JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 104, NO. B1, PAGES 749–766, JANUARY 10, 1999

Lithologic tomography: From plural geophysical data to lithology estimation

Miguel Bosch, [1]

Institute de Physique du Globe de Paris, France

Abstract. The information provided by different geophysical data sets (gravimetric, magnetic, seismic, etc.) can be used, together with petrophysical and geostatistical information, to estimate the major lithological properties of the rocks within the studied volume. The formalization of this inverse problem requires a joint representation and parameterization of the different media properties in the model. The information relating rock properties together couples the inversion of the plural geophysical data sets and allows one to relate the observations with the lithological parameters of the model. The representation by probability density functions (*pdf*s) of the different types of information entering the problem is also required and provides the mathematical framework to formulate their combination. The resulting joint posterior *pdf* is composed of two factors: the joint likelihood function, which is the product of independent likelihood functions associated with each geophysical data set, and the joint prior *pdf*. The latter decomposes, following a partition of the model parameter space in a primary (lithological) subspace and secondary (physical) subspace, as a marginal *pdf* over the primary model parameter space and a conditional *pdf* over the secondary model parameter space. A Markov-chain Monte Carlo method was adapted to sample joint models from the posterior *pdf*: (1) the method starts with a Markov-chain sampling primary models from the marginal prior *pdf*, (2) the chain is extended to the joint model space, by sampling from the conditional *pdf* of the secondary parameters with respect to the primary parameters, and (3) it is modified to sample from the posterior *pdf*, by applying the Metropolis rule, which uses the evaluation of the joint likelihood function to accept or reject model transitions in the sampling chain. Finally, posterior marginal or posterior conditional *pdf*s for the model parameters or the model properties can be straightforwardly calculated from the set of joint models sampled by the chain.

1. Introduction

Interpretation of geophysical data is a complex process integrating many kinds of information. The quantitative treatment of the geophysical data produces images of the studied Earth volume that must be matched with the geological model of the area and with petrophysical information. This combination of information, commonly solved by the expert using qualitative criteria, is an inverse problem that can be formalized into a method for probabilistic estimation of the lithology. Such an integrated and quantitative approach to interpretation is of a major importance in the present situation of geophysical methods.

Geophysical exploration has developed a series of powerful tools for tomography, commonly consisting of the inversion of a single geophysical data set to estimate the geophysical related property: compressional velocity from arrival times [*Bosch*, 1997], conductivity from electromagnetic data [*Newman and Alumbaugh*, 1997], mass density from gravity data [*Camacho et al.*, 1997], magnetic susceptibility from magnetic data [*Li and Oldenburg*, 1996], etc. Nevertheless, conventional geophysical tomographic methods face the problem of irregular data coverage over the surface of the studied volume, producing irregular image resolution. This problem is difficult to address for each isolated geophysical technique and demands an effort on the integration of different geophysical methods into a single inversion scheme.

The objective of this work is to present a general method to invert several geophysical data sets, obtaining joint information on several properties characterizing the physics and the lithology of the media. The method requires the modeling of the media by joint description of several media properties. The importance of this realistic media representation is that the relations between properties can be described. This allows the explicit introduction in the inverse problem of valuable information on petrophysics, geostatistics, and the geology of the region, as a way to couple the inversion of the plural geophysical data set and as a way to estimate the lithological media properties represented in the model.

[1] Also at Departamento de Física Aplicada, Universidad Central de Venezuela, Caracas.

Copyright 1999 by the American Geophysical Union.

Paper number 1998JB900014.
0148-0227/99/1998JB900014$09.00

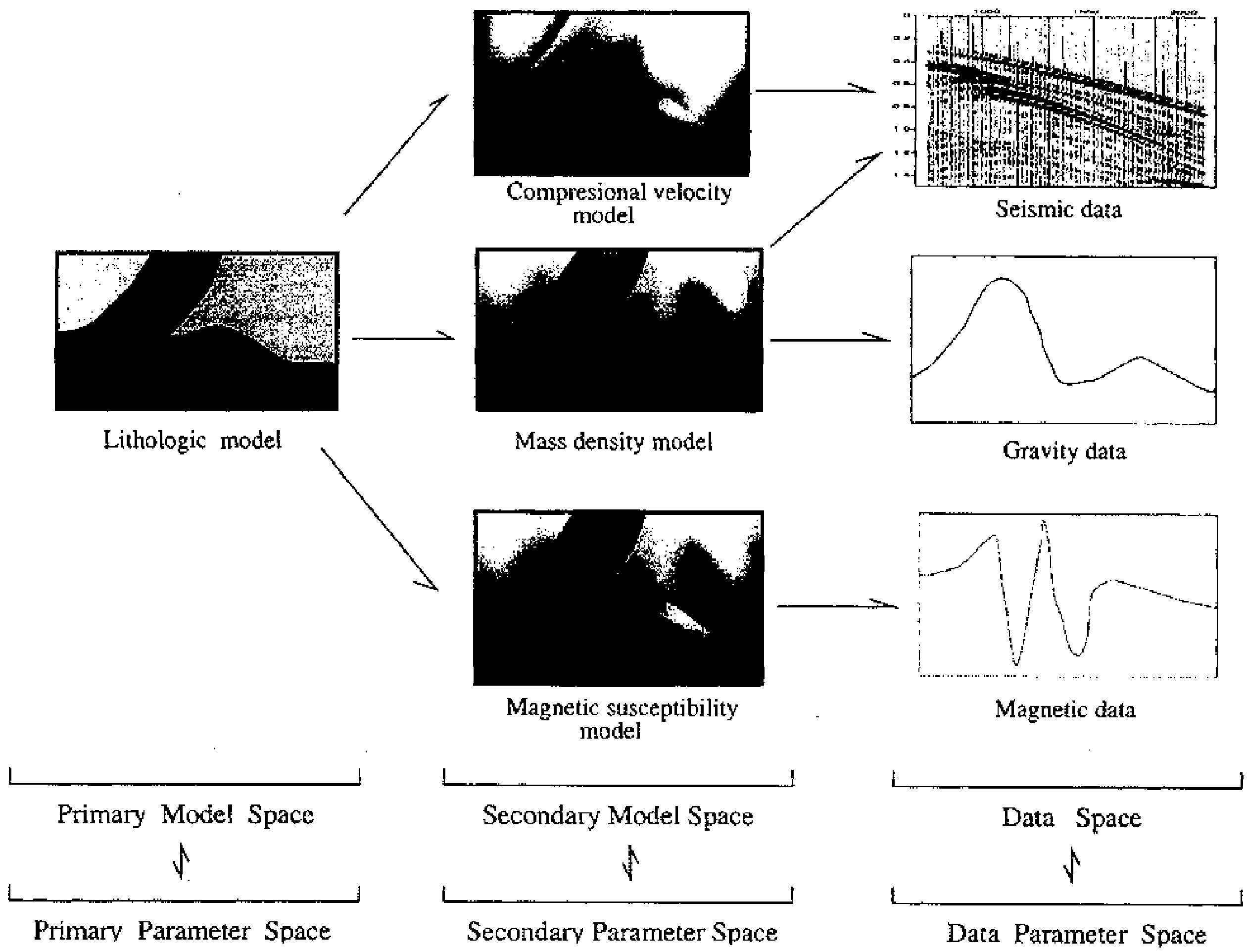


Figure 1. Parameter spaces and model spaces in the multiple data and multiple property geophysical problem. The link between the primary and the secondary model space is provided by the petrophysical and the geostatistical information. The link between the secondary model space and the data space is established by the forward geophysical calculation. The arrows indicate the forward direction from the primary model space to the data space.

Petrophysical studies show that Earth rocks exhibit strong correlation between their properties and strong dependence of their physical properties on lithology. A large amount of work has been devoted to the experimental and theoretical study of these relations. Also, the description and simulation of spatially dependent properties have been largely considered in geostatistical work. These studies provide a background for the joint modeling of several Earth media properties within a volume under exploration.

A probabilistic inference approach [*Press,* 1968; *Wiggins,* 1969; *Jackson and Matsu'ura,* 1985; *Tarantola,* 1987; *Mosegaard and Tarantola,* 1995; *Mosegaard et al.,* 1997] is used here to combine the prior information provided by petrophysics, geostatistics, and the geology of the area with the information provided by the plural geophysical observations. The general mathematical framework used to formulate the problem requires the representation of the information by probability density functions (here after *pdfs*, probability densities or simply densities), defined over the space of model parameters. The information resulting from the combination is represented by the posterior *pdf.*

In the interdisciplinary context already described, the present contribution consists in formalizing the problem of multiple-property multiple-data inversion, formulating the structure of the correspondent posterior *pdf* and assembling a Markov-chain sampling method adapted to this structure. This theoretical work is presented in the two next sections. Section 4 illustrates and discuses the methodology for inferring the conditional *pdfs* from petrophysical and geostatistical information. Finally, section five presents a synthetic example of the application of the method to a geologic media consisting of a peridotite nape surrounded by granite rocks.

2. General Formulation of the Multiple-Property Inversion

The objective of the method is to estimate different media properties (see Figure 1) from the inversion of a plural data set containing several types of geophysical data (seismic, gravimetric, electric, etc.). The media properties considered in the theoretical formulation could be very general. They include (1) those explicitly related by a physical theory to the geophysical data sets (e.g., mass density for gravimetric data, electric parameters for electric data), (2) those characterizing the lithology of the rocks (e.g., lithotype, mineralogical composition), or (3) other properties characterizing rock structure or rock conditions (porosity, fluid content, temperature, pressure, etc.). These variables could be categorical as lithotype or continuous as temperature.

In principle, the media properties may be represented by continuous fields defined over a studied volume Ω. However, the inference of these functions involves an infinite dimensional problem which is, in general, difficult to handle (see, for instance, the work of *Stark* [1992] considering linear problems and simple priors). For practical purposes, the common approach to follow consists in the parameterization (discretization) of the property fields. With this operation, the inference problem is translated into a finite dimensional space.

To formalize the inference of the properties within Ω, consider property fields following a parametrical model $z_i(\mathbf{x}, \mathbf{m}_i)$, defining the ith property as a function of a finite number of model parameters $\mathbf{m}_i$ and the position $\mathbf{x} \in \Omega$. Common field models define the property field according to blocks in Ω, interpolation from a set of points in Ω, or linear combinations of a base of functions. For example, for whole Earth tomography, most researchers either divide Earth into a set of cells and assume that seismic velocity is constant in these cells, or they use a truncated spherical harmonic expansion in the angular variables, tensored with some kind of polynomial in radius. In the first case, the property field is a piecewise constant function, discontinuous in the border of the cells, and the model parameters are the velocity for each cell. In the second case, the model parameters are the coefficients for each orthogonal function in the base, and the property field is the weighted superposition of these functions. Another example is the parameterization (discretization) of the property field by giving the values of the field over a finite set of points (control points) and choosing an interpolation rule. In this case, the model parameters would be the property values at the control points and the property field the interpolated field.

Let me call the property model parameter space $\mathcal{M}_i$, the space containing the array of model parameters $\mathbf{m}_i$, and the property model space $\mathcal{Z}_i$ the space containing the property fields calculated from the parameters. The joint parameterization requires a joint model parameter space being the product space, $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2 \times \ldots \times \mathcal{M}_k$, and a joint model space being the product space, $\mathcal{Z} = \mathcal{Z}_1 \times \mathcal{Z}_2 \times \ldots \times \mathcal{Z}_k$. The distinction between these two spaces is relevant in this work because the prior information directly concerns the property fields and has to be translated in prior information about the model parameters.

Another terminology to be used along the work is the array $\mathbf{m} = \{\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_k\}$ called the joint model parameter array; the array of property fields originated from $\mathbf{m}$ via the parameterization, $\mathbf{z}(\mathbf{x}, \mathbf{m}) = \{z_1(\mathbf{x}, \mathbf{m}_1), z_2(\mathbf{x}, \mathbf{m}_2), \ldots, z_k(\mathbf{x}, \mathbf{m}_k)\}$ is called a joint property model; each individual $z_i(\mathbf{x}, \mathbf{m}_i)$ is called a property model.

