $\delta(x)$. Integrating over a small area around the edge, we obtain

$$\int_{z_j}^{z_{j+1}} \int_{x_i-\varepsilon}^{x_i+\varepsilon} \left| \frac{\partial m}{\partial x} \right| dx\, dz$$
$$= |m_{i,j} - m_{i-1,j}| \int_{z_j}^{z_{j+1}} \int_{x_i-\varepsilon}^{x_i+\varepsilon} \delta(x - x_i)\, dx\, dz$$
$$= |m_{i,j} - m_{i-1,j}| \Delta z_j \tag{7}$$

and similarly the contributions of all other edges. Due to the absolute value used, $f_2^{(b)}$ yields the same result for transitions between two blocks with different conductivities either taking one large step or a number of smaller steps. Only the effective step magnitude is penalized.

The two objective functions data misfit $f_1(\mathbf{m})$ and model restriction norm $f_2(\mathbf{m})$ are to be minimized simultaneously. In general, this can be expressed as a multi-objective optimization task or as the task of optimizing a vector-valued functional

$$\mathbf{f}(\mathbf{m}) = (f_1(\mathbf{m}), f_2(\mathbf{m})) \rightarrow \min!. \tag{8}$$

$f_1(\mathbf{m})$ as well as $f_2(\mathbf{m})$ depend non-linearly on their arguments since the mapping of model conductivity $\boldsymbol{\sigma} = \exp(\mathbf{m})$ to synthetic data $\boldsymbol{\rho_a} = \exp(\mathbf{d})$ is non-linear and norms other than the ℓ_2-norm lead to non-linear optimization tasks.

The most popular approach for solving eq. (8) is to introduce a regularization parameter λ and to minimize a scalar-valued functional of the type $f(\mathbf{m}) = f_1(\mathbf{m}) + \lambda f_2(\mathbf{m})$. For Newton-type minimization, $f(\mathbf{m}) = [f_1(\mathbf{m})]^2 + \lambda[f_2(\mathbf{m})]^2$ is usually used in case f_1 and f_2 denote ℓ_2-norm quantities (Constable *et al.* 1987; Park & Van 1991; Farquharson & Oldenburg 2004). This can be interpreted just as a linear combination of both objective functions with linear coefficients 1 and λ. By varying λ from 0 to ∞, a family of solutions can be obtained with f_1 and f_2 given different weights. Various approaches exist to choose an optimal value of λ like, for example, the L-curve criterion, discrepancy principle, and generalized cross-validation (Engl *et al.* 2000; Günther 2004) whereas the particular choice is still a subject of investigation (Farquharson & Oldenburg 2004). Nevertheless, a better overview over the variety of possible inverse solutions and a deeper insight into the nature of the inverse problem can be achieved if not just one optimal solution is considered but a complete set.

To obtain the whole family of solutions, minimization for a number of fixed values of λ has to be performed. Using a genetic algorithm for the minimization task, we wish to obtain the solution of problem (8) independently of λ, that is, to tackle problem (8) directly. Therefore, it is necessary to define the comparison of vectors. Vector $\mathbf{f}^* = (f_1^*, f_2^*, \ldots, f_n^*)$ is said to be less than vector $\mathbf{f} = (f_1, f_2, \ldots, f_n)$,

$$\mathbf{f}^* \prec \mathbf{f}, \quad \text{if} \begin{cases} f_i^* \leq f_i & \text{for } i = 1, \ldots, n \text{ and} \\ f_i^* < f_i & \text{for at least one } i. \end{cases} \tag{9}$$

Let $S \subseteq \mathcal{D}(\mathbf{f})$ where $\mathcal{D}$ denotes the domain of $\mathbf{f}$ and $\mathbf{m}^* \in S$. If there exists no $\mathbf{m} \in S$ such that $\mathbf{f}(\mathbf{m}) \prec \mathbf{f}(\mathbf{m}^*)$, then $\mathbf{m}^*$ is called a Pareto-optimal solution of problem (8) (Fig. 1). The set of all Pareto-optimal solutions

$$\{\mathbf{m}^* \in S : \nexists \mathbf{m} \in S \text{ such that } \mathbf{f}(\mathbf{m}) \prec \mathbf{f}(\mathbf{m}^*)\} \tag{10}$$

is called Pareto set. The Pareto set is termed global if $S \equiv \mathcal{D}(\mathbf{f})$ and local otherwise. If objective functions are contradictory to each other, the Pareto set contains a multitude of solutions, which represent optimal compromises between conflicting objective functions.

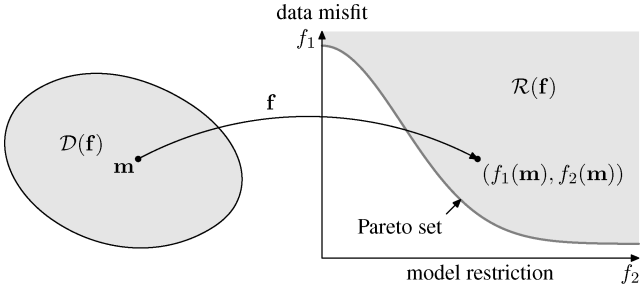

Figure 1. $\mathbf{f}$ is a mapping from model space to the vector space spanned by data misfit f_1 and model restriction norm f_2, more precisely from the domain $\mathcal{D}(\mathbf{f})$ to the range $\mathcal{R}(\mathbf{f})$.

An approximation to the global Pareto set—a local Pareto set—can be obtained with the *Non-dominated Sorting Genetic Algorithm* NSGA-2 (Deb *et al.* 2000) described in the next section. Thus, linear combination of the objective function's vector components and linearization of non-linear objective functions—particularly of the non-linear model-data mapping—can be avoided.

3 THE GENETIC ALGORITHM

Genetic algorithms were developed by J. Holland in the 1960/1970s (Holland 1975). They can be classified as one species of stochastic optimization methods known as evolutionary algorithms. These algorithms transfer principles and processes of biological evolution, like Darwin's survival of the fittest, genetics and adaptation of life to its environment, onto artificial systems. Genetic algorithms operate on a number of encoded representatives of points in model space which form a population of strings. One string is, for example, the concatenation of 8-bit sequences where each sequence represents the model parameter of one model block and encodes 2^8 resistivity values. Strings are assigned a fitness value that is calculated decoding the string, evaluating the objective function and comparing the resulting values of all strings in the current population. A number of new strings is formed by the application of the genetic operators selection, crossover and mutation (Fig. 2). The selection operator chooses strings for further operation giving strings with higher fitness values a higher probability to be selected. During crossover, pairs of selected strings exchange parts of their information. Mutation finally alters bits of information within a typically small number of strings. All newly created strings again are decoded, evaluated and assigned a fitness value. Taking individuals from the pool of old and new strings, a new population is formed and the whole process is repeated. Since survival of the fittest strings of the old or combined population cannot be guaranteed using stochastic genetic operators only, elitist strategies may be employed to ensure this. The artificial evolution is terminated if sufficiently good solutions are found, a predefined number of generations or objective function evaluations is reached, or search stagnates.

