Geophys. J. Int. (1998) **132,** 458–468

Optimized and robust experimental design: a non-linear application to EM sounding

Hansruedi Maurer[1] and David E. Boerner[2]

[1] *Institute of Geophysics, ETH-Hönggerberg,* 8093 *Zürich, Switzerland. E-mail: maurer@augias.ig.erdw.ethz.ch*
[2] *Geological Survey of Canada,* 1 *Observatory Cres., Ottawa, ON, Canada,* K1A 0Y3

Accepted 1997 August 14. Received 1997 July 17; in original form 1997 March 10

SUMMARY

Pragmatic experimental design requires objective consideration of several classes of information including the survey goals, the range of expected Earth responses, acquisition costs, instrumental capabilities, experimental conditions and logistics. In this study we consider the ramifications of maximizing model parameter resolution through non-linear experimental design. Global optimization theory is employed to examine and rank different EM sounding survey designs in terms of model resolution as defined by linearized inverse theory. By studying both theoretically optimal and heuristic experimental survey configurations for various quantities of data, it is shown that design optimization is critical for minimizing model variance estimates, and is particularly important when the inverse problem becomes nearly underdetermined. We introduce the concept of robustness so that survey designs are relatively immune to the presence of potential bias errors in important data. Bias may arise during practical measurement, or from designing a survey using an appropriate model.

Key words: electromagnetic sounding, experimental design, inversion, robust statistics.

1 INTRODUCTION

Geophysical experimental design is fundamentally based on our ability to understand the physical laws relating Earth properties to measurable data. This experience is usually gained through a combination of theoretical investigations, repeated simulations with simple numerical or analogue models and interpretation of previous field surveys. While extremely powerful, heuristic experimental design can be limited, particularly by the degree of specialization required to understand the subtleties of applying geophysical techniques in complicated environments. Heuristic designs are further complicated by variations in socio–environmental, logistical or instrumental constraints between surveys. The ultimate goal of all geophysical surveys is that resolution of Earth structures should be limited only by the capabilities of the geophysical method, and not by inappropriate survey layouts or insufficient data. Philosophically, we desire an 'optimal' data set. Understanding that no better data could have been acquired (within the experimental constraints) is extremely powerful as it dramatically expands the scope and importance of forward modelling by providing additional rigour to feasibility studies, instrumentation design and survey costing.

In this paper, we consider the issues surrounding the quantitative design of an 'optimal' survey. Our particular goal is to maximize the formal resolution of model parameters. Designing experiments that enhance model resolution is not new, and the published studies illustrate the universality of this concept within geophysics. For example, Glenn & Ward (1976) and Jones & Foster (1986) employed linearized inverse theory for designing electromagnetic surveys over layered conductivity structures. Barth & Wunsch (1990) designed marine seismic experiments, and Hardt & Scherbaum (1994) determined optimal earthquake network configurations. More recently, Curtis & Snieder (1997) determined the model parametrization that led to the optimal model resolution. We augment these previous studies by considering that pragmatic designs should also allow constraints due to cost and logistics, while admitting the possibility of data error. Our aim here is to design a practical experiment that optimally resolves important features in the Earth (within the limits imposed by the physical method). We examine the influence of minimizing data acquisition costs (that is collecting less data) on the design process by quantifying the corresponding degradation of model resolution. Moreover, we develop a method to design experiments that are robust in the sense that small variations in the design, or the presence of bias errors in the data, do not unduly compromise the survey objectives.

Experimental design is based on an *expected* inverse problem (that is some knowledge of the Earth is assumed, but there is no data to interpret). Conducting geophysical surveys admits that the *a priori* information is incomplete, thus raising important concerns about inappropriate models or assumptions affecting the design. For this reason, experimental design is

perhaps best viewed as an iterative exercise in hypothesis testing (Fig. 1). The design process involves the delineation of data space regions that are either critical or unimportant for resolving the model parameters, presumably for a range of models. The decisive advantage of planning surveys with experimental design techniques is that information and assumptions can be examined, tested and considered before incurring data acquisition expenses. However, non-linear relationships between the observed data and the causative model can limit the applicability of experimental design. For this reason we consider the issue of non-linearity in our example design study.

Following, Barth & Wunsch (1990), Hardt & Scherbaum (1994) and Curtis & Snieder (1997), we design the experiment objectives by formulating a constrained *global optimization problem* (Fig. 1). The formal resolution of model parameters is determined through linearized inversion and forms the basis of quantifying acceptable survey designs. Thus any geophysical technique that is Fréchet differentiable can be subjected to this experimental design procedure. Our test case is a frequency domain electromagnetic (EM) survey of layered Earth structures. EM methods are strongly non-linear and admirably suited to investigate the robustness of design studies to inappropriate expectations of the Earth. Diffusive EM methods pose problems for model resolution and yet, for layered earth models, are simple enough to understand the results of the optimization in detail. In addition to studying model resolution, we address the important issue of designing robust experiments. Pragmatic survey design must constrain subsurface models while not being susceptible to gross data errors or model inadequacies at potentially important data points.

2 THEORY

In establishing the experimental design procedure, we assume that the model parameter resolution is expressed in terms of linearized inverse theory. In contrast, the non-linear optimization process of actually selecting the experimental layout is performed using a genetic algorithm. As both techniques are well described in the literature, the cursory introduction that follows serves only to introduce notation and key concepts as applied to experimental design. Additional details about the individual methods can be found in several excellent texts (e.g. Menke 1984; Tarantola 1987; Sen & Stoffa 1995).

2.1 Linearized inverse theory

Non-linear physical experiments are generally governed by a functional $\mathbf{g}(\mathbf{m})$:

$$\mathbf{d}^{\text{obs}} = \mathbf{g}(\mathbf{m}) + \mathbf{e} \tag{1}$$

relating observed data $\mathbf{d}^{\text{obs}}$, model parameters $\mathbf{m}$, and errors $\mathbf{e}$. The well-known solution of the linearized inverse problem corresponding to eq. (1) (e.g. Menke 1984) can be written as

$$\mathbf{m} = \mathbf{m}_{\text{o}} + \mathbf{G}^{-1}\Delta\mathbf{d}, \tag{2}$$

where $\mathbf{m}_{\text{o}}$ is an initial model, $\mathbf{G}^{-1}$ a generalized inverse matrix and $\Delta\mathbf{d}$ the difference between observed and predicted data. The generalized inverse $\mathbf{G}^{-1}$ can be, for example, determined using singular value decomposition (Lanczos 1961; Lawson & Hanson 1974):

