

3

Inference Networks in Earth Models with Multiple Components and Data

Miguel Bosch

ABSTRACT

The integration of information for the inference of earth structure and properties can be treated in a probabilistic framework by considering a posterior probability density function (PDF) that combines the information from a new set of observations and a prior PDF. To formulate the posterior PDF in the context of multiple datasets, the data likelihood functions are factorized assuming independence of uncertainties for data originating across different surveys. A realistic description of the earth medium requires the modelization of several properties and other structural parameters, which relate to each other according to dependency and independency notions. Thus, conditional probabilities across model components also factorize. The relationships across model components can be described via a direct acyclic graph. The basic rules for factorization of the posterior PDF are easily obtained from the graph organization. Once the posterior probability has been formulated, realizations can be obtained following a sampling approach or searching for a maximum posterior probability earth medium configuration. In the first case, sampling algorithms will adapt to the factorized structure of the posterior PDF. In the second case, iterative second- or first-order approximations of the objective function conduce to the solution of a system of equations for the model update.

3.1. INTRODUCTION

The appraisal of solid Earth structure and properties requires modeling the medium's heterogeneous composition and lithotypes, the morphology of geological bodies, pore fluids, fractures, knowledge, and information that are linked at various scales, and the medium's physical response to field measurements. This implies the description of various types of (1) medium properties and structural parameters, (2) observations provided by surveys, well-logs, and rock sample studies, and (3) knowledge to establish relationships across the model properties and observations. The full multiplicity of components and intervening relationships could be

considered too large to be workable. Nevertheless, given the goal of a specific study and the available information, a relevant set of such components and relationships can be retained for modeling, while the rest is neglected. The appropriate selection conforms to the pertinence of the model component to the phenomenon observed and the goal of the study.

Integration of multiple data, information, and knowledge has been considered a key issue in natural sciences, and particularly in Earth sciences. The relevant information is heterogeneous in its nature (properties, objects, phenomena, scales) and available at diverse treatment levels: The data are the raw support of the information (processed and interpreted data), which is understood in terms of knowledge (a successful theory). With the advent of larger computational possibilities in the past decades, the quantitative treatment of multicomponent

Applied Physics Department, Engineering Faculty, Universidad Central de Venezuela, Caracas, Venezuela

Integrated Imaging of the Earth: Theory and Applications, Geophysical Monograph 218, First Edition.

Edited by Max Moorkamp, Peter G. Lelièvre, Niklas Linde, and Amir Khan.

© 2016 American Geophysical Union. Published 2016 by John Wiley & Sons, Inc.

complex models and observations is in progress to provide (a) more realism to the description of the Earth media and (b) more accuracy and precision to the estimates [Linde *et al.*, 2006a; Bosch *et al.*, 2010; Torres-Verdin *et al.*, 2012]. Examples of joint inversion of multiple geophysical datasets, in various Earth sciences inference contexts, are described in the works by Lines *et al.* [1988], Haber and Oldenburg [1997], Bosch [1999], Bosch *et al.* [2001, 2004], Tiberi *et al.* [2003], Gallardo *et al.* [2003], Gallardo and Meju [2004], Linde *et al.* [2006b], Guillen *et al.* [2007], Khan *et al.* [2007], Alpak *et al.* [2008], Doetsch *et al.* [2010], Buland and Kolbjomsen [2012], and Chen and Hoversten [2012].

The present work focuses on the formulation of inverse problems in Earth sciences, for the case of models configured with multiple spatially distributed properties, subjected to different types of physical observations and/or embodying multiple relationships across properties and observations. The nature of the formulation and approaches to solve inverse problems is not different for this scenario. The issues to solve consists in (1) how to compose relationships and information across the model components and (2) how to draw from the composed information realizations of the joint model, or maximum posterior probability model configurations, according to the solution approach followed. To explain these issues, I will not follow a rigorous deductive path. I will mention the basic principles and illustrate their formulation for various common examples of inferential interest in earth sciences.

The first section of this chapter describes the formulation of posterior probability densities for complex models, for the case where their components are structured in hierarchical layers, and for multiple sets of data. The second section unfolds the structure of the posterior probability density for less structured models, via the support of direct acyclic graphs [Pearl, 1986; Thulasiraman and Swamy, 1992] that describe the relations across the model components and datasets. The third section describes the stochastic approach to the solution of the inverse problem by sampling the posterior density with Markov Chains, following its factor structure. The fourth section describes the optimization approach to the solution of the inverse problem, which consists of searching for the model configuration that maximizes the posterior density. Finally, the discussion and conclusion sections close the chapter.

3.2. MULTIPLE PHYSICAL OBSERVATIONS AND MODEL COMPONENTS

The general formulation of inverse problems is outlined here in a probabilistic framework, within the scope of Bayesian inference. The term Bayesian refers to the

interpretation of probabilities as a description of the information, knowledge, or uncertainty on parameter spaces that support a model of a natural object or phenomenon. A state of information about the modeled object is described by the *probability density function* (PDF) defined over the model parameter space or, equivalently, by the *cumulative distribution function* (CDF). The variables in the parameter space are considered *random* as they take different values each time they are evaluated and represent the corresponding state of information. Each time the model parameters are evaluated, their outcome is drawn in proportion to the corresponding PDF realizing a different configuration of the modeled earth medium. The discussion section expands the analysis of the relationship between the model parameter space and the modeled object space.

The common formulation of inverse problems considers a *prior* state of information, a set of observations related with the modeled object, and the improved *posterior* state of information resulting from the combination of the prior and the new information provided by the data interpretation. The *posterior* probability density is given by

$$\sigma(\mathbf{m}) = c \rho(\mathbf{m}) L_{\text{data}}(\mathbf{m}), \quad (3.1)$$

where \mathbf{m} is a multivariate random variable in the model parameter space, c is a normalization constant, $\rho(\mathbf{m})$ is the prior PDF, and $L_{\text{data}}(\mathbf{m})$ is the data likelihood function, which embodies the new information provided by the observations. Normalization constants will be included in the following equations of the posterior PDF but not further identified.

As we know, observations are not commonly made directly on the model parameters, but in terms of additional related parameters that we refer to as *data*. The knowledge of the relationship between the data and the model parameters, including the associated uncertainties, provides the means to transform the former into information of the latter—that is, interpret the data in terms of the modeled object information. The formal derivation of the data likelihood function in (3.1) depends on the formulation of the relationship between the model and data spaces and the associated uncertainties. In the general case [Tarantola, 2005], the information provided by the theory and the observations are modeled independently and combined in a joint model-data space. The data likelihood function is calculated as a marginal non-normalized probability in the model parameter space,

$$L_{\text{data}}(\mathbf{m}) = \int \theta_{\text{theory}}(\mathbf{d} | \mathbf{m}) \rho_{\text{obs}}(\mathbf{d}) d\mathbf{d}. \quad (3.2)$$

Above, \mathbf{d} are the true data parameters here considered with uniform homogeneous probability density. It is the

data that would have been observed in the absence of observational and data processing errors. The probability density $\rho_{\text{obs}}(\mathbf{d})$ describes the information on the true data provided by the observation experience, $\mathbf{d} = \mathbf{d}_{\text{obs}} + \Delta\mathbf{d}_{\text{obs}}$, via the corresponding observed data \mathbf{d}_{obs} and the associated uncertainties $\Delta\mathbf{d}_{\text{obs}}$. The conditional probability density $\theta_{\text{theory}}(\mathbf{d}|\mathbf{m})$ describes the data–model relationship, based on theoretical or empirical knowledge, including the data modeling uncertainties. The homogeneous probability density [Mosegaard, 2011; Mosegaard and Tarantola, 2002; Tarantola, 2005] describes the state of null information about the parameter. The homogeneous PDF of the data should be included in the denominator of the integrand in (3.2), in the case where it is not modeled as uniform (constant).

It is also common to model the observational uncertainties within the data–model conditional, instead of providing the independent observational PDF, $\rho_{\text{obs}}(\mathbf{d})$, present in the general expression (3.2). In this case the data likelihood function can be formulated straight forwardly as

$$L_{\text{data}}(\mathbf{m}) = \Theta_{\text{theory}}(\mathbf{d}_{\text{obs}}|\mathbf{m}), \quad (3.3)$$

where \mathbf{d}_{obs} are the observed data. Both formulations are equivalent if uncertainties are appropriately modeled. The reader is referred to the work by Tarantola [2005] for a general derivation of the above expressions from the theory of combination of information states, generalized for parameters with nonuniform homogeneous probability densities. Equations (3.1)–(3.3) are the common basis for statistical inference, although derived from different theoretical approaches.