To reduce the objective function values of a multi-objective optimization problem to a scalar fitness value, Deb *et al.* (2000) apply a scheme which sorts strings according to their degree of local non-domination. A string S_i is said to be non-dominated if its corresponding objective function $\mathbf{f}(S_i) \prec \mathbf{f}(S_j)$ for all other strings S_j, $j \neq i$, of the considered set of strings (Fig. 3). In fact, S_i is locally non-dominated with respect to the domain $\mathcal{D}(\mathbf{f})$ since the population represents only a subset of $\mathcal{D}(\mathbf{f})$. To perform the sort, first, all non-dominated strings of the population are determined. They are given the rank '1' and called the first non-dominated set (S_2, S_3, S_6 in

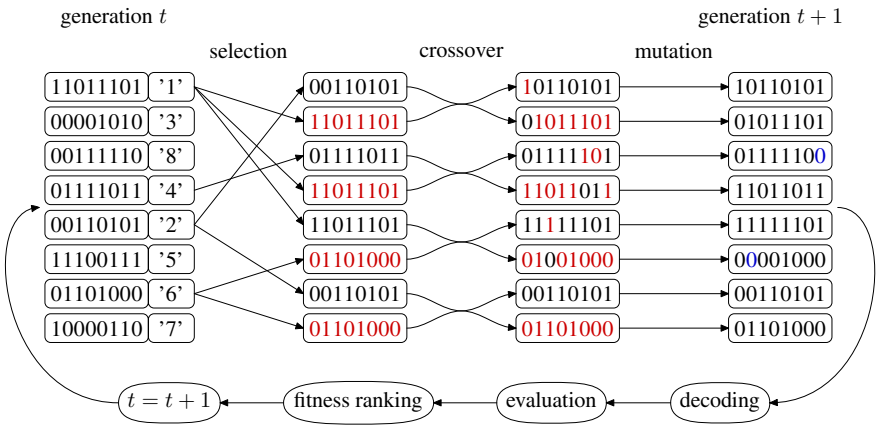

Figure 2. Scheme of a genetic algorithm applied to a population of eight binary encoded strings. Fitness values enclosed in inverted commas are each string's rank. Ranks are assigned according to the objective function value(s) calculated after decoding each string. Note that a particular, on average fitter string like no. '3' is not guaranteed to be selected since the selection operator used is a stochastic process as are crossover and mutation as well.

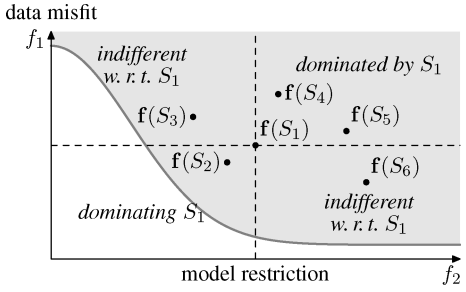

Figure 3. Subdivision of the $f_1 - f_2$-plane into dominated, dominating, and indifferent parts with respect to string S_1. Among the six strings S_1 to S_6, strings S_2, S_3 and S_6 are non-dominated and, therefore, form the local Pareto set. Note that even though string S_6 is non-dominated it does not dominate any of the other strings. The non-dominated sorting procedure would assign rank '1' to strings S_2, S_3, S_6, rank '2' to string S_1, and rank '3' to strings S_4, S_5.

Fig. 3). This procedure is repeated for the remaining strings which are not assigned a rank, yet. Thus, rank '2' is given for the second non-dominated set (S_1 in Fig. 3), rank '3' for the third one (S_4, S_5 in Fig. 3), and so on until all strings are assigned a rank. In addition to the ranks assigned according to non-domination, a second, auxiliary fitness value is calculated to distinguish between individuals of the same rank and to obtain as diverse as possible solutions in model or objective function space. The auxiliary fitness is either the distance of a string to its neighbours in objective function space or to all other strings in model space. Comparing two strings, first, the ranks are considered, and second, the auxiliary fitness is used if ranks are equal. Using this procedure coupled with an elitist strategy, the genetic algorithm generates a series of local Pareto sets—local with respect to all strings considered so far—which approximate the global Pareto set increasingly well.

Based on the NSGA-2 (Deb *et al.* 2000) just described, we implemented our genetic inversion algorithm. Apart from coding the problem-specific objective functions (Section 2), restructuring, tuning and porting the C-code of NSGA-2 to C++, a number of principal changes were made to the genetic algorithm to improve efficiency. Integration of the controlled elitism described in Deb & Goel (2000) guarantees an appropriate distribution of strings within each rank, especially within the first one, as well as among the different ranks. Binary Gray-encoding replaces the standard binary encoding and helps to avoid artifacts which arise due to the discrete approximation of the underlying continuous optimization problem (Rana 1999). Parallelization was crucial to render the genetic inversion algorithm feasible for 2-D inversion at the current state of computer power. To obtain solutions of sufficiently high quality, it is necessary to use relatively large populations within the genetic algorithm. Unfortunately, large populations require not only a large number of objective function evaluations per generation but also a large number of generations to reach convergency. Therefore, the global population is split into a number of subpopulations, which evolve independently and exchange their best individuals in regular intervals. Thus, a faster convergency rate is obtained using relatively small subpopulations whereas the larger total population size guarantees sufficient sampling of model space. Each subpopulation can additionally be forced to focus its search on a part of the Pareto set (Deb *et al.* 2002). Further, all n subpopulations are assigned an even share p/n of the number p of available processing units—considering only the case where p is a multiple of n. Since the actual genetic algorithm, that is, genetic operators and non-dominated sorting, has neglible run times compared to the objective function evaluations, it is only carried out on one of each subpopulation's processors (sequentially with respect to each subpopulation but in parallel with respect to the global population) while all processors are used to evaluate the objective functions. Exchange of individuals and solutions between subpopulations is implemented using the *Message Passing Interface* MPI (Gropp *et al.* 1996).

For the purpose of migration between the subpopulations, they are assigned vertices on a hypercube of dimension $\log_2 n$ where the number of subpopulations n is an integer power of 2. During migration, each subpopulation exchanges non-dominated individuals with the subpopulation at the neighbouring vertex in the direction of dimension i where i cycles through all dimensions of the hypercube. Since it is difficult to make the optimal choice for the genetic algorithm's parameters crossover and mutation rate and since long run times of real world problems prohibit appropriate tests to find optimal values, they are generated at random. Each subpopulation gets its own set of parameters and is thus exposed to slightly different environmental conditions. To extract the last locally non-dominated set of solutions, all subpopulations are gathered to form one large population, which is sorted according to the degree of non-domination.

Tests regarding different numbers of subpopulations and the focussed search were not conclusive (Section 5). This is probably due to a still too small population size compared with the number of unknowns.

