Geophys. J. Int. (1994) **119**, 353–368

Bayesian statistics of non-linear inverse problems: example of the magnetotelluric 1-D inverse problem

Pascal Tarits,[1,*] Virginie Jouanne,[1] Michel Menvielle[2] and Michel Roussignol[3]

[1]*Institut de Physique du Globe, 4 Place Jussieu, F-75252 Paris, France*
[2]*Laboratoire de Physique de la Terre et des Planètes, Batiment 504, Université Paris Sud, F-91405 Orsay, France*
[3]*Laboratoire de Statistique et Probabilités, UST Lille, F-59655 Villeneuve d'Ascq, France*

Accepted 1994 February 18. Received 1993 November 29; in original form 1992 March 25

SUMMARY

We present a practical algorithm for determining the Bayesian solution of non-linear inverse problems with a limited number of parameters. This approach allows the use of very general conditional probability density functions (pdfs of the data given the model parameters) and *a priori* pdfs (prior beliefs upon the parameters). The results consist in the *a posteriori* marginal pdfs of the model parameters. The marginal pdfs describe the additional information (if any) brought by the data to the prior knowledge about the model parameters.

We have paid a special attention to the numerical calculation of the *a posteriori* pdfs and we propose a solution which enables the simultaneous determination of the numerical estimates of the *a posteriori* pdfs and their uncertainties.

With the help of a practical example (the 1-D magnetotelluric inverse problem) with synthetic and real data, we show that in most cases, the *a posteriori* marginal pdfs are complicated functions and cannot be predicted from the data pdf or the *a priori* pdfs. Consequently, the standard analyses applied to the same data sets such as the maximum likelihood techniques or the asymptotic estimation technique would lead to severely biased estimates of the parameters. This remark is likely to be true for any non-linear inverse problem.

Key words: *a posteriori* pdf, *a priori* probability density function (pdf), Bayesian inversion, magnetotelluric inversion, marginal pdf.

1 INTRODUCTION

The solution of an inverse problem first entails the determination of an acceptable model which verifies a given criterion (e.g. χ^2 tests, maximum likelihood solutions...). The second step is the estimation of the reliability of the inverse solution given the data set and the class of models used to simulate the data. This step is essential to interpret quantitatively the results of the inversion. The reliability of the inverse solution may be discussed in the classical frame of resolution-uncertainty trade-offs (e.g. Backus 1970) or with a statistical approach which may allow the determination of an uncertainty on the inverse solution with the help of assumptions such as models described by a Gaussian probability density function (pdf) (e.g. Tarantola & Valette 1982; Duijndam 1988b). These approximations are necessary in many cases because of the complexity of the model, the large number of data involved and the equally large number of parameters to be determined. There is no guarantee, however, that the solutions obtained are not biased by the assumptions made.

We have explored this question with the help of a Bayesian description of the inverse problem. The Bayesian inversion has been discussed during the last 10–20 years (e.g. Backus 1970, 1988; Tarantola & Valette 1982; Jackson & Matsu'ra 1985) and the method is now used more frequently because of the availability of powerful computers (e.g. Duijndam 1988a,b; Mackie, Bennet & Madden 1988). Bayes's theorem (1763) expresses the solution of an inverse problem in terms of the statistical distribution of a set of random data given a model and prior assumptions upon its parameters. This solution is well suited for appropriate classes of model with data that can be described in terms of a statistical distribution and when *a priori* information—even vague—is available. The *a posteriori* solution provides a complete statistical description of the model parameters given *a priori* information.

*Now at: Laboratoire de Géosciences Marines (URA 1278) U.B.O., 6 Av Le Gorgeu, F-29275 Brest cedex, France.

From this formal point of view, the Bayesian approach is very attractive for Earth Sciences investigations. Practically, the Bayesian solutions involve a tremendous amount of calculation. To overcome this difficulty, linear or linearizable problems with Gaussian data and priors have been studied thoroughly. For these problems, the *a posteriori* pdf is also Gaussian. Analytical (for linear problems) or iterative (for certain non-linear problems) solutions may be found (e.g. Jackson & Matsu'ura 1985) and this type of approach has been used by several authors in various fields (e.g. Duijndam 1988a; Backus 1988; Mackie *et al.* 1988). These authors also showed that the Bayesian solutions based on the Gaussian approximation were similar to other stochastic methods, although the interpretation was different because of the introduction of the *a priori* pdf concept.

One of the limitations of the classical stochastic inverse techniques is that Gaussian priors are rarely adequate to describe our *a priori* beliefs about the model parameters. Furthermore, for non-linear problems, the Gaussian hypothesis does not lead to an *a posteriori* Gaussian pdf (Jackson & Matsu'ura 1985). When the non-linear problem is linearized, the solution cannot converge towards an *a posteriori* model that is very different from the original and much of the richness of the Bayesian approach is lost.

We have restricted our study to non-linear problems with a small number of parameters to be determined. This allows the complete calculation of the parameters's *a posteriori* pdfs. Although most geophysical inverse problems are theoretically infinite dimensional, geophysical and geological evidence shows that some are in practice well approximated using finite-dimensional models, with a limited number of relevant parameters. For these problems, a general *a posteriori* solution may be found numerically at a reasonable cost (e.g. Tarits & Jouanne 1990; Provost, Garcia & Carçon 1992) and without *a priori* assumptions about its nature. The *a posteriori* pdf is analysed with the help of marginal pdfs. We illustrate the complexity of the statistics of the *a posteriori* model parameters with an application to the case of the 1-D magnetotelluric inverse problem.

2 BAYESIAN STATISTICS

2.1 Data and models

Consider a random data vector $\mathbf{Z}$ with nf data points $(Z_1, Z_2, \ldots, Z_{nf})$. The data are noised measurable physical quantities and are considered to be random variables. They are the observations of some physical process defined by a set $\mathbf{P}$ of np parameters $(P_1, P_2, \ldots, P_{np})$ which may not be directly measured. Let $f(\mathbf{Z} \mid \mathbf{P})$ be the pdf for the random data vector $\mathbf{Z}$ given the parameters $\mathbf{P}$. The expected value of $\mathbf{Z}$ is the solution of the forward problem given the parameter vector $\mathbf{P}$ and is denoted by $\mathbf{F}$. We assume that the relation between the expected value $\mathbf{F}(\mathbf{P})$ of the data vector $\mathbf{Z}$ and the parameters $\mathbf{P}$ is known (i.e. the forward problem has a solution).

2.2 Bayesian estimators

Let $g(\mathbf{P})$ be the *a priori* pdf for the parameter vector $\mathbf{P}$ over some finite domain $D \in \mathcal{R}^{+(np)}$. The function $g(\mathbf{P})$ may be any pdf simulating our *a priori* beliefs about the parameters.

In the Bayesian context, $f(\mathbf{Z} \mid \mathbf{P})$ is the conditional pdf of the random data vector $\mathbf{Z}$ given the parameters $\mathbf{P}$. The joint pdf for the two random vectors $\mathbf{Z}$ and $\mathbf{P}$ is $f(\mathbf{Z} \mid \mathbf{P}) \cdot g(\mathbf{P})$. The conditional pdf of $\mathbf{P}$ given $\mathbf{Z}$, $f(\mathbf{P} \mid \mathbf{Z})$, is the *a posteriori* pdf. It is given by Bayes's theorem (e.g. Bayes 1763; Tarantola 1987):

$$f(\mathbf{P} \mid \mathbf{Z}) = \frac{f(\mathbf{Z} \mid \mathbf{P}) \cdot g(\mathbf{P})}{\int_{(D)} \cdots \int f(\mathbf{Z} \mid \mathbf{P}) \cdot g(\mathbf{P})\, d\mathbf{P}}. \tag{1}$$

The *a posteriori* pdf $f(\mathbf{P} \mid \mathbf{Z})$ may be used to determine various types of estimates. For instance we may calculate the *a posteriori* maximum likelihood which is equivalent to the classical likelihood estimate if $f(\mathbf{Z} \mid \mathbf{P})$ and $g(\mathbf{P})$ were Gaussian and $\mathbf{F}$ linearly related to $\mathbf{P}$. In that case, it is straightforward to prove that the *a posteriori* pdf is also Gaussian (e.g. Backus 1988).

The Bayesian estimator $\langle \mathbf{P} \rangle$ of the expected value of $\mathbf{P}$ is given by the relation:

$$\langle \mathbf{P} \rangle = \int_{P_1} \cdots \int_{P_{np}} \mathbf{P} \cdot f(\mathbf{P} \mid \mathbf{Z})\, dP_1, \ldots, dP_{np}. \tag{2}$$

The higher order moments may be determined similarly.

For an inverse problem with a large number of parameters, the direct numerical calculation of relations (1) or (2) is not possible and the common practice is to restrict the analysis to the calculation of the maximum likelihood and asymptotic solutions (e.g. Jackson & Matsu'ura 1985). In some cases however, specific algorithms may be found to determine the Bayesian estimators (e.g. Roussignol *et al.* 1993). The direct numerical approach is realistic only for a limited number of parameters.

2.3 Marginal probability density functions

Although eq. (1) contains all the information about the *a posteriori* parameter vector $\mathbf{P}$, it does not provide a practical description of the individual pdfs of the parameters. This information may be obtained with the marginal pdfs of the *a posteriori* pdf.

The marginal pdf for a given paremater P_i is the integral of the pdf over the whole domain of the parameters less P_i (e.g. Tarantola 1987; Duijndam 1988a; Press 1989). The marginal pdf $f^i_m(P_i)$ is given by:

$$f^i_m(P_i) = \int_{P_1} \cdots \int_{P_{i-1}} \int_{P_{i+1}} \cdots \int_{P_{np}} f(\mathbf{P} \mid \mathbf{Z})\, dP_1 \cdots dP_{i-1} dP_{i+1} \cdots dP_{np}. \tag{3}$$

The expected values $\langle \mathbf{P} \rangle (\langle P_1 \rangle, \ldots, \langle P_{np} \rangle)$ and the variances $\mathbf{S}(S_1, \ldots, S_{np})$ may be derived from the definition of the marginal pdfs and are given by:

$$\langle P_i \rangle = \int_{P_i} P_i f^i_m(P_i)\, dP_i, \tag{4a}$$

$$S_i = \int_{P_i} P_i^2 f^i_m(P_i)\, dP_i - \langle P_i \rangle^2. \tag{4b}$$

The marginal pdf f^i_m may be interpreted as the pdf of the parameter P_i. Any departure of the *a posteriori* marginal

pdfs from the *a priori* beliefs is an indication of new knowledge of the parameters. For general non-linear problems, its distribution cannot be predicted from the *a priori* pdf or from the data pdf. It must be calculated from eq. (3) and examined for each parameter.