3.2.1. Likelihood Function Factorization

Let us now consider a model with various inner components of model parameters, $\mathbf{m} = \{\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_K\}$, influenced by various types of data observations, $\mathbf{d} = \{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_N\}$, as required in integrated data modeling. The inner components of the model parameters, \mathbf{m}_k , could correspond to different property fields, the same property fields at different scales, geometric boundaries of geological objects, and other subsets of the model parameters specific case. Each of these multiple components is commonly of high dimensionality, such as a property field distributed on a spatial grid. The data components are partitioned according to different observational phenomenon (gravity, seismic, electric), derived data (seismic travel times, amplitudes, frequencies), or field surveys. We can decompose the observational PDF in the joint data space by the product of marginals,

$$\begin{aligned} \rho_{\text{obs}}(\mathbf{d}) &= \rho_{\text{obs}N}(\mathbf{d}_N)\rho_{\text{obs}N-1}(\mathbf{d}_{N-1})\cdots\rho_{\text{obs}1}(\mathbf{d}_1) \\ &= \prod \rho_{\text{obs}n}(\mathbf{d}_n), \end{aligned} \quad (3.4)$$

under the assumption of independent observational data uncertainties across the surveys. Recall that the observational PDF, $\rho_{\text{obs}}(\mathbf{d})$, only embodies information about the measurement process and does not anticipate the posterior information in data space—that is, as present in the posterior marginal of (3.1). The assumption is well justified for different surveys, which commonly use different instrumentation and field teams, and for different observational phenomenon (rock samples, well-logs, seismic experiment, gravity measurement). Nevertheless, in a strict sense, Eq. (3.4) is an approximation, as measurements across different surveys may be affected by commonly used information (a common digital elevation model, seasonal terrain conditions, or other factors). We assume herein that possible correlated factors are minor compared with the total observational uncertainty. The above decomposition of the observational PDF can in some cases be applied to data components derived from the same survey. As an example, various types of partial data can be obtained from seismic surveys, such as phase travel times, reflection amplitudes, and frequency content, which are commonly interpreted independently assuming unrelated uncertainties.

Similarly, the theoretical conditional probability is composed by the product of conditional marginals,

$$\begin{aligned} \Theta_{\text{theory}}(\mathbf{d}|\mathbf{m}) &= \Theta_{\text{theory}N}(\mathbf{d}_N|\mathbf{m})\Theta_{\text{theory}N-1}(\mathbf{d}_{N-1}|\mathbf{m})\cdots \\ \Theta_{\text{theory}1}(\mathbf{d}_1|\mathbf{m}) &= \prod \Theta_{\text{theory}n}(\mathbf{d}_n|\mathbf{m}), \end{aligned} \quad (3.5)$$

when assuming independence of the modeling uncertainties. Modeling of different phenomenon (seismic, electric, gravity) and different components of the same phenomenon (seismic travel times, amplitude reflections, frequency decay) are based on different types of theoretical knowledge and are likely to have independent modeling uncertainties, thereby supporting the stated assumption. Again, the expression is an approximation that neglects possible related factors emerging from common modeling choices (e.g., common spatial property discretization for instance).

By substitution of (3.4) and (3.5) in (3.2) followed by integration, we obtain the joint likelihood function as the product of the data likelihood functions for each of the data components,

$$L_{\text{data}}(\mathbf{m}) = L_{\text{data}1}(\mathbf{m})L_{\text{data}2}(\mathbf{m})\cdots L_{\text{data}N}(\mathbf{m}) = \prod L_{\text{data}n}(\mathbf{m}), \quad (3.6)$$

with

$$L_{\text{data}n}(\mathbf{m}) = \int \Theta_{\text{theory}n}(\mathbf{d}_n|\mathbf{m})\rho_{\text{obs}n}(\mathbf{d}_n)d\mathbf{d}_n. \quad (3.7)$$

A similar result is obtained by substitution of (3.5) in (3.3). The factorization of the joint data likelihood

according to the data subsets is equivalent to the addition of data objective function terms in the framework of deterministic solutions to the inverse problem, as will be explained in the corresponding section of this chapter.

3.2.2. Prior PDF Factorization

At the multiple components model, the prior PDF on the model parameters can be decomposed by following the rule of conditional probabilities:

$$\rho(\mathbf{m}) = \rho(\mathbf{m}_K | \mathbf{m}_{K-1}, \dots, \mathbf{m}_1) \times \rho(\mathbf{m}_{K-1} | \mathbf{m}_{K-2}, \dots, \mathbf{m}_1) \dots \rho(\mathbf{m}_1). \quad (3.8)$$

Given the knowledge on the relationships across the model components, some of these conditionals could be simplified. Causal or empirical statistical relationships impose the relevant dependencies, and independencies, enforcing a hierarchy to the model components. In figure 3.1, we present a common scheme for a multicomponent model with multiple data observations, structured in two layers of properties: *primary*, $\mathbf{m}_{\text{pri}} = \{\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_M\}$, and *secondary*, $\mathbf{m}_{\text{sec}} = \{\mathbf{m}_{M+1}, \mathbf{m}_{M+2}, \dots, \mathbf{m}_K\}$, and one observed data layer. In this setting, the secondary properties are dependent on the primary properties, while the observed data are dependent on the secondary properties. The data are not directly (explicitly) dependent on the

primary properties. With the imposed hierarchy and model layers, the prior information can be satisfactorily decomposed by

$$\rho(\mathbf{m}) = \rho(\mathbf{m}_{\text{sec}} | \mathbf{m}_{\text{pri}}) \rho(\mathbf{m}_{\text{pri}}) = \rho(\mathbf{m}_{M+1}, \dots, \mathbf{m}_K | \mathbf{m}_M, \dots, \mathbf{m}_1) \rho(\mathbf{m}_M, \dots, \mathbf{m}_1), \quad (3.9)$$

and the posterior PDF for the situation in Figure 3.1 takes the form

$$\sigma(\mathbf{m}) = c \rho(\mathbf{m}_{M+1}, \mathbf{m}_K | \mathbf{m}_M, \dots, \mathbf{m}_1) \rho(\mathbf{m}_M, \dots, \mathbf{m}_1) \times L_{\text{data1}}(\mathbf{m}_{M+1}) L_{\text{data2}}(\mathbf{m}_{M+2}) L_{\text{data3}}(\mathbf{m}_{M+2}, \mathbf{m}_K), \quad (3.10)$$

with the factorization of the likelihood function as previously explained and including the explicit dependency of each data component on the corresponding secondary model component as indicated in Figure 3.1. It is common to define an objective function proportional to the logarithm of the posterior PDF, as will be described in the optimization approach section of this chapter. The factorization of the posterior PDF, shown in expression (3.10), is in this setting equivalent to the addition of terms in the objective function, each one corresponding to a particular factor of the posterior PDF.

In Figure 3.1, I have drawn separate data component boxes to indicate independence across the data

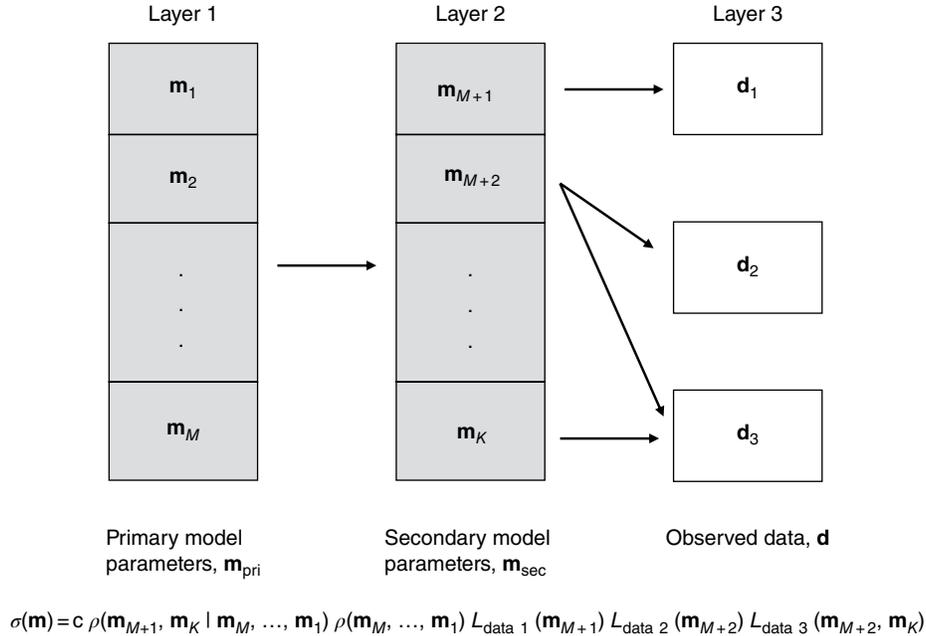


Figure 3.1 Random variables organized in hierarchical layers describing model parameters and data in an inference problem. Bold arrows indicate dependencies across random variables and its modeling sense. Variables in common blocks are modeled jointly. Gray boxes indicate model parameters describing Earth medium properties and structure, while white boxes indicate parameters describing experimental observations and measurements. The composition of the posterior PDF is shown at the bottom as a product of data likelihood functions, priors, and conditional PDFs.