4 MODELLING

To evaluate the first objective function, data misfit, calculation of synthetic data is required. To make feasible the large number of model calculations that are required for the genetic inversion algorithm, an especially adapted 2-D finite difference direct current forward operator has been created. First, a singularity removal technique is combined with boundary conditions imposed on the secondary potential. Second, Gauss-quadrature formulae are applied to compute the inverse Fourier transform. Third, source location loop, wavenumber loop and solver for the system of linear equations are arranged in such a way that optimum speed and accuracy for multiple sources are obtained. See Dey & Morrison (1979) for the basic finite difference formulation.

4.1 Singularity removal technique

We apply the singularity removal technique (Lowry *et al.* 1989; Zhao & Yedlin 1996; Spitzer *et al.* 1999; Li & Spitzer 2002) described for the 3-D case therein to the 2-D case here. Assuming that the conductivity σ is constant with respect to one spatial direction, say y, the quasistatic electrical potential φ is searched in the form of a Fourier cosine transform

$$\varphi(x, y, z) = \frac{2}{\pi}\int_0^\infty \tilde{\varphi}(x, k_y, z)\cos(k_y y)\, dk_y. \tag{11}$$

The transformed potential $\tilde{\varphi}$ is split into a normal and an anomalous part, $\tilde{\varphi} = \tilde{\varphi}_n + \tilde{\varphi}_a$. The boundary value problem is formulated in terms of the anomalous potential $\tilde{\varphi}_a$ for a finite domain Ω with homogeneous Neumann boundary conditions at the air–earth interface Γ_0 and homogeneous mixed boundary conditions on the subsurface boundary Γ_1 ($\Gamma_0 \cup \Gamma_1 = \partial\Omega$)

$$-\nabla\cdot(\sigma\nabla\tilde{\varphi}_a) + k_y^2\sigma\tilde{\varphi}_a = \nabla\cdot(\sigma_a\nabla\tilde{\varphi}_n) - k_y^2\sigma_a\tilde{\varphi}_n \quad \text{on } \Omega \tag{12}$$

$$\frac{\partial\tilde{\varphi}_a}{\partial n} + \alpha\,\tilde{\varphi}_a = 0 \quad \text{on } \partial\Omega, \tag{13}$$

where $\sigma_a = \sigma - \sigma_n$, $\alpha = 0$ on Γ_0, and

$$\alpha = k_y\frac{K_1(k_y\sqrt{(x-x_s)^2+(z-z_s)^2})}{K_0(k_y\sqrt{(x-x_s)^2+(z-z_s)^2})}\cos\theta \ \text{ on } \Gamma_1. \tag{14}$$

K_0 and K_1 are the modified Bessel functions of order 0 and 1. θ denotes the angle between the outward normal direction on the boundary and the vector from the source position (x_s, z_s) to the boundary point. The normal conductivity σ_n is chosen such that the normal potential $\tilde{\varphi}_n$ can be computed analytically and that $\sigma_n = \sigma$ in the neighbourhood of the source point. For a surface source placed on the node of a rectangular, 2-D finite difference grid, the normal potential is given by

$$\tilde{\varphi}_n = \frac{I}{2\pi\bar{\sigma}}K_0(k_y\sqrt{(x-x_s)^2+(z-z_s)^2}), \tag{15}$$

where I denotes current and $\bar{\sigma} = (\sigma_1 + \sigma_2)/2$ with the quarter space normal conductivities σ_1, σ_2 to the left and right of the source, respectively. This choice of background conductivity guarantees that the singular behaviour of the anomalous potential at the source is exactly removed: The potential is split into a singular part with respect to the source location, $\lim_{(x,z)\to(x_s,z_s)}\tilde{\varphi}_n = \infty$, and a non-singular part, $\lim_{(x,z)\to(x_s,z_s)}\tilde{\varphi}_a < \infty$. Thus, the numerical effort to discretize large potential gradients at the source is reduced. As the anomalous potential defined this way is more regular than the total potential, it is better suited to numerical approximation. Apart from this, numerical errors are reduced since they are constrained to the anomalous potential, which is usually significantly smaller in magnitude than the normal potential.

Various finite difference discretizations have been applied to the 2-D and 3-D resistivity problems (for example Dey & Morrison 1979; Mundry 1984; Spitzer & Wurmstich 1999). We adapted the area discretization approach of Dey & Morrison (1979) to the problem (12), (13) stated in terms of the split potential. The resulting system of linear equations can be written as

$$\mathbf{C}_\sigma\tilde{\boldsymbol{\varphi}}_a = \mathbf{C}_{\sigma_a}\tilde{\boldsymbol{\varphi}}_n + \mathbf{b}, \tag{16}$$

where $\tilde{\boldsymbol{\varphi}}_a$ and $\tilde{\boldsymbol{\varphi}}_n$ are vectors of anomalous and normal potentials, respectively. The matrices $\mathbf{C}_\sigma$ and $\mathbf{C}_{\sigma_a}$ contain the coupling coefficients in terms of conductivities σ and σ_a. Vector $\mathbf{b}$ contains contributions of the non-vanishing normal derivative of the normal potential at the subsurface boundary giving rise to additional terms at the boundary nodes. Eq. (16) is solved for the anomalous potential by exploiting sparsity, symmetry and positive definiteness of matrix $\mathbf{C}_\sigma$.

4.2 Inverse Fourier transform

The inverse Fourier transform eq. (11) is applied to the anomalous potential only. The analytically computed normal potential is added at the last step. Thus, integration errors for the normal potential are avoided. Further, we consider only the most common case of measurements along a profile perpendicular to the strike direction. In this case, $y = 0$ can be chosen such that the cosine term in eq. (11) reduces to one and only a simple integration needs to be carried out. Similarly to LaBrecque *et al.* (1996), this integration is split into two parts. For small wavenumbers $0 \le k_y \le 1/(2\Delta_{\min})$, where $\Delta_{\min}$ denotes the smallest cell size, the anomalous potential $\tilde{\varphi}_a$ as a function of wavenumber k_y shows steep gradients and oscillations. This part is approximately integrated by Gauss–Legendre quadrature. For $1/(2\Delta_{\min}) \le k_y \le \infty$, $\tilde{\varphi}_a$ decreases exponentially and is integrated by Gauss–Laguerre quadrature.

4.3 Implementation—speed and accuracy

Gauss quadrature formulae provide both quadrature weights and abscissas, that is, weights and the values of wavenumbers k_y. The total number of wavenumbers as well as partitioning wavenumbers between Gauss–Legendre and Gauss–Laguerre quadrature control accuracy of integration and speed of the algorithm. Sixteen wavenumbers, eight Gauss–Legendre abscissas and eight Gauss–Laguerre abscissas proved to be an acceptable trade-off between accuracy and speed. An efficient arrangement of wavenumber loop, source location loop and solver for the system of linear equations is essential for optimum speed for multiple sources. If the loop over all source locations is placed within the wavenumber loop, the matrix of coupling coefficients $\mathbf{C}_\sigma$ needs to be pre-conditioned or factored only once per wavenumber. However, using the mixed boundary condition as given by eqs (13) and (14) the coupling coefficients depend on the particular source location. We relaxed this boundary condition and replaced the source location (x_s, z_s) in α by a source-independent, common reference point to exploit the effort of the matrix factorization.