At best, we expect the marginal pdfs to be unimodal (i.e. the pdf has a single and well delimited maximum value) when the parameter is well resolved. In that case the *a posteriori* expected values and the associated variances may be interpreted by analogy to a nearly Gaussian distribution. On the contrary, when the marginal pdf is not unimodal, the expected values and the variances of the parameters do not have a simple interpretation. We shall see in the examples studied in the next sections that the nature of the marginal pdfs varies significantly from one parameter to the other depending upon the data. Eventually, when the *a posteriori* marginal pdf is identical to the *a priori* pdf, the data have provided no additionnal information.

3 COMPUTATION OF THE BAYESIAN ESTIMATORS

Although a very powerful tool in analysing the inference problem, the Bayesian approach suffers from the requirement to evaluate multiple integrals (as many as the number of parameters). These calculations are time consuming and the numerical application of such algorithms has to be studied carefully. The Monte Carlo method is in practice efficient in computing multiple integrals of large order (typically seven and more) when the function to be integrated is not easily approximated by functions for which the integration is easy to perform (e.g. Tarantola 1987; Press *et al.* 1989). We use this method to calculate simultaneously the numerical estimates of the marginal pdfs, the *a posteriori* expected values and the variances of the parameters along with the numerical error of these estimates.

3.1 Numerical estimates of the marginal pdfs

Let a given *a priori* model have np parameters. Each parameter P_i $(1 \le i \le np)$ has a pdf over a finite interval. Let α_i (and β_i) be the minimum and the maximum bounds, respectively, for this parameter. Let D be a domain of $\mathcal{R}^{np}$. D is the cartesian product of the intervals $[\alpha_i, \beta_i]$, $(1 \le i \le np)$. We define the *a posteriori* marginal pdf for a parameter $P_i \in [\alpha_i, \beta_i]$ by:

$$f^i_m(P_i) = \frac{\int_{D_i} \cdots \int f(\mathbf{Z} \mid \mathbf{P}) g(\mathbf{P})\, dP_1, \ldots, dP_{i-1} dP_{i+1}, \ldots, dP_{np}}{\int_{D} \cdots \int f(\mathbf{Z} \mid \mathbf{P}) g(\mathbf{P})\, dP_1, \ldots, dP_{np}} . \tag{5}$$

D_i is a domain of $\mathcal{R}^{np-1}$:

$$D_i = [\alpha_1, \beta_1] \times [\alpha_2, \beta_2] \times , \ldots , \times [\alpha_{i-1}, \beta_{i-1}] \times [\alpha_{i+1}, \beta_{i+1}] \times , \ldots , \times [\alpha_{np}, \beta_{np}].$$

The expected value $\langle P_i \rangle$ and the variance S_i for the parameter P_i are given by:

$$\langle P_i \rangle = \int_{[\alpha_i, \beta_i]} P_i f^i_m(P_i)\, dP_i, \tag{6}$$

$$S_i = \int_{[\alpha_i, \beta_i]} P_i^2 f^i_m(P_i)\, dP_i - \langle P_i \rangle^2. \tag{7}$$

We use a discrete approximation $f_i(k)$, $(1 \le k \le K)$, of the *a posteriori* marginal pdf f^i_m to calculate these quantities. The discretization consists of dividing the intervals $[\alpha_i, \beta_i]$, $(1 \le i \le np)$, in K equal segments $I_{i1}, \ldots, I_{iK}$ ($I_{i1} \cup , \ldots , \cup I_{iK} = [\alpha_i, \beta_i]$). The discrete pdf $f_i(k)$ over the interval I_{iK} is:

$$f_i(k) = \int_{I_{ik}} f^i_m(P_i) dP_i \qquad 1 \le k \le K. \tag{8a}$$

With the discrete form of the marginal pdf, we define discrete approximations of the expected value and of the variance that will be used to replace eqs (6) and (7):

$$\langle P_i \rangle = \sum_{k=1}^{K} p_{ik} f_i(k) \approx \langle P_i \rangle, \tag{8b}$$

$$S_i = \sum_{k=1}^{K} p_{ik}^2 f_i(k) - \langle P_i \rangle^2 \approx S_i, \tag{8c}$$

where p_{ik} is the central value in the segment I_{ik}.

Let

$$D_{ik} = [\alpha_1, \beta_1] \times [\alpha_2, \beta_2] \times , \ldots , \times [\alpha_{i-1}, \beta_{i-1}] \times I_{ik} \times [\alpha_{i+1}, \beta_{i+1}] \times , \ldots , \times [\alpha_{np}, \beta_{np}]$$

be a domain of $\mathcal{R}^{np}$. The pdf $f_i(k)$ is given by:

$$f_i(k)_{1 \le k \le K} = \frac{\int_{D_{ik}} \cdots \int f(\mathbf{Z} \mid \mathbf{P}) g(\mathbf{P})\, dP_1, \ldots, dP_{np}}{\int_{D} \cdots \int f(\mathbf{Z} \mid \mathbf{P}) g(\mathbf{P})\, dP_1, \ldots, dP_{np}} . \tag{9}$$

The computation of each pdf $f_i(k)$ involves the calculation of two np-order integrals E and E_{ik} defined by:

$$E_{ik} = \int_{D_{ik}} f(\mathbf{Z} \mid \mathbf{P}) g(\mathbf{P})\, d\mathbf{P}, \tag{10a}$$

$$E = \int_{D} f(\mathbf{Z} \mid \mathbf{P}) g(\mathbf{P})\, d\mathbf{P}. \tag{10b}$$

The procedure is described below.

We draw at random L independent variables $\mathbf{X} = \{X_1, \ldots, X_L\}$ in $\mathcal{R}^{np}$ with the same pdf $h(\mathbf{X})$ strictly positive over D. We rewrite eqs (10a) and (10b) in order to interpret E_{ik} and E as the expected values over D_{ik} and D, respectively, of the function $f(\mathbf{Z} \mid \mathbf{X}) \cdot g(\mathbf{X})/h(\mathbf{X})$:

$$E_{ik} = \int_{D_{ik}} \left[f(\mathbf{Z} \mid \mathbf{X}) \frac{g(\mathbf{X})}{h(\mathbf{X})} \right] h(\mathbf{X})\, d\mathbf{X}, \tag{11a}$$

$$E = \int_{D} \left[f(\mathbf{Z} \mid \mathbf{X}) \frac{g(\mathbf{X})}{h(\mathbf{X})} \right] h(\mathbf{X})\, d\mathbf{X}. \tag{11b}$$

We define the quantities denoted $Y_{ik}^{(L)}$ and $Y^{(L)}$ by:

$$Y_{ik}^{(L)} = \frac{1}{L}\sum_{l=1}^{L} f(\mathbf{Z} \mid \mathbf{X}_l)\frac{g(X_l)}{h(X_l)}\mathcal{J}_{(X_l \in D_{ik})}, \tag{12a}$$

$$Y^{(L)} = \frac{1}{L}\sum_{l=1}^{N} f(\mathbf{Z} \mid \mathbf{X}_l)\frac{g(\mathbf{X}_l)}{h(X_l)}, \tag{12b}$$

with $\mathcal{J}_{\{X_l \in D_{ik}\}} = 1$ for $X_l \in D_{ik}$ but is 0 otherwise.
As a consequence of the large number law, it comes to the limit:

$$Y_{ik}^{(L)} \xrightarrow[L\to\infty]{\text{a.s.}} E_{ik},$$

$$Y^{(L)} \xrightarrow[L\to\infty]{\text{a.s.}} E.$$

Note here that the abbrevation 'a.s.' stands for 'almost surely' and indicates that we exclude exotic situations like zero probabilities.

Then almost surely (in physically realistic situations at least):

$$F_{ik}^{(L)} = \frac{Y_{ik}^{(L)}}{Y^{(L)}} \xrightarrow[L\to\infty]{\text{a.s.}} f_i(k). \tag{13}$$

Hence the expected value $\langle P_i^{(L)}\rangle$ and the variance $S_i^{(L)}$ are the estimates of $\langle P_i\rangle$ and S_i, respectively and, are given by:

$$\langle P_i^{(L)}\rangle = \sum_{k=1}^{K} p_{ik}F_{ik}^{(L)} \xrightarrow[L\to\infty]{\text{a.s.}} \langle P_i\rangle, \tag{14a}$$

$$S_i^{(L)} = \sum_{k=1}^{K} p_{ik}^2 F_{ik}^{(L)} - \langle P_i^{(L)}\rangle^2 \xrightarrow[L\to\infty]{\text{a.s.}} S_i. \tag{14b}$$

A systematic sampling of D with K intervals for np parameters would need the calculation of $f(\mathbf{Z} \mid \mathbf{P})$ K^{np} times. This may turn out to be an enormous number. On the other hand the random sampling of D with a number L of trials leads to numerical estimates of the multiple integrals with an accuracy which may be determined statistically and the value of L may be tuned to reach a given precision in the computation. We show in section that we get a reasonable estimate of the multiple integrals and hence of the *a posteriori* Bayesian estimators, $F_{ik}^{(L)}$, $\langle P_i^{(L)}\rangle$ and $S_i^{(L)}$ with L much less than K^{np}.

3.2 Numerical uncertainty on the estimates of the marginal pdfs

In this section we present an estimation of the errors related to the Monte Carlo calculation of the pdf $f_i(k)$ and of the expected value $\langle P_i\rangle$ for each parameter P_i ($1 \le i \le np$, $1 \le k \le K$). The calculation of the errors is based on the multiple dimension central limit theorem. The detail of the derivations is given in Appendix A.

3.2.1 *Error in the marginal pdf estimates*

We prove in Appendix A that the pdf of $\sqrt{L}[F_{ik}^{(L)} - f_i(k)]$ converges as $L \to \infty$ towards a normal pdf of zero mean and of variance U_{ikk} given by (see Appendix A):

$$U_{ikk} = \frac{V_{ik}(E - E_{ik})^2 + E_{ik}^2(V - V_{ik})}{E^4}, \tag{15}$$

with V_{ik} and V defined by:

$$V_{ik} = \int_{D_{ik}} \left[\frac{f(\mathbf{Z} \mid \mathbf{P})g(\mathbf{P})}{h(\mathbf{P})}\right]^2 h(\mathbf{P})\, d\mathbf{P}, \tag{16a}$$

$$V = \int_{D} \left[\frac{f(\mathbf{Z} \mid \mathbf{P})g(\mathbf{P})}{h(\mathbf{P})}\right]^2 h(\mathbf{P})\, d\mathbf{P}. \tag{16b}$$

The standard deviation $\sigma[F_i^{(L)}]$ of $|F_{ik}^{(L)} - f_i(k)|$ therefore is:

$$\sigma[F_i^{(L)}] = \frac{\sqrt{U_{ikk}}}{\sqrt{L}}. \tag{17}$$

This formula shows that the error decreases as the square root of the number of random trials increases. It is possible to tune the optimum number L of random trials at a given error level with few runs. The error also depends on the model parameters and on the data through U_{ikk}. The consequence is that the number of random trials necessary to reach a certain error level will also have to be adjusted according to the data sets and the *a priori* models.