$$\mathbf{G}^{-1} = \mathbf{V}\Lambda^{-1}\mathbf{U}^{\text{T}}, \tag{3}$$

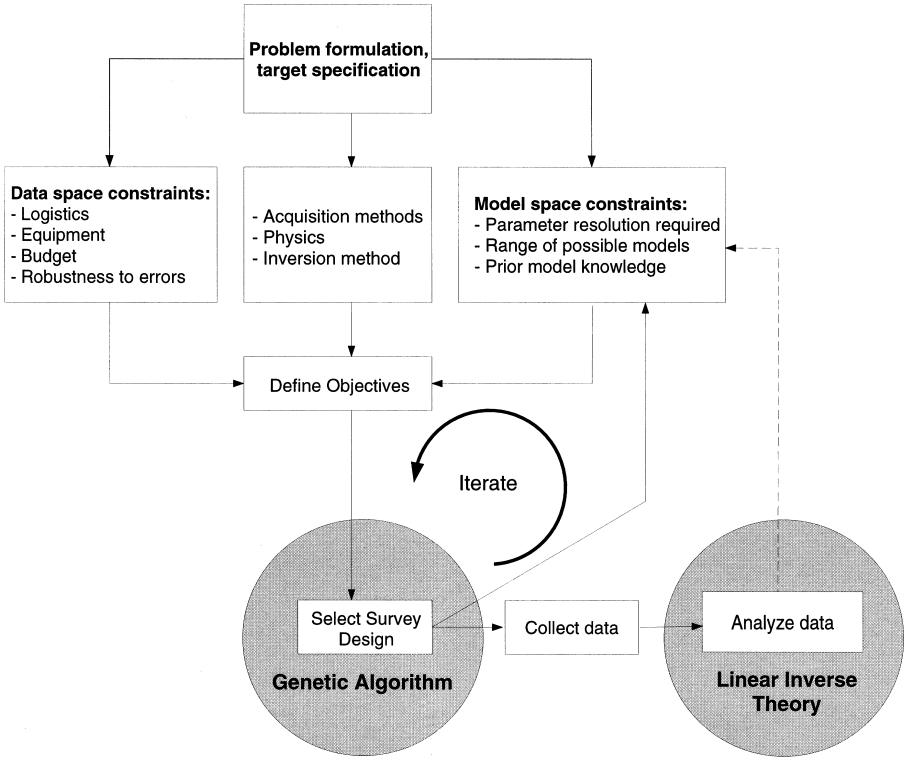

Figure 1. Schematic overview of general experimental design, showing components involved and their relationships. Notice that the genetic algorithm (GA) is used only to select the optimal experimental design, while standard linearized inverse theory is used to characterize the formal model resolution. In this paper we consider only model resolution, costs and robustness in defining the design objective.

where the matrix **U** contains a set of orthonormal vectors that span the data space, and the columns of matrix **V** are orthonormal vectors that span the model space (Menke 1984). The diagonal matrix Λ contains the singular values describing the relationship between the data space and the model space. Singular value decomposition also results in a simple form of the model covariance matrix:

$$\mathbf{C_M} = \mathbf{C_D V} \Lambda^{-2} \mathbf{V}^{\mathsf{T}}, \tag{4}$$

which describes the model-space mapping of the data variances $\mathbf{C_D}$ into *a posteriori* model parameter uncertainties (e.g. Menke 1984).

The matrix **H** describes the data-space mapping between the observed data $\mathbf{d}^{\mathrm{obs}}$ and the predicted data $\mathbf{d}^{\mathrm{pre}}$ (e.g. Hoaglin & Welsch 1978; Menke 1984):

$$\mathbf{d}^{\mathrm{pre}} = \mathbf{H}\mathbf{d}^{\mathrm{obs}}, \quad \text{where} \quad \mathbf{H} = \mathbf{U}\mathbf{U}^{\mathsf{T}}. \tag{5}$$

In the geophysical literature, **H** is also known as the *data resolution matrix* (Menke 1984) or the *information density matrix* (Wiggins 1972). **H** is a projection matrix and is symmetric and idempotent i.e. $\mathbf{H} = \mathbf{H}^2$ (Staudte & Sheather 1990). It follows from these properties that the length squared of the ith column vector $\mathbf{H}_i$ is equal to h_{ii}. Furthermore, the trace of **H** is

$$\mathrm{tr}(\mathbf{H}) = \sum_{i=1}^{N} h_{ii} = M, \tag{6}$$

where N is the number of data points and M is the number of unknown model parameters. Since the diagonal elements of **H** describe the relative importance (that is the ability of influencing model parameters by small perturbations of the data value) of a particular data point, diag(**H**) is called *data importance* (Menke 1984). From eq. (6) it follows that the average importance of a data point is M/N. When an element of diag(**H**) is significantly larger than the remaining h_{ii} values, the associated datum is called a *leverage point*. Data errors at a leverage point can strongly affect the inversion result (e.g. Staudte & Sheather 1990). In robust regression analysis, leverage points are avoided whenever possible by adopting a process called *equileverage design* (Staudte & Sheather 1990).

2.2 Genetic algorithms

For strongly non-linear optimization problems, or for situations in which the partial derivatives of $\mathbf{g}(\mathbf{m})$ cannot be easily formed, global optimizers provide a useful alternative to linear inverse methods (e.g. Sen & Stoffa 1995). An important class of global optimizers, genetic algorithms (GA), were originally proposed by Holland (1975). To some extent a genetic algorithm simulates biological evolution. The goal is to ensure the progeny of the most fit members survive through successive generations. The algorithm is formulated as follows:

(1) initiate a random population, each of which represents a particular geophysical survey design (encoded as a bit string);
(2) evaluate the 'fitness' (misfit between predicted and observed data $|\mathbf{d}^{\mathrm{obs}} - \mathbf{g}(\mathbf{m})|^k$, where k is an arbitrary constant) of the population members via an objective function defined in terms of model resolution and any additional constraints;
(3) procreate a new generation after selecting the most fit parents and applying genetic operators (crossover, mutation and replication);
(4) repeat steps 2 to 4 until the convergence criteria are met.

Darwinian selection should evolve the whole population to a higher degree of fitness. The definition of a single objective function means that genetic algorithms tend to reduce diversity within the population, and so the output of the GA runs was carefully examined to ensure convergence had been reached.

2.3 Experimental design

Physical laws, past experience with data inversion, and/or simulations with synthetic data govern experimental design. However, quantifying this procedure in a numerical procedure by iterating on a single design component is inefficient and may not lead to an optimal design (Fisher 1925). As an alternative, we follow the suggestion of Fisher (1925) and impose some randomness on the design specifications through the genetic algorithm formulation. Evolution starting from a broad range of diverse designs should lead to the culling of the unfit and selection of those designs that have optimal components (that is leading to the smallest objective function).

Statistical experimental design identifies data acquisition parameters (for example source–receiver configurations, bandwidth, acceptable signal to noise ratios) that 'optimally' determine a particular subsurface model. Our approach is shown schematically in Fig. 1. The first step is to construct a hypothetical *a priori* subsurface model (or range of models, if sufficient uncertainty regarding the Earth exists). The desired model resolution then becomes a constraint that defines the space (genetic algorithm population) available to search for the best data acquisition configuration.

The objective function, Γ, defines 'optimal' in a mathematical sense. In our case, Γ shows how the data from a particular survey layout resolve the parameters of a hypothetical *a priori* subsurface model. An appropriate objective function definition is the critical element of experimental design. We must characterize, with a single number, the quality of a particular design configuration. As the data will be analysed with linearized inversion theory, one reasonable objective function could minimize the *a posteriori* covariances $\mathbf{C_M}$ for all subsurface model parameters (eq. 4). Since $\mathbf{C_M}$ is a function of Λ^{-2}, we have defined an objective function Γ as

$$\Gamma = \sum_{i=1}^{M} \frac{1}{\Lambda_i^2 + \delta}, \tag{7}$$

where δ is a positive constant. This definition has several desirable properties.