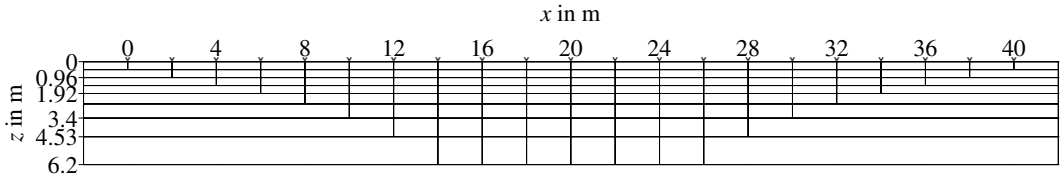

Figure 4. Subdivision of the subsurface into 120 blocks as used for the inversion. The 21 electrodes are indicated by '∨'.

The most widely used solver for the symmetric positive definite system of linear equations is the SSOR- (Spitzer 1995) or IC- (Zhang *et al.* 1995) pre-conditioned conjugate gradient method (see also Spitzer & Wurmstich 1999). Although not suitable for larger systems of linear equations arising in 3-D problems, for our 2-D problem, a direct solver consisting of in-place Cholesky decomposition and backward-/forward substitution proved to be superior to pre-conditioned conjugate gradients in case of multiple sources, that is, a number of right-hand sides.

5 APPLICATION

In this section, we present inversion results obtained by the genetic inversion algorithm that was applied to two synthetic data sets and one field data set. All data sets consist of a complete set of 114 dipole–dipole measurements using the first eight dipole separations on a linear 21-electrodes array of 40 m length.

For inversion, a finite area of the subsurface was parameterized by 120 blocks (Fig. 4). Each block was assigned a single value of constant resistivity which was to be found by the inverse process. To improve accuracy of the forward calculation, the grid given by the block parameterization was refined, and cells with increasing size were added at the subsurface boundaries. Therewith, a grid of 110×12 cells was used for calculating the synthetic data. Within the genetic algorithm, each block's resistivity was encoded by an 8-bit sequence of the binary string such that the 2^8 representable values were equidistantly spaced between 10 and 10 000 Ωm on a logarithmic scale. Thus, block resistivity is discretized such that adjoining values differ by about 1 per cent. This provides a discretization which approximately fits the resolutional power of data and is fine enough not to suppress important details of the objective function. At the same time, search is restricted to a reasonable number of points in model space. The resulting total string length is 120×8 bits $=$ 960 bits.

With one exception, the genetic algorithm was run with a total population size of 8192 individuals in all cases. The initial population was generated at random and evolved for 2048 generations. Thus, at most 8192 individuals $\times$ 2048 generations $= 2^{24}$ points were tested by the genetic algorithm, that is, a fraction of ca. 1.7×10^{-282} of all 2^{960} different conductivity distributions that a binary string of length 960 can encode. One inversion required a total of 1580–1740 hr CPU time and 490–690 MBytes memory on an SGI® Altix™ 3700 (32 Intel®-Itanium™-2 processors at 1.3 GHz, 128 GB shared memory). By using 24 or 16 processors, the processing time could be reduced to 66–100 hr. In the one exceptional case a twice as large population was evolved for 1024 generations. Even though this resulted in about the same number of tested points, CPU time was increased to 2880 hr.

Two special choices of objective functions have been considered. First, the familiar ℓ_2-measures of data misfit, eq. (1) with $p = 2$, and model restriction, eq. (2), have been used. To illustrate the variety of possible solutions to the inverse problem, we further present results obtained by ℓ_1-measures of data misfit, eq. (1) with $p = 1$, and model restriction, eq. (4). Both measures required roughly the same numerical effort demonstrating the potential of the genetic algorithm as a non-linear optimizer.

5.1 Synthetic data examples

Two synthetic data sets were generated using the same modelling code and parameters as for the inversion. The first model is a half-space of 100 Ωm containing a resistive and a conductive block of 200 Ωm and 50 Ωm, respectively, which have an extension of one model cell length and two cells height (Fig. 5a). The second model consists of two quarterspaces of 100 Ωm and 200 Ωm, respectively, with an overburden of 50 Ωm (Fig. 6a). In the following, these models will be referred to by the term synthetic models. The synthetic data—apparent resistivity ρ_a—generated from the synthetic models were disturbed by Gaussian noise of zero mean and standard deviation $\Delta\rho_a$ according to Friedel (2003),

$$\Delta\rho_a = \mu\rho_a + k\frac{U_{\min}}{I}, \tag{17}$$

where $\mu = 1$ per cent is a constant noise level, k the configuration factor, $U_{\min} = 0.1$ mV the voltage noise level, and $I = 100$ mA the driving current.

Inversion using the ℓ_2-norm objective functions was performed with the global population split into eight subpopulations with 1024 individuals each. Migration took place every sixteenth generation. For the quarterspace model, each subpopulation was set to search only for a part of the Pareto set (focussing search, *cf.* Section 3) during the first 1024 generations. No focussing was accomplished for the other 1024 generations as well as for the block model. The other option in each case and, additionally, trying one large population without subdivision produced slightly inferior results. In contrast to that, best results were obtained for the ℓ_1-norm inversion of the block model data set using one single population of 8192 individuals. For the quarterspace model and the ℓ_1-norm, we present results from a population of size 16 384 at generation 1024. In spite of the doubled population size, these results represent suboptimal solutions as will be discussed later on.

All non-dominated solutions from the final generation—the last local Pareto set—are taken as an approximation to the unknown global Pareto set. Fig. 5(b) plots the objective function values—data misfit f_1 versus model restriction f_2—of the approximated Pareto set obtained from the block model data ℓ_2-norm inversion, Fig. 5(c) correspondingly for the ℓ_1-norm inversion. A star further indicates those objective function values which are obtained evaluating the respective objective functions with the synthetic model from Fig. 5(a) as used to generate the synthetic data. In the ℓ_2-norm case, data misfit $f_1^{\ell_2}$ evaluated for the synthetic model $\mathbf{m}^*$ reduces to the sample standard deviation of the pseudo-random sequence $\{\nu_1, \ldots, \nu_n\}$ of Gaussian distributed numbers which has been used to disturb the data. This can be seen substituting the noisy data $d_i = F_i(\mathbf{m}^*) + \nu_i s_i$ $(i = 1, \ldots, n)$ into eq. (1). With $f_1^{\ell_2} = 0.599$, the sample standard deviation $[f_1^{\ell_2}]^2 = 0.359$ clearly deviates from the expectation value of 1. This fact is attributed to the statistically small number of samples (114 data points). The gaps within the Pareto sets of Figs 5(b) and (c) (similarly for Figs 6c and 7c) might be