3.2.2 *Uncertainty in the estimates of the expected value* $\langle P_i\rangle$

We denote $\mathbf{p}_i$ and $\mathbf{F}_i^{(L)}$ the column vectors $(p_1, p_2, \ldots, p_{ik}, \ldots, p_{iK})$ and $[F_{i1}^{(L)}, F_{i2}^{(L)}, \ldots, F_{ik}^{(L)}, \ldots, F_{iK}^{(L)}]$, respectively. Eq. (14a) may be written as $\langle P_i^{(L)}\rangle = {}^t\mathbf{p}_i \cdot \mathbf{F}_i^{(L)}$. ${}^t\mathbf{p}_i$ is the transposed vector of $\mathbf{p}_i$.

In Appendix A we prove that the pdf of $\sqrt{L}(\langle P_i^{(L)}\rangle - \langle P_i\rangle)$ converges as $L \to \infty$ towards a normal pdf of zero mean and of variance ${}^t\mathbf{p}_i \cdot \mathbf{U}_i \cdot \mathbf{p}_i$ where $\mathbf{U}_i$ is the variance–covariance matrix. Its diagonal terms U_{ikk} are given by eq. (15) and its non-diagonal terms are given by (see Appendix A):

$$U_{ikj} = \frac{E_{ik}E_{ij}V - V_{ik}E_{ij}E - E_{ik}EV_{ij}}{E^4} \quad k \neq j. \tag{18}$$

The standard deviation $\sigma(\langle P_i^{(L)}\rangle)$ of $|\langle P_i^{(L)}\rangle - \langle P_i\rangle|$ is therefore given by:

$$\sigma(\langle P_i^{(l)}\rangle) = \frac{\sqrt{{}^t\mathbf{p}_i \cdot \mathbf{U}_i \cdot \mathbf{p}_i}}{\sqrt{L}}. \tag{19}$$

Again, the error decreases as the square root of the number of random trials increases.

Equation (19) leads to the estimation of the expected value of the variance $S_i^{(L)}$, $E[S_i^{(L)}] = S_i - \mathbf{p}_i^t \cdot \mathbf{U}_i \cdot \mathbf{p}_i/L$. The expected value is asymptotically unbiased because $\mathbf{p}_i^t \cdot \mathbf{U}_i \cdot \mathbf{p}_i/L$ vanishes as $L \to \infty$.

4 AN EXAMPLE OF THE 1-D MT INVERSE PROBLEM

We consider the example of magnetotelluric (MT) soundings limited to the 1-D approximation (i.e. the electrical resistivity varies only with a depth). The MT technique is used to determine the distribution of electrical resistivity of the Earth which in turn provides constraints on the Earth's structure and its thermodynamics (e.g. mineralogy, amount of fluids, temperature at depth, and partial melting).

MT data sets consist of complex coefficients (also called MT impedances) determined from the frequency analysis of simultaneous records of the natural fluctuations of the

electric and magnetic fields. The frequency range of the fluctuations is quite broad (thousands of hertz to less than one cycle per year). The frequency range measured during a MT sounding experiment is fixed by the objectives: for crustal investigation, the bandwith is approximately 1000–0.001 Hz. Deeper investigation involves data acquisition at lower frequencies and accordingly shallower investigations imply data acquisition at higher frequency. The MT data set is generally limited to five to 10 data points per decade for a single MT sounding. Each coefficient is an independent statistical estimate. In most cases, the pdf of the MT coefficients is Gaussian or may be approximated by a Gaussian distribution after robust processing (i.e. Chave *et al.* 1987) and is characterized by a mean and a variance. Note that this is not essential in our discussion and other distributions could be used as well.

4.1 Numerical considerations

The choice of the pdf $h(.)$ of the random variables $X_1, \ldots, X_L$ is based upon two contradictory criteria: the necessity to simulate a random selection of parameters according to $h(.)$ at a good computation rate and to maintain the rate of convergence of the variance–covariance matrix $\mathbf{U}/L$ as high as possible. The latter criterion leads to select functions $h(.)$ which minimize $\mathbf{U}/L$. In practice, this criterion is very difficult to fulfil with functions $h(.)$ which also satisfy the first criterion.

The uniform pdf $h(.)$ over D has the advantage of being very efficient in terms of computation time, for simulating the L random sets of parameters. However, this choice is poor for calculating the *a posteriori* pdfs, hence the necessity to be able to determine their numerical error.

The choice of D is closely linked to the choice of the model parameters and will be discussed later on.

4.2 Synthetic data analysis

We have selected a synthetic set of MT data in order to illustrate the Bayesian approach to the inference problem. The results presented depict the complexity of the marginal pdfs of the model parameters with respect to their *a priori* pdf.

4.2.1. A priori *information*

As stated in Section 2.1, we need to specify the nature of the forward modelling technique used to calculate the expected values of the data set $\mathbf{Z}$ and accordingly the conditional pdf of the data given the model parameters. This step already includes an *a priori* hypothesis common to any inverse problem—the nature of the model class. In the example studied, we used models which consisted of stacked homogeneous layers over a homogeneous half-space. The parameters $\mathbf{P}$ are the resistivities, ρ in each layer and the depths, z of each interface (Fig. 1).

To build up the *a priori* pdf $g(\mathbf{P})$, we need the following *a priori* information: (1) the number of homogeneous layers and (2) the *a priori* pdfs for the parameters.

For each parameter P_i, the pdf $h(.)$ is taken uniform over a given interval $[P_{\min}, P_{\max}]$. The domaine D defined in Section 2 is the union of the intervals $[P_{\min}, P_{\max}]$. The range for the resistivities was limited to four to five orders of magnitude for numerical reasons. The ranges of depths of the layer interfaces were taken such that no overlap from one interface to the next could be possible.

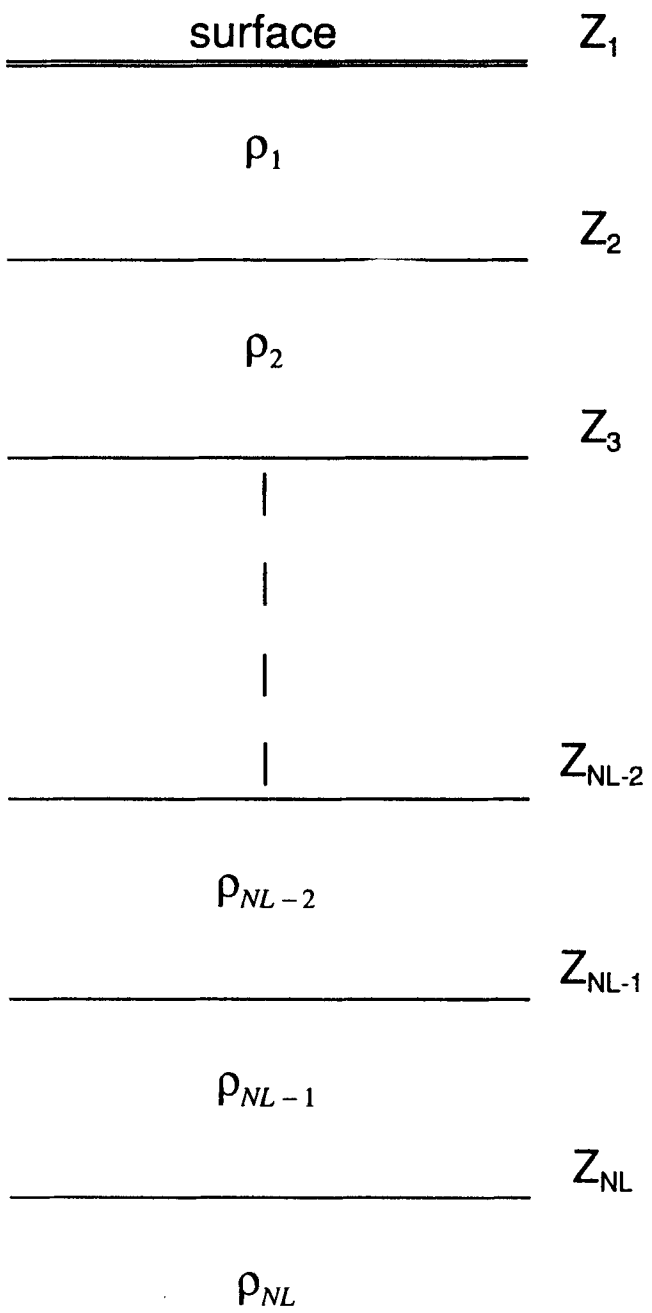


Figure 1. MT parameters for the 1-D model of electric resistivity vs. depth. The z_is are the depths of the interfaces between two successive homogeneous layers and $z_1 = 0$.

For the present discussion, we will consider successively cases where the number of homogeneous layers in the *a priori* model is correct, overestimated, and underestimated.

4.2.2 The data set

Among the infinite range of possible models which could be used for testing, we selected a model simulating data which

Table 1. Parameters of the synthetic five-layer model P_0.

Layers	Resistivity Ω − m	Depth interface
1	200	0
2	10	20
3	100	70
4	300	140
5	0.1	300

could have been recorded on the sea-bottom. This choice was guided by the fact that in the real world, the frequency range of ocean-bottom MT data is quite narrow and on the lowest side of the spectrum (typically 10^{-6} to 10^{-3} Hz, e.g. Filloux 1987). Furthermore, the natural signal energy is weaker on the sea-floor than inland because of the ocean cut-off. This is therefore one of the most difficult conditions for inverting MT data sets.

The selected synthetic five-layer model P_0 is presented in Table 1. The model is characterized by a resistive top layer overlying a conductive intermediate medium. The bottom half-space is conductive. The resistivity of the intermediate medium increases from layer 2 to layer 3. The vertical scales and the values for the resistivities are realistic for a young oceanic lithosphere (e.g. Tarits 1986) although the details of the resistivity profile may not be. Their purpose are to test the ability of the algorithm to describe *a posteriori* parameter pdfs in different situations.