In the methodology followed here, the inverse problem is formulated as an inference problem, consisting of updating the prior knowledge of the models, taking into account the results of the geophysical experiences. The inference problem is formulated in the joint model parameter space $\mathcal{M}$, and the knowledge about the joint model parameters is described by a *pdf* defined over $\mathcal{M}$. In the initial state of the inference, the knowledge about the model parameters is described by a prior probability density $\rho(\mathbf{m}_1, \ldots, \mathbf{m}_k)$. The updated knowledge is described by a posterior probability density $\sigma(\mathbf{m}_1, \ldots, \mathbf{m}_k)$, which is obtained as [*Tarantola and Valette*, 1982; *Tarantola*, 1987]

$$\sigma(\mathbf{m}_1, \ldots, \mathbf{m}_k) = c\,\rho(\mathbf{m}_1, \ldots, \mathbf{m}_k)\; L(\mathbf{m}_1, \ldots, \mathbf{m}_k),$$

with $L(\mathbf{m}_1, \ldots, \mathbf{m}_k)$ being the likelihood function and c being a normalization constant. The likelihood function measures (in probability terms) the misfit between the geophysical data calculated from the model and the observations. In the next subsections, the structure of the the prior density and the likelihood function in the multiple-data and multiple-parameter inversion problem is considered.

2.1. The Prior Probability Density

The prior *pdf* over the joint model space is, in general, a complicated function. It represents prior knowledge about model properties at individual locations, spatial relations of property values, cross-relations between different properties, and spatial dependence of cross-relations. As a consequence, in most real problems it would be difficult to formulate this density function directly in the joint model space. A straightforward formulation is decomposing the density function in two (or more) convenient factors by the rule of conditional probabilities,

$$\rho(\mathbf{m}_1, \ldots, \mathbf{m}_k) = \theta_{s|p}(\mathbf{m}_{n+1}, \ldots, \mathbf{m}_k | \mathbf{m}_1, \ldots, \mathbf{m}_n)\, \rho_p(\mathbf{m}_1, \ldots, \mathbf{m}_n). \quad (1)$$

Above, a partition of the joint model space in two subspaces is considered (Figure 1): the subspace of primary properties ($\mathcal{M}_{pri} = \mathcal{M}_1 \times \ldots \times \mathcal{M}_n$) and the subspace of secondary properties ($\mathcal{M}_{sec} = \mathcal{M}_{n+1} \times \ldots \times \mathcal{M}_k$). The term $\rho_p(\mathbf{m}_1, \ldots, \mathbf{m}_n)$ is a marginal density in $\mathcal{M}_{pri}$; it contains information about the primary media properties. The term $\theta_{s|p}(\mathbf{m}_{n+1}, \ldots, \mathbf{m}_k | \mathbf{m}_1, \ldots, \mathbf{m}_n)$ is a conditional density in $\mathcal{M}_{sec}$; it contains information about secondary properties and the dependence of secondary properties on primary properties.

Such strategy for decomposing the joint prior *pdf* is known in statistical inference work, where $\mathbf{m}_{pri} = \{\mathbf{m}_1, \ldots, \mathbf{m}_n\}$ are sometimes called "hyperparameters" [see *Besag et al.*, 1995]. For the present problem, it has the following advantages.

1. Often we have privileged properties better related with the structure of the media and more relevant to the determination of the rest of the properties. Primary simulation of these properties is needed to simulate the rest of them [*Deutsch and Journel*, 1992].

2. The conditional probability density is particularly convenient to introduce petrophysical laws (empirical or theo-

retical) relating rock properties together. On the other hand, the marginal probability density is convenient to describe the properties better constrained by the prior information. The formulation accommodates these two different sources of information.

These advantages are particularly significant if lithology is introduced in the model as a primary property. Physical rock properties used for the geophysical calculations of the observed fields are naturally dependent on lithology. They are macroscopic consequences of the structure of the rock (composition, texture, and genesis), which is described by lithological properties. This dependency has been largely studied for different geological environments [*Domenico*, 1984; *Han et al.*, 1986; *Christensen and Mooney*, 1995] and can be experimentally studied for a particular region [see *Christensen and Mooney*, 1995] by collection and analysis of samples. With the approach presented here, this information can be included in the inverse problem by the conditional *pdf* on secondary properties.

Also, prior information based in the geologic knowledge of the area is commonly available in lithologic terms (geologic surface charts, probable lithotypes, geometric relations between lithotypes, stratigraphic directions in sedimentary formations, etc.). This information can be included in the problem by the marginal *pdf* on the primary properties.

Important primary properties for Earth media description depend on the scale and nature of the exploration problem. Here are some examples: (1) for mantle structure, temperature, iron content [see *Jordan*, 1979], and phase changes;(2) at the crustal scale, lithotypes, or silicium content [see *Christensen and Mooney*, 1995]; for upper crust description, lithotypes; and for sedimentary basins description, lithotypes, porosity, fluid content, and stratigraphic direction.

Depending on the properties and the information available, this principle of decomposition could be applied again within each subspace, producing a larger partition of the model space. For instance, in a problem including lithotypes, porosity and elastic parameters, a primary simulation step could concern the parameters of the lithotype model, a secondary step could superpose a porosity field within each region, conditioned by the lithotype, and a tertiary step could simulate elastic parameters according to the lithotype and porosity.

The consideration of lithologic properties for a realistic modeling of Earth media has been well established in geosciences. In common exploratory situations, the physical rock properties and its spatial laws vary significantly between the different lithologies present in the region. However, inside each lithology, statistical relations between properties (e.g., mean values and variograms, or more generally marginal and conditional *pdfs*) can be better described and can be assumed as statistically homogeneous (statistics are invariant by translations) in favorable situations. This assumption is very important because most models for estimation and simulation of media properties are based on the hypothesis of statistical homogeneity, also called spatial stationarity. In consequence, when heterogeneous lithology is present in a volume under study, a two-step simulation is mandatory [*Deutsch and Journel;* 1992]: a primary simulation of major lithological domains and a secondary simulation of the rest of the properties within the domain. In this way, within the major lithological domains, the statistical homogeneity can often be assumed, and statistical homogeneous models can be used to simulate the secondary properties within the lithotype domains.

2.2. The Likelihood Function

Consider m different geophysical methods used to explore the studied region (see Figure 1) and data parameter arrays $\mathbf{d}_1, \ldots, \mathbf{d}_m$ describing the observations of each one correspondingly. The joint data parameter space is in this case the product of the data parameter spaces for each geophysical method considered in the problem, $\mathcal{D} = \mathcal{D}_1 \times \mathcal{D}_2 \times \ldots \times \mathcal{D}_m$. The information provided by these observations can be described by a joint probability density $\nu(\mathbf{d}_1, \ldots, \mathbf{d}_m)$ defined over the joint data space, with a product structure,

$$\nu(\mathbf{d}_1, \ldots, \mathbf{d}_m) = \prod_{i=1,m} \nu_i(\mathbf{d}_i), \qquad (2)$$

justified with the assumption that the observations are independent across different geophysical methods. Consider also that the geophysical forward problem can be solved exactly and is represented by functions $\{\mathbf{g}_i(\mathbf{m}_1, \ldots, \mathbf{m}_k) = \mathbf{d}_i^{cal}\}$ for each geophysical technique. To avoid further complications in the notation, these function are defined on the joint model space $\mathcal{M}$, although only a part of the model parameters would be relevant for each simulation method (e.g., a gravity calculation needs only a mass density model). The likelihood function is

$$L(\mathbf{m}_1, \ldots, \mathbf{m}_k) = \nu(\mathbf{g}_1(\mathbf{m}_1, \ldots, \mathbf{m}_k), \ldots, \mathbf{g}_m(\mathbf{m}_1, \ldots, \mathbf{m}_k)).$$

Because of the independence of observational uncertainties between the geophysical methods (2), it follows immediately that the joint likelihood function is the product of independent likelihood functions for each geophysical method,

$$\begin{aligned} L(\mathbf{m}_1, \ldots, \mathbf{m}_k) &= \textstyle\prod_{i=1,m} \nu_i(\mathbf{g}_i(\mathbf{m}_1, \ldots, \mathbf{m}_k)) \\ &= \textstyle\prod_{i=1,m} L_i(\mathbf{m}_1, \ldots, \mathbf{m}_k). \end{aligned} \qquad (3)$$

This result also holds with forward calculation errors independent across the different geophysical methods. The assumption of independence of the observational uncertainties across different geophysical methods is realistic in most cases. Sources of error affecting geophysical data (instrumental, environmental, human, station positioning, corrections) are commonly different across geophysical methods and surveys.

2.3. The Posterior Probability Density

Summarizing results of the previous subsections, the posterior *pdf* in the multiple-data multiple-property problem should be, in most cases,

$$\sigma(\mathbf{m}_{pri}, \mathbf{m}_{sec}) = \underbrace{\theta_{s|p}(\mathbf{m}_{sec}|\mathbf{m}_{pri})\, \rho_p(\mathbf{m}_{pri})}_{\rho(\mathbf{m}_{pri},\mathbf{m}_{sec})} \underbrace{\prod_{i=1,m} L_i(\mathbf{m}_{pri}, \mathbf{m}_{sec})}_{L(\mathbf{m}_{pri},\mathbf{m}_{sec})}; \qquad (4)$$

$\mathbf{m}_{\rm pri}$ denotes the array $\{\mathbf{m}_1, \ldots, \mathbf{m}_n\}$ of primary model parameters, and $\mathbf{m}_{\rm sec}$ denotes the array $\{\mathbf{m}_{n+1}, \ldots, \mathbf{m}_k\}$ of secondary model parameters.

To keep generality in expression (4), the likelihood function, $L(\mathbf{m}_{\rm pri}, \mathbf{m}_{\rm sec})$, is defined over the joint model parameter space, although it would likely depend only on the secondary parameters if primary parameters are associated with the lithologic description of the media. Similarly, the likelihood function for a particular geophysical method, $L_i(\mathbf{m}_{\rm pri}, \mathbf{m}_{\rm sec})$, depends on one (or a few) model property parameter arrays (e.g., density model parameters for gravity data).

The integration of the joint posterior density over the secondary parameter space gives the marginal posterior density for the primary model parameters,

$$\sigma_{\rm p}(\mathbf{m}_{\rm pri}) = c\,\rho_{\rm p}(\mathbf{m}_{\rm pri}) \int_{\mathcal{M}_{\rm sec}} \theta_{\rm s|p}(\mathbf{m}_{\rm sec}|\mathbf{m}_{\rm pri}) \prod_{i=1,m} L_i(\mathbf{m}_{\rm pri}, \mathbf{m}_{\rm sec})\, d\mathbf{m}_{\rm sec}. \quad (5)$$

In the above expression, the integral term is a likelihood function for the primary model parameters. It is a combination of the joint likelihood function, $L(\mathbf{m}_{\rm pri}, \mathbf{m}_{\rm sec})$, with the conditional density $\theta_{\rm s|p}(\mathbf{m}_{\rm sec}|\mathbf{m}_{\rm pri})$.

A method to sample the joint posterior density is given in the following section; it is based on a Markov-chain sampler adapted to the structure of the posterior *pdf* given by (4). Another approach may be useful to estimate the characteristics of the true properties in the explored region from expressions (4) or (5). Using a gradient method or another optimization method a model maximizing the posterior probability density could be searched. A maximum posterior *pdf* model, together with uncertainties on its parameters, is an acceptable approximation of monomodal posterior *pdf*s. However, it provides an incomplete representation of the posterior probability density, and it is a wrong representation of probability densities having complex shapes and being multimodal.

The posterior probability density of the multiple-data multiple-parameter inverse problem results from the combination of several nonlinear geophysical simulations and complicated priors, and it is likely to be a complicated function. Although the sampling approach commonly demands a larger calculation effort than the optimization approach, it will be developed here because it is a more general method and provides more information about the posterior knowledge of the studied volume. In addition, the sampling approach proceeds without problem categorical parameters, whereas a gradient based optimization approach is limited to continuous parameters.

3. Sampling the Posterior Probability Density

As the posterior probability density resulting from a geophysical inverse problem is commonly a complicated function defined over a large and high-dimensional space, it cannot be analytically integrated. Hence posterior probabilities have to be calculated by statistical integration methods, consisting of approximating the *pdf* integral by a summation over a sample taken from the support $\mathcal{M}$ of the *pdf*.