components with respect to the observational and modeling data uncertainties. The juxtaposition of boxes shown for model components at the primary and secondary model layers indicate retention of dependencies across these components in the prior, $\rho(\mathbf{m}_M, \dots, \mathbf{m}_1)$, and conditional, $\rho(\mathbf{m}_{M+1}, \mathbf{m}_K | \mathbf{m}_M, \dots, \mathbf{m}_1)$; these properties should be modeled jointly considering their cross-relations, and spatial relations could be accounted for.

3.2.3. Examples in Common Inferential Settings

When an inferential problem is analyzed, the first step is to define the model components, their internal relations, and their data. This should be done by an expert, or a team of experts, in order to ensure the pertinence of the model and data to satisfy the goal of the inference. The network design introduced above involves retaining model components, their relevant relations, their sets of data, and their relations with the model components. Dependencies and independencies need to be defined.

In this section, I will illustrate the structure of common inferential problems in Earth sciences and the appropriate

formulation of the posterior PDF. In Figures 3.2 and 3.3, I show layered multicomponent models that are useful in inverse problems at local, regional and planetary scales. In Figure 3.2 the setting for an integrated description of a siliciclastic sedimentary medium is depicted by four parameter layers: three model parameter layers and one data layer.

In sedimentary basins, the spatial statistical characteristics of the medium properties is at large scope heterogeneous, whereas within the same formation or units the statistics can be analyzed as spatially homogeneous. Thus, for appropriate statistical modeling, a primary space describing the formation delineation and their sequence is needed. This information can be parameterized by the formation category sequence (formation identification) and a geometrical framework delimiting the statistically homogeneous medium regions by the corresponding horizons. Prior information on these primary parameters, $\mathbf{m}_{\text{formation}}$, is usually obtained via interpreted seismic horizons, well-log data, and geological knowledge of the area.

Within formations, several types of lithology can be present (carbonates, igneous intrusions, siliciclastic

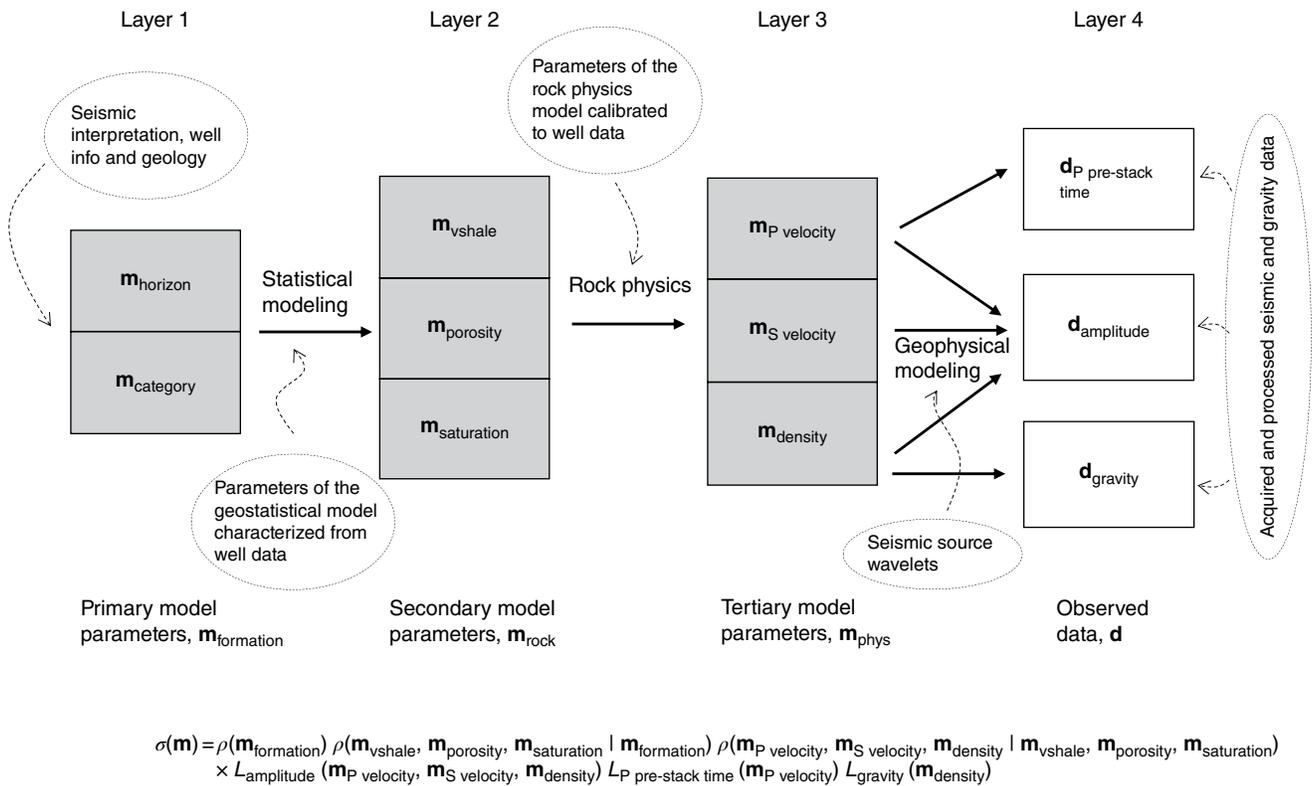
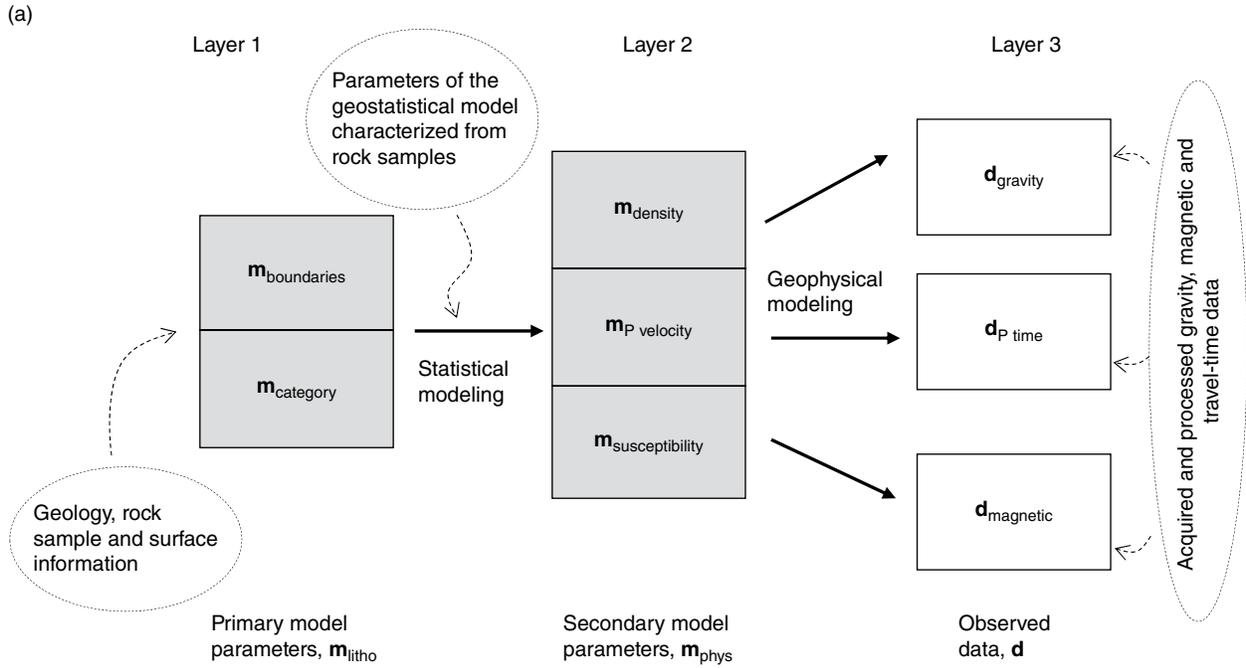
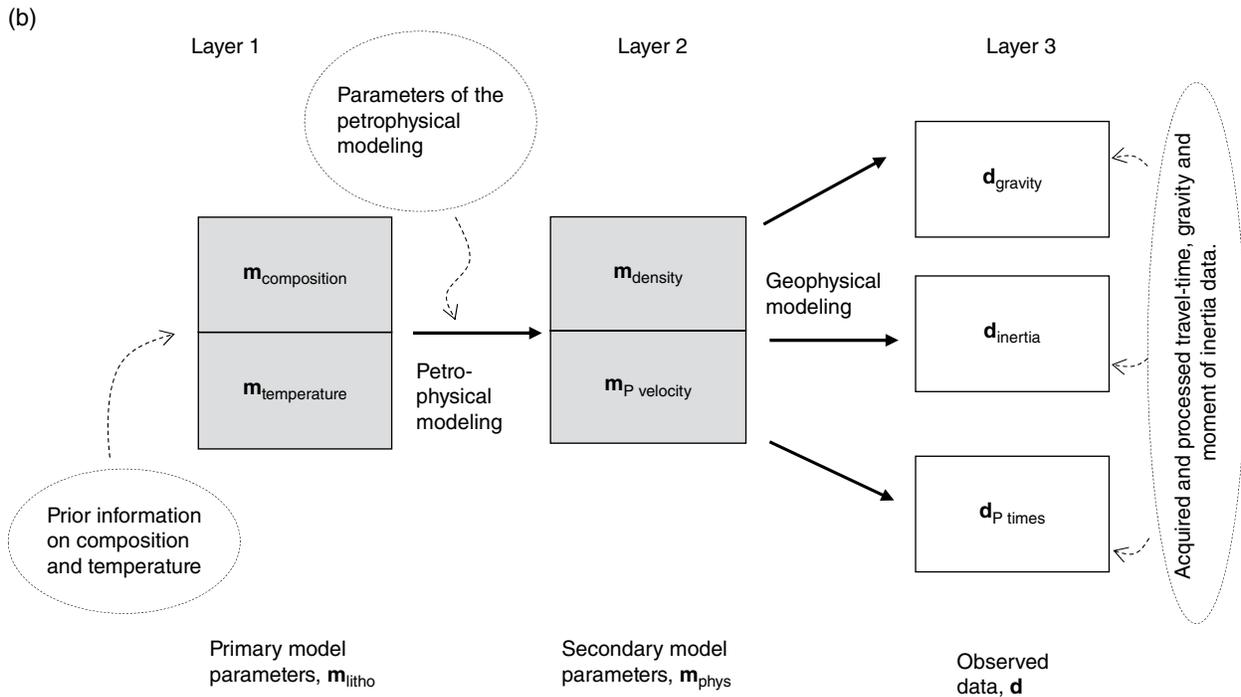