The impedances have been computed at 22 frequencies and randomly tried within an interval equal to standard deviations obtained at the same frequencies for real sea-floor data processed with a band-averaging technique. The final error bars on the synthetic data set are typically two to four times larger than on natural data processed with more sophisticated techniques such as robust processing (e.g. Chave, Thomson & Ander 1987). The set of sea-floor impedances (Z_{ref}) are presented in Table 2.

4.2.3 *The* a posteriori *pdfs*

The computations of the *a posteriori* pdfs were carried out with from 50 000 (for preliminary adjustments) to up to 5×10^6 (for final analysis for models with up to 13 parameters) random sets of independent decimal logarithm values for the parameters $\mathbf{P} = (\rho, z)$. For each parameter P_i, the *a priori* interval $[P_{i\min}, P_{i\max}]$ was divided into $K = 20$ equal segments I_{ik}, $k = 1, 2, \ldots, K$. This discretization was used for all the examples studied including the example with real MT data.

The Bayesian analysis of the data was always performed with a number of parameters lower than the number of data. For numerical reasons, we restricted ourselves to *a priori* models with a number of layers less than or equal to seven (13 parameters).

For each run, we calculated the *a posteriori* marginal pdfs and discussed their nature in order to characterize the additional information from the data for each parameter.

Table 2. The synthetic impedances with noise and their 95 per cent error bars (in percentage).

Period (s)	Real (μV/m/nT)	Imaginary (μV/m/nT)	Error (%)
930	0.14038	0.22518	16
1100	0.12151	0.19440	18
1300	0.18345	0.21854	13
1800	0.13640	0.13878	15
2900	0.15104	0.11441	12
4100	0.11355	0.12149	12
4700	0.12527	0.09115	13
5200	0.14081	0.11563	12
5800	0.10545	0.10565	13
6200	0.10073	0.10522	11
6700	0.10839	0.08938	9.2
7200	0.10051	0.08689	10
7300	0.10136	0.09374	7.9
8500	0.09504	0.10096	11
10000	0.08176	0.09241	8.8
12000	0.07632	0.10250	12
16000	0.05846	0.09247	21
21000	0.04056	0.08604	28
27000	0.02659	0.05385	23
31000	0.02694	0.05740	17
38000	0.00244	0.05502	27
120000	0.00007	0.01596	27

4.2.4 *Analysis of synthetic data*

4.2.4.1 Five-layer a priori *model.* For this example, the *a priori* model has the same number of layers as P_0 and therefore has nine parameters. The *a priori* pdfs for the resistivities and the depths of the interfaces are shown in Fig. 2(a). For this run, we randomly tried 1.6×10^6 sets of resistivities and depths of the interfaces.

The marginal pdfs for this run are shown in Fig. 3(a). The horizontal axis represents the interval $[P_{\min}, P_{\max}]$ for each parameter. The amplitude of the marginal pdf—normalized to a unit surface—is along the vertical axis. The crosses represent ±1 standard deviation. Note how accurately the pdfs are determined for a number of random trials much less than the number of discrete models necessary to sample regularly the nine intervals (5×10^{11}).

The marginal pdfs of ρ_2, z_2 and z_5 are unimodal though the pdf for z_2 spreads over most of the *a priori* interval. The data resolve these parameters which in turn may be characterized by a mean value and a standard deviation (very large in the case of z_2).

The marginal pdfs of ρ_1, ρ_3, ρ_4 and ρ_5 have an approximate step function shape. This indicates that the data only allow the determination of either a minimum (ρ_1, ρ_3, ρ_4) or a maximum (ρ_5) value for the parameters. Finally, the marginal pdf for z_4 is uniform over $(z_{4\min}, z_{4\max})$. The data provide no additional information to the *a priori* belief. Although the original model had five layers, the imperfect (i.e. noised and discrete over a narrow frequency band) data set does not resolve this many parameters.

The *a posteriori* model deduced from examination of the marginal pdfs is presented in Fig. 4. We represented the

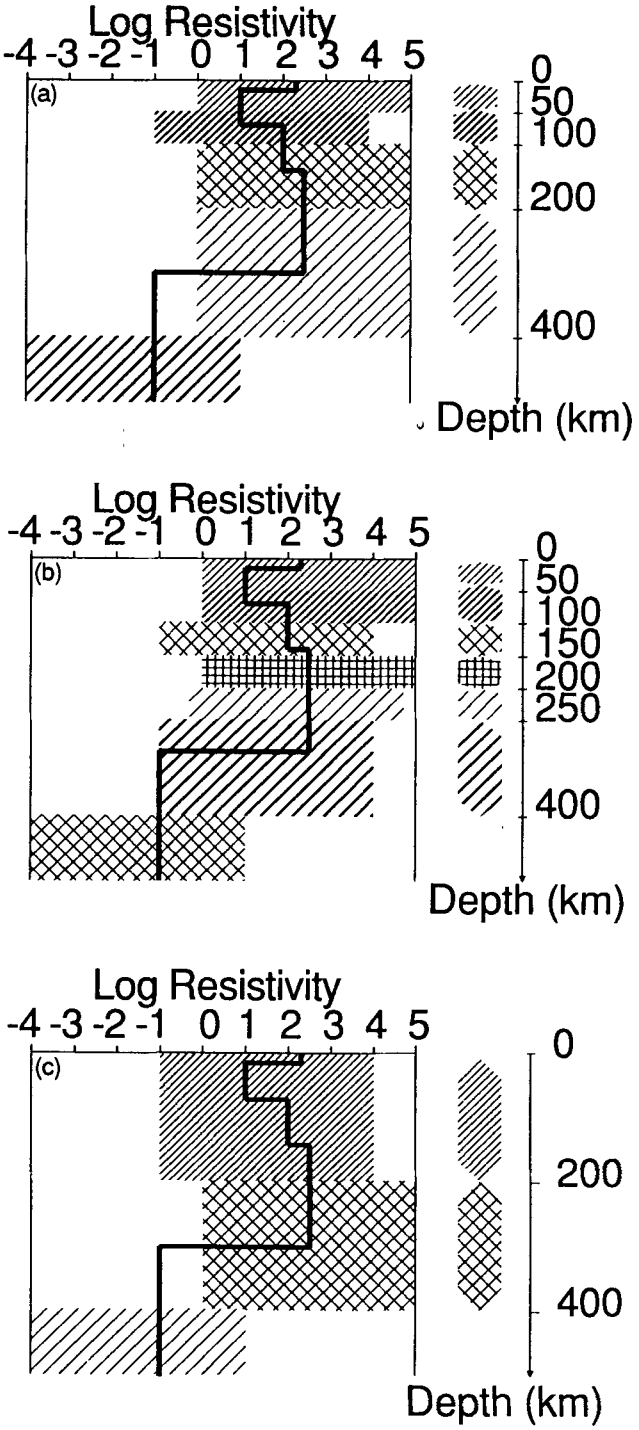


Figure 2. *A priori* pdfs for the inversion of the synthetic data set (see text for details). (a) Five-layer model, (b) seven-layer model, (c) three-layer model. The pdf is uniform over the intervals defined by the shaded areas and is zero outside. The vertical shaded areas represent the range of possible values for the depth of each interface. Each random trial samples a set of resistivities and depths within the shaded intervals.

mean of the parameters plus or minus 1 standard deviation when the parameters had a unimodal marginal pdf. We estimated graphically the upper (or the lower) bounds for the rest of the parameters. The black thick line depicts the original model P_0.

We also considered the importance of the number of data points as well as the noise in the data, first by taking half the data set (the low-frequency half, or the high-frequency half or alternate frequencies) and second by taking the original error bars divided by four.

Taking half the data set led to only small changes in the marginal pdfs. They were smoothed out but their main characteristics did not change. Dividing the error bars by four led to a better resolution on some of the parameters (but not all). The marginal pdfs are more complicated with an increase of the error (Fig. 3b) because the number of random trials (1.6×10^6) was too small to sample correctly the *a priori* intervals to get numerical estimates with the same accuracy as shown in Fig. 3(a).

4.2.4.2 Seven-layer a priori *model.* We now consider an *a priori* model with more parameters than in the model P_0. The *a priori* model is presented in Fig. 2(b). 5×10^6 random trials were used. The marginal pdfs are shown on Fig. 5(a). Note the overall decrease in the resolution, particularly in the depths of the interfaces. Only the interface z_7 and the resistivity ρ_2 are resolved by the data; z_2 is almost unresolved. z_3 and ρ_1 are more or less step-like as well as ρ_3–ρ_6 which are almost identical. The corresponding interfaces depth z_4–z_6 are not resolved.

Analysis of the marginal pdfs suggests that the number of parameters in the *a priori* model is too large. A parameter in excess is either unresolved or has a marginal pdf very similar to the pdf of one of the adjacent parameters.

4.2.4.3 Three-layer a priori *model.* Consider now an *a priori* model with less parameters than the original P_0 model. The *a priori* model is presented in Fig. 2(c). 10^6 random trials were used for the computation of the marginal pdfs presented in Fig. 5(b). In this example, we observe that both interfaces are resolved as well as the resistivity ρ_1. The pdfs for ρ_2 and ρ_3 still have a step-like shape. Note that z_3 is centred over an intermediate value (120 km) which does not correspond to any of the P_0s. This example illustrates the bias introduced by the lack of degrees of freedom (i.e. of parameters) in the *a priori* model.

4.3 Real MT data analysis

In this section, we present an application of the Bayesian analysis to real MT data. We have selected data sets from very different origins, namely a sea-floor MT sounding (long-period MT data) and two MT soundings carried out for crustal and subsurface investigations, respectively. In this paper, we restrict our attention to the Bayesian inversion of these data. We do not intend to describe the geophysical implications. The results presented hereafter demonstrate that the pdfs of the *a posteriori* inverse solution based on real data sets have the same characteristics and variability as the pdfs obtained with the synthetic data.