Consider any sequence $S = \{\mathbf{m}^1, \mathbf{m}^2, \ldots, \mathbf{m}^N\}$ of joint model parameter arrays; hereafter $\mathbf{m}^t$ denotes the joint model parameter array $\{\mathbf{m}_1^t, \ldots, \mathbf{m}_k^t\}$ for the position t of the sequence. Consider also an indicator function, $X(\mathbf{m}) = \{1$, if $\mathbf{m} \in \mathcal{A}$; 0, if $\mathbf{m} \notin \mathcal{A}\ \}$, with $\mathcal{A}$ being a subset of the model space. If S is a sample from $\sigma(\mathbf{m})$, it is well known that

$$\int_{\mathcal{A}} \sigma(\mathbf{m})\, d\mathbf{m} = \frac{1}{N} \sum_{t=1,N} X(\mathbf{m}^t) + \epsilon, \quad (6)$$

with an error arbitrarily small, $\lim_{N\to\infty} \epsilon = 0$.

The sample S allows fast and straightforward probability calculations about any kind of request about media properties in the area (marginal and conditional probabilities) and other studies as the identification of modes and the relative probabilities of the modes. For example, summation (6) is used in section 5 to calculate marginal histograms. To produce a sample from the posterior *pdf*, I use here a Markov-chain Monte Carlo method combining (1) a conditional sampling [*Gelfand and Smith*, 1990; *Besag et al.*, 1995] of the joint prior *pdf* and (2) a Metropolis sampling [*Metropolis et al.*, 1953; *Hastings*, 1970; *Mosegaard and Tarantola*, 1995] of the posterior *pdf*.

A Markov-chain over the space $\mathcal{M}$ can be intuitively understood as the record of a walk in $\mathcal{M}$, $S^{\rm T} = \{\mathbf{m}^1, \mathbf{m}^2, \ldots, \mathbf{m}^{\rm T}\}$, visiting a point $\mathbf{m}^t$ at each step t of the walk, until a final step T. Mathematically, a chain, $\mathbf{M}^{(t\to T)} = \{\mathbf{M}^{(1)}, \mathbf{M}^{(2)}, \ldots, \mathbf{M}^{(T)}\}$, is a sequence of random functions indexed by a variable $t = 1, 2, \ldots, T$, usually called "time." A Markov chain is said to be ergodic to a probability density $\pi(\mathbf{m})$ if any outcome set of points, $S^{\rm T} = \{\mathbf{m}^1, \mathbf{m}^2, \ldots, \mathbf{m}^{\rm T}\}$, converges to a sample from $\pi(\mathbf{m})$ as the number of steps increases, no matter what the initial point $\mathbf{m}^1$ of the chain is. There are different methods to construct Markov chains ergodic to an arbitrary density $\pi(\mathbf{m})$; the two better studied methods are the Metropolis sampling and the Gibbs sampling (for a review on Markov chain samplers, see the work of *Tierney* [1994] and *Besag et al.* [1995]).

Assume that a Markov chain, $\mathbf{M}^{(t\to T)}_{\rm primary}$, ergodic to the marginal prior density $\rho_{\rm p}(\mathbf{m}_{\rm pri})$ has been defined. Let us call this Markov chain the primary chain. The practical way to produce such an algorithm is by following perturbation rules for model parameters, each perturbation representing a step in the chain. For simple probability densities $\rho_{\rm p}(\mathbf{m}_{\rm pri})$ the appropriate random functions can be easily defined. Complicated probability densities can be treated as well, using sampling algorithms as Metropolis or others (see work already referenced). Given the primary chain, the method here developed describes the way to construct a Markov chain, $\mathbf{M}^{(t\to T)}_{\rm posterior}$, ergodic to the posterior probability density given by (4), in two stages.

First, the primary chain is extended to the joint model parameter space, taking advantage of the conditional density $\theta_{\rm s|p}(\mathbf{m}_{\rm sec}|\mathbf{m}_{\rm pri})$. An outcome $\mathbf{m}^t_{\rm sec}$ of this *pdf*, conditioned by the primary chain value $\mathbf{m}^t_{\rm pri}$, provides secondary model parameters for any step t of $\mathbf{M}^{(t\to T)}_{\rm primary}$. These secondary model parameters complete a joint model parameter array $\mathbf{m}^t = \{\mathbf{m}^t_{\rm sec}, \mathbf{m}^t_{\rm pri}\}$ defining a Markov chain in the joint model parameter space $\mathcal{M}$; let us call this Markov chain in $\mathcal{M}$ the prior chain, $\mathbf{M}_{\text{prior}}^{(t\to T)}$. Due to this construction, the prior chain is ergodic to the prior density $\rho(\mathbf{m})$. The method described consists of sampling from a joint *pdf* by sampling from its marginal and conditional *pdfs*; it is a method extensively used in statistics [e.g., *Gelfand and Smith*, 1990].

Second, to generate a Markov chain $\mathbf{M}_{\text{posterior}}^{(t\to T)}$ sampling the joint posterior density $\sigma(\mathbf{m})$, the prior chain is modified following the Metropolis rule at all steps of the chain [*Hastings*, 1970; *Mosegaard and Tarantola*, 1995]; I follow the formulation of *Mosegaard and Tarantola* [1995]. Suppose that the posterior chain is in the step t, with an outcome $\mathbf{m}^t$ of $\mathbf{M}_{\text{posterior}}^{(t)}$; to determine the outcome $\mathbf{m}^{t+1}$ of $\mathbf{M}_{\text{posterior}}^{(t+1)}$ the Metropolis rule is as follows.

1. Consider $\mathbf{m}^t$ to be the current joint model parameter array visited by the prior chain and take an outcome $\mathbf{m}_{\text{can}}^{t+1}$ from $\mathbf{M}_{\text{prior}}^{(t+1)}$ following the rules defining the prior chain. Let us call this outcome the candidate joint model parameter array.
2. Make $\mathbf{m}^{t+1} = \mathbf{m}_{\text{can}}^{t+1}$ with probability,

$$p = \text{Min}[1, L(\mathbf{m}_{\text{can}}^{t+1})/L(\mathbf{m}^t)].$$

3. Otherwise make $\mathbf{m}^{t+1} = \mathbf{m}^t$.
4. Continue with (1).

Above, $L(\mathbf{m})$ is the joint likelihood function for the multiple data according to expression (3). It is well known that a chain generated with this procedure is stationary to the posterior probability density $\sigma(\mathbf{m})$. A detailed description of the Metropolis algorithm is presented in the work of *Mosegaard and Tarantola* [1995] and an application to the analysis of seismic reflections is presented in the work of *Mosegaard et al.* [1997] including interesting prior information and the characterization of data uncertainties.

4. Inferring the Prior Conditional Density From Field Samples

In order to reproduce the real behavior of the media properties in the studied region, $\theta_{s|p}(\mathbf{m}_{n+1}, \ldots, \mathbf{m}_k|\mathbf{m}_1, \ldots, \mathbf{m}_n)$ should be inferred from petrophysical and geostatistical data. This section presents an example, to be used later in a synthetic test, together with an overview of the common geostatistical methods that can be used to this purpose.

4.1. From Rock Samples to Global Probability Densities

Figure 2 shows an example of the information provided by plots of rock properties. The logarithmic transformation of the mass density and the magnetic susceptibility was used all along this work; for the two cases this transformation showed to be helpful to better model the property variability. Logarithmic transformed properties are identified with an asterisk following the symbol of the nontransformed property and named, for simplicity, in the same way of the nontransformed property.

Figures 2a and 2b show representations of granite and peridotite samples in the bidimensional space of mass density ρ^* and magnetic susceptibility k^*. These properties have been measured at surface temperature and surface pressure conditions by *Horen* [1997]. The plots show marked differences between the distributions of the physical properties for the two types of rocks: (1) the granite data show, on average, smaller magnetic susceptibility and mass density than the peridotite data, (2) the dispersion of the mass density is much larger for the peridotite samples than for the granite sample, and (3) the mass density and the magnetic susceptibility are anticorrelated in peridotite samples, whereas they are positively correlated for granite samples. The positive correlation of magnetic susceptibility and mass density in granite, and their anticorrelation with the silicium weight content of the rock, has been shown with a larger data set by *Bourne* [1993]. There is also a physical reason for the anticorrelation between magnetic susceptibility and mass density in peridotite. *Horen* [1997] shows that generation of magnetic minerals in peridotites is associated with alteration processes (serpentinization) which also have an effect on reducing the mass density.

Assuming statistical homogeneity, each of these plots can be regarded as a sample from a large population of granite and peridotite rocks characterized by conditional densities $f(\rho^*, k^* \,|\, \text{lit} = \text{granite}, \text{z} = \text{surface})$ and $f(\rho^*, k^* \,|\, \text{lit} = \text{peridotite}, \text{z} = \text{surface})$. These *pdfs* describe the probabilities for the secondary properties (ρ^* and k^*) conditioned by the primary properties (lithotype and depth) in field rock samples. Figures 2c and 2d show a grayscale graph of the conditional *pdfs* inferred from the samples using a bivariate Gaussian model. The centroid and the covariance matrix were calculated from the samples shown in Figures 2a and 2b, respectively. The parametrical approach followed is better adapted for scarce sampling than a nonparametrical approach, as it provides reasonable interpolation to the poorly sampled regions of the support space; complex or multimodal probability densities can also be modeled as a mixture of monomodal ones.

The estimation of the conditional *pdfs* from the samples may be complicated by several problems, like spatial clustering (for an overview on declustering techniques see the work of *Isaaks and Srivastava* [1989]). Also, complete representation of lithotypes and petrophysical properties may be unlikely, and additional information (data sets for other regions or proposed petrophysical laws) could be used. For example, in the absence of additional data, the conditional *pdfs* of Figure 2 (corresponding to z=surface) could be generalized to other depths, by correcting the centroid according to reported [e.g., *Christensen and Mooney*, 1995] effects of depth on rock properties.

Another problem to be aware of is the regularization of property statistics due to spatial averages in the model. Laboratory measurements of the sample rock properties are averages in centimetric scale, whereas the model parameterization of the three-dimensional volume under study is likely to represent a much larger scale in spatial resolution. The effect of this change of scale in petrophysical data support has a systematic influence on the property dispersion and on the form of the distribution. Averaging values together has the effect of reducing the variance of the data and making their distribution more symmetric (more Gaussian) while keeping the

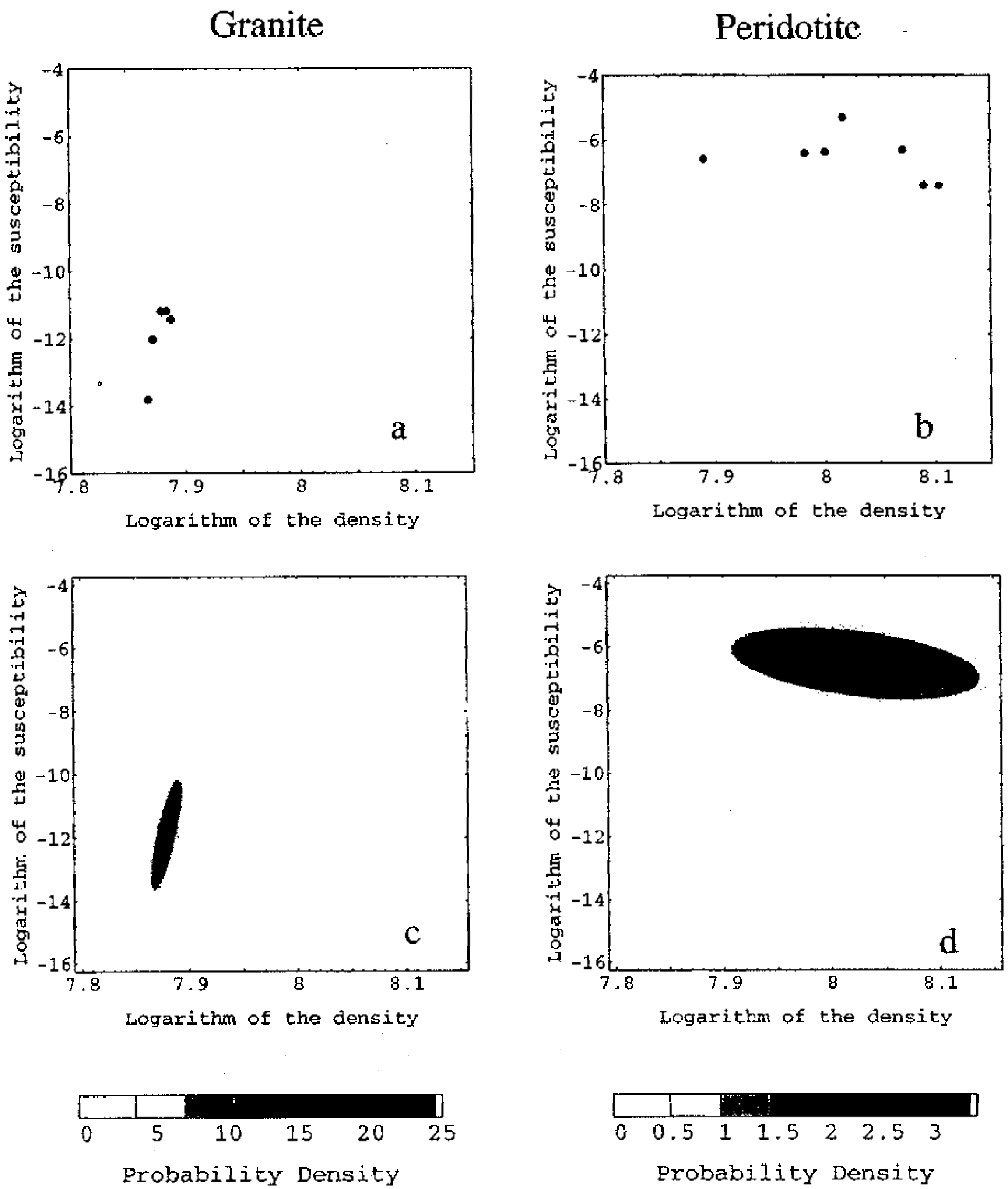


Figure 2. Cross plots of the logarithm of the magnetic susceptibility and the logarithm of the density for (a) granite and (b) peridotite rock samples. The magnetic susceptibility is in SI units, and the density is in kg/m^3 units before the logarithmic transformation. Grayscale plots represent contour levels of a bivariat Gaussian model for (c) the conditional *pdf* $f(\rho^*, k^* |$ lit = granite, z = surface) and (d) the conditional *pdf* $f(\rho^*, k^* |$ lit = peridotite, z = surface), with parameters calculated from the data shown in Figures 2a and 2b respectively.