(1) Γ is directly and simply related to $\mathbf{C_M}$.
(2) Since Γ is most sensitive to small singular values, the algorithm represents an attempt to reduce the underdetermined components in the inverse problem (eq. 7).
(3) In case of an inherent ambiguity in the inverse problem (zero or nearly zero singular values), Γ ensures that any *resolvable* subsurface model parameters are well constrained (but perhaps not optimally resolved).
(4) The parameter δ in eq. (7) can be selected to match the expected noise characteristics of the observed data. In the case of error-free data, δ should be so small that it only prevents Γ from going to infinity and keeps the algorithm numerically stable. If the expected noise level is significant, δ should be chosen large enough such that the objective function Γ ignores singular values associated with unresolvable components of the model space.

There are many other objective functions that minimize model covariances. Appendix A provides a comparison of our choice in eq. (7) with some other definitions found in the literature. For minimizing the objective function Γ (eq. 7) we have used a library of genetic algorithms developed by Hunter (1995). Appendix B provides more information on parameter settings, convergence speed and other properties of the individual GA runs.

3 EXPERIMENTAL DESIGN OF A FREQUENCY-DOMAIN EM SURVEY

3.1 The experiment

Simulating a simple controlled-source EM survey is useful for exploring the characteristics of experimental design. A horizontal electric dipole (HED) source is placed at the surface of a layered half-space. The data comprise vertical magnetic field measurements at various positions between 0.1 and 1000 m from the source (Fig. 2). The frequency range lies between 1 Hz and 1 MHz, although the quasi-static limit has been imposed, thus preventing radiative propagation of the high-frequency EM fields. This source–receiver configuration has a cylindrical symmetry modulated by $\sin\theta$ (Fig. 2), so we consider that the experimental design parameters are simply frequency (f) and distance (r). Three simple 1-D earth models (Fig. 3) are used to illustrate the experimental design procedure. Sampling is logarithmic with 10 samples per decade, giving a total of 2400 possible receiver configurations (60 frequencies and 40 distances). Actually acquiring such a large number of frequency–distance estimates would be impractical, but is necessary to provide a reasonably large model space for the optimization algorithm to explore. Fréchet differentiability for this non-linear class of models and EM methods was established by Chave (1984).

3.2 Minimization of *a posteriori* error estimates

The data in this experiment are assumed to be error-free, of unlimited bandwidth and accurate to computer precision. Thus, the δ parameter in eq. (7) is kept small (10^{-20}). We seek to find the optimal data subset consisting of 20 points in (f,r) space. When distributed adequately, this subset size should result in a moderately overdetermined system of equations, sufficient to constrain each of the test models (which have at most five free parameters).

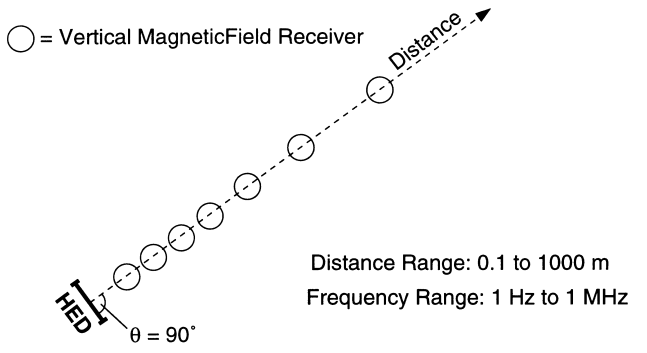

Figure 2. Schematic representation of the simulated EM experiment. HED denotes horizontal electric dipole.

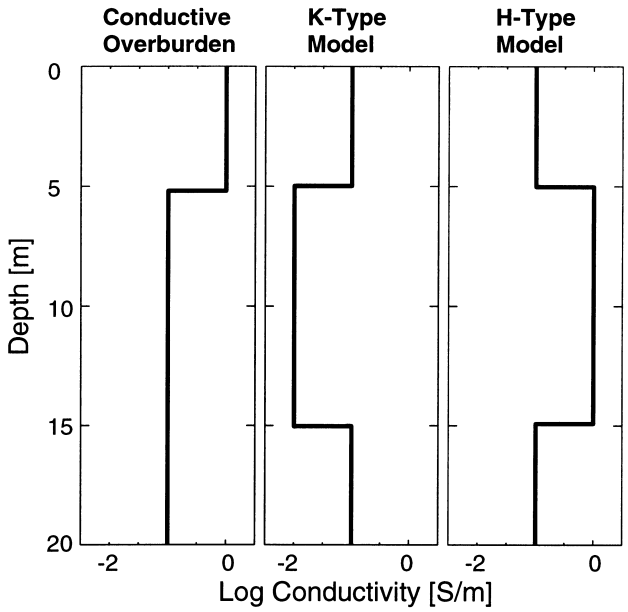

Figure 3. Conductivity models used in this study.

Figs 4(a)–(c) show the best receiver configurations in (f,r) space selected by the genetic algorithm for each of the test models. In other words, these locations are optimal for resolving the model parameters according to the objective function (7) and for 20 data. Desirable measurement locations for all models tend to cluster along relatively narrow diagonal bands in frequency–distance space, roughly along lines of constant induction number (e.g. Ward & Hohmann, 1988). For the conductive overburden and the K-type model, the information is mostly concentrated at higher frequencies (above 10 kHz), reflecting the presence of a relatively conductive surface layer, whereas the H-type model also includes lower-frequency regions.

Although the earth models are quite different, the data sampling configurations in Figs 4(a)–(c) exhibit relatively consistent and confined patterns, delineating well-defined regions that contain most of the information about the subsurface. This observation is reinforced by examining the fittest individuals of the final generation. Figs 4(d)–(f) show areas that include 99.9 per cent of (f,r) pairs belonging to the fittest 10 per cent of the final population. Although genetic algorithms tend to homogenize populations (Hunter 1995), there is some diversity shown in Figs 4(d)–(f), yet the delimited areas cluster within and around those already identified by the fittest structures (Figs 4a–c).

A physical interpretation of the regions delineated by the genetic algorithm is straightforward. The selected (f,r) points fall in the transition between the electromagnetic near field and far field, or at the electromagnetic inductive limit, the data regions most sensitive to changes in the conductive parts of the models (Boerner & West 1989). Note that the genetic algorithm solution considers the resolution of *all* model parameters simultaneously, and thus is a more refined sampling of (f,r) space than the single-parameter example shown by Boerner & West (1989). As the transition zone is narrow, the genetic algorithm tends to duplicate important data points, particularly in the case of the conductive overburden model