4.3.1 Low-frequency sea-floor MT sounding

The data set (Fig. 6c) comes from a sea-floor MT site about 500 km south-east–east off Hawaii (Tarits *et al.* 1987). It consists of 12 apparent resistivities and phases and the periods range from 668 to 57 600 s. These data resulted from a short experiment (20 days). As a consequence, the

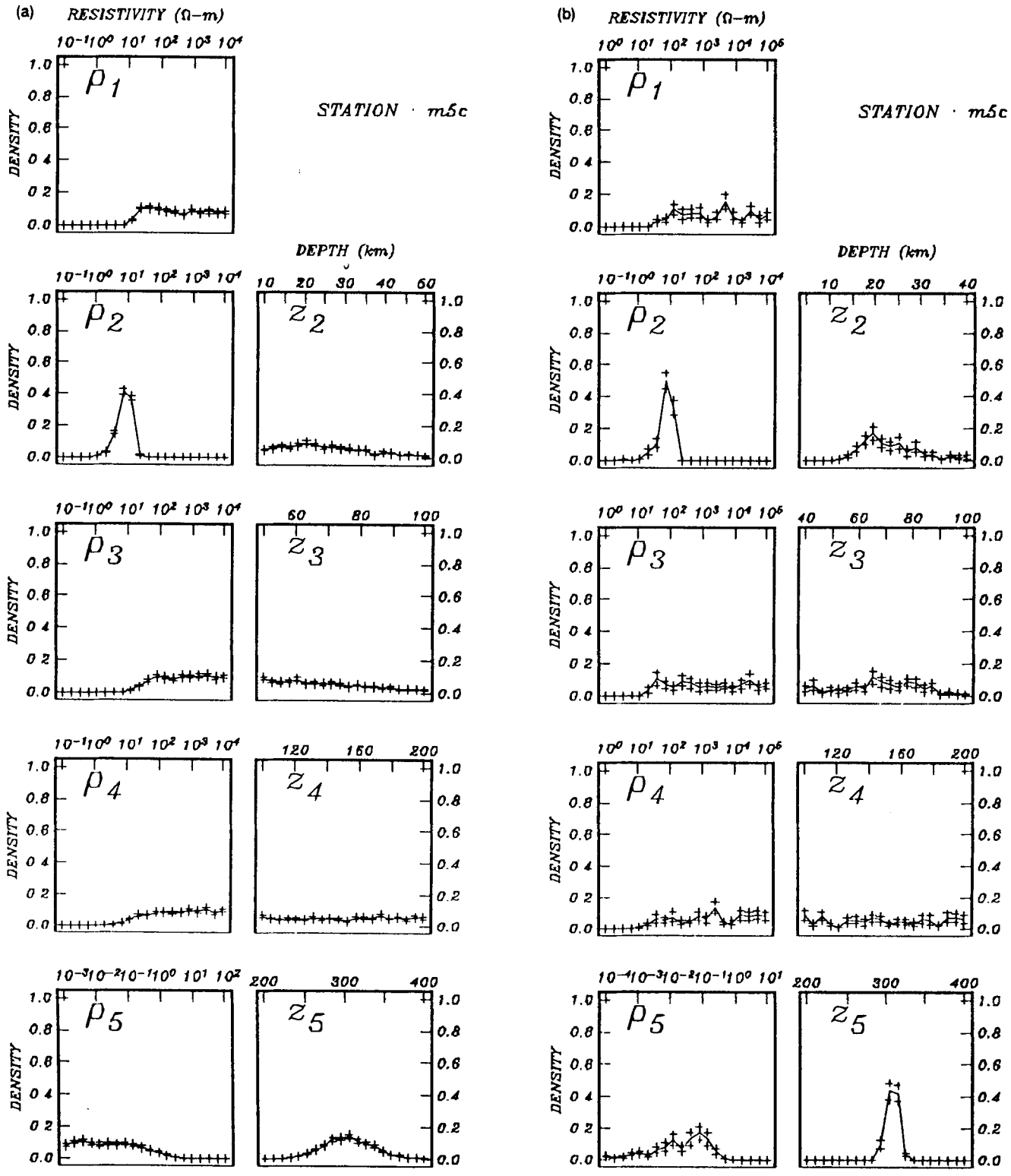


Figure 3. Marginal pdfs of the *a posteriori* law of the five-layer model parameters. The *a priori* intervals for each parameter are along the horizontal axes. The vertical axes represent the density of the marginal pdf normalized so the surface is one. The broken line joins the marginal pdf mean values and the crosses represent ±1 standard deviation. (a) Solution for the synthetic data presented in Table 1. (b) Solution for the same data set but with the error bars divided by four.

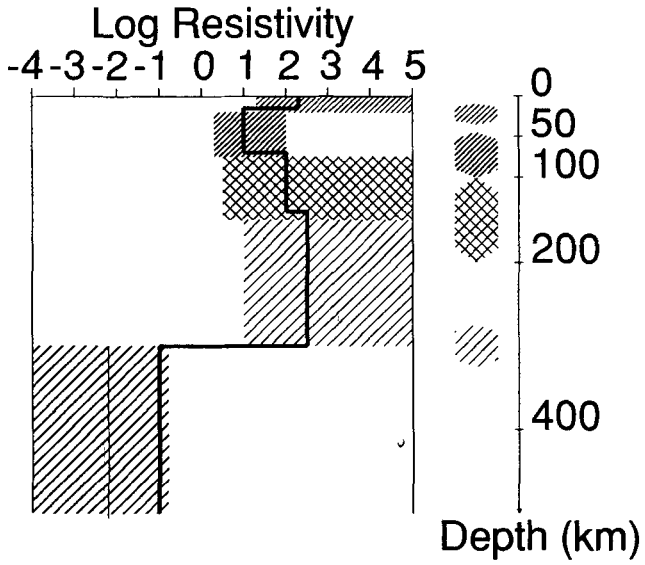


Figure 4. *A posteriori* model for the five-layer case (see text for details) deduced from analysis of the marginal pdfs presented Fig. 3(a). The shaded areas represent the intervals of possible values of the parameters. The thick black line is the synthetic five-layer model P_0.

frequency range is narrow and the 95 per cent confidence limit intervals depicted on Fig. 6(c) are large.

The *a priori* model used is shown in Fig. 6(a). It has three layers. 4×10^5 random trials were used for this example. The marginal pdfs are presented Fig. 6(b). Four out of the five parameters are resolved (i.e. they have a unimodal pdf) though the first interface has a large variance. Only a maximum value for the half-space resistivity is resolved by the data. The fit between the data and the impedances for the model made with the expected values of the parameters is shown in Fig. 6(c).

4.3.2 MT sounding of the upper crust in the Couy region (France)

This data set comes from a site in the south of the Parisian Basin as part of an experiment carried out to study the Parisian Basin Magnetic Anomaly (Dupis *et al.* 1990). It consists of 16 apparent resistivities and phases in the period range 7×10^{-3} to 100 s (Fig. 7c).

In this example, we tried to decrease the number of parameters until all of them were resolved. The *a priori* model used is shown in Fig. 7(a). It has two layers. 10^6 random trials were used for this example. The marginal pdfs are presented Fig. 7(b). The marginal pdfs are all unimodal. The fit between the data and the impedances for the model made with the expected values of the parameters is shown in Fig. 7(c).

We also used a three layer *a priori* model in order to estimate whether the two-layer *a posteriori* model would be biased. The resulting marginal pdfs are presented Fig. 8. The resistivities ρ_1 and ρ_2 are almost identical and the interface z_2 is not resolved. This clearly indicates that a two-layer model is a better representation of the distribution of resistivity with depth given the available data set.

4.3.3 Audio-frequency magnetotelluric (AMT) sounding in the Bresse region (France)

Fifteen AMT soundings have been carried out in the Bresse region in order to study the Rhône–Saône transfer zone, a particular segment of the West European rift (Grandis *et al.* 1991). We selected a data set at a site where the 1-D approximation held. The data consists of 10 apparent resistivities in the frequency range 10–1000 Hz (Fig. 9c).

The *a priori* model used is shown in Fig. 9(a). It has two layers. 10^6 random trials were used for this example. The marginal pdfs are presented in Fig. 9(b). Only the marginal pdfs for z_2 and ρ_1 are unimodal. The pdf for ρ_2 is only defined by a maximum value.

The fit between the data and the impedances for the model made with the expected values is shown in Fig. 9(c). Note that the calculated impedance is not biased by the abnormal data at 50 Hz.

5 CONCLUSION

We have analysed the ability of Bayesian statistics to solve non-linear inverse problems with a limited number of parameters and to provide a detailed description of the model parameters after inversion, by means of their individual marginal probability density functions.

We paid special attention to the numerical calculation of the *a posteriori* pdfs. We developed an algorithm which enabled the simultaneous determination of the numerical estimates of the *a posteriori* pdfs and of their uncertainties.

The numerical algorithm proposed may be generalized to any pdf which may be simulated. The marginal pdfs enable the characterization of the *a posteriori* statistics of the model parameters and may be used directly for evaluation of the range of possible values for the parameters for further interpretation of the Earth's structure.

With the help of practical examples chosen in magnetotellurics, for synthetic and real data, we showed that in most cases, the *a posteriori* parameters have complicated pdfs which cannot be described by a mean value and a standard deviation (i.e. we almost never obtain *a posteriori* Gaussian pdfs). Some of the parameters are only bounded by a maximum or a minimum value or may not be resolved at all by the data (i.e. the *a posteriori* pdf for that parameter is identical to *a priori* pdf).

The variability of the marginal pdf from one parameter to another illustrates well the difficulty of properly assessing the reliability of a given parameter resulting from an inversion process. Therefore, any inversion scheme that implied given features for the pdf of the inverse solution would have led to a severely biased model.

The determination of the Bayesian *a posteriori* pdf without major assumptions about its nature requires a complete description of the actual *a posteriori* pdf given the data. It provides an interesting insight into the complexity of the relationship between the data uncertainties and the model parameter uncertainties as the result of an inversion process. The approach used in this study is of course only practicable for a limited number of cases. Nevertheless, any inverse solution of a non-linear problem is necessarily complex and an accurate estimate of the inverse solution reliability is inevitable.