same mean. The covariance function is also systematically affected by increasing its range. Whether the sample *pdf* and the covariance function need to be corrected from this effect or not depends on several factors, like the covariance range, the model spatial resolution, and the form of the covariance function. In general, if the covariance range is larger than the model resolution distance and the covariance is high within the model resolution distance, the property will be very continuous, so rock sample properties and model properties will have the same distribution. In other situations, the smoothing effect due to spatial averaging in the model can be significant and the *pdf* may be better inferred from block sample averages than from the original samples. The methods providing the regularization of the *pdf* and the covariance function are well known in geostatistical work [*Journel and Huijbregts*, 1978; *Isaaks and Srivastava*, 1989].

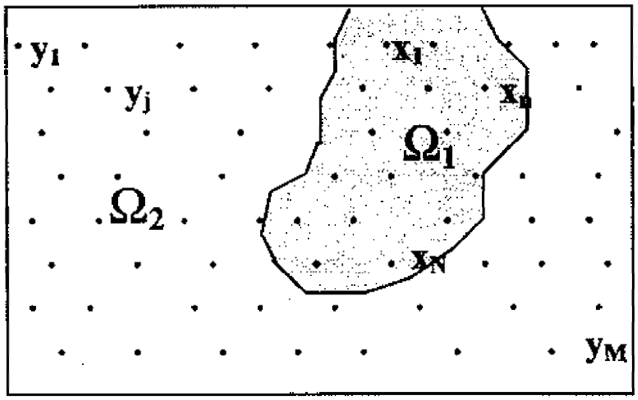


Figure 3. Parameterization of the secondary properties at points $\mathbf{x}_n$ inside region Ω_1, and points $\mathbf{y}_j$ inside region Ω_2.

4.2. Spatial Correlation of the Properties

Let us consider here that the lithotype field is parameterized given a geometrical description of the region Ω filled by each lithotype and the secondary properties within each region are parameterized given its values in a series of control points (Figure 3). Taking as an example a volume filled with a granite region and a peridotite region, the secondary parameters would be the values of the density ρ^* and the magnetic susceptibility k^* for all control points within the lithotype region.

$$\begin{aligned} \mathbf{s}_{\text{gra}} &= \{\rho^*(\mathbf{x}_1), \ldots, \rho^*(\mathbf{x}_N), k^*(\mathbf{x}_1), \ldots, k^*(\mathbf{x}_N)\}, \\ \mathbf{s}_{\text{per}} &= \{\rho^*(\mathbf{y}_1), \ldots, \rho^*(\mathbf{y}_M), k^*(\mathbf{y}_1), \ldots, k^*(\mathbf{y}_M)\}. \end{aligned} \quad (7)$$

with $\mathbf{x}_n \in \Omega_{\text{gra}}$ and $\mathbf{y}_j \in \Omega_{\text{per}}$.

Properties at the control points are spatially related, and this relation can be introduced within the model parameters given by (7). If a multivariate Gaussian model is chosen to relate secondary parameters in granite and peridotite regions, the conditional *pdf*s would be

$$\begin{aligned} &\theta_{\text{s|p}}(\mathbf{s}_{\text{gra}}|\ \text{lit} = \text{granite}) \propto \\ &\qquad \exp[-\tfrac{1}{2}\ (\mathbf{s}_{\text{gra}} - \overline{\mathbf{s}}_{\text{gra}})^t \mathbf{C}_{\text{gra}}^{-1}(\mathbf{s}_{\text{gra}} - \overline{\mathbf{s}}_{\text{gra}}]), \\ &\theta_{\text{s|p}}(\mathbf{s}_{\text{per}}|\ \text{lit} = \text{peridotite}) \propto \\ &\qquad \exp[-\tfrac{1}{2}\ (\mathbf{s}_{\text{per}} - \overline{\mathbf{s}}_{\text{per}})^t \mathbf{C}_{\text{per}}^{-1}(\mathbf{s}_{\text{per}} - \overline{\mathbf{s}}_{\text{per}}]), \end{aligned} \quad (8)$$

with $\mathbf{C}_{\text{gra}}$ and $\mathbf{C}_{\text{per}}$ being the covariance matrices and $\overline{\mathbf{s}}_{\text{gra}}$ and $\overline{\mathbf{s}}_{\text{per}}$ being the mean values of the secondary parameters; the mean values of ρ^* and k^* and the elements of the covariance matrix can be inferred, in favorable situations, from the geostatistical analysis of field data samples.

The sampling process of the secondary model parameters may be arranged in such a way that only one parameter is perturbed at a time. In this situation the conditional *pdf* needed to simulate a secondary model parameter $\rho^*(\mathbf{x}_N)$ is

$$\theta(\rho^*(\mathbf{x}_N)|\rho^*(\mathbf{x}_1), \ldots, \rho^*(\mathbf{x}_{N-1}), k^*(\mathbf{x}_1), \ldots, k^*(\mathbf{x}_N)) \propto \exp[\tfrac{-(\rho^*(\mathbf{x}_N) - \rho_c^*)^2}{2\sigma_c^2}], \quad (9)$$

with ρ_c^* being the simple cokriging estimate of $\rho^*(\mathbf{x}_N)$ and σ_c^2 being the simple cokriging variance of the estimation [*Deutsch and Journel*, 1992]. Both parameters can be found by solving a linear system of equations.

Other methods and more complicated spatial models can be found, for instance, in the work of *Davis* [1987], *Isaaks and Srivastava* [1989], *Deutsch and Journel* [1992], *Dietrich* [1993], *Almeida and Journel* [1994], and *Oliver* [1995].

5. A Synthetic Example of Lithologic Tomography

The method presented in this work can be applied to different scales and environments of the geophysical prospection. Here, a test based on synthetic data calculated from two-dimensional (2-D) models is presented. In the test, gravity and magnetic data are integrated to discriminate a peridotite nape from a background granite media. The properties considered in the inversion are the lithotype (granite or peridotite), the density, and the magnetic susceptibility.

Figure 4a shows the "true" model of the peridotite nape used to calculate the "true" gravimetric and magnetometric data for 41 stations spaced 0.5 km apart. Gravimetric stations were considered located at the surface of the model and mag-

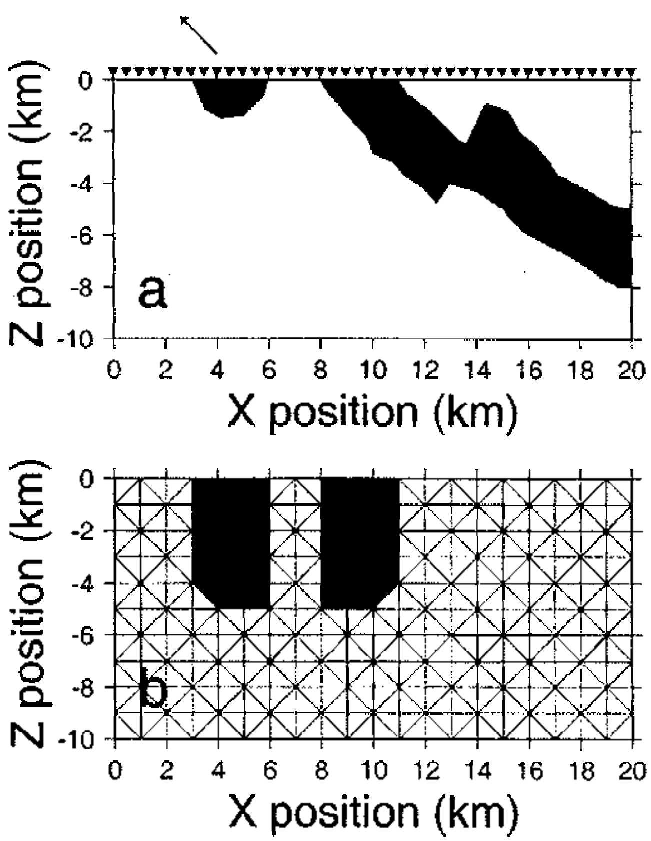


Figure 4. (a) True and (b) initial lithotype models used in the inversion synthetic test. Peridotite regions are shown in dark gray, whereas the granite region is shown in light gray. Bold triangles in Figure 4a show the horizontal position of gravimetric and magnetometric stations; the vector represents the direction of Earth's magnetic field used for magnetic data calculations. Lines in Figure 4b show the triangularization of the 2-D section.