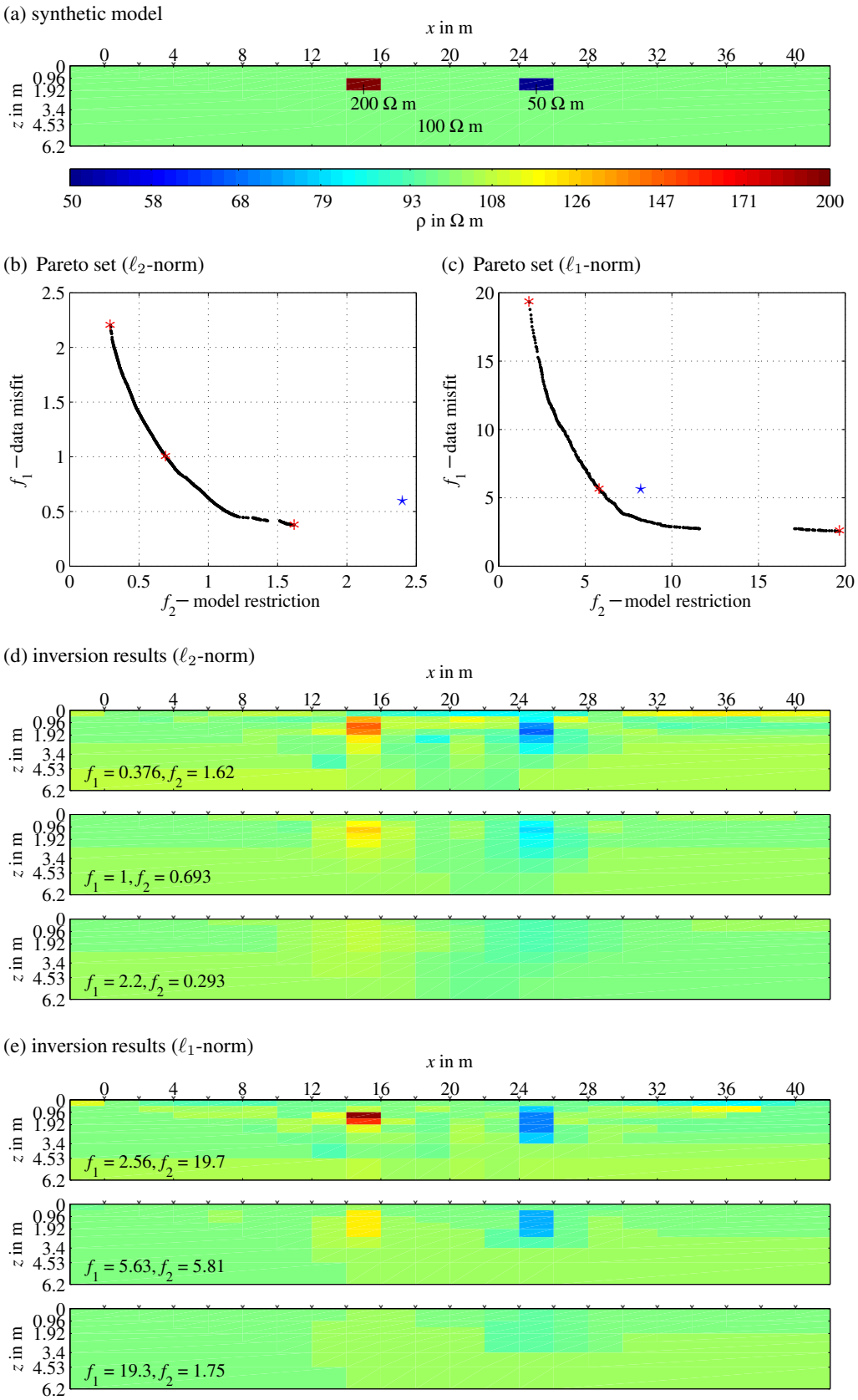

Figure 5. 'Block model' synthetic data inversion example. (a) Model to generate the data. Electrode positions are indicated by '∨'. (b), (c) Objective function values of the Pareto sets obtained from the genetic inversion algorithm using ℓ_2- and ℓ_1-norm measures of data misfit f_1 and model restriction f_2. The star '⋆' marks the value obtained by evaluting the objective function with model (a) and the noisy data. (d), (e) Selected inversion results corresponding to the three asterisks '∗' in (b), (c). Same colourscale as in (a).