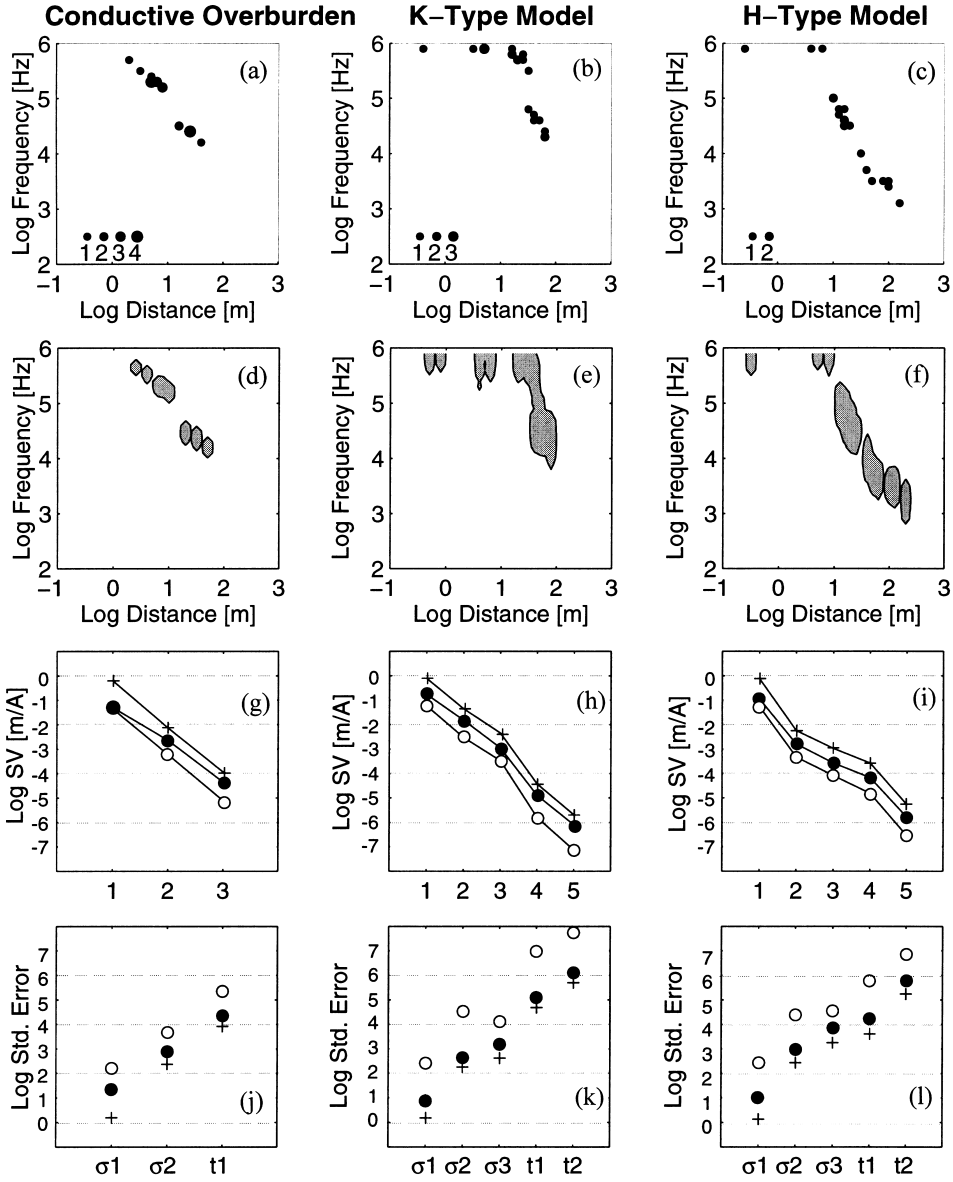

Figure 4. Experimental design through minimization of *a posteriori* model variances. Plots (a) to (c) show the optimal distribution of (*f,r*) pairs for the different conductivity models. Duplicated data points are represented with larger dots (see legend). (d)–(f) Regions of desirable locations for the fittest 10 per cent of the final GA population. The shaded areas delimit regions containing 99.9 per cent of all (*f,r*) pairs. (g)–(i) Singular-value spectra for the complete data set (+), the random solution (○) and the GA solution (●). (j)–(l) *A posteriori* error estimates for the complete data set (+), the random solution (○) and the GA solution (●). $\sigma 1$ to $\sigma 3$ denote layer conductivities and t1 and t2 layer thicknesses.

(Fig. 4a). While perfectly legitimate in synthetic studies, duplication is not desirable for real survey designs, as discussed below.

To study quantitatively the properties and reliability of the genetic algorithm, we compare the singular-value spectrums and *a posteriori* covariance estimates based on the best genetic algorithm solution with those of other survey designs. The complete data set of all 2400 (*f,r*) pairs approximates the best attainable experiment, while a randomly selected subset of 20 data points should approach the worst case (completely naive sampling) limit. Figs 4(g)–(i) show that the shapes of the singular value spectra are generally similar for the three cases, but there are large differences in the magnitudes. Note that the largest singular values for the genetic algorithm solution approach those of the random solution for all models, as expected from the definition of the objective function (eq. 7).

While singular-value spectra form an interesting basis for appraising the properties of the genetic algorithm, the corresponding standard errors of model parameters allow a formal comparison based on the objective function. Figs 4(j)–(l) depict the standard errors of the model parameters (defined as the square root of the diagonal elements of the model covariance matrix, assuming the data variance $\mathbf{C}_D = \sigma^2\mathbf{I}$ in eq. 4). The individual panels of Figs 4(j)–(l) reveal that by judicious selection of <1 per cent of the complete data set, the corresponding standard errors differ on average by less than a factor of five. Genetic algorithm solutions are better, on average, by a factor of 30 compared with the random solutions.

As already noted for the singular-value spectra, we also observe considerable deviations from these average values.

3.3 Influence of the subset size

Since the amount of data acquired is usually related to the survey costs, it is often desirable to minimize the data volume. Experimental design offers an opportunity to examine the trade-off between data quantity and model parameter resolution. As a test case, we consider the H-type model using the complete and genetic algorithm data sets together with a random data set.

Figs 5(a)–(e) show the *a posteriori* standard error estimates for the individual model parameters as functions of subset size. For larger subset sizes, the standard errors of the random and the genetic algorithm solutions exhibit a power-law decay proportional to the square root of the subset size, as expected for a purely overdetermined inverse problem. Note that duplication of important data points in larger subsets allows the error estimated from the genetic algorithm solution to be smaller than that of the complete data set (Fig. 5e).

Deviations from the power-law behaviour occur for the random solutions at smaller subset sizes, indicating the occurrence of significant underdetermined components in the inverse problem. Interestingly, the standard errors of the genetic algorithm solutions are inversely proportional to the square root of N (subset size) over the entire subset size range. This observation indicates that design optimization has succeeded in keeping the inverse problem for the conductivity structure overdetermined, almost independently of subset size. *A posteriori* errors for the first layer conductivity do deviate from a power-law decay for smaller subsets, approaching the error level of the random solution (Fig. 5a). Because the first-layer conductivity is the best-resolved parameter (*cf.* y-axis in Fig. 5), it is primarily determined by the largest values in the singular-value spectrum. Conversely, the objective function preferentially selects configurations that improve the resolution of poorly resolved model parameters. When the well-resolvable parameters are the primary survey target, the objective function should be modified accordingly.