Figure 3.2 Example of model parameter structure in Earth science inference settings: the case of siliciclastic sedimentary basin description based on seismic reflection amplitudes, seismic pre-stack P arrival times and gravity observations. Dependencies across random variables and their hierarchies are shown by bold arrows. Nonrandom parameters required for the conditionals, priors, and likelihoods are indicated in dashed ellipses. Bold boxes show the random parameters for description of the earth medium (gray) and observations (white). The corresponding structure of the posterior PDF is shown as the bottom.



$$\sigma(\mathbf{m}) = \rho(\mathbf{m}_{\text{litho}}) \rho(\mathbf{m}_{\text{density}}, \mathbf{m}_{\text{P velocity}}, \mathbf{m}_{\text{susceptibility}} | \mathbf{m}_{\text{litho}}) L_{\text{gravity}}(\mathbf{m}_{\text{density}}) L_{\text{P time}}(\mathbf{m}_{\text{P velocity}}) L_{\text{magnetic}}(\mathbf{m}_{\text{susceptibility}})$$



$$\sigma(\mathbf{m}) = \rho(\mathbf{m}_{\text{litho}}) \rho(\mathbf{m}_{\text{density}}, \mathbf{m}_{\text{P velocity}} | \mathbf{m}_{\text{litho}}) L_{\text{gravity}}(\mathbf{m}_{\text{density}}) L_{\text{P time}}(\mathbf{m}_{\text{P velocity}}) L_{\text{inertia}}(\mathbf{m}_{\text{density}})$$

Figure 3.3 Example of model parameter structure in earth science inference settings: (a) Lithotype geobody description at crustal regional scale based on gravity, magnetic, and seismic travel-time data. (b) Planet scale composition and temperature description constrained by seismic travel times, gravity data, and the inertia moment. Symbols are the same as in previous figures. The corresponding structure of the posterior PDF is shown at the bottom of each figure.

sedimentary rocks). We will consider here the case of siliciclastic sedimentary rocks, where lithology can be described by the shale volume fraction. Another secondary parameter is the total porosity that influences the elastic medium properties and density, and finally the pore fluid volume fraction (saturation) is important in systems with two or more fluids. Conditioned to the formation and well-log information, geostatistical parameters describing the rock matrix and fluid properties are commonly characterized to model this secondary layer of parameters, $\mathbf{m}_{\text{rock}} = \{\mathbf{m}_{\text{vshale}}, \mathbf{m}_{\text{porosity}}, \mathbf{m}_{\text{saturation}}\}$. In this case we refer to the use of well data to calibrate spatially homogeneous property statistics (means, covariances). An example with spatially localized well-log information will be considered in the next section of this chapter.

According to the rock matrix and fluid configurations, rock physical models are used to calculate properties that characterize the mechanical behavior of the medium, like compressional seismic velocity and shear seismic velocity as well as the mass density, $\mathbf{m}_{\text{phys}} = \{\mathbf{m}_{\text{Pvelocity}}, \mathbf{m}_{\text{Svelocity}}, \mathbf{m}_{\text{density}}\}$. This set of parameters represents the third layer of model parameters.

Finally, Figure 3.2 includes in the fourth layer data from common geophysical surveys that provide information to interpret sedimentary basin stratification, depth, and structure. Interpretation of the seismic data can be in various ways, either in full wave form (full data) or by separating data components. It is common to interpret in a separate manner the pre-stack P-wave travel time for major well-identified reflectors and the reflection amplitudes after migration (spatial repositioning of the seismic data). We conform the data parameter layer in this example with these seismic partial data and the gravity data, $\mathbf{d} = \{\mathbf{d}_{\text{amplitude}}, \mathbf{d}_{\text{Ppre-stack time}}, \mathbf{d}_{\text{gravity}}\}$.

Notice that seismic data subsets of different nature contribute at the right-hand side of the figure to the observed data and at the left-hand side to the prior information. The prior information is based on interpreted horizons in migrated and stacked data, whereas the data to be modeled at the right-hand side correspond to (1) seismic reflection amplitudes and (2) travel times of major events in pre-stack domain. There is no redundancy or cyclicity in the problem definition.