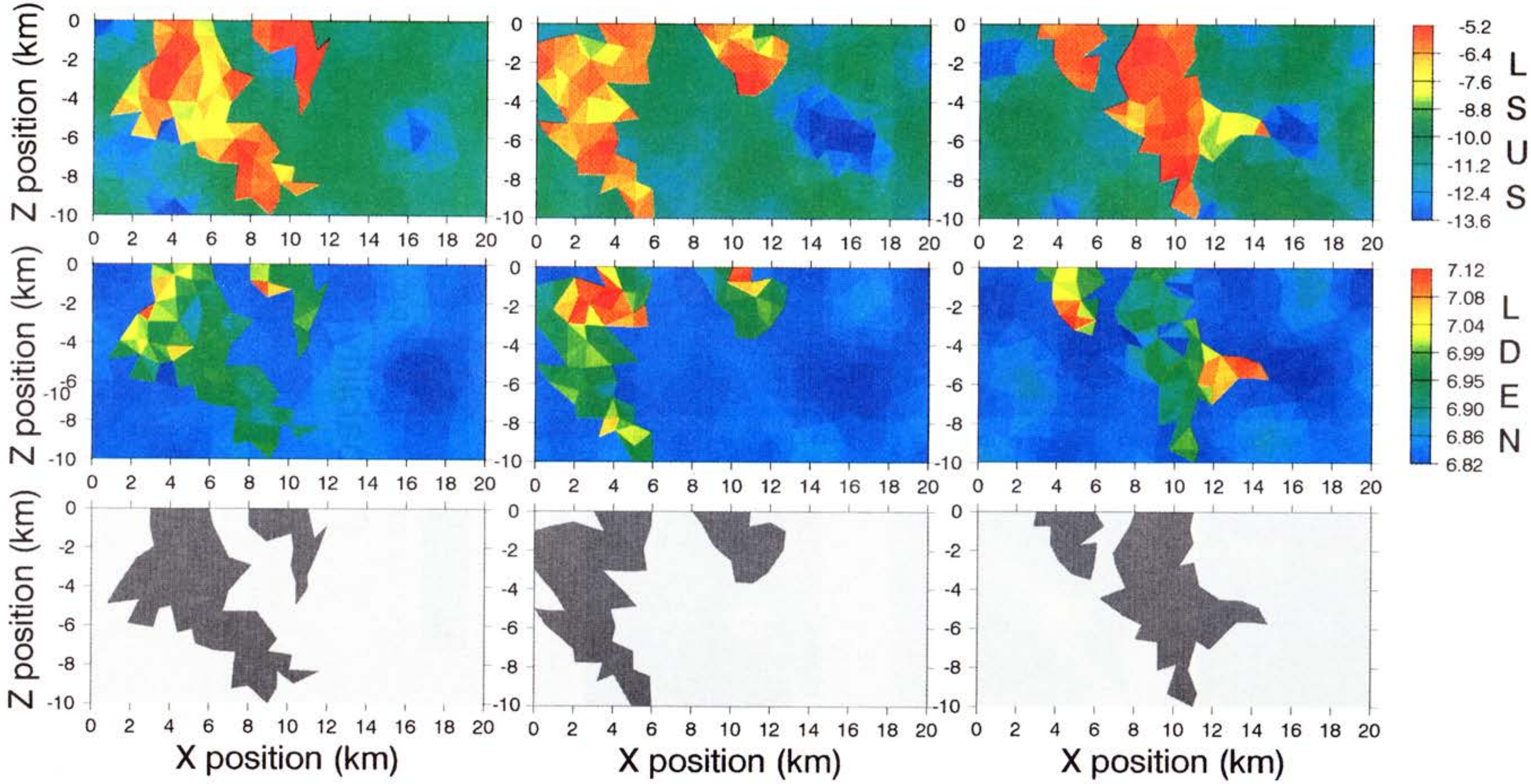


Plate 1. The lithotype model (bottom), the density model (center), and the magnetic susceptibility model (top) for three joint models (model 300,000; model 600,000; and model 900,000) pulled at regular intervals from the model set generated by the Markov chain sampler from the prior *pdf*. LDEN is the logarithm of the density, and LSUS is the logarithm of the magnetic susceptibility; the modeling was performed on the logarithmic transformed properties. Peridotite regions are shown in dark gray, whereas the granite region is shown in light gray.

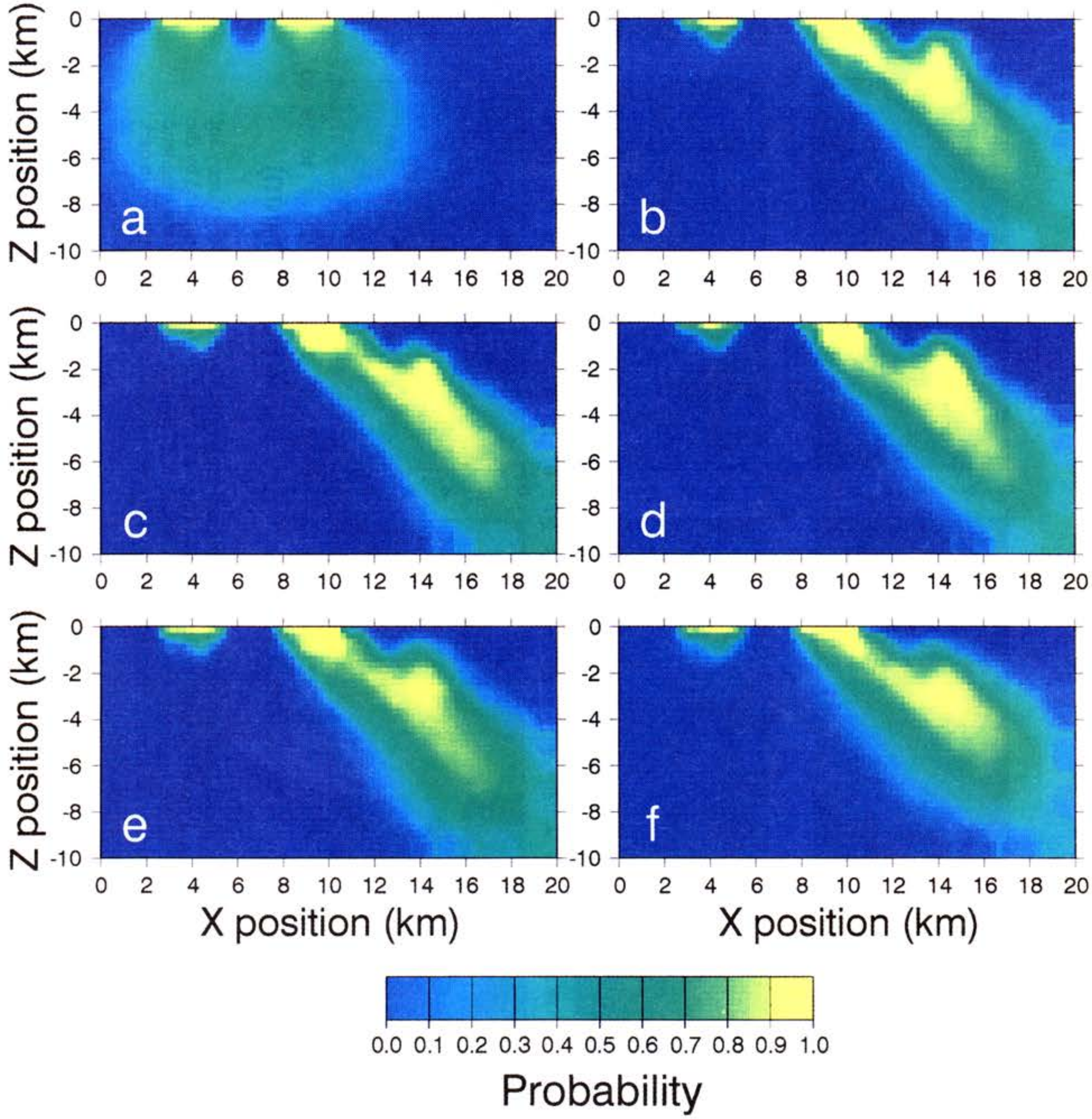


Plate 2. Peridotite frequency at 80 × 40 grid points within the 2-D section. The frequency is calculated over all lithotype models generated by the Markov chain sampler, between model 20,000 and model 1,000,000, from (a) the prior *pdf*, (b) all data posterior *pdf*, (c, d) all data posterior *pdf* using other random number sequences, (e) the magnetic data posterior *pdf*, and (f) the gravimetric data posterior *pdf*.

netometric stations at an elevation of 0.35 km (airborne). The data fields were the vertical component of the gravity acceleration and the total magnitude of the magnetic intensity. The uncertainties of these observations were considered to be 2 nT for the magnetic data and 2 mGal for the gravimetric data.

The model was parameterized by dividing the space into prisms of triangular section, with axes perpendicular to a 2-D section and infinite lateral (axial) extension. The lithotype, the magnetic susceptibility, and the mass density were constant within the prisms. The forward calculation of the gravitational field of a prism was done by the method described by *Chapman* [1979], and the forward calculation of the magnetic field by the method of *Bott* [1963]. Both methods provide the physical exact solution of the fields due to homogeneous prisms. The Earth magnetic field was considered inclined 135° in counterclockwise rotation from the $\hat{x}$ axis of the models.

The 10 km × 20 km model section was divided into 400 triangles connected as shown in Figure 4b. The parameters of the model were (1) the position of the triangle vertices (fixed for the four boundaries of the section), (2) the lithotype of each triangle (granite or peridotite), and (3) the logarithm of the density and the logarithm of the magnetic susceptibility for each triangle. The division of the 2-D section in triangulated elements was chosen because this parameterization is often used for geological modeling.

5.1. Simulation of the Lithotype Regions

The parameters affecting the lithotype regions were taken as the primary set of model parameters. They correspond with parameters 1 and 2 of the above enumeration. The prior information included in the simulation of the lithologic models was as follows.

1. The lithotype at the surface ($z = 0$) was considered to be known (geological chart), so the lithotype for triangles bounding the surface ($z = 0$) was fixed following the surface lithology given by the initial model.

2. The topology of the regions was maintained constant, following the topology of the initial model, in order to keep the consistency with a prior geological model. To maintain this condition, regions were not allowed to be disconnected, and the creation of new regions was not allowed.

3. An estimation of the total volume (area in 2-D) fraction of peridotite was considered to introduce prior information about the relative amount of the two lithotypes in prior models. Also, an estimation of the expected surface to volume ratio (area to perimeter in 2-D) was used; this global variable has influence on the frontier appearance (smooth or dendritic).

To generate lithological models, an initial Markov chain was constructed perturbing the lithological model parameters. A step in the chain consisted of one of the following perturbations.

1. Select a vertex in the model and change its position. In this operation, conditions were verified to avoid triangle overlapping.

2. Select a triangle in the frontier of lithotype regions and change its lithology. In this operation, conditions were verified to avoid region disconnection.

These two types of operations were arranged in a regular sequence along the chain. The initial Markov chain generated in this way was modified using the Metropolis rule to sample models with the expected distribution of the volume fraction of peridotite and the volume to surface ratio. The peridotite expected volume was provided as 0.35 ± 0.05 of the total volume in the model and the expected peridotite surface area to volume ratio was given as 1 ± 0.1 km^{-1}. The resulting primary chain summarizes prior information about the lithotype models.

5.2. Simulation of the Secondary Properties

The secondary model parameters in this example were the density ρ^* and the magnetic susceptibility k^* for each prism of triangulated section. To simulate these parameters, a Gaussian model for the conditional *pdf* was adopted, as described by (8). The covariance function for ρ^* and k^*, as well as a cross-covariance function , also followed a Gaussian model,

$$\begin{aligned} C_{\rho^*\rho^*}(h) &= C_{\rho^*\rho^*}(0)\exp[-3h^2/a^2], \\ C_{k^*k^*}(h) &= C_{k^*k^*}(0)\exp[-3h^2/a^2], \\ C_{k^*\rho^*}(h) &= C_{k^*\rho^*}(0)\exp[-3h^2/a^2], \end{aligned} \tag{10}$$

with a being the range, h the separation distance, and $C_{\rho^*\rho^*}(0)$, $C_{k^*k^*}(0)$ and $C_{k^*\rho^*}(0)$ the density ρ^* variance, the magnetic susceptibility k^* variance and the covariance inferred from sample data in Figure 2. The range was chosen to be 4 km in the granite region and 2 km in the peridotite region; the same range was used for ρ^* and k^*. The range and the covariance model were not inferred from geostatistical field data. They wer, however, chosen to follow the characteristics of the density and the magnetic susceptibility in Canadian granites [*Bourne*, 1993] that show smooth variations and covariance ranges of several kilometers.

For the simulation, property control points were considered at the center of each triangle. The model resolution distance (separation between triangle centers) was of the order of 0.5 km, several times smaller than the assumed covariance range for the media properties. Hence the estimated variation of the properties within the triangular sections of the model could be considered small, as additionally, the covariance function model was Gaussian. In consequence, no regularization of the *pdfs* shown in figure 2 was needed.

Also, no depth correction to the centroid of the *pdfs* was introduced, as it turned out to be irrelevant for the types of rocks and depths under consideration. Density data for peridotite and granite were presented by *Christensen and Mooney* [1995] at different depths, showing that the variation of the mean at the surface and at 10 km depth were smaller than the standard deviation of the mean.

The cosimulation of the secondary parameters ρ^* and k^* of a triangle was demanded by two types of independent operations along the chain: (1) each time the triangle changed its lithology (a change in the lithological model) and (2) selecting at random any triangle and simulating again its secondary properties.

The cosimulation, at a control point $\mathbf{x}_N$, was performed by taking in sequence outcomes from the conditional *pdfs*,

$$\theta_1(k^*(\mathbf{x}_N)|\rho^*(\mathbf{x}_1),\ldots,\rho^*(\mathbf{x}_{N-1}),k^*(\mathbf{x}_1),\ldots,k^*(\mathbf{x}_{N-1})),$$
$$\theta_2(\rho^*(\mathbf{x}_N)|\rho^*(\mathbf{x}_1),\ldots,\rho^*(\mathbf{x}_{N-1}),k^*(\mathbf{x}_1),\ldots,k^*(\mathbf{x}_N));$$

constructed from the simple cokriging estimate and variance as in (9). These two parameters are obtained as solution of a simple cokriging system of equations.

Plate 1 shows three models pulled from a chain of 1 million models generated by the prior Markov chain constructed in this way. It can be seen that the models follow the prior information: (1) the fixed lithology at the surface and the fixed topology, (2) the total volume of peridotite within the expectation, (3) the magnetic susceptibility and the density show the spatial correlation range of the properties (4 km in the granite and 2 km in the peridotite) and (4) the density and magnetic susceptibility are anticorrelated in the peridotite region and correlated in the granite region, as expected (see Figure 2).