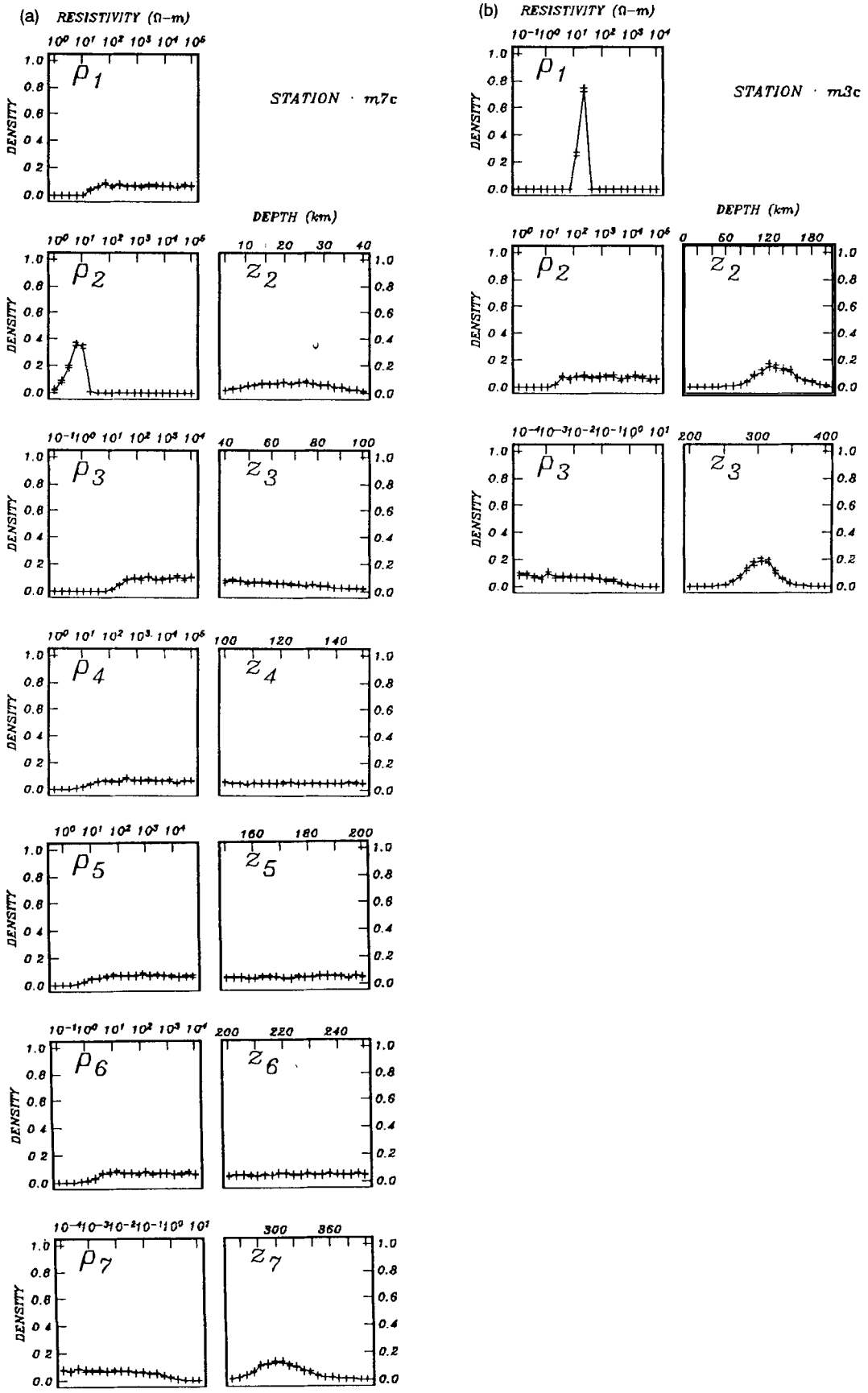


Figure 5. Marginal pdfs for: (a) the seven-layer model parameters; (b) the three-layer model parameters. The symbols are described in Fig. 3.

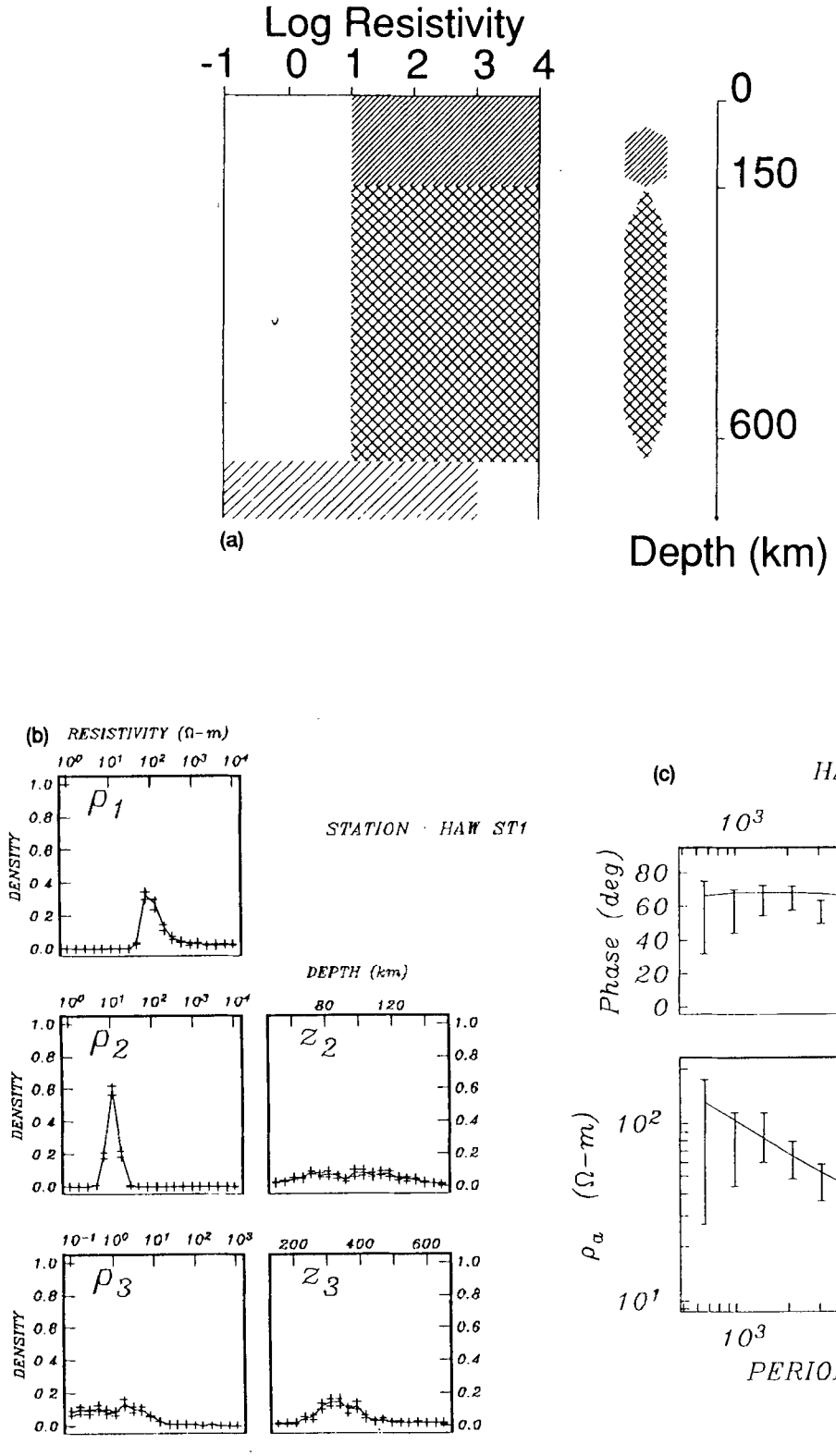


Figure 6. Sea-floor MT data set. (a) The *a priori* model used in the inversion. (b) Marginal pdfs. (c) Fit (apparent resistivity and phase) between the data (with their 95 per cent error bars) and the impedances calculated for a model produced with the *a posteriori* expected values of the parameters (full line).

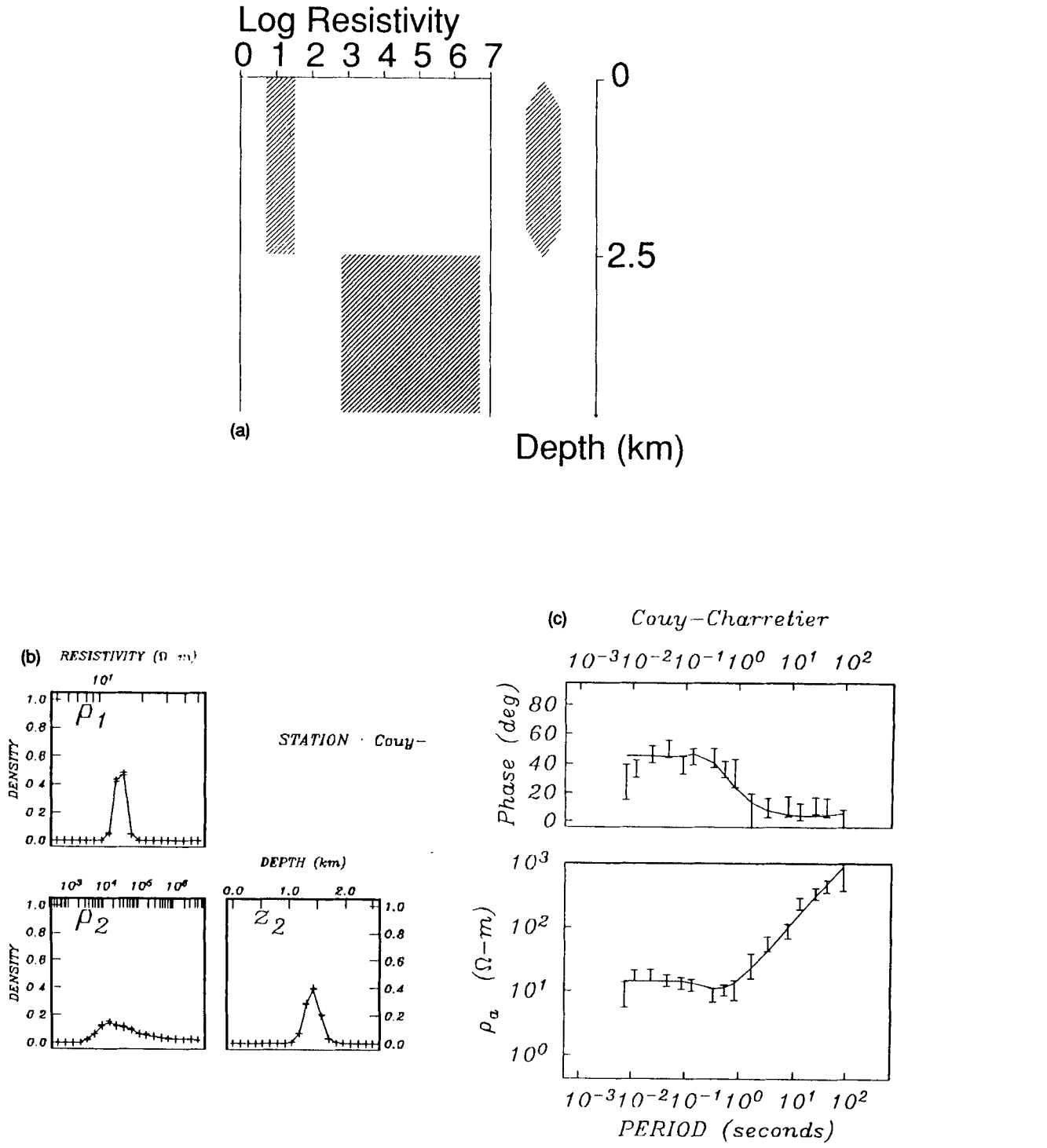
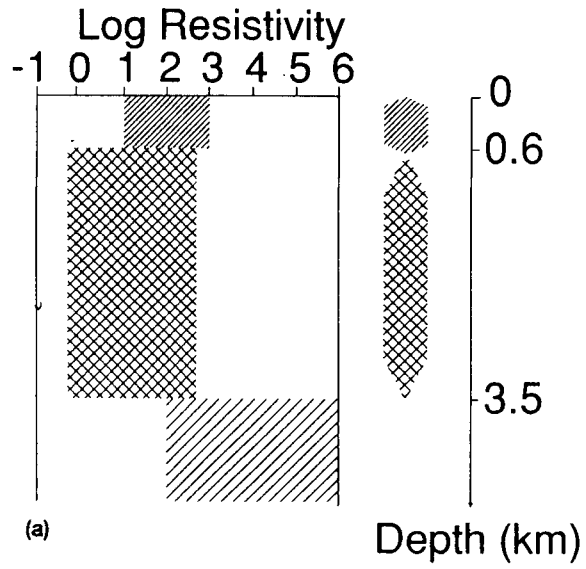