According to the layered model and the data relations in Figure 3.2, and applying the previous concepts for composition of the posterior density, we have

$$\begin{aligned} \sigma(\mathbf{m}) = & c \rho(\mathbf{m}_{\text{formation}}) \rho(\mathbf{m}_{\text{vshale}}, \mathbf{m}_{\text{porosity}}, \mathbf{m}_{\text{saturation}} | \mathbf{m}_{\text{formation}}) \\ & \times \rho(\mathbf{m}_{\text{Pvelocity}}, \mathbf{m}_{\text{Svelocity}}, \mathbf{m}_{\text{density}} | \mathbf{m}_{\text{vshale}}, \mathbf{m}_{\text{porosity}}, \mathbf{m}_{\text{saturation}}) \\ & \times L_{\text{amplitude}}(\mathbf{m}_{\text{Pvelocity}}, \mathbf{m}_{\text{Svelocity}}, \mathbf{m}_{\text{density}}) L_{\text{Ppre-stack time}} \\ & \times (\mathbf{m}_{\text{Pvelocity}}) L_{\text{gravity}}(\mathbf{m}_{\text{density}}). \end{aligned} \quad (3.11)$$

where, $\rho(\mathbf{m}_{\text{formation}})$ is the prior PDF on the formation sequence and delimiting horizon boundaries, $\rho(\mathbf{m}_{\text{vshale}}, \mathbf{m}_{\text{porosity}}, \mathbf{m}_{\text{saturation}} | \mathbf{m}_{\text{formation}})$ is the PDF of the rock matrix and fluid parameters conditioned by the formation, and $\rho(\mathbf{m}_{\text{Pvelocity}}, \mathbf{m}_{\text{Svelocity}}, \mathbf{m}_{\text{density}} | \mathbf{m}_{\text{vshale}}, \mathbf{m}_{\text{porosity}}, \mathbf{m}_{\text{saturation}})$ is the PDF of the physical rock properties conditioned by the rock matrix and fluid parameters. The likelihood functions are identified according to the set of observations and the related physical model argument. Figure 3.2 also shows some of the information needed for the definition of the priors, conditionals, and likelihoods, which is employed in the modeling as nonrandom parameters: the seismic source wavelets, the parameters of the rock physics model, and parameters of the conditional geostatistical models. These parameters will not vary in the inference and are previously estimated from the analysis of the data and additional information.

Similar components to the example shown in Figure 3.2 can be found in the papers by *Bosch* [2004], *Larsen et al.* [2006], *Bosch et al.* [2007], *Bosch et al.* [2009], *Grana and Della Rosa* [2010], and *Grana et al.* [2012] with details on how to model the specific prior, conditionals, and observational PDFs.

Considering now a larger scale, Figure 3.3a shows a setting for the inference of the geological structure at the crust, similar to the one employed by *Bosch* [1999], *Bosch et al.* [2001, 2004] and *Guillen et al.* [2007]. In this case, the primary frame is given by the description of the geometry of major lithotype geobody boundaries (e.g., gabbro, granite, sediments). The physical medium properties are modeled conditioned to the geobody lithotype by an empirical joint physical property density that is derived based on laboratory rock measurements for each lithotype. Finally, the observed data corresponds to common survey observations that provide information at large regional/crustal scales: gravity, magnetic, seismic refraction, and/or earthquake travel times. The resulting posterior density according to the model and data structure is given in the figure.

A similar parameter structure, shown in Figure 3.3b, was used at global satellite and planetary scale by *Khan et al.* [2006] to infer the thermal and compositional structure of the moon from available data on P-wave travel times, gravity observations, and inertia moment. In this case the primary parameters were the temperature and composition, and the secondary parameters were the seismic compressional velocity and mass density. A petrological model, based on computations of mineral phase proportions in the mantle, was used for the prediction of the mass density and seismic velocity conditioned to the mantle composition and temperature. The same approach was applied, with differences in the constraining geophysical data, to infer the composition and temperature of Mars [*Khan and Connolly*, 2008] and the Earth's mantle [*Khan et al.*, 2008].

3.2.4. The Role of Rock Physics and Dynamic Models as Coupling Information

In the examples presented above, an important role is given to relationships across model components, which are described by inner model conditional PDFs. Multiple properties defined at the same points are naturally linked. In solid Earth models, relationships are imposed by rock physics, but also geology, sedimentology, mineralogy, and chemistry can provide relational information depending on the setting.

Approaches to model the conditionals between the physical rock properties (e.g., elastic moduli, seismic velocities, mass density, viscosity, electrical resistivity) from basic rock frame constitution and fluids (e.g., matrix lithology, porosity, fluid types, and fractions) are multiple. I will refer below to empirical and rock physics model-based approaches.

An *empirical approach* to the formulation of the conditional of physical medium parameters to lithotype categories can be illustrated, for example, by the work of *Bosch* [1999] and *Bosch et al.* [2001]. The characterization of the mass density and magnetic susceptibility was based on laboratory rock sample measurement data for each one of the involved geobody lithotypes of the studied area. An empirical spatial (geostatistical) simulation model was elaborated for the conditional $\rho(\mathbf{m}_{\text{density}}, \mathbf{m}_{\text{susceptibility}} | \mathbf{m}_{\text{litho}})$ by using mixtures of multivariate Gaussian functions. The PDF for the mass density and magnetic susceptibility in a node depended on the lithotype of the geobody (categorical variable) and the mass density and magnetic susceptibility values at the other nodes within the geobody, according to the mentioned model. Because of the scarcity of data in some applications, the covariance (or equivalently semivariogram) ranges need to be assumed; in the work by *Bosch et al.* [2001] it was based on the geostatistical characterization of similar areas from field measurements [*Bourne*, 1993]. Additional examples of the application of an empirical approach to formulate probabilities of the physical rock properties, conditioned to lithology and reservoir properties, are described in the works by *Mukerji et al.* [2001], *Larsen et al.* [2006], and *Ulymoen et al.* [2010].

Relationships between the elastic moduli, mass density, and other physical rock properties have been studied for various types of rocks and Earth media, within the domain of rock physics. Common models of rock physics for relating acoustic and elastic properties to rock matrix and fluid components are described in detail by *Mavko et al.* [2003] and *Hilterman* [2001]. For sedimentary rocks, and particularly for the most common siliciclastic sedimentary rocks, a large set of modeling tools are available. Predictive models for elastic and other physical properties for mantle material have also been studied. A second type of approach, less dependent on a specific set of data

than the empirical approach, consists in using appropriate rock physics models for the prediction of the physical properties from rock matrix and fluid properties,

$$\mathbf{m}_{\text{phys}} = \mathbf{f}(\mathbf{m}_{\text{rock}}) + \Delta\mathbf{m}_{\text{phys}}. \quad (3.12)$$

where, $\mathbf{f}(\mathbf{m}_{\text{rock}})$ is a rock physics function that models the involved physical rock properties and $\Delta\mathbf{m}_{\text{phys}}$ are the deviations from the prediction. It is important to mention that no rock physics model has universal validity. Hence, all rock physics models should be evaluated and calibrated against actual property data from the application area. The statistics for the deviations $\Delta\mathbf{m}_{\text{phys}}$ can be characterized by comparing the rock physics model prediction and property measurements for the area (usually well-log or core data). The statistical model for the deviations,

$$\rho(\Delta\mathbf{m}_{\text{phys}}) = \rho(\mathbf{f}(\mathbf{m}_{\text{rock}}) - \mathbf{m}_{\text{phys}}) = \rho(\mathbf{m}_{\text{phys}} | \mathbf{m}_{\text{rock}}), \quad (3.13)$$

conduce straightforwardly to the conditional PDF for the physical model parameters. In addition to the rock physics model deviations, measurement uncertainties could also be accounted for in $\rho(\Delta\mathbf{m}_{\text{phys}})$ when relevant.

Examples of the *rock physics-based* approach outlined above for modeling the dependency of acoustic and elastic properties from rock matrix and fluid properties are described by *Mavko and Mukerji* [1998], *Bosch* [2004], *Bosch et al.* [2007], *Spikes et al.* [2007], *Bosch et al.* [2009], *Grana et al.* [2012], *Suman and Mukerji* [2013], and *Grana* [2014] for modeling the dependency of acoustic and elastic properties from rock matrix and fluid properties.

At the planetary scale, physical medium properties, such as seismic velocities and density, are a function of the mantle composition and temperature. The problem is nonlinear due to mineral phase changes. Nevertheless, it is fully tractable via petrophysical models, such as the one described by *Connolly* [2005] that is based on the minimization of the free energy associated with the mixture of mantle minerals. This approach was followed by *Khan et al.* [2006] and *Khan and Connolly* [2008] to model petrophysical conditionals for inferring the thermal and compositional configuration of the moon and Mars. Also, *Hacker et al.* [2003] elaborated a model for the prediction of the compressional velocities and mass density in the Earth mantle.

When modeling phenomena evolving in time, dynamic models are the natural link between the various time-lapse observations or velocity observations. In the case of the mantle description, flow equations can be useful to link temperature fields and mantle kinematics. In the case of time-lapse seismic in reservoirs under production, fluid flow modeling can be also used as inner link between the time-lapse configurations. Applications under various approaches are shown in the papers by *Huang* [2001], *Mezghani et al.* [2004], and *Dadashpour et al.* [2009].