Plate 2 shows a histogram of the peridotite lithotype in a regular grid on the 2-D section, for the same run used to pull the models shown Plate 1. The frequency of the peridotite (1 or 0) at each position of the grid is accumulated for outcomes 20,000 to 1,000,000 of the chain and averaged at the end of the chain. This histogram represents the marginal prior probability density of the lithotype for each grid point. The frequency of the peridotite is dominated, as expected, by the prior information: the fixed lithotypes at the surface, the limited volume, as well as the fixed topology make the peridotite lithotype highly probable below its outcrop in the surface.

The calculated data for 20 models pulled from the same chain are shown in Figures 5a and 5c. The misfit from the calculated data and the true data indicate that the geophysical data should significantly contribute to the posterior probability density.

5.3. Sampling Models From the Joint Posterior Density

The Markov chain process sampling the prior *pdf* is modified with the Metropolis rule in order to generate a process ergodic to the joint posterior *pdf*, as described in section 3. The geophysical likelihood has been calculated using spatially independent data uncertainties of $\sigma_m = 2$ nT for the total magnetic intensity anomaly and $\sigma_g = 2$ mGal for the vertical gravity acceleration. The L_1 norm was used to measure the data misfit, and the likelihood function was taken as

$$L(\mathbf{m}) \propto \exp[-(\sum \frac{|d_i^{\mathrm{cal}} - d_i^{\mathrm{obs}}|}{\sigma_g}) - (\sum \frac{|d_j^{\mathrm{cal}} - d_j^{\mathrm{obs}}|}{\sigma_m})],$$

with d_i being the gravity data and d_j being the magnetic data. The choice of norm is arbitrary in this synthetic example; the L_2 was also used with adequate results.

The progress in data misfit reduction is shown in Figure 6 for a 1 million sample long chain. Reduction to a stable level is achieved in less than 15,000 iterations. In the phase of misfit reduction, the models sampled by the chain are still influenced by the choice of initial model. Hence this initial phase, commonly called "burn-in period," is not taken into account for the inference, as usual in Markov chain Monte Carlo [*Besag et al.*, 1995]. All calculations of estimates in this work exclude this phase (the first 20,000 iterations) and are using the other 980,000 model outcomes from the chains.

Plate 3 shows three models taken at large regular intervals from the posterior chain. In all models there are general structural features that are constant: the peridotite region at the left is small and centered below its outcrop, and the peridotite region at the right is large and dips to the right, as for the true model. Figures 5b and 5d show the calculated data and the true data as well as the uncertainties for 20 models pulled from the same chain. They indicate that models in the chain correctly explain the true data, although the models are still very variable in some aspects, as shown in Plate 3.

Plate 2 shows the accumulated histogram for the peridotite from model 20,000 to model 1,000,000. This histogram represents the posterior marginal probability density for the peridotite over the grid points. The histogram is very different from the prior histogram (Plate 2a), and the area of high peridotite probability resembles the form of the true peridotite nape. It shows some details that could not be seen in individual models because they are very variable. For instance, the fault step in the top of the peridotite nape is clearly identified from the posterior histogram.

Independent executions with different random number sequences were performed to analyze the stationarity of the posterior histogram. The results are shown in Plates 2c and 2d, presenting variations in details but maintaining the major characteristics.

The peridotite histogram obtained using only magnetic data (Plate 2e) and only gravity data (Plate 2f) show more spatial dispersion of the peridotite and indicate that both types of data provide enough information to distinguish the large-scale structure of the nape in this example.

Common qualitative criteria to examine the convergence of the chain in applications to physics are based on (1) regarding the stabilization of estimated values for the parameters (or the functions of parameters) of major importance to the interpretation and (2) comparing the similarity of estimated values across independent chains. These are necessary conditions for the convergence, providing in most cases adequate confidence in the estimates. As the form of the posterior *pdf* is not known, no sufficiency conditions exist for the convergence of the chain in a finite number of iterations [*Smith and Roberts*, 1993; *Robert*, 1995], unless some assumptions are made about the regularity of the posterior *pdf*.

An additional criterion should be included in the case of inference using prior information. As the prior information introduced in the problem is known, an important diagnostic step is to (3) verify that the prior Markov chain is, in fact, reproducing the adequate prior features, hence indicating the ability of the chain to adequately explore the model space constrained by the prior information. This is an important issue as, in common applications, the prior density is flatter than the posterior (the likelihood with observations is sup-

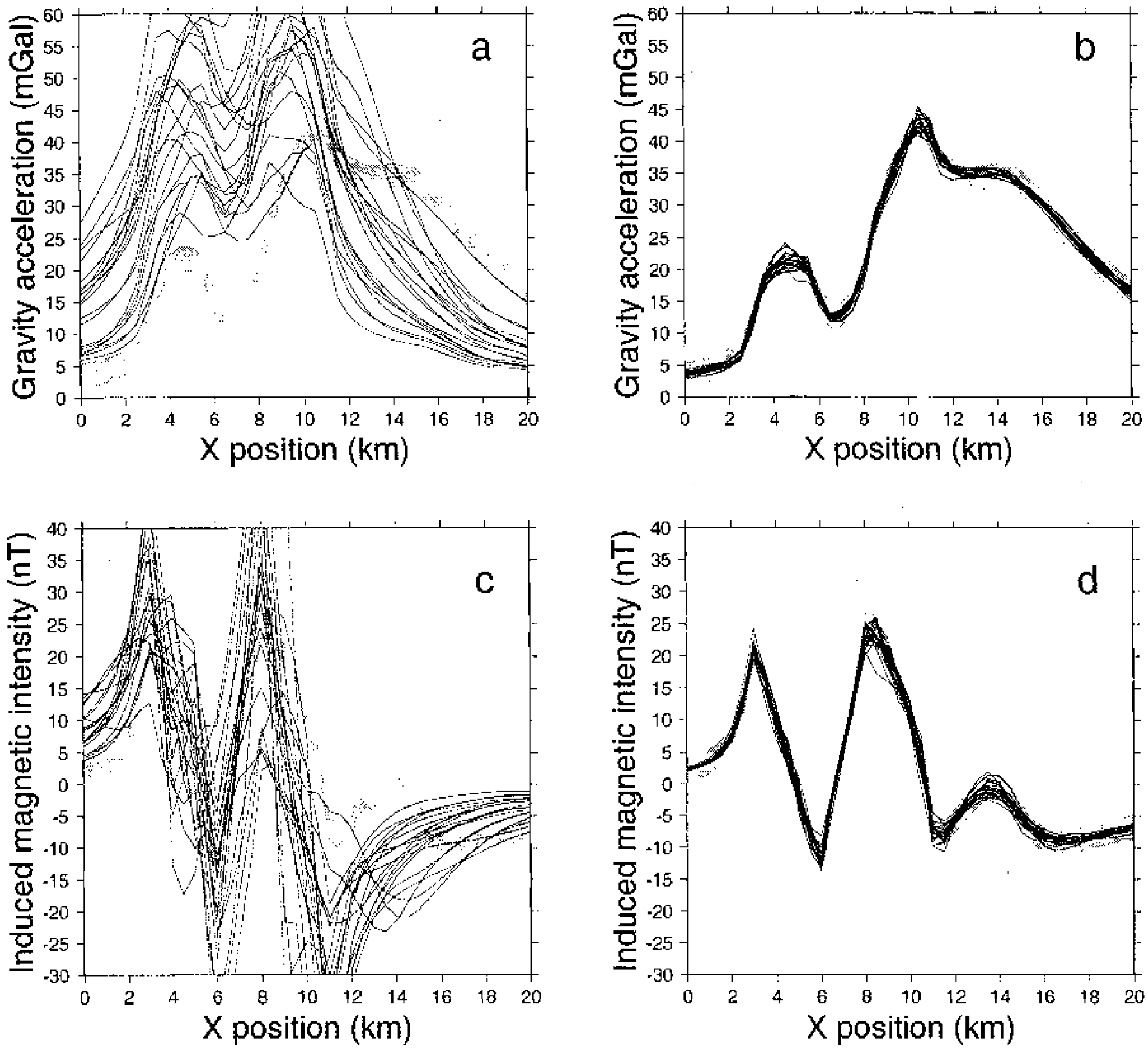


Figure 5. (a) Calculated gravity data and (c) calculated magnetic data for 20 models (including cases of Plate 1) pulled at regular intervals of 50,000 models from the 1,000,000 model set generated by the Markov chain sampler from the prior *pdf*. The data calculated from the true model are plotted in white, and the observation uncertainties (± 1 standard deviation) are represented by a light gray band. (c) Calculated gravity data and (d) calculated magnetic data for 20 models (including cases of Plate 3) pulled at regular intervals of 50,000 models from the 1,000,000 model set generated by the Markov chain sampler from the posterior *pdf*.

posed to further constrain probable models to a smaller region of the model space). Again, the convergence of the prior chain, in say, N iterations, does not warrants that the posterior should converge in the same iteration length. A major difficulty in convergence arises if the likelihood function is very irregular, introducing "barrier" problems in the posterior density (isolating modes surrounded with regions of very low probability). In this case, independent chains should be trapped by these barriers, hence allowing a diagnostic of the problem by using criterion 2 enumerated above.

These three criteria have been used in this example. Elaborated quantitative techniques have been proposed to monitor Markov chain convergence but their generality and reliability are under discussion (see critical work of *Cowles and Carling* [1996] and the debate about the use of multiple chains [*Gelman and Rubin,* 1992] or a single long chain [*Geyer,* 1992]).

Another interesting aspect of Markov chains is the dependence between consecutive samples. Consecutive joint model outcomes from the Markov chain are highly correlated, as only one parameter or a small number of parameters are changed at each step. The correlation distance (in iterations) is different, depending on the particular parameter (or function of parameters) of interest, on the form of the posterior density, and on the walk design (step schedule, step directions, step lengths). One can be interested, for instance, in the total volume of granite, in the total data misfit, in the depth of a particular frontier, or in the values of a particular model parameter. The relevant parameters or functions can be recorded along the chain to further complete statistical study (autocorrelation function, probability bounds, posterior marginal histogram, correlation with another parameter or function).

The estimation of parameters or functions can be performed using consecutive samples from the chain or using selected samples pulled at regular intervals from the chain. The autocorrelation of the series is sometimes used to estimate a correlation distance and pull approximate independent samples of the relevant parameter from the complete chain. As long as there is evidence on the chain convergence, both consecutive dependent samples and selected approximate independent samples provide unbiased estimates. The practical difference between the two approaches consists in the way to calculate the uncertainties of the estimation (see the work of *Hastings* [1970] or the work of *Tierney* [1994] for appropriate formulas taking into account the correlation). Nevertheless, the evaluation of uncertainties of the estimation can be addressed in a simple and general way by running independent chains and comparing the resulting estimates, as presented here in Plates 2b, 2c and 2d for the lithotype spatial histograms.

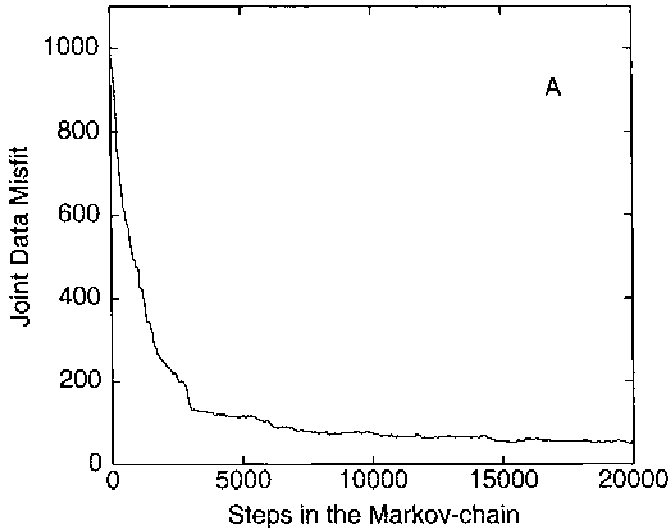
6. Discussion

The general strategy of the lithologic inversion method presented here is schematically illustrated in Figure 1. In comparison with the conventional geophysical inversion, a more complete parameterization of the media is considered, particularly by the introduction of model parameters describing the lithology. Lithologic properties of the rocks are not explicitly present in the physical laws governing the measured geophysical fields, but they are relevant to the observations as long as the physical behavior of the media is strongly determined by lithology.