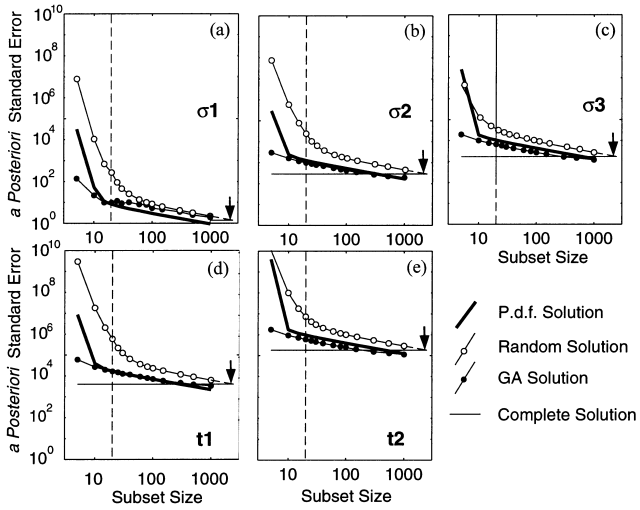

Figure 5. Influence of the data acquisition subset size (i.e. cost). Linear extrapolation of the curve for random survey designs intersects that of the complete data set at a subset size of 2400 (= size of the complete data set) and is marked with an arrow. The vertical dashed line indicates a subset size of 20. $\sigma 1$, $\sigma 2$ and $\sigma 3$ denote layer conductivities and t1 and t2 layer thicknesses. The heavy lines represent the solution derived by selecting a design using **H** as a probability density function.

3.4 Analysis of data importance

The objective function developed above is based solely on resolving the details of the model space. An alternative, but important, context is to consider objective functions in the data space. In particular, we examine the role of *data importance*, expressed as the diag(**H**) (eqs 5 and 6). Comparisons of the genetic algorithm solutions (Figs 4a–c) with the importance values for the complete data set in frequency–distance space are shown in Fig. 6. While the real and imaginary parts of the vertical magnetic field are treated equally in the genetic algorithm objective function, it is interesting to identify the contributions of the individual components to the experimental design. Most of the information from the real part of H_z is contained in a relatively narrow distance range (Figs 6a, d and g), whereas the imaginary parts provide additional information in the high-frequency range at shorter distances (Figs 6b, e and h).

Figs 6(c), (f) and (i) display the genetic algorithm solutions superimposed on the data importance contours, averaged for the real and imaginary parts. The strong correlation between large importance values and the GA solution suggests that the importance distribution could be used as a probability density function (pdf) for selecting an optimal 'random' data set. Since the calculation of diag(**H**) involves only the calculation of N scalar vector products with length M, experimental design based on data importance would be computationally much less demanding than the GA solution (requiring multiple singular-value decompositions). Using an importance pdf is much like defining a new objective function, focused primarily on resolving parameters associated with large singular values. In fact, exploiting the important data space in this fashion is quite close to the process of heuristic experimental design. An experienced geophysicist would have discovered that the transition zone of the EM fields is the most important region of the data space, and could design an experiment to sample this space.

Selecting a 'random' design using the **H** matrix as a pdf for determining measurement points was tested for the subset size experiment (also shown in Fig. 5). Generally, the results are comparable with those of the GA solution, but with two significant differences.

(1) When the subset sizes approach the underdetermined case (< 5 frequency–receiver pairs = 10 constraining equations), the pdf solution degrades markedly. Lacking information about interrelationships between data points limits the usefulness of the pdf experimental design to purely overdetermined problems.

(2) The first-layer conductivity is better resolved for larger subsets with the importance pdf than by the genetic algorithm solution. Since data importance is governed mainly by well-resolvable parameters, the pdf approach is most effective for constraining these parameters, whereas the objective function in the genetic algorithm tries to resolve all parameters. Conversely, *a posteriori* errors of potentially poorly resolvable

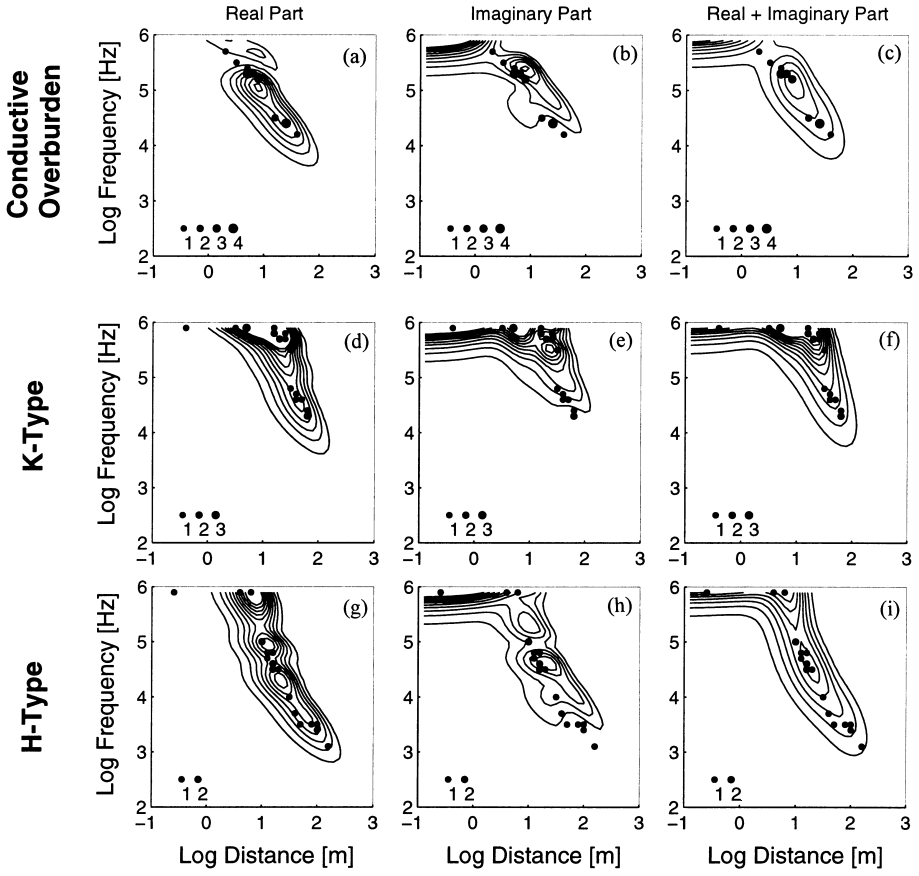

Figure 6. Contour plots of the data importance superimposed on the best configuration determined with the GA (Figs 4a–c). The contour lines are linearly spaced between 0.02 and 0.2. See text for further explanations.

parameters are slightly higher than those of the GA solution (Figs 5c and e).

The limitations observed for the pdf solution probably reflect the results of heuristic experimental design.

3.5 Equileverage design

Experimental design, as outlined so far, is basically a process of locating data points important to the model resolution, or, in terms of robust statistics, selecting *leverage points* (e.g. Staudte & Sheather 1990). However, severe repercussions in terms of parameter estimation could result from even minor data errors at a leverage point. Alternatively, an inappropriate expected model could introduce an effective bias into the design, and this bias would be most evident at leverage points. Consequently, it is desirable to find survey designs that minimize *a posteriori* variances, but that do not rely exclusively on leverage points.

Equileverage constraints require a redefinition of the objective function to penalize leverage points (Staudte & Sheather 1990). We define a modified objective function:

$$\Gamma^* = \log(\Gamma) + \alpha(H_{ii}^{\max} - H_{ii}^{\min}), \tag{8}$$

where $(H_{ii}^{\max} - H_{ii}^{\min})$ denotes the importance spread and α is an arbitrary positive constant. Possible values of the importance spread lie between 0 and 1, whereas Γ can vary over several orders of magnitudes. Accordingly, we restrict the numerical range of Γ through the use of $\log(\Gamma)$ (eq. 8).