In addition to rock physics and dynamic models, structural features such as the location of geologic boundaries [Lines et al., 1988; Haber and Oldenburg, 1997; Bosch et al., 2001; Gallardo et al., 2003; Guillen et al., 2007] or the requirement of similar directions for property spatial gradients [Gallardo and Meju, 2004; Doetsch et al., 2010] are also used for model conditioning across medium property fields.

3.3. GRAPHS AND POSTERIOR PROBABILITY DENSITIES

As shown in the previous section, the information about parameter components and dependencies are easily presented in graphical form, as in Figures 3.1–3.3, facilitating a straightforward definition of the posterior probability densities. So far we have presented examples with hierarchical model parameters layers and a final data layer: Priors are given for the first model layer, conditionals at the intermediate model layers and data likelihoods at the final layer. However, more heterogeneous networks of model and data subspaces can be analyzed for inference following the underlying principles.

A graphical structure known as *direct acyclic graph* (DAG) [Thulasiraman and Swamy, 1992] is useful to describe relationships across model parameters that are less structured than a hierarchical sequence of model layers.

The DAG is defined by a set of nodes, which are here the model and data components, and a set of directed arrows that link the nodes, which will represent direct dependency relationships. In the DAG, it is required that no closed directed path exists in the graph. If a parameter subspace \mathbf{m}_n points in the graph to a subspace \mathbf{m}_k , the latter is considered a *descendant* of the former, and the former an *ascendant* of the latter. Acyclicity warrants that no node can be its own descendant or ascendant; Figure 3.4 shows an example of a DAG relating model and data components. Notice that we have data that are dependent on different model component generations and direct influences (arrows) across model components separated by more than one generation; data nodes do not have descendants. A given DAG and the PDF defined over the joint parameter space defines what is called a *Bayesian network*, sometimes also called belief network or simply inference network [Pearl, 1986, 1994; Ben-Gal, 2007; Griffiths et al., 2008].

The same principles applied in (3.6) and (3.9) produce the following rules for the factorization of the posterior PDF over the DAG:

1. Each model node with no ascendants introduces a prior PDF factor.
2. Each arrow across model nodes contributes with a conditional PDF factor.
3. Each data node contributes with a likelihood function factor.

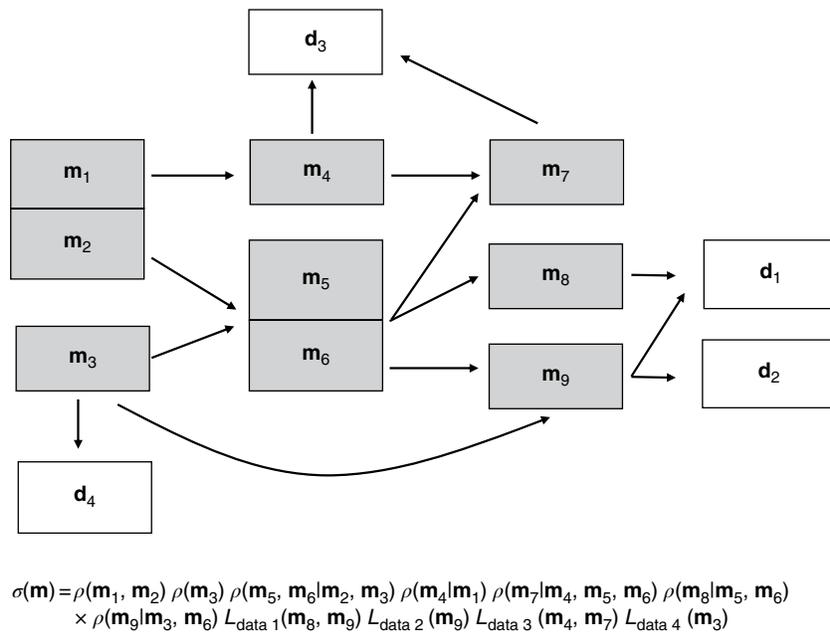


Figure 3.4 Example of an inference network defined over a direct acyclic graph (DAG). Dependencies across random variables are shown by bold arrows. The corresponding structure of the posterior PDF is shown at the bottom of the figure. Model components within common boxes are jointly modeled. Gray boxes indicate model parameters describing the earth medium properties and structure, while white boxes indicate model parameters describing experimental observations and measurements.

According to these rules, the posterior PDF for the DAG in Figure 3.4 is

$$\begin{aligned} \sigma(\mathbf{m}) = & c\rho(\mathbf{m}_1, \mathbf{m}_2) \rho(\mathbf{m}_3) \rho(\mathbf{m}_5, \mathbf{m}_6 | \mathbf{m}_2, \mathbf{m}_3) \\ & \times \rho(\mathbf{m}_4 | \mathbf{m}_1) \rho(\mathbf{m}_7 | \mathbf{m}_4, \mathbf{m}_5, \mathbf{m}_6) \rho(\mathbf{m}_8 | \mathbf{m}_5, \mathbf{m}_6) \\ & \times \rho(\mathbf{m}_9 | \mathbf{m}_3, \mathbf{m}_6) L_{\text{data1}}(\mathbf{m}_8, \mathbf{m}_9) L_{\text{data2}}(\mathbf{m}_9) \\ & \times L_{\text{data3}}(\mathbf{m}_4, \mathbf{m}_7) L_{\text{data4}}(\mathbf{m}_3). \end{aligned} \quad (3.14)$$

The sense of some of the DAG relations depends on the modeling decisions made. When the relationships are based on theoretical models, the nature of the theory commonly imposes the sense of the direct (easier) modeling, complying with a causality notion. Examples of Bayesian model networks used for oil reservoir description, with illustration of the specific graph structure defined, are described in the works by *Eidsvik et al.* [2004],

Bosch et al. [2007], *Rimstad et al.* [2012] and *Chen and Hoversten* [2012]. Examples of similar networks applied to decision making can be found in the work by *Bhattacharjya and Mukerji* [2006] and *Martinelli et al.* [2013].

Figure 3.5 shows an application example of the inference with a DAG relational description across model and data components, for sedimentary strata description. In the setting of Figure 3.5, the formational random parameters of Figure 3.2 are simplified to be a known (known horizons and formations in reflection seismic time) part of the prior information. Additional data have been included consisting of well-log observations in given locations (well paths) for the porosity, shale fraction, water saturation, and elastic medium parameters (P-wave and S-wave velocities and mass density). The seismic source wavelet, considered a nonrandom parameter in Figure 3.3, has been randomized in order to adjust the seismic source