Depending on the geological setting, the exploration interests, and the geophysical observations, the lithologic media characterization could be different. As the upper crust is very heterogeneous, the lithotypes are a straightforward way to characterize the media. Classification techniques such as the discriminant analysis or the clustering analysis can help to define the lithological categories to be represented in the model according to the available data. In the exploration of sedimentary environments, the porosity and the fluid content are strongly correlated with the physical rock properties; these continuous properties may be included as lithologic parameters in the joint model and estimated from geophysical data.

At a crustal scale, the silicium weight content of the rock may be another good parameter to be estimated from geophysical data. Silicium is the principal mineral component of crustal rocks with a significant variation, ranging from 40% to 80% in most common rocks, and it is strongly correlated with several physical rock properties [*Miller and Christensen*, 1994; *Christensen and Mooney*, 1995].

Under particular conditions, the joint posterior *pdf* formulation given in section 2 reduces to simpler inversion problems. If properties are assumed constant within each lithotype region, the conditional prior *pdf* on secondary parameters would be trivial, and the inversion would only be related to the estimation of the primary model parameters. On the contrary, if a constant lithological model, or a homogeneous lithology, is assumed, no simulation of primary parameters would be needed and the inversion would be only related to the joint estimation of the secondary model parameters. If only one media property and one geophysical data set is considered, the joint posterior *pdf* reduces to the simple form of the posterior *pdf* for a conventional geophysical inverse problem.



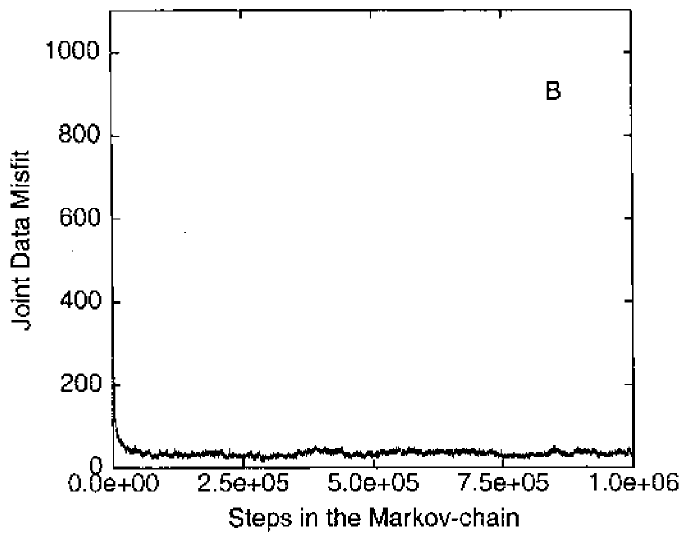


Figure 6. Progress of the misfit between calculated and true data as a function of iterations of the posterior Markov chain, for the same run presented in Plate 3. (a) The first 20,000 iterations. (b) The complete chain of 1 million iterations. Misfit reduction is achieved within the first 15,000 iterations to maintain a stable level. Note that the first 20,000 models are excluded in the lithologic estimation of Plate 2.

The generation of lithological models in the example of section 5 followed a simple method based on random perturbations of the primary model parameters with some prior constraints and conditions. More sophisticated methods for the parameterization and the simulation of lithotype regions can naturally be considered and used to construct the primary Markov chain. There are generally two major approaches for the simulation of categorical property models (like the lithotype models): the indicator approach and the object oriented approach.

The indicator approach is based in a grid type parameterization to describe the lithotype field. It associates to any cat-

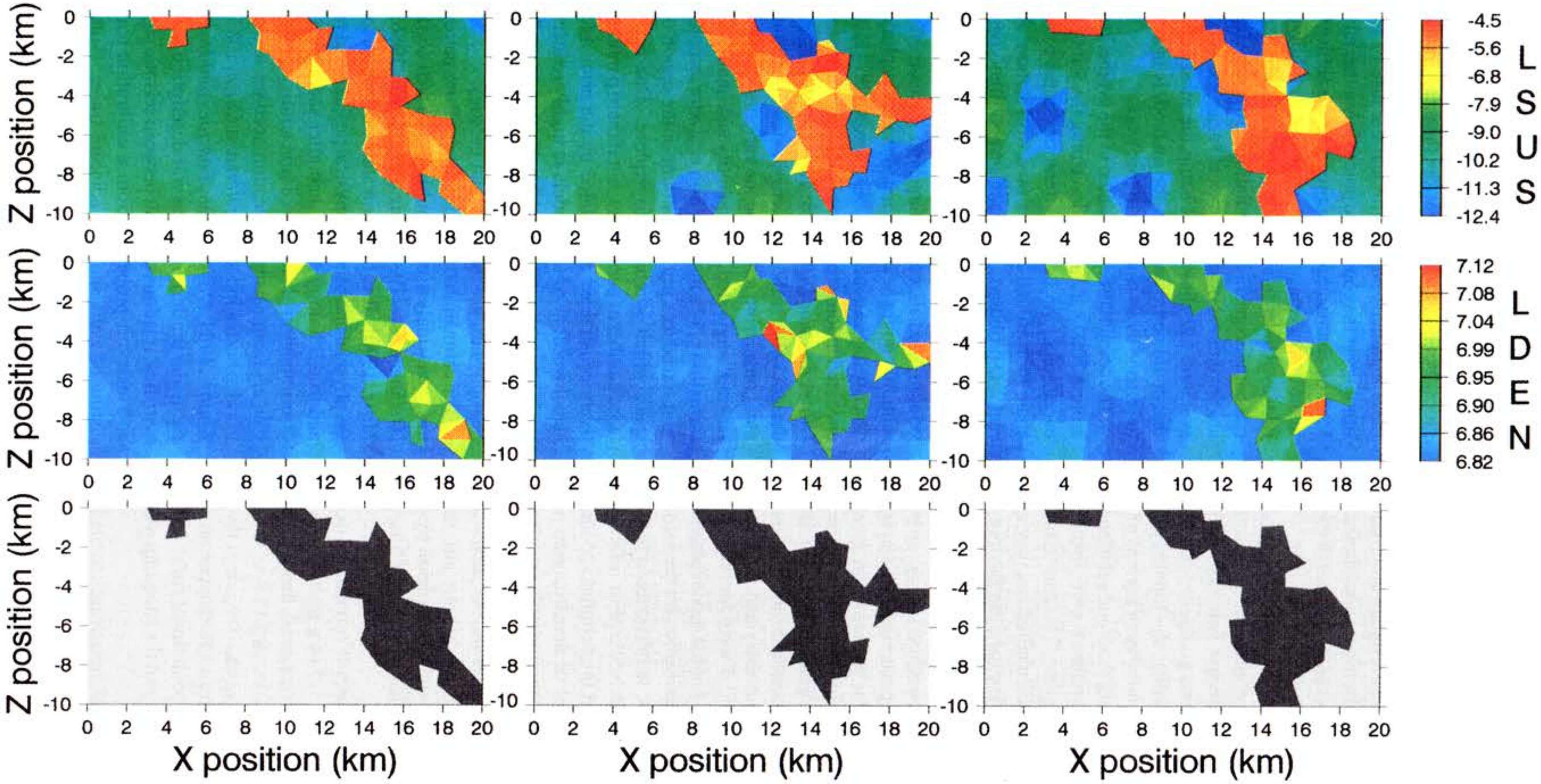


Plate 3. The lithotype model (bottom), the density model (center), and the magnetic susceptibility model (top) for three joint models (model 300,000; model 600,000; and model 900,000) pulled at regular intervals from the model set generated by the Markov chain sampler from the posterior *pdf*. LDEN is the logarithm of the density, and LSUS is the logarithm of the magnetic susceptibility. Peridotite regions are shown in dark gray, whereas the granite region is shown in light gray. Modeling was performed on the logarithmic transformed properties.

egory of rock an indicator continuous variable ranging from 0 to 1. This variable is interpreted as the probability of the category and it is interpolated between the points of the grid according to prior information on the spatial covariance of the indicator variable and according to conditioning lithotype data [*Deutsch and Journel*, 1992; *Mallet and Shtuka*, 1997]. The object-oriented approach is based on the description of the lithotype regions by parameterizing its boundaries. The parameterization is adapted to the expected shapes of the modeled objects (layers, faults, domes, etc.).

The different methods allow us to introduce geological prior information concerning geometrical relations between lithotypes, lithotype probabilities, expected volumes, topology, etc. In the example presented, some geological prior information was considered. In particular, the topology of the model was fixed (which fixes the number of regions of each lithotype in the model and its neighborhood relations). It maintained compact the peridotite regions, avoiding the dispersion of peridotite in multiple regions. This constraint, used in the example, is not inherent to the method and may not be used in other applications, or could be used in a flexible manner by letting the number of connected regions vary between some bounds according to prior probabilities.

A Markov chain Monte Carlo method was described in this work to sample from the joint posterior *pdf* given by (4). Based on the same formulation of the posterior *pdf*, optimization methods may also be developed to search for models maximizing the *pdf*. In particular, gradient methods may be used when the concerned model parameters are continuous.

The joint posterior probability density and the sampling method presented in sections 2 and 3 were formulated with sufficient generality to accept any kind of model parameterization. A particular model parameterization for the secondary properties was chosen in section 4, to introduce in the simulation the petrophysical and geostatistical prior information. This parameterization consisted in the definition of the secondary properties over a finite set of nonstructured points within each lithotype region. Then properties in these control points were simulated by a Monte Carlo method according to geostatistical information. Another way to parameterize and simulate the secondary media properties may also be useful, like describing the properties within the region by coefficients of a base of continuous fields (polynomial functions, harmonic functions).

In the synthetic test the conditional probability density $\theta_{s|p}(\mathbf{m}_{sec}|\mathbf{m}_{pri})$ was assumed to follow a multivariate spatial Gaussian model. It is important to remark that the Markov chain sampling method described in section 3 is valid for any kind of conditional density $\theta_{s|p}(\mathbf{m}_{sec}|\mathbf{m}_{pri})$ on the secondary model parameters, including nonhomogeneous models. As long as any model for the conditional *pdf* is defined, the conditionals can be calculated and the sampling method can be performed.

The use of a Markov chain sampling approach in the illustrative example presented is justified by several reasons. Although the conditional density $\theta_{s|p}(\mathbf{m}_{sec}|\mathbf{m}_{pri})$ model was simple (Gaussian), the joint prior *pdf*, which results from a product with the prior *pdf* over the primary model parameters (1), does not follow such a simple Gaussian model. The resulting joint priors obey complex spatially nonhomogeneous statistics incorporating the information of fixed lithology in the surface and topologic constraints for the lithologic regions. For illustration, consider the prior marginal of the secondary properties for a particular position $\mathbf{x}$ in the section: these marginals are multimodal because of the changes of lithology at the coordinate point $\mathbf{x}$. Another advantage of the Markov chain sampling approach is the ability to incorporate in the model categorical parameters (as lithotypes) in a natural way. Under this realistic parameterization, simpler methods based in gradient calculations are not useful, as gradients are not defined over categorical variables. In conclusion, the synthetic test shown, not being the most complicated case that could be treated with the method, is interesting enough to be used as an example in this work.

The computations presented in the synthetic tests involved around 2000 parameters describing the joint model, and the execution of the 1 million iterations of the posterior Markov chain took around 2 hours on a workstation. Application to real 2-D problems, including several lithotypes, is on course with execution times of the order of 4 hours. For very highly dimensional model parameter spaces, as for realistic 3-D inversion, specialized techniques to increase efficiency of the sampling might be incorporated to the sampling scheme here described. I can mention two approaches: (1) adapting variance reduction techniques such as importance sampling [*Kloek and van Dijk*, 1978; *Tierney*, 1994] and (2) carefully designing the walk (step length, step directions, step schedule) considering simultaneous variations of several parameters in a single step. The quantity and quality of the prior information play a major role on posterior sampling efficiency, as it introduces prior constraint in the region of the volume space which is relevant to explore. One may have a very large number of parameters, but if these parameters are effectively constrained and related by the prior information, the actual freedom of the models is highly reduced. This is an implementation advantage for the present method, which is able to introduce strong realistic prior information from the petrophysics, geostatistics, and geology.