Figure 7. Data set from the Couy region. (a) The two-layer *a priori* model used in the inversion. (b) Marginal pdfs. (c) Fit (apparent resistivity and phase) between the data (with their 95 per cent error bars) and the impedances calculated from the *a posteriori* expected values of the parameters (full line).



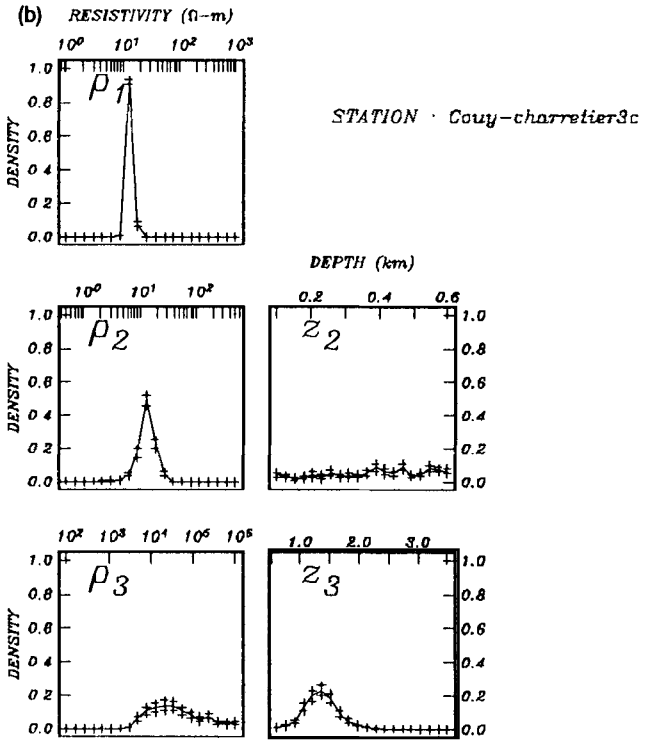


Figure 8. Data set from the Couy region. (a) The three-layer *a priori* model used for the inversion. (b) Marginal pdfs.

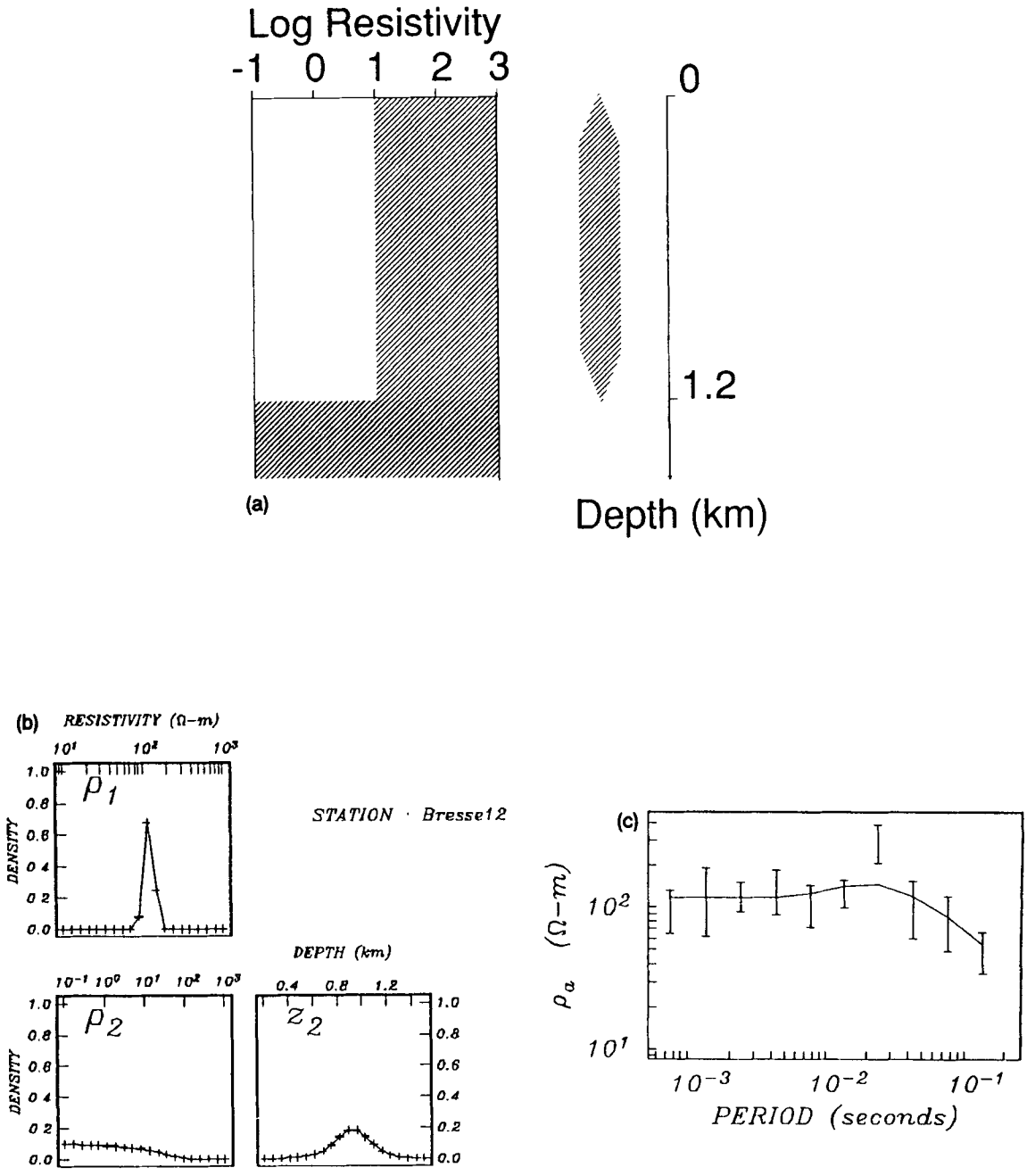


Figure 9. Data set from the Bresse region. (a) The *a priori* model used in the inversion. (b) Marginal pdfs. (c) Fit (apparent resistivity and phase) between the data (with their 95 per cent error bars) and the impedances calculated from the *a posteriori* expected values of the parameters.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for their useful comments. Virginie Jouanne was supported by a grant from MRT. Computer facilities were those of the Institut de Physique du Globe de Paris.

REFERENCES

Backus, G.E., 1970. Inference from inadequate and inaccurate data, I, *Proc. Nat. Acad. Sci. U.S.A.*, **65**, 1–7.

Backus, G.E., 1988. Bayesian inference in Geomagnetism, *Geophys. J.*, **92**, 125–142.

Bayes, T., 1763. An essay toward solving a problem in the doctrine of chances, *Phil. Trans. R. Soc. Lond.*, **53**, 370–418.

Chave, A.D., Thomson, D.J. & Ander, M.E., 1987. On the robust estimation of power spectra, coherences, and transfer functions, *J. geophys. Res.*, **92**, 633–648.

Constable, S.C., Parker, R.L. & Constable, C.G., 1987. Occam's inversion: a practical algorithm for generating smooth models from EM sounding data, *Geophysics*, **52**, 1603–1620.

Dacunha-Castelle, D. & Duflo, M., 1982. *Probabilités et Statistiques*, Vol. 1, Masson, Paris.

Dacunha-Castelle, D. & Duflo, M., 1983. *Probabilités et Statistiques*, Vol. 2, Masson, Paris.

Duijndam, A.J.W., 1988a. Bayesian estimation in seismic inversion. Part I: Principles, *Geophys. Prospect.*, **36**, 878–898.

Duijndam, A.J.W., 1986b. Bayesian estimation in seismic inversion. Part II: Uncertainty analysis, *Geophys. Prospect.*, **36**, 899–918.

Dupis, A., Baltenberger, P., Fabriol, H., Gasmi, M., Ghorberl, N., Shout, H. & Thera, A., 1990. Contribution de la magnéto-tellurique et de la sismique à l'étude de l'anomalie magnétique de Bassin Parisien, *Bull. Soc. geol. France*, (8), VI (5): 749–766.

Filloux, J.H., 1987. Instrumentation and experimental methods for oceanic studies, in *Geomagnetism*, Vol. 1, pp. 143–248, ed. Jacobs, J.A., Academic, San Diego, CA.

Grandis, H., Jouanne, V., Menvielle, M., Villemin, T., Dupis, A. & Terra, A., 1991. The Rhine–Saone transfer zone: preliminary results from audio-MT investigations (abstract), in *XX IUGG General Assembly*, Vienna, Austria.

Jackson, D.D. & Matsu'ura, M., 1985. A Bayesian approach to nonlinear inversion, *J. geophys. Res.*, **90**, 581–591.

Mackie, R.L., Bennet, B.R. & Madden, T.R., 1988. Long period magnetotelluric measurements near the Central California coast: a land locked view of the conductivity structure under the Pacific Ocean, *Geophys. J.*, **95**, 181–194.

Press, S.J., 1989. *Bayesian Statistics: Principle, Models and Applications*, John Wiley & Sons, London.

Press, W.H., Flannery, B.P., Teukolsky, S.A. & Vetterling, W.T., 1989. *Numerical Recipes*, Cambridge University Press, Cambridge.