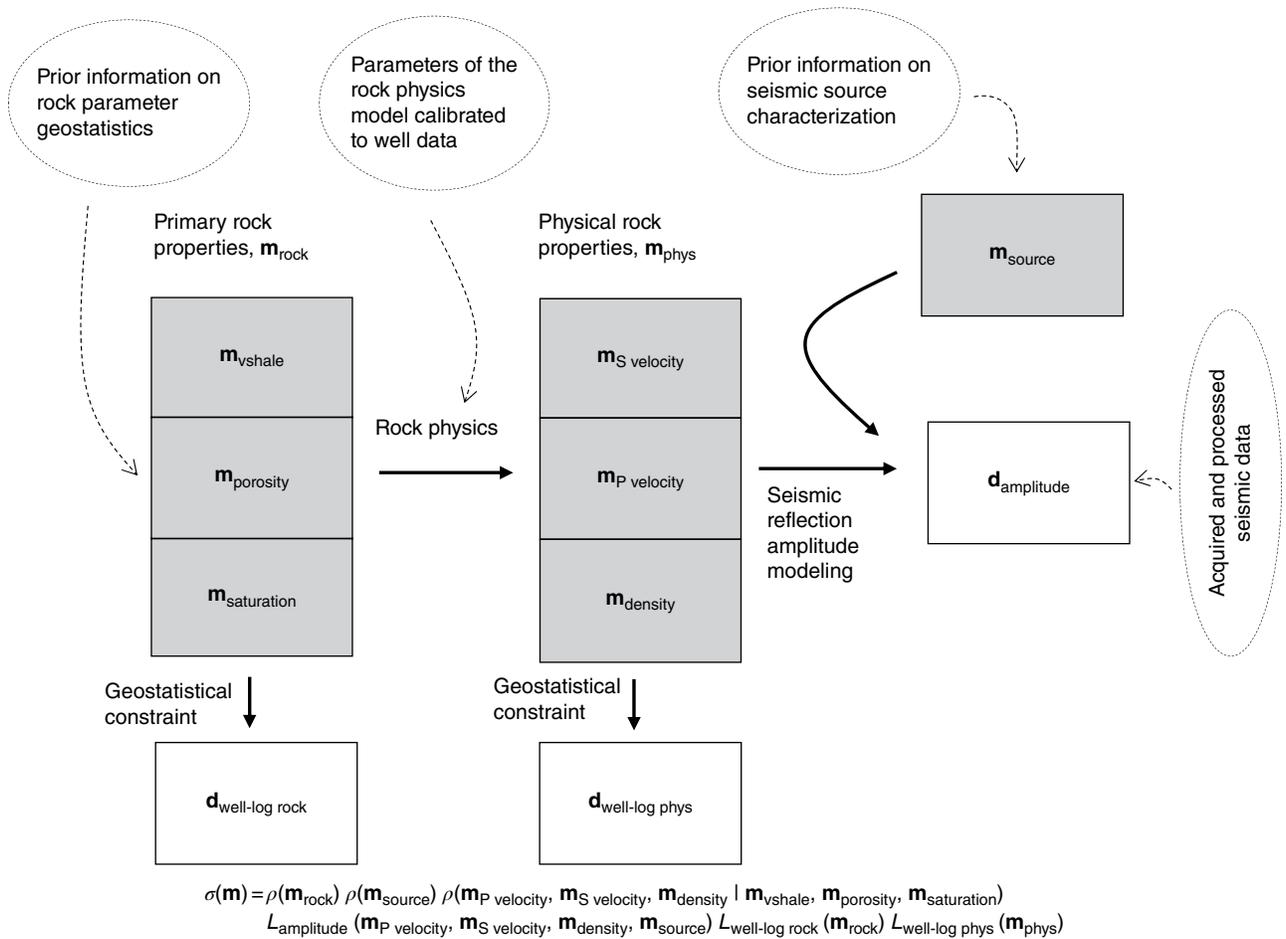


Figure 3.5 Inference network for siliciclastic sedimentary basin description based on pre-stack seismic reflection data, including the estimation of the source wavelet and well-log data priors on the medium properties. Symbols are the same as in previous figures. The corresponding structure of the posterior PDF is shown at the bottom.

wavelet within the inference process. With the relationships shown in the figure, the posterior PDF is

$$\begin{aligned} \sigma(\mathbf{m}) = & c \rho(\mathbf{m}_{\text{rock}}) \rho(\mathbf{m}_{\text{source}}) \rho(\mathbf{m}_{\text{phys}} | \mathbf{m}_{\text{rock}}) \\ & \times L_{\text{amplitude}}(\mathbf{m}_{\text{phys}}, \mathbf{m}_{\text{source}}) L_{\text{well-log rock}}(\mathbf{m}_{\text{rock}}) \\ & \times L_{\text{well-log phys}}(\mathbf{m}_{\text{phys}}), \end{aligned} \quad (3.15)$$

where $\rho(\mathbf{m}_{\text{source}})$ is the prior information on the source wavelet, usually obtained by preliminary well to seismic tie, and $\rho(\mathbf{m}_{\text{rock}})$ is the prior information on the rock matrix and fluid parameters (total porosity, shale fraction, and water saturation). The seismic amplitude likelihood, $L_{\text{amplitude}}(\mathbf{m}_{\text{phys}}, \mathbf{m}_{\text{source}})$, is now dependent on the source wavelet in addition to the elastic medium parameters, and there are likelihood functions, $L_{\text{well-log rock}}(\mathbf{m}_{\text{rock}})$ and $L_{\text{well-log phys}}(\mathbf{m}_{\text{phys}})$, corresponding to the well-log measurements of the rock and physical properties at specific well-path locations. More details about the source

wavelet inference are given by *Bosch et al.* [2007], and for well-log data inclusion see *Bosch et al.* [2009].

Figure 3.6 shows another inference network defined over a DAG. I am showing in this figure a network proposal for the inference of mantle properties and dynamics, by coupling seismic tomography, gravity, and plate velocity observations. In this case, dynamic models for the mantle and plates are part of the inner coupling of the model components. Primary parameters are the mineralogical mantle composition, together with the temperature and pressure. A descendant set of parameters are the compressional seismic velocity, the shear velocity, the mass density, and the viscosity, which should be modeled by composing mineral fractions, temperature, and pressure dependencies. Another set of descendant parameters are the mantle dynamics given by stress and velocities, dependent on the mantle configuration and physical properties. Mantle velocity imposes anisotropy in the

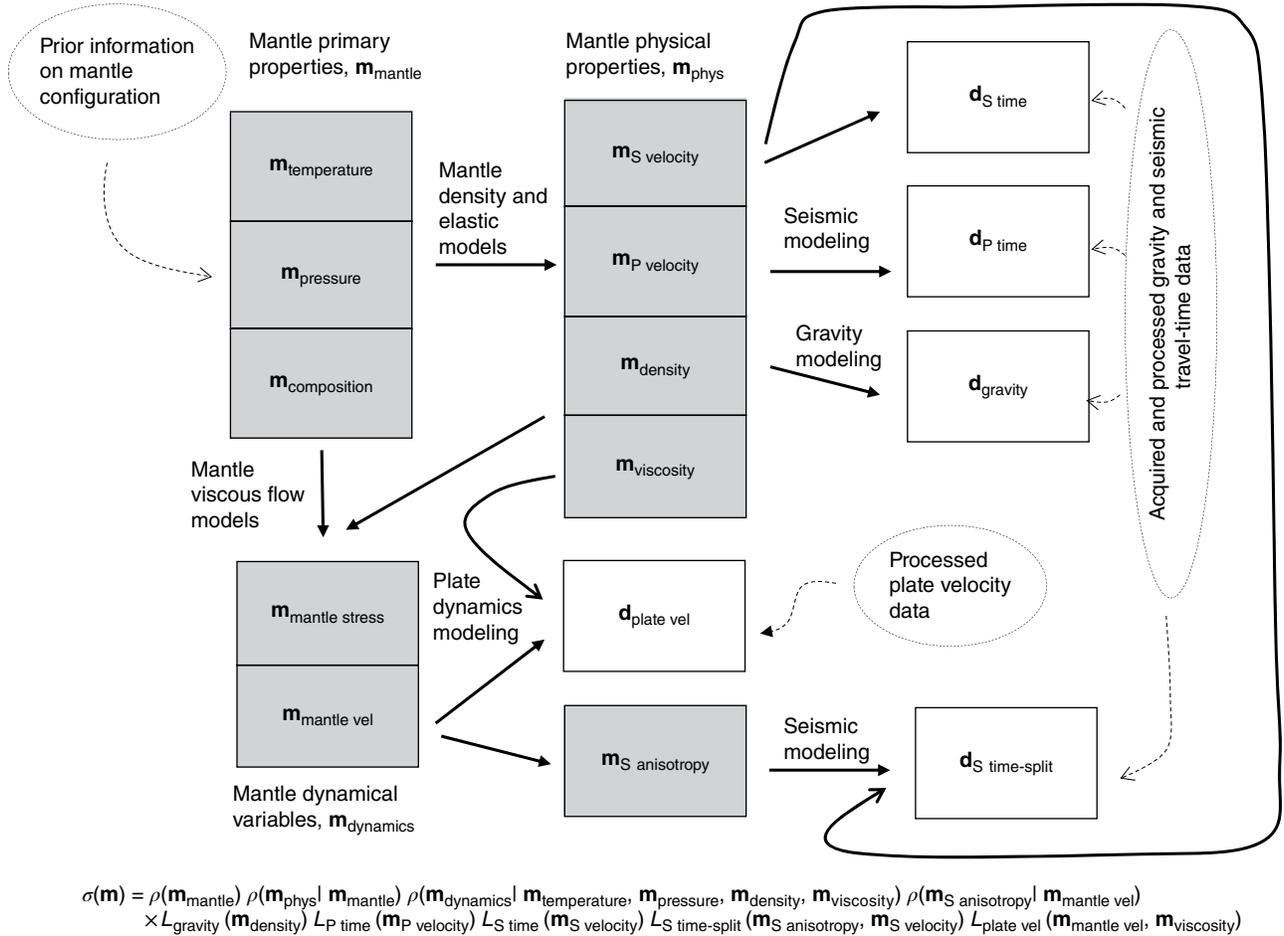


Figure 3.6 Example of a proposed inference network for mantle configuration coupling seismic tomography, mantle flow, plate dynamics, and gravity data constraints. The corresponding structure of the posterior PDF is shown at the bottom; compact notation $\mathbf{m}_{\text{mantle}}$, \mathbf{m}_{phys} , and $\mathbf{m}_{\text{dynamics}}$ have been used to abbreviate the expression. Same symbols as in previous figures.

propagation of seismic shear waves (birefringence phenomenon), which is measurable from seismological observations. Mantle velocities at the surface are also related to the plate kinematics, through dynamic plate models. The observed data includes travel-time seismological observations for P and S phases, for the travel-time split of the S phase due to anisotropy, plate measured velocities, and gravity data.