The efficiency of equileverage design for the case of the H-type model is shown in Fig. 7(f). There is a decrease of importance spread as α increases. Even for small values of α the occurrence of leverage points is effectively suppressed. Figs 7(a)–(e) show standard errors as a function of α and demonstrate that equileverage configurations constrain potentially well-resolvable parameters even better than non-robust designs (Fig. 4). This apparent paradox can be explained by considering eq. (6), which explains why penalizing a large importance spread $(H_{ii}^{\max} - H_{ii}^{\min})$ not only avoids leverage points, but also prevents the incorporation of unimportant data. Consequently, minimizing the importance spread in the data space can be a powerful tool for improving the experimental design.

Figs 8(a)–(c) display optimal (f,r) points superimposed on the importance distribution for different values of α, showing that the equileverage design algorithm tends to duplicate data (that is such that all data have equal importance). In the presence of unbiased random data errors, repeat recordings of the same data point improve data precision. However, this concept is already applied in standard instrumentation and recording practices. It may be better, from the design point of view, to assume that data are recorded to the best obtainable precision and that repeat measurements do not reduce random error. Whereas data duplication in design studies indicates

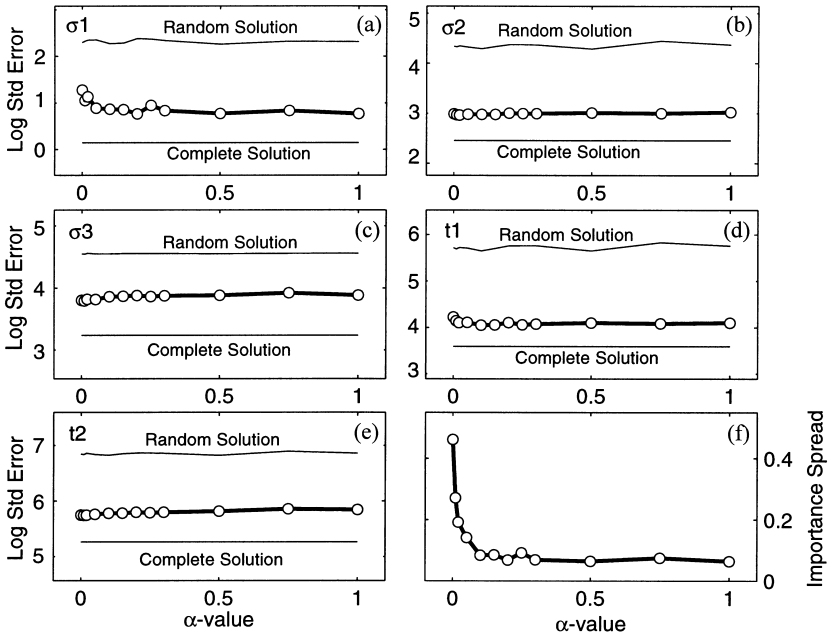

Figure 7. (a)–(e) *A posteriori* error estimates of an equileverage design for the H-type model. Error estimates are plotted as a function of the α value defined in eq. (8). Error levels of the complete data set and the random solution are shown as a reference. Note the different ordinate scales of the individual panels. (f) Importance spread plotted as a function of α (see eq. 8). $\sigma 1$ to $\sigma 3$ denote layer conductivities and t1 and t2 layer thicknesses.

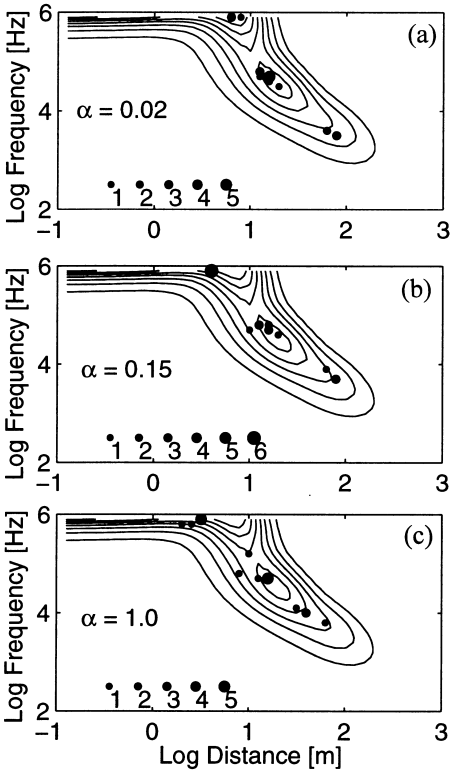

Figure 8. Best equileverage design solutions for some selected α values superimposed on the data importance of the complete data set.

which data require high precision and accuracy, it is also critical to recognize that we have limited ability in the field to improve arbitrarily the precision of an individual data point.

3.6 Restricting duplication

Data duplication can be suppressed in experimental design. We have chosen to multiply the objective function by a positive constant $\beta > 1$,

$$\Gamma^{**} = [\log(\Gamma) + \alpha(H_{ii}^{\max} - H_{ii}^{\min})]\beta^{p}, \tag{9}$$

raised to the power p, where p is the total number of data point duplications. For the H-type model, an empirically determined value of $\beta = 1.1$ has proved suitable for eliminating all duplicate points. Fig. 9 shows the equileverage design obtained by employing eq. (9) instead of (8), and Fig. 10 depicts the corresponding error estimates. Data duplication has been curtailed, and the resulting configurations continue to correlate well with the importance distribution. As shown in Figs 9(b) and (c), increasing α tends to spread out the clustering observed around the importance maxima.

There are no significant differences in *a posteriori* error estimates for the configurations shown in Figs 7 and 10, demonstrating the experimental design optimization problem may be highly ambiguous. In contrast to conventional inversion problems, non-uniqueness in experimental design is desirable because it offers the possibility of imposing additional constraints without degrading the estimated model variances.

4 DISCUSSION AND CONCLUSIONS

Inversion to recover a causative model from measured data is often mathematically ill-posed (Backus & Gilbert 1968; Jackson 1979) and thus inherently non-unique. The problem is exacerbated by band-limited and inaccurate data. Solution non-uniqueness in inverse problems is usually addressed by regularization (e.g. Tikhonov 1963), such that some knowledge of the required models is implicitly or explicitly introduced.

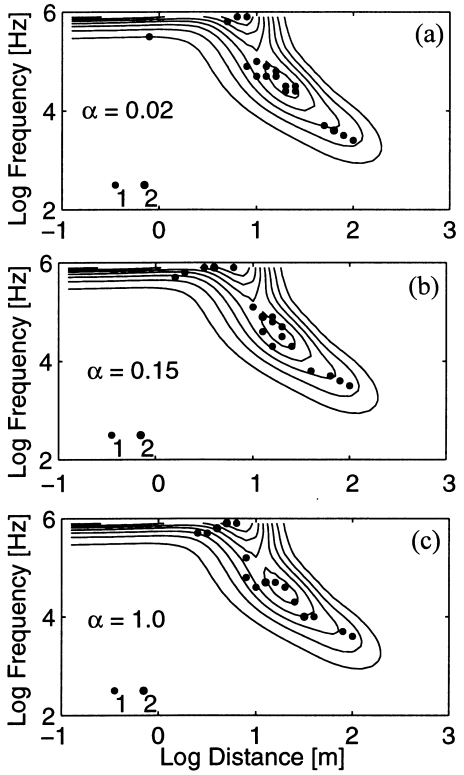

Figure 9. Best solutions of the modified equileverage design (eq. 9) for some selected α values superimposed on the data importance of the complete data set. Compare with Figs 6 and 8.