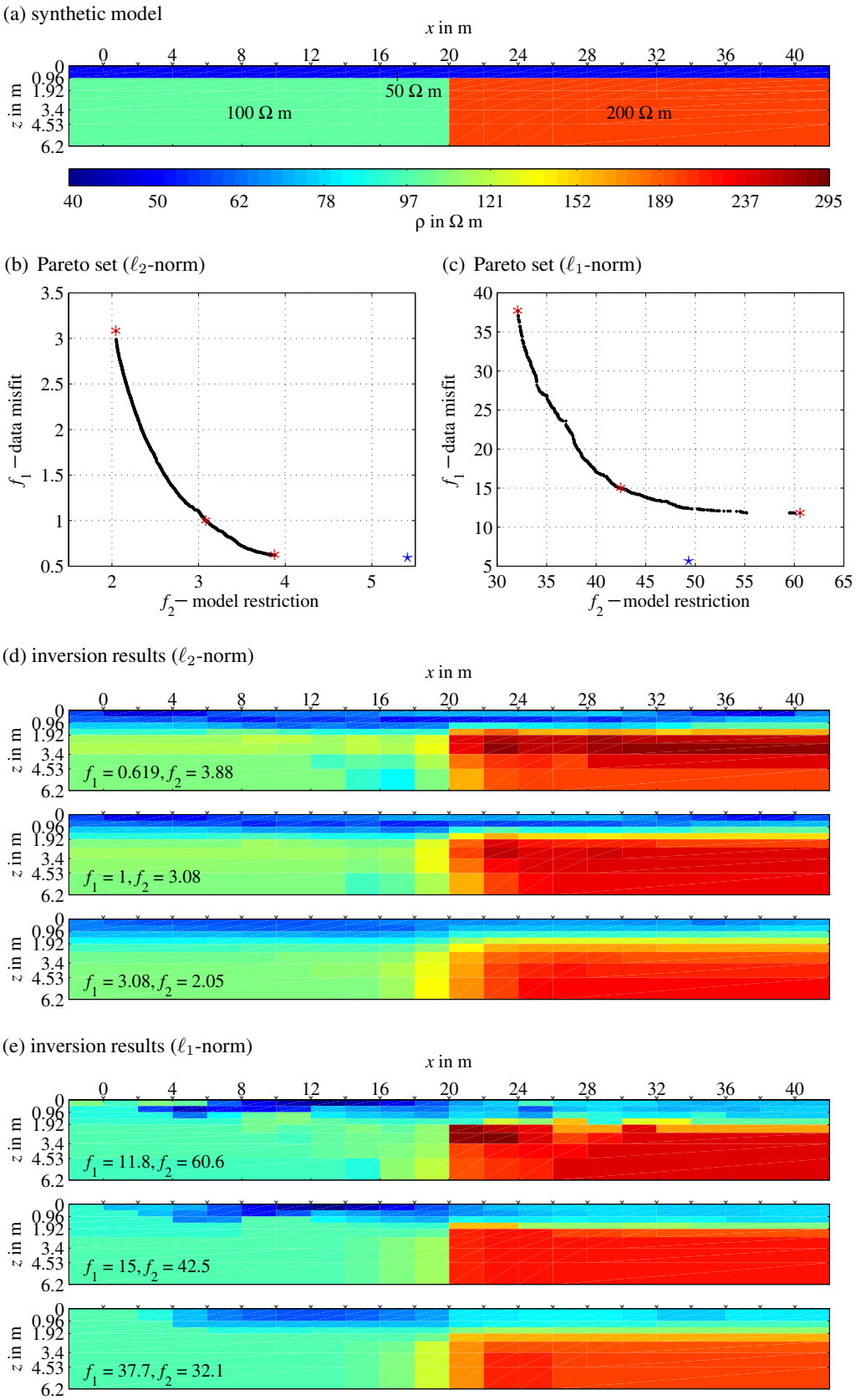

Figure 6. 'Quarterspace model' synthetic data inversion example. (a) Model to generate the data. Electrode positions are indicated by '∨'. (b), (c) Objective function values of the Pareto sets obtained from the genetic inversion algorithm using ℓ_2- and ℓ_1-norm measures of data misfit f_1 and model restriction f_2. The star '⋆' marks the value obtained by evaluating the objective function with model (a) and the noisy data. (d), (e) Selected inversion results corresponding to the three asterisks '∗' in (b), (c). Same colourscale as in (a).

caused by the discrete sampling of model space. More probably, it indicates that evolution has not reached its equilibrium state, yet, since auxiliary fitness (*cf.* Section 3) measuring distances between neigbouring points in objective function space should have produced more or less evenly spaced points.

Fig. 5(d) shows three selected solutions of the ℓ_2-norm inversion for the block model data set. These solutions are marked in Fig. 5(b) by asterisks. They comprise the two extreme points of the Pareto set as well as an intermediate one which was chosen such that data misfit f_1 equals 1 (expectation of $[f_1]^2$). The three inversion results are arranged in such a way that, from top to bottom, data misfit increases and model restriction norm decreases. Note the decreasing range of resistivity values with increasing model restriction norm as revealed by the common colourscale shown in Fig. 5(a). We emphasize that only the trough-shaped region of inner cells of Fig. 4 should be taken into account for interpretation. The boundary cells not only lie in the less sensitive area of our electrode layout but may additionally be subject to artifacts that arise when truncating the infinite subsurface region for inversion and modelling purposes. Nevertheless, these cells are shown for completeness as they were inversion parameters as well. Similarly to Fig. 5(d), Fig. 5(e) displays three inversion results corresponding to Fig. 5(c), that is to the block model data set inverted using ℓ_1-norm measures. In both cases, inversion results do not reconstruct the absolute values of the blocks. While the position and horizontal extension of the blocks is detected, their vertical extension is poorly resolved. If terms of the model restriction objective function f_2 with respect to vertical and horizontal discontinuities were weighted according to the finer vertical model discretization, this problem might be overcome. Remarkably, the inverse process found models that fit the data better than the synthetic model does. This is attributed to the effects of data noise in connection with the small-scale structure of the synthetic model.

In contrast to that, inversion of the quarterspace model data set failed to even reach the synthetic model with respect to data fit as can be seen from Figs 6(b) and (c). Inspecting Figs 6(d) and (e), which depict selected inversions results for the quarterspace model data set, we connect the failure to overestimating the thickness of the overburden by one cell layer in the ℓ_2-norm case and two cell layers in the ℓ_1-norm case. Therefore, the degree of failure is more pronounced in Fig. 6(c) than in Fig. 6(b) as the synthetic model even dominates part of the approximated Pareto set in the ℓ_1-norm case. We suspect the influence of a suboptimal region in objective function space the genetic algorithm was trapped in and unable to escape. The topmost inversion results in Figs 6(d) and (e) (best data fit) further exhibit a partial overshooting of resistivity values of 40–50 per cent in the resistive quarterspace to the right. Apart from that, the other features of the synthetic model are reconstructed fairly well.

Comparing the inversion results with regard to the ℓ_2- and ℓ_1-norm type model restriction functions in general, we observe besides a strong similarity of general features some well-known, characteristical differences (Farquharson & Oldenburg 1998): The ℓ_2-norm seems to promote a 'smoothed' appearance of structural features, staircase-like transitions between regions of different resistivity, and the ℓ_1-norm a 'blocky' appearance with sharp transitions and a flat distribution within regions of similar resistivity. Since both synthetic models are of the sharp contrast type, the ℓ_1-norm model restriction is expected to be more suitable for successful reconstruction than its ℓ_2-norm counterpart. This is indeed reflected in Figs 5(b), (c), 6(b) and (c): In the ℓ_2-case, the synthetic model seems to represent an extreme value of the model restriction norm f_2 that was not reached by the inversion. In the ℓ_1-case, inversion produced results with greater as well as smaller values of model restriction norm f_2.

5.2 The Rothschönberg drainage gallery

The field data example is taken from a survey which was carried out to detect the course of the Rothschönberg drainage gallery. The gallery was built about 150 years ago to drain the Freiberg/Saxony (Germany) mining field. Measurements were made on a profile perpendicularly to the gallery's strike near the gallery mouth at Rothschönberg. There, the gallery is supposed to have a width of 2.5 m, a height of 2.3 m, and the top to be situated about 2 m below the earth's surface. The data are displayed in Fig. 7(a) as a pseudosection of the apparent resistivity. Data noise was estimated using eq. (17) and the parameters described in the synthetic data examples section.

For inversion in the ℓ_2-norm case, the population was split into eight subpopulations of 1024 individuals each, which searched for the whole Pareto set and exchanged individuals every 16 generations. In the ℓ_1-norm case, one large population of 8192 individuals produced better results than splitting it into eight smaller ones.

Figs 7(b) and (c) plot objective function values of all non-dominated solutions of the final generation. We can assess the quality of results in the ℓ_2-case by observing that a data misfit of 1 corresponds to the expectation of $[f_1]^2$. Inversion slightly failed to reach this level. This indicates that explicit and implicit model restrictions prevent overfitting the data. Implicit model restriction summarizes effects of the relatively coarse model parameterization and of mapping arbitrarily placed, sized, and shaped underground features to a fixed, rectangular grid with piecewise constant resistivity values. Further, the noise level may be estimated inaccurately and, as noted earlier, the assumption of Gaussian distributed noise is not valid for logarithmized data. The correspondence between data misfit 1 and expectation of f_1 cannot be established for the ℓ_1-case. Nevertheless, applying the equivalence relations of norms in $\mathbb{R}^n$, data misfit measures are related by $f_1^{(\ell_2)} \leq f_1^{(\ell_1)} \leq \sqrt{n} f_1^{(\ell_2)}$ where n denotes the number of data points. Considering the best fit models we clearly see that data fit in the ℓ_1-case is not as good as in the ℓ_2-case since $f_1^{(\ell_1)} = 15.7 > 13.5 = \sqrt{n} f_1^{(\ell_2)}$.