Provost, C., Garcia, O. & Garçon, V., 1992. Analysis of satellite SST time series in the Brazil Malvinas current confluence region: dominance of the annual and semi-annual periods, *J. geophys. Res.*, **97**, 17 841–17 858.

Roussignol, M., Jouanne, V., Menvielle, M. & Tarits, P., 1993. Bayesian electromagnetic imaging, in *Computer Intensive Methods in Statistics*, pp. 85–87, eds Hardle, W. & Simar, L., Physica-Verlag.

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

Tarantola, A. & Valette, B., 1982. Generalized nonlinear inverse problems solved using the least squares criterion. *Rev. Geophys. Space Phys.*, **20**, 219–232.

Tarits, P., 1986. Conductivity and fluids in the oceanic upper mantle, *Phys. Earth planet. Inter*, **42**, 215–226.

Tarits, P. & Jouanne, V., 1990. Résultats de sondages magnétotelluriques sous-marins et structure thermique des points chauds, *Bull. Soc. géol. France, VI* (6): 921–931.

Tarits, P., Filloux, J.H., Larsen, J.C. & Klein, D.P., 1987. Electromagnetic investigation of the Hawaiian swell (abstract), in *XIX IUGG General Assembly*, Vancouver, Canada.

APPENDIX A

In this appendix, we prove the validity of estimations of the numerical uncertainties in the calculation of the marginal pdf and of the expected value of a parameter P_i for a given number L of random trials.

The derivation of the uncertainties is based on the multiple dimension central limit theorem (e.g. Dacunha-Castelle & Duflo 1982, p. 133). This theorem states that the probability density function (pdf) of the random vector

$$\sqrt{L}\begin{bmatrix} Y_{ik}^{(L)} - E_{ik} \\ Y^{(L)} - E \end{bmatrix} \tag{A1}$$

over an interval I_{ik} converges as $L \to \infty$ towards the probability density function of a Gaussian centred random vector $\begin{pmatrix} \Delta_{ik} \\ \Delta \end{pmatrix}$ the variance–covariance matrix $\mathbf{C}_2$ of which is:

$$\mathbf{C}_2 = \begin{pmatrix} V_{ik} - E_{ik}^2 & V_{ik} - E_{ik}E \\ V_{ik} - E_{ik}E & V - E^2 \end{pmatrix}, \tag{A2}$$

with V_{ik} and V defined by:

$$V_{ik} = \int_{D_{ik}} \left[\frac{f(\mathbf{Z}/\mathbf{P})g(\mathbf{P})}{h(\mathbf{P})}\right]^2 h(\mathbf{P})\, d\mathbf{P}, \tag{A3a}$$

$$V = \int_{D} \left[\frac{f(\mathbf{Z}/\mathbf{P})g(\mathbf{P})}{h(\mathbf{P})}\right]^2 h(\mathbf{P})\, d\mathbf{P}. \tag{A3b}$$

The estimates of V and V_{ik} are denoted by $W^{(L)}$ and $W_{ik}^{(L)}$, given by:

$$W_{ik}^{(L)} = \frac{1}{L_l}\sum_{l=1}^{L}\left[\frac{f(\mathbf{Z}/X_l)g(X_l)}{h(X_l)}\right]^2 \mathcal{I}_{(X_l \in D_{ik})}, \tag{A3c}$$

$$W^{(L)} = \frac{1}{L_l}\sum_{l=1}^{L}\left[\frac{f(\mathbf{Z}/X_l)g(X_l)}{h(X_l)}\right]^2, \tag{A3d}$$

with $\mathcal{I}_{\{X_l \in D_{ik}\}} = 1$ for $X_l \in D_{ik}$ but 0 otherwise.

From the δ method (e.g. Dacunha-Castelle & Duflo 1983), we know that when a function $G(x, y)$ is differentiable in $(x = E_k, y = E)$, the pdf of the random variable $\sqrt{L}\{G[Y_{ik}^{(L)}, Y^{(L)}] - G(E_{ik}, E)\}$ converges toward the pdf of the random variable

$$\frac{\partial G}{\partial x}(E_{ik}, E)\,\Delta_k + \frac{\partial G}{\partial y}(E_{ik}, E)\,\Delta.$$

In the present study, $G(E_{ik}, E) = E_{ik}/E$. Hence, the random variable $\sqrt{L}\left[\frac{Y_{ik}^{(L)}}{Y^{(L)}} - \frac{E_{ik}}{E}\right]$ converges toward the pdf of the random variable:

$$\frac{\Delta_{ik}}{E} - \frac{E_{ik}\Delta}{E^2}.$$

This variable has a centred Gaussian pdf the variance of which is:

$$U_{ikk} = \operatorname{var}\left(\frac{\Delta_{ik}}{E} - \frac{E_{ik}\Delta}{E^2}\right) = \frac{\operatorname{var}(\Delta_{ik})}{E^2} + \frac{E_{ik}^2 \operatorname{var}(\Delta)}{E^4} - 2\frac{E_{ik}\operatorname{cov}(\Delta_{ik}, \Delta)}{E^3},$$

$$U_{ikk} = \frac{V_{ik}(E - E_{ik})^2 + E_{ik}^2(V - V_{ik})}{E^4}. \tag{A4}$$

We estimate U_{ikk} by replacing E, E_{ik} and V, V_{ik}, respectively, with $Y^{(L)}$, $Y_{ik}^{(L)}$ and $W^{(L)}$, $W_{ik}^{(L)}$ in eq. (A4) and it becomes:

$$U_{ikk} \approx \frac{W_{ik}^{(L)}}{[Y^{(L)}]^2}\left[1 - \frac{Y_{ik}^{(L)}}{Y^{(L)}}\right]^2 + \left[\frac{Y_{ik}^{(L)}}{Y^{(L)}}\right]^2 \frac{W^{(L)} - W_{ik}^{(L)}}{[Y^{(L)}]^2}. \tag{A5}$$

Determination of the variance of the pdf of the uncertainty in the numerical estimation of the expected value is a generalization of the precedent derivation. The expected value is given by eq. (14a) which shows that it is a linear combination of the $F_{ik}^{(L)}$s.

The pdf of the random vector

$$\sqrt{L}\begin{bmatrix} Y_{il}^{(L)} - E_{il} \\ \cdots \\ Y_{ik}^{(L)} - E_{ik} \\ \cdots \\ Y_{iK}^{(L)} - E_{iK} \\ Y^{(L)} - E \end{bmatrix}$$

over an interval I_{ik} converges as $L \to \infty$, towards a centred Gaussian $K+1$ vector:

$$\begin{pmatrix} \Delta_{il} \\ \cdots \\ \Delta_{ij} \\ \cdots \\ \Delta_{ik} \\ \cdots \\ \Delta_{iK} \\ \Delta \end{pmatrix}$$

of variance–covariance matrix $\mathbf{C}_{K+1}$:

$$\mathbf{C}_{K+1} = \begin{pmatrix} V_{il} - E_{il}^2 & \cdots & -E_{il}E_{ij} & \cdots & -E_{il}E_{ik} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & V_{ij} - E_{ij}^2 & \cdots & -E_{ij}E_{ik} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & -E_{ik}E_{ij} & \cdots & V_{ik} - E_{ik}^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

$$\begin{pmatrix} \cdots & -E_{il}E_{iK} & V_{il} - E_{il}E \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & V_{ij} - E_{ij}E \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & V_{ik} - E_{ik}E \\ \cdots & \cdots & \cdots \\ \cdots & V_{iK} - E_{iK}^2 & V_{iK} - E_{iK}E \\ \cdots & \cdots & V - E^2 \end{pmatrix} \qquad \text{(A6)}$$

We define the vectors $\mathbf{F}_i^{(L)}$ and $\mathbf{f}_i$ by:

$$\mathbf{F}_i^{(L)} = \begin{bmatrix} \cdots \\ \dfrac{Y_{ik}^{(L)}}{Y^{(L)}} \\ \cdots \end{bmatrix}_{1 \le k \le K}, \qquad \mathbf{f}_i = \begin{bmatrix} \cdots \\ f_i \quad (k) \\ \cdots \end{bmatrix}_{1 \le k \le K} \qquad \text{(A7)}$$

The generalization of the δ-method to the random vector $\sqrt{L}[\mathbf{F}_i^{(L)} - \mathbf{f}_i]$ leads to the evaluation of the variance–covariance matrix $\mathbf{U}_i = \{U_{ijk}\}$, $(j, k = 1, \ldots, K)$ for the random vector:

$$\begin{pmatrix} \dfrac{\Delta_{il}}{E} - \dfrac{E_{il}\Delta}{E^2} \\ \cdots \\ \dfrac{\Delta_{ij}}{E} - \dfrac{E_{il}\Delta}{E^2} \\ \cdots \\ \dfrac{\Delta_{ik}}{E} - \dfrac{E_{ik}\Delta}{E^2} \\ \cdots \\ \dfrac{\Delta_{iK}}{E} - \dfrac{E_{iK}\Delta}{E^2} \end{pmatrix} \qquad \text{(A8)}$$

the vector pdf of which is Gaussian and centred. The diagonal terms U_{ikk} of $\mathbf{U}_i$ are given by eq. (A4) above. The non-diagonal terms are the covariances between two different elements of eq. (A8):

$$U_{ijk} = \frac{1}{E^2}\operatorname{cov}(\Delta_{ij}, \Delta_{ik}) + \frac{E_{ij}E_{ik}}{E^4}\operatorname{var}(\Delta) - \frac{E_{ik}}{E^3}\operatorname{cov}(\Delta_{ij}, \Delta) - \frac{E_{ij}}{E^3}\operatorname{cov}(\Delta_{ik}, \Delta),$$

$$\underset{j \ne k}{U_{kij}} = \frac{E_k E_j V - V_k E_j E - E_k E V_j}{E^4} \qquad k \ne j. \qquad \text{(A9)}$$

Taking into account eq. (14a), the pdf of the expected value $\langle P_i^{(L)} \rangle$ may be identified to the following 1-D normal pdf:

$$\langle P_i^{(L)} \rangle \sim N_1\left(\langle P_i \rangle, \frac{{}^t\mathbf{pUp}}{L}\right) \qquad \text{(A10)}$$

of mean $\langle P_i \rangle$ and variance ${}^t\mathbf{p} \cdot \mathbf{U} \cdot \mathbf{p}/L$.

We estimate U_{ijk} by replacing E, E_{ik} and V, V_{ik}, respectively, with $Y^{(L)}$, $Y_{ik}^{(L)}$ and $W^{(L)}$, $W_{ik}^{(L)}$.