Examples of regularization include step-length minimization (Marquardt 1970), ad hoc model smoothness (Constable, Parker & Constable 1987), and enforcing stochastic properties on the medium (e.g. Pilkington & Todoeschuck 1991). Regularization can be based either on *information* or *assumptions* about the subsurface structure. While additional data in the inverse problem can improve model resolution substantially, assumptions do not formally contract the model space allowed by the data. In fact, portions of an inverted model may reflect only the subjective regularization constraints. Experimental design is one means of ensuring that the most appropriate data are acquired, mitigating the requirement for regularization assumptions.

Optimal experimental design complements and extends our heuristic design capabilities by treating operational constraints objectively and reducing personal biases. The benefit of experimental design is clearly demonstrated by associating heuristic and data importance pdf experimental design. Experienced experimentalists implicitly understand the role of data importance, yet experiments designed using this approach can unnecessarily degrade model resolution for slightly overdetermined problems, or for weak eigenparameters. Ensuring an overdetermined inverse problem requires not only knowledge of potentially important areas in the data space, but also of the relative interactions of the individual data points. We expect that experimental design will become an important issue in surveys of multidimensional Earth structures, where data acquisition costs serve to make the corresponding inverse problems underdetermined, or only slightly overdetermined.

Heuristic designs cannot include a data robustness criterion. In fact, past experience probably biases the investigator to select leverage points (that is selecting important data regions). When the important regions of the data space can be adequately sampled, heuristic designs will work quite well. However, bias in leverage points can potentially disrupt any ability to resolve the subsurface, particularly when the data are sufficiently limited that the inverse problem becomes underdetermined. For small data sets, the relative contribution of each data point is critical in resolving the model, and the importance pdf-guided (heuristic) design is inappropriate. It is important to point out that robustness applies to data errors

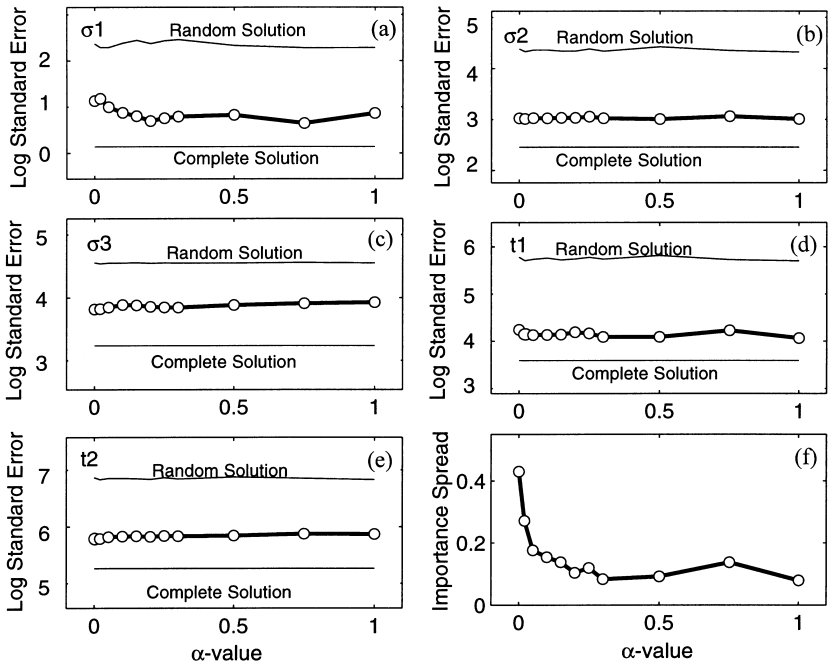

Figure 10. *A posteriori* estimates for the modified equileverage design (eq. 9). Compare with Fig. 7.

as well as inadequacies of the expected model the design was based on.

Experimental design is based on expectations on the subsurface structure, that is it requires a clear intent and a certain amount of *a priori* information. Quantitative designs are thus best suited to feasibility studies or follow-up surveys where increasing the benefit/cost ratio is highly important. It is certainly expedient to consider experimental design when planing additional measurements, or perhaps contemplating the use of another geophysical method. On the other hand, experimental design is probably an effective method of transferring knowledge about the applicability of particular geophysical methods to non-specialists.

ACKNOWLEDGMENTS

We thank Alan Green, Klaus Holliger, Alan Jones and Mark Pilkington for constructive reviews of the manuscript. Comments by the two anonymous journal reviewers and editor G. Müller were valuable in clarifying the presentation. We also thank Andrew Hunter for generously making his genetic algorithm code available to us. Geological Survey of Canada publication number 1997058 and Institute of Geophysics, ETH Zürich, contribution number 979.

REFERENCES

Backus, G. & Gilbert, F., 1968. The resolving power of gross earth data, *Geophys. J. R. astr. Soc.,* **16,** 169–205.

Barth N. H. & Wunsch, C., 1990. Oceanographic experiment design by simulated annealing, *J. phys. Ocean.,* **20,** 1249–1263.

Boerner, D.E. & West, G.F., 1989. Spatial and spectral sensitivity analysis, *Geophys. J. Int.,* **98,** 11–21.

Chave, A.D., 1984. The Fréchet derivatives of electromagnetic induction, *J. geophys. Res.,* **89,** 3373–3380

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,** 279–288.

Curtis, A. & Snieder, R., 1997. Reconditioning inverse problems using the genetic algorithm and revised parameterization, *Geophysics,* **62,** 1524–1532.

Fisher, R.A., 1925. *Statistical Methods for Research Workers,* Oliver & Boyd, Edinburgh.

Glenn, W.E. & Ward, S.H., 1976. Statistical evaluation of electrical sounding methods. Part I: experimental design, *Geophysics,* **41,** 1207–1221.

Hardt, M. & Scherbaum, F., 1994. The design of optimum networks for aftershock recordings, *Geophys. J. Int.,* **117,** 716–726.

Hoaglin, D.C. & Welsch, R., 1978. The hat matrix in regression and ANOVA, *Am. Stat.,* **32,** 17–22.

Holland, J.H., 1975. *Adaptation in Natural and Artificial Systems,* MIT Press, Boston, MA.

Hunter, A., 1995. *SUGAL User Manual,* www:http://osiris.sunderland.ac.uk./ahu/sugal/home.hrml.

Jackson, D., 1979. The use of *a priori* data to resolve non-uniqueness in linear inversion, *Geophys. J. R. astr. Soc.,* **57,** 137–157.

Jones, A.G. & Foster, J.H., 1986. An objective real-time data-adaptive technique for efficient model resolution improvement in magnetotelluric studies, *Geophysics,* **51,** 90–97.

Kijko, A., 1977. An algorithm for the optimum distribution of a regional seismic network-1, *Pageoph,* **115,** 999–1009.