Figs 7(d) and (e) display the inversion results using the ℓ_2- and ℓ_1-norm objective functions, respectively. All models exhibit essentially a lower resistive background of increasing resistivity with depth and two higher resistive features. The first feature extends from $x = 18$ to 22 m and $z = 1.44$ to 2.88 m. Note that in the topmost plots of Figs 7(d) and (e), the values of cell $x = 18$–20 m, $z = 1.44$–1.92 m are clipped by the colourscale from remarkably large values of 1968 Ωm in the ℓ_2-case and 6140 Ωm in the ℓ_1-case. The second resistive feature covers the area $x > 26$ m, $z > 0.96$ m. The first feature is identified as the Rothschönberg drainage gallery, the second one is supposed to be a basement feature. Both are confirmed by results of further, extensive investigation (Günther 2004).

6 CONCLUSIONS

We have shown that 2-D direct current resistivity inversion using a genetic algorithm is possible at the current state of computer performance. This premised a fast and accurate forward operator, parallelization and an efficient treatment of regularizing the inverse problem. Regularization was cast into the formulation of a multi-objective minimization task. Even though plausible results have been obtained, we note that the solutions may be further improved if still larger population sizes are used and evolution is run till equilibrium state is reached. Population size and stopping criterion were restricted by computing times, which were already extremely long for the cases considered.

The advantage of our approach is, however, that an ensemble of different solutions to the inverse problem is obtained. Currently,

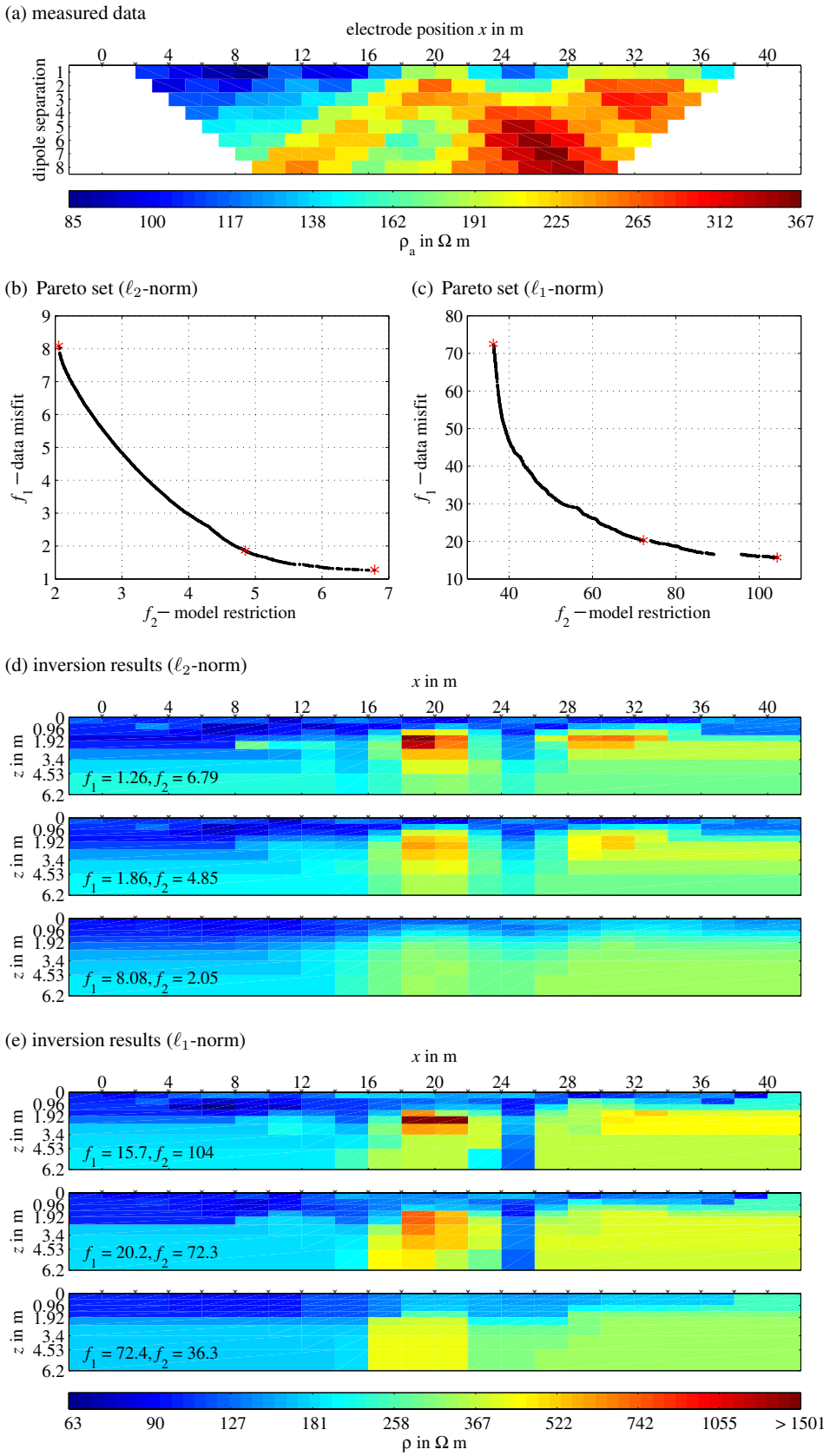

Figure 7. Field data inversion example. (a) Dipole-dipole pseudosection of apparent resistivity across the Rothschönberg drainage gallery. Electrode positions are indicated by '∨'. (b), (c) Objective function values of the Pareto sets obtained from the genetic inversion algorithm using ℓ_2- and ℓ_1-norm measures of data misfit f_1 and model restriction f_2. (d), (e) Selected inversion results corresponding to the three asterisks '∗' in (b), (c). Same colourscale in (d) and (e).

this set of inverse models is undergone an optical inspection. Even though this gives an instructive picture of the range of possible inverse models—compared to taking only one solution into account as it is usually done with classical approaches—a more sophisticated appraisal procedure would be desirable. Nevertheless, as Sambridge & Mosegaard (2002) point out, it is not advisable to use results calculated by a genetic algorithm for a statistical characterization of the set of inverse models since solutions may be biased and are distributed with unknown probability density.

Another advantage of our genetic inversion algorithm is the direct approach to the non-linear and multi-objective minimization task, which avoids linearizations and the choice of appropriate starting models as necessary for classical minimization methods. This opens the possibility to apply unorthodox objective functions and to study the inverse problem in a wider context.

ACKNOWLEDGMENTS

This work was partially funded by the German Research Foundation under grant no. Ja 590/18-2. CS likes to thank the Studienstiftung des deutschen Volkes for financial support. We would also like to thank Anders Vest Christiansen, an anonymous reviewer and the Editor Karsten Bahr for their helpful comments on our manuscript.

REFERENCES

Boschetti, F., Dentith, M. & List, R., 1997. Inversion of potential field data by genetic algorithms, *Geophys. Prosp.,* **45**(3), 461–478.