The estimation of the lithologies from petrophysical data has been widely used in multiple-data well-log interpretation [*Delfiner et al.*, 1984; *King and Quirein*, 1986; *King*, 1990; *Moss*, 1990]. It is based on statistical associations between the physical and pseudo-physical properties given by the logs and true lithology established after sample extraction. The discrimination of subsurface lithostratigraphy combining attributes from seismic reflection sections and well-log data has also been applied [*Angeleri and Carpi*, 1982; *Sinvhal and Khattri*, 1983; *Doyen*, 1988; *Fournier and Derain*, 1995], in particular for deriving reservoir description. In these works, statistical relations are built between seismic signal attributes and reservoir properties from a calibration population consisting of wells and their adjacent traces. This information is used to condition the interpolation of media properties between the wells. *Berkhout and Wapenaar* [1990] and *Lortzer and Berkhout* [1992] described in a theoretical work a stepwise scheme to arrive at lithological estimation from seismic reflection data. As in the above referenced works, the lithologic estimation for oil reservoir description is commonly based on the previous results of the inversion or processing of the geophysical data. This is one difference with the methodology presented here, in which these two stages are integrated in the inversion.

7. Summary and Conclusions

Earth models in geophysics have been commonly parameterized in terms of, or to describe, physical properties explicitly related with the data by a physical law. However, the ultimate aim of the data acquisition and analysis is, usually, to estimate the composition, lithology or petrology of the rocks at locations where they have not been sampled. In this work, a general methodology was presented formulating the problem of lithologic estimation as an integrated inverse problem, having as input plural geophysical data, petrophysical and geostatistical information, and to some extent, geological information describing the structure of the model.

The different types of information involved in the problem of lithologic estimation can be combined following a probabilistic inference approach. A joint posterior *pdf* was formulated (4), considering multiple properties in the model and multiple types of geophysical data. In this formulation the joint prior *pdf* was decomposed over a partition of the model parameter space in a primary subspace and a secondary subspace, according to a hierarchy in the properties, and the joint likelihood function was straightforwardly obtained as a product of the likelihood functions associated to each geophysical data set. The formulation is able to incorporate any kind of geophysical data and any kind of forward modeling of the geophysical observations.

A Markov chain Monte Carlo method was developed to sample joint models according to the joint posterior probability density. The organization of the sampling algorithm is adapted to the structure of the posterior *pdf* itself: (1) an initial Markov chain should be designed to sample from the primary prior *pdf*, (2) it is generalized to the joint model space sampling from the prior conditional *pdf* (describing secondary parameters conditional to primary parameters), and (3) it is modified to sample from the posterior *pdf* using the Metropolis rule.

The inference of the conditional prior density is based on petrophysical and geostatistical information and it follows common geostatistical methods, in most of its aspects. Some important problems to deal with are the representation of the lithologies and the spatial clustering of the field samples, the parameterization of the *pdfs*, the modeling of the spatial correlation between the properties and the regularization of sample *pdfs* according to the scale of the model spatial resolution.

It is shown in this work that the information describing rock property relations has a relevant contribution in the structure of the joint posterior *pdf*. This information allows us to infer conditional *pdfs* indispensable for the joint model simulations. Following this view, the acquisition and analysis of petrophysical and geostatistical data in exploration campaigns may be better valorized and considered as important to interpretation as the acquisition of the geophysical data itself.

Acknowledgments. This work would have not been possible without the contribution of Albert Tarantola (Institut de Physique du Globe de Paris) by significant discussions about the theory presented here. Maria Zamora (Institut de Physique du Globe de Paris) kindly provided the petrophysical data used in this work. Significant discussion about sampling algorithms was held with Klaus Mosegaard (Niels Bohr Institut, Denmark) at the initial stage of this project. I acknowledge Klaus Mosegaard (Associate Editor of JGR), Olivier Dubrule (Elf Research Center), and an anonymous referee for their valuable reviews of the manuscript. This research was partially funded by program GeoFrance-3D (France), program TEK-VOLC (European Union), and the Consejo de Desarrollo Científico y Humanístico de la Universidad Central de Venezuela. This paper is the publication n. 26 of GeoFrance-3D.

References

Almeida, A. S., and G. Journel, Joint simulation of multiple variables with a Markov-type coregionalization model, *Math. Geol., 26,* 565-588, 1994.

Angeleri, G. P., and R. Carpi, Porosity prediction from seismic data, *Geophys. Prospect., 30,* 580-607, 1982.

Berkhout, A. J., and C. P. A. Wapenaar, Delphi: Delft philosophy on acoustic and elastic inversion, *Leading Edge, 9,* 30-33, 1990.

Besag, J., P. Green, D. Higdon, and K. Mengersen, Bayesian computation and stochastic systems, *Stat. Sci., 10,* 3-66, 1995.

Bosch, M., P-wave velocity tomography of the Venezuelan region from local arrival times, *J. Geophys. Res., 102,* 5455-5472, 1997.

Bott, M. H. P., Two methods applicable to computers for evaluating magnetic anomalies due to finite three-dimensional bodies, *Geophys. Prospect., 11,* 292-299, 1963.

Bourne, J. H., Use of magnetic susceptibility, density and modal mineral data as a guide to composition of granitic plutons, *Math. Geol., 25,* 1993.

Camacho, A. G., F. G. Montesinos, and R. Vieira, A three-dimensional gravity inversion applied to Sao Miguel Island (Azores), *J. Geophys. Res., 102,* 7717-7730, 1997.

Chapman, M. E., Techniques for interpretation of geoid anomalies, *J. Geophys. Res., 84,* 3793-3801, 1979.

Christensen, N. I., and W. E. Mooney, Seismic velocity structure and composition of the continental crust: A global view, *J. Geophys. Res., 100,* 9761-9788, 1995.

Cowles, M. K., and B. P. Carling, Markov Chain Monte Carlo convergence diagnostics: A comparative review, *J. Am. Stat. Assoc., 91,* 883-904, 1996.

Davis, M., Production of conditional simulations via the LU triangular decomposition of the covariance matrix, *Math. Geol., 19,* 91-98, 1987.

Delfiner, P., O. Serra, and O. Peyret, Automatic determination of lithology from Well-logs, Paper SPE 13290, presented at the SPE Annual Technical Conference and Exhibition, Soc. of Pet. Exp., 1984.

Deutsch, C. V., and A. G. Journel, *GSLIB, Geostatistical Software Library and User's Guide,* 340 pp., Oxford Univ. Press, New York, 1992.

Dietrich, C. R., Computationally efficient generation of gaussian conditional simulations over regular sample grids, *Math. Geol., 25,* 439-451, 1993.

Domenico, S. N., Rock lithology and porosity determination from shear and compressional wave velocity, *Geophysics, 49,* 1188-1195, 1984.

Doyen, P. M., Porosity from seismic data: A geostatistical approach, *Geophysics, 53,* 1263-1275, 1988.

Fournier, F., and J. F. Derain, A statistical methodology for deriving reservoir properties from seismic data, *Geophysics, 60,* 1437-1450, 1995.

Gelfand, A. E., and A. F. M. Smith, Sampling-based approaches to calculating marginal densities, *J. Am. Stat. Assoc., 85,* 398-409, 1990.

Gelman, A., and D. B. Rubin, Inference from iterative simulation using multiple sequences, *Stat. Sci., 7,* 457-511, 1992.

Geyer, C.J., Practical Markov Chain Monte Carlo, *Stat. Sci., 7,* 473-511, 1992.

Han, D., A. Nur, and D. Morgan, Effects of porosity and clay content on wave velocities in sandstones, *Geophysics, 51,* 2093-2107, 1986.

Hastings, W. K., Monte Carlo sampling method using Markov chains and their applications, *Biometrika, 57,* 97-109, 1970.

Horen, H., Proprietés magnetiques et sismiques de la lithosphere oceanique, a partir d'échantillons marins et d'analogues ophiolitiques, Doctoral thesis, 350 pp., Univ. de Paris VII, 1997.

Isaaks, E. D., and R. M. Srivastava, *Applied Geostatistics,* 561 pp., Oxford Univ. Press, New York, 1989.

Jackson, D. D., and M. Matsu'ura, A Bayesian approach to nonlinear inversion, *J. Geophys. Res.,* 90, 581-591, 1985.

Jordan, T. H., Mineralogies, densities and seismic velocities of granet lherzolites and their geophysical implications, in *The Mantle Sample: Inclusions in Kimberlites and Other Volcanics, Proceedings 2nd International Kimberlite Conference*, vol. 2, edited by F.R. Boyd and H.O.A. Meyer, pp. 1-14, AGU, Washington, D.C., 1979.

Journel, A. G., and C. Huijbregts, *Mining Geostatistics,* 600 pp., Academic, San Diego, Calif., 1978.

King, D. E., Incorporating geological data in well-log interpretation, in *Geological Applications of Wireline Logs,* edited by A. Hurst, M.A. Lovell and A.C. Morton, pp. 45-55, *Geol. Soc. Spec. Publi., N. 48,* 1990.

King, D. E., and J. Quirein, Wireline log modeling of geological data: Application to automated lithology identification, *AAPG Bull., 70,* 607, 1986.

Kloek, T., and H. K. van Dijk, Bayesian estimates of equation system parameters: An application of integration by Monte Carlo, *Econometrica, 46,* 1-19, 1978.

Li, Y., and D. W., Oldenburg, 3-D inversion of magnetic data, *Geophysics, 61,* 394-408, 1996.

Lortzer, G. J. M., and A. J. Berkhout, An integrated approach to lithologic inversion, Part I, Theory, *Geophysics, 57,* 233-244, 1992.

Mallet, J. L, and A. Shtuka, Stratigraphic inversion integrating wells, seismic and sedimentology data, *Expanded Abstr. SEG Annual Meeting, 67th,* 1632-1635, 1997.

Metropolis, N. , A. W. Rosenbluth, M. N Rosenbluth, A. H. Teller, and E. Teller, Equation of state calculations by fast computing machines, *J. Chem. Phys., 21,* 1087-1092, 1953.

Miller, D. J., and N. I. Christensen, Seismic signature and geochemistry of an island arc: A multidisciplinary study of the Kohistan accreted terrane, northern Pakistan, *J. Geophys. Res., 99,* 11,623-11,642, 1994.

Mosegaard, K., and A. Tarantola, Monte Carlo sampling of solutions to inverse problems, *J. Geophys. Res., 100,* 12,431-12,447, 1995.

Mosegaard, K., S. Singh, D. Snyder, and H. Wagner, Monte Carlo analysis of seismic reflections from Moho and the W reflector, *J. Geophys. Res., 102,* 2969-2981, 1997.

Moss, B. P., Stochastic reservoir description: A methodology, in *Geological Applications of Wireline Logs,* edited by A. Hurst, M.A. Lovell, and A.C. Morton, pp. 57-75, *Geol. Soc. Spec. Publi. N. 48,* 1990.

Newman, G. A., and D. L. Alumbaugh, Three-dimensional massively parallel electromagnetic inversion, II, Analysis of a crosswell electromagnetic experiment, *Geophys. J. Int., 128,* 355-363, 1997.

Oliver, D. S., Moving averages for gaussian simulation in two and three dimensions, *Math. Geol., 27,* 939-960, 1995.

Press, F., Earth models obtained by Monte Carlo inversion, *J. Geophys. Res., 73,* 5223-5234, 1968.

Robert, C. P., Convergence control methods for Markov Chain Monte Carlo Algorithms, *Stat. Sci., 10,* 231-253, 1995.

Sinvhal, A., and K. N. Khattri, Application of seismic reflection data to discriminate subsurface lithostratigraphy, *Geophysics, 48,* 1498-1513, 1983.

Smith, A. F., and G. O. Roberts, Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods, *J. R. Stat. Soc. B, 55,* 3-23, 1993.

Stark, P. B., Inference in infinite-dimensional inverse problems: discretization and duality, *J. Geophys. Res., 97,* (B10), 14,005-14,082, 1992.

Tarantola, A., *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation,* 613 pp., Elsevier, New York, 1987.

Tarantola, A., and B. Valette, Inverse problems = Quest for information, *J. Geophys., 50,* 159-170, 1982.

Tierney, L., Markov-chains for exploring posterior distributions, *Ann. Stat., 22,* 1702-1762, 1994.

Wiggins, R. A., Monte Carlo inversion of body wave observations, *J. Geophys. Res., 74,* 3171-3181, 1969.

M. Bosch, Departamento de Física Aplicada, Facultad de Ingeniería, Universidad Central de Venezuela, Los Chaguaramos, Caracas, Venezuela. (mbosch@reacciun.ve)

(Received January 8, 1998; revised July 15, 1998;
accepted September 9, 1998.)