Lanczos, C., 1961. *Linear Differential Operators,* Van Nostrand, London.

Lawson, C.L., Hanson, R.J., 1974. *Solving Least-Squares Problems,* Prentice-Hall, Englewood Cliffs, NJ.

Marquardt, D.W., 1970. Generalized inverse, ridge regression, biased linear estimation and nonlinear estimation, *Technometrics,* **12,** 591–612.

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

Pilkington, M. & Todoeschuck, J.P., 1991. Naturally smooth inversions with *a priori* information from well logs, *Geophysics,* **56,** 1811–1818.

Rabinowitz, N. & Steinberg, D.M., 1990. Optimal configuration of a seismographic network: a statistical approach, *Bull. seism. Soc. Am.,* **80,** 187–196.

Sen, M.K. & Stoffa, P.L., 1995. *Global Optimization Methods in Geophysical Inversion,* Elsevier Science, Amsterdam.

Staudte, R.G. & Sheather, S.J., 1990. *Robust Estimation and Testing,* John Wiley, New York, NY.

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

Tikhonov, A.N., 1963. Regularization of ill-posed problems, *Dokl. Akad. Nauk SSR,* **153,** 1–6.

Ward, S.H. & Hohmann, G.W., 1988. Electromagnetic theory for geophysical applications, in *Electromagnetic Methods in Applied Geophysics,* Vol. 1, pp. 131–311, ed. Nabighian, M.N., Soc. Expl. Geophys., Tulsa, OK.

Wiggins, R.A., 1972. The general linear inverse problem: implication of surface waves and free oscillations for earth structure, *Rev. Geophys.,* **10,** 251–285.

APPENDIX A: ALTERNATIVE FORMULATIONS OF THE OBJECTIVE FUNCTION

There are many different objective functions that can be used instead of eqs (7), (8) or (9). A simple alternative to our choice in eq. (7) is to maximize the condition number, defined as $\Lambda_{ii}^{\min}/\Lambda_{ii}^{\max}$. Unfortunately, the condition number can be maximized by decreasing the largest SV (singular value) as well as by increasing the smallest SV. A better solution would therefore be to consider only $\Lambda_{ii}^{\min}$. However, this is not appropriate for problems with undetermined components, i.e. if $\Lambda_{ii}^{\min}=0$.

Other objective functions found in the literature operate on the inverse of the Hessian matrix $(\mathbf{G}^{\mathsf{T}}\mathbf{G})^{-1}$. Kijko (1977) and Rabinowitz & Steinberg (1990) proposed the 'D-criterion', in which the volume of the *a posteriori* model error ellipsoid, described by $1/\sqrt{\det(\mathbf{G}^{\mathsf{T}}\mathbf{G})}$, is minimized. Hardt & Scherbaum (1994) implemented an experimental design equivalent to the D-criterion by minimizing the product of the eigenvalues of $(\mathbf{G}^{\mathsf{T}}\mathbf{G})^{-1}$. As a variant it is also possible to consider a functional that is proportional to the *sum* of the eigenvalues of $(\mathbf{G}^{\mathsf{T}}\mathbf{G})^{-1}$. Curtis & Snieder (1997) have shown that such an approach is computationally efficient by virtue of avoiding SVD computations.

All these objective functions seek to enhance the magnitude of the eigenvalue related to the most poorly resolved model parameter. In contrast, eq. (7) is most sensitive to the least-determined eigenparameter of the inverse problem, and this could be a function of several model parameters. If the elimination of underdetermined components is the goal of experimental design, eq. (7) would be the preferred choice. On the other hand, choosing an objective function based on the Hessian matrix is advantageous in experimental designs where a particular model parameter has to be resolved with high accuracy. This can be achieved by introducing a weighting function for the eigenvalues.

APPENDIX B: THE GENETIC ALGORITHM

In this study we employed SUGAL, the SUnderland Genetic ALgorithm package, developed by Hunter (1995), which

Table B1. Relevant input parameters for the SUGAL package. See User Manual for further information.

Parameter Name	Parameter Setting
annealing_temperature	10
annealing_decay	0.95
crossover	onepoint
crossover_points	3
elitism	on
generations	2000
init	uniform
mutation	uniform
population	5000
replacement	uniform
selection	roulette
stop	generations

provides wide and easy-to-use support for most of the genetic algorithm parametrizations found in the literature. SUGAL also provides an annealing scheme that may improve the convergence speed. Table B1 summarizes the most important SUGAL settings and options that are common to all our studies. More information on the individual parameters can be found in the SUGAL User Manual located at www : http://osiris.sunderland.ac.uk./ahu/sugal/home.html.

Convergence behaviour is an important aspect of global optimization. In order to determine whether the procedure has converged or not, we analysed the fitness of each generation. Statistical parameters included the fitness of the best and the worst population member as well as an average fitness of the whole population. When all of these quantities are similar, convergence has been achieved. Fig. B1 shows fitness minimum, maximum and average as a function of the generation number for the subset size experiment depicted in Fig. 5. The individual panels reveal that convergence speed is a function of the subset size—the larger the subset, the faster the convergence. This probably reflects the degree of freedom of the different experimental designs. Having only a few data points available, there are only a few configurations that are 'optimal' according to the definition of the objective function. With increasing subset size the variability of desirable configurations increases.

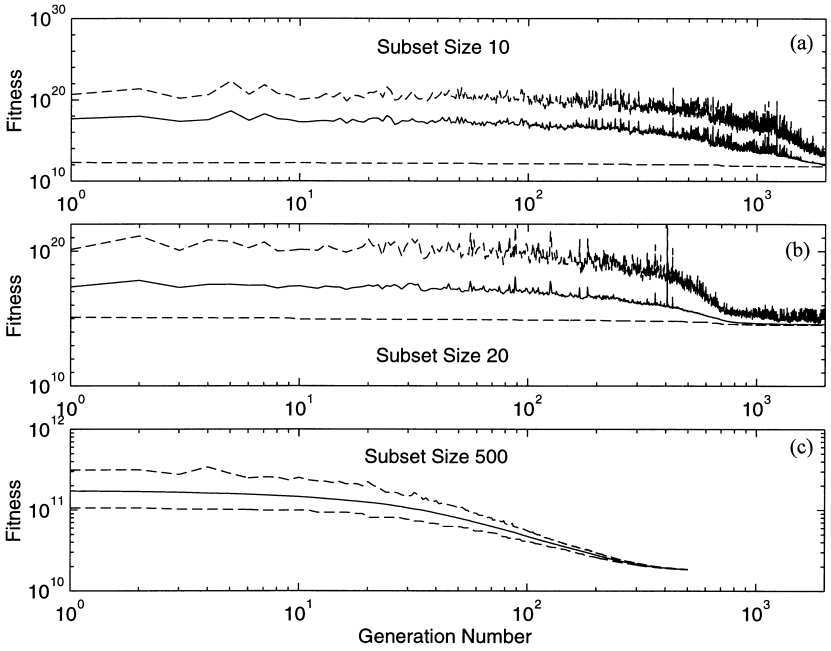

Figure B1. Convergence behaviour of the GA for different subsets (Fig. 5). Fitness of the best and the worst configurations for each generation is plotted with dashed lines, and the average fitness is represented with a solid line.