Chunduru, R.K., Sen, M.K. & Stoffa, P.L., 1996. 2-D resistivity inversion using spline parameterization and simulated annealing, *Geophysics,* **61**(1), 151–161.

Chunduru, R.K., Sen, M.K. & Stoffa, P.L., 1997. Hybrid optimization methods for geophysical inversion, *Geophysics,* **62**(4), 1196–1207.

Constable, S.C., Parker, R.L. & Constable, C.G., 1987. Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics,* **52**(3), 289–300.

Davis, L., 1996. *Handbook of genetic algorithms*, International Thomson Computer Press, London.

Deb, K., 2001. *Multi-Objective Optimization Using Evolutionary Algorithms*, Wiley, Chichester.

Deb, K. & Goel, T., 2000. Controlled elitist non-dominated sorting genetic algorithms for better convergence, Tech. Rep. 2000004, Kanpur Genetic Algorithms Laboratory (KanGAL), Kanpur, India.

Deb, K., Agrawal, S., Pratap, A. & Meyerivan, T., 2000. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II, Tech. Rep. 2000001, Kanpur Genetic Algorithms Laboratory (KanGAL), Kanpur, India.

Deb, K., Zope, P. & Jain, A., 2002. Distributed computing of Pareto-optimal solutions using multi-objective evolutionary algorithms, Tech. Rep. 2002008, Kanpur Genetic Algorithms Laboratory (KanGAL), Kanpur, India.

Dey, A. & Morrison, H.F., 1979. Resistivity modelling for arbitrarily shaped two-dimensional structures, *Geophys. Prosp.,* **27,** 106–136.

Engl, H.W., Hanke, M. & Neubauer, A., 2000. *Regularization of inverse problems*, Kluwer Acad. Publ., Dordrecht.

Everett, M.E. & Schultz, A., 1993. Two-dimensional nonlinear magnetotelluric inversion using a genetic algorithm, *Journal of Geomagnetism and Geoelectricity,* **45**(9), 1013–1026.

Farquharson, C.G. & Oldenburg, D.W., 1998. Non-linear inversion using general measures of data misfit and model structure, *Geophys. J. Int.,* **134,** 213–227.

Farquharson, C.G. & Oldenburg, D.W., 2004. A comparison of automatic techniques for estimating the regularization parameter in non-linear inverse problems, *Geophys. J. Int.,* **156**(3), 411–425.

Friedel, S., 2003. Resolution, stability and efficiency of resistivity tomography estimated from a generalized inverse approach, *Geophys. J. Int.,* **153,** 305–316.

Gallagher, K. & Sambridge, M., 1994. Genetic algorihtms: a powerful tool for large-scale nonlinear optimization problems, *Computers &Geosciences,* **20**(7/8), 1229–1236.

Gallagher, K., Sambridge, M.S. & Drijkoningen, G., 1991. Genetic algorithms: an evolution on Monte Carlo methods in strongly non-linear optimisation problems, *Geophys. Res. Lett.,* **18**(12), 2177–2180.

Gropp, W., Lusk, E., Doss, N. & Skjellum, A., 1996. A high-performance, portable implementation of the MPI Message-Passing Interface standard, *Parallel Computing,* **22**(6), 789–828.

Günther, T., 2004. Inversion methods and resolution analysis for 2d/3d reconstruction of resistivity structures from dc measurements, *PhD thesis*, TU Bergakademie Freiberg.

Holland, J., 1975. *Adaption in natural and artificial systems*, The University of Michigan Press, Ann Arbor, Michigan.

Kennett, B.L.N. & Sambridge, M.S., 1992. Earthquake location: Genetic algorithms for teleseisms, *Phys. Earth planet. Inter.,* **75,** 103–110.

LaBrecque, D.J., Miletto, M., Daily, W., Ramirez, A. & Owen, E., 1996. The effects of noise on Occam's inversion of resistivity tomography data, *Geophysics,* **61**(2), 538–548.

Li, Y. & Spitzer, K., 2002. Three-dimensional DC resistivity forward modelling using finite elements in comparison with finite-difference solutions, *Geophys. J. Int.,* **151,** 924–934.

Lowry, T., Allen, M.B. & Shive, P.N., 1989. Singularity removal: a refinement of resistivity modeling techniques, *Geophysics,* **54**(6), 766–774.

Mundry, E., 1984. Geoelectrical model calculations for two-dimensional resistivity distributions, *Geophys. Prosp.,* **32,** 124–131.

Park, S.K. & Van, G.P., 1991. Inversion of pole-pole data for 3-D resistivity structure beneath arrays of electrodes, *Geophysics,* **56**(7), 951–960.

Pérez-Flores, M.A. & Schultz, A., 2002. Application of 2-D inversion with genetic algorithms to magnetotelluric data from geothermal areas, *Earth Planets Space,* **54,** 607–616.

Rana, S., 1999. Examining the role of local optima and schema processing in genetic search, *PhD thesis*, Department of Computer Science, Colorado State University, Fort Collins, Colorado.

Sambridge, M. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.,* **40**(3), 1–29.

Scales, J.A. & Snieder, R., 2000. The anatomy of inverse problems, *Geophysics,* **65**(6), 1708–1710.

Sen, M.K. & Stoffa, P.L., 1992. Rapid sampling of model space using genetic algorithms: examples from seismic waveform inversion, *Geophys. J. Int.,* **108,** 281–292.

Sen, M.K., Bhattacharya, B.B. & Stoffa, P.L., 1993. Nonlinear inversion of resistivity sounding data, *Geophysics,* **58**(4), 496–507.

Spitzer, K., 1995. A 3D finite difference algorithm for dc resistivity modeling using conjugate gradient methods, *Geophys. J. Int.,* **123,** 903–914.

Spitzer, K. & Wurmstich, B., 1999. Speed and accuracy in 3D resistivity modeling, in *Three-Dimensional Electromagnetics, SEG Book Series 'Geophysical Developments',* Vol. 7, pp. 161–176, eds Oristaglio, M.L. & Spies, B.R., Society of Exploration Geophysicists.

Spitzer, K., Chouteau, M. & Boulanger, O., 1999. Grid-independent electrode positioning for 3D DC resistivity and IP forward modeling, in *Extended Abstracts Book, 2nd International Symposium on Three-Dimensional Electromagnetics, October 27–29, Salt Lake City, Utah, USA*, pp. 189–192.

Stoffa, P.L. & Sen, M.K., 1991. Nonlinear multiparameter optimization using genetic algorithms: inversion of plane-wave seismograms, *Geophysics,* **56**(11), 1794–1810.

Wilson, W.G. & Vasudevan, K., 1991. Application of the genetic algorithm to residual statics estimation, *Geophys. Res. Lett.,* **18**(12), 2181–2184.

Zhang, J., Mackie, R.L. & Madden, T.R., 1995. 3-D resistivity forward modeling and inversion using conjugate gradients, *Geophysics,* **60,** 1313–1325.

Zhao, S. & Yedlin, M.J., 1996. Some refinements on the finite-difference method for 3-D dc resistivity modeling, *Geophysics,* **61**(5), 1301–1307.