3.4 SAMPLING IN MODEL NETWORKS

Once the posterior PDF has been formulated in a model network, the solution of the inverse problem consists in drawing realizations from the posterior PDF, or alternatively solving for maximum posterior model configurations. We will discuss the first option in this section.

The structure of the model components, data components, and relationships is established via the factors of the combined PDF and, as explained before, satisfactorily depicted by the associated model network graph. The straightforward approach to sample from the posterior PDF is to construct a Markov Chain sampler following the sequence imposed by the graph, from the ascendants to the descendants. There are many techniques to sample from the priors, conditionals, and posterior PDFs [Geyer, 1992; Smith and Roberts, 1993; Tierney, 1994; Gelfand and Smith, 1990; Liu, 1998]. The procedure I recommend here is sufficiently general to successfully adapt to any model network, and it is efficient for sampling. The procedure can be separated in two major phases for (a) sampling the joint prior PDF and (b) sampling the joint posterior PDF.

3.4.1. Sampling the Joint Prior PDF

The prior PDF is an important factor of the posterior PDF and is equal to the posterior if the likelihood data factors are ignored (no observations or infinite data uncertainties). To sample a realization from the joint prior PDF:

1. Draw a realization from each one of the component priors included in the PDF.
2. According to the realization of the ascendant model parameters, draw realizations of the first generation of descendants according to the conditional PDFs.
3. Continue the procedure through all the *generations* until the last descendant has been realized. Recall that to generate a descendant realization, all its ascendants should be realized, as warranted by the acyclic graph configuration.

The appropriate technique to produce prior and conditional realizations varies according to the nature of each PDF. In the case of continuous variables, often the PDF can be formulated as a multivariate Gaussian function. In this case, standard sampling methods are well known, like

the square root of the covariance matrix method or Gibbs sampling through Gaussian conditionals. Some parameters are not Gaussian distributed, but may be transformed to Gaussian after appropriate change of variables. For the case of categorical multivariate parameters or non-Gaussian continuous parameters, Gibbs sampling from the univariate conditionals is often convenient.

3.4.2. Sampling the Joint Posterior PDF

Likelihood function evaluations corresponding to geophysical observations are commonly costly (in terms of computation times), difficult to calculate (in terms of elaborated numerical nonlinear computations), and not represented by parameterized continuous PDF models. The Metropolis–Hastings sampler is an appropriate technique to account for this type of likelihood. In the inferential setting, this sampler uses as candidate outcome a realization of the prior PDF and proceeds by accepting or rejecting the realization by testing the likelihood function ratio between the candidate and the current realization in the chain. The recommended procedure is as follows:

1. Generate a candidate realization following the joint prior sampling chain rules described in the previous subsection.
2. Evaluate the joint data likelihood for this realization.
3. Calculate the joint data likelihood ratio between the candidate realization and the current realization.
4. Accept the candidate realization as the next step of the posterior chain with probability equal to the minimum between the likelihood ratio and one.
5. If the candidate is rejected, assign the current realization in the posterior chain.
6. Iterate the procedure from the first step.

The Metropolis sampler warrants the convergence of the chain to a sample of the posterior PDF in long enough runs. A description of the Metropolis sampler applied to posterior PDFs in geophysical inverse problems can be found in the work by Mosegaard and Tarantola [1995].

Likelihoods associated with property sampling in specific locations (e.g., well-logging or surface rock sampling) are less difficult to evaluate and can be in most cases related to geostatistical Gaussian spatial PDFs (Kriging and Gaussian simulation). These likelihoods may be either (a) included within the likelihood evaluated by the Metropolis sampler or (b) used as additional modeling constraint to the prior information chain as shown by Bosch *et al.* [2009]. For a review in statistical spatial models the reader is referred to the works by Dubrule [2003], Chiles and Delfiner [2009] and Deutsch and Journel [1992]. Multipoint statistics [Caers *et al.*, 2000; Strebelle, 2002] allows for prior PDF sampling with improved description of morphological features. Examples of their application to inversion of seismic data in complex models

are described in the works by *González et al.* [2008] and *Grana et al.* [2012]. Object oriented modeling of fluvial sedimentary systems have been demonstrated by *Holden et al.* [1998] and *Deutsch and Wang* [1996].

3.4.3. Efficient Sampling Through Factorized Likelihoods

When several data likelihoods are present, the Metropolis sampler may be applied in cascade to each one of the likelihood factors; using partial posteriors as prior sampling PDFs for the consecutive data likelihood, as shown by *Bosch et al.* [2000] and *Tarantola* [2005]. The procedure is as follows:

1. Generate a candidate realization following the joint prior sampling chain procedure.
2. Evaluate one data likelihood ratio between the candidate and the current realization.
3. Retain the candidate realization for the next likelihood factor test with probability equal to the minimum between the likelihood ratio and one.
4. If the candidate is not retained, accept the current realization and go to 1.
5. If the candidate is retained, repeat from 2 following with the next likelihood factor. If retained after the last likelihood factor test, accept the model realization and go to 1.

Two criteria should be used when ordering the likelihood factors, leaving with preference at the beginning: (1) smoother likelihoods in terms of information (larger uncertainties) and (2) likelihoods with smaller computational cost. The former condition (smoothness) allows avoiding unwanted barrier problems (i.e., inability to mix the sampling across modes separated by very low probability zones) in the preliminary likelihood evaluations, which can potentially affect the efficiency of the method. In any case, the efficiency between various likelihood sequences should be tested for evaluation and should be compared with the option of single joint likelihood evaluation.

3.5. MAXIMUM POSTERIOR PROBABILITIES IN MODEL NETWORKS

Another alternative in realizing a solution to the inverse problem is searching for the configuration that maximizes the posterior probability density (MAP) and calculating the local posterior covariance matrix. The MAP search commonly converges to the nearest mode, although there are methods to search for the global MAP. In the case of multimodal PDFs a single MAP configuration is not a complete description of the problem solution; identification of major modes and the corresponding local MAP configuration may be an alternative. It is common in geophysical inference that the data likelihoods are complex multimodal functions. Nevertheless, if the prior

information is monomodal and highly informative, and such that it circumscribes the posterior into the region of one of the modes of the data likelihood, the posterior is then close to monomodal and the MAP constitutes an acceptable description of the problem solution.

A classic method for searching a MAP configuration is the Gauss–Newton method [*Tarantola*, 2005], which requires the gradient and the approximate Hessian of the natural logarithm of the posterior PDF. By defining the *objective function*, $S(\mathbf{m}) = -\ln(\sigma(\mathbf{m}))$, we obtain

$$\sigma(\mathbf{m}) = \exp(-S(\mathbf{m})). \quad (3.16)$$

As the exponential is a positive monotonically increasing function, a MAP configuration corresponds to a minimal value of the objective function, and, neglecting third-order derivatives, the posterior local covariance matrix is the Hessian of the objective function evaluated at the MAP. The model parameters update, $\Delta\mathbf{m} = \mathbf{m}^{n+1} - \mathbf{m}^n$, for a step $n+1$ in the iterative search towards the \mathbf{m}_{MAP} satisfies

$$\mathbf{Hess}[S(\mathbf{m}^n)]\Delta\mathbf{m} = -\mathbf{Grad}[S(\mathbf{m}^n)], \quad (3.17)$$

where **Hess** symbolizes the Hessian operator and **Grad** the gradient operator. If multiplying by the prior model covariance matrix, $\mathbf{C}_m^{\text{prior}}$, the linear system matrix gets dimensionless and more stable for the numerical solution,

$$\underbrace{\mathbf{C}_m^{\text{prior}} \mathbf{Hess}[S(\mathbf{m}^n)]}_{\text{Curvature}} \Delta\mathbf{m} = -\underbrace{\mathbf{C}_m^{\text{prior}} \mathbf{Grad}[S(\mathbf{m}^n)]}_{\text{Steepest descent direction}}. \quad (3.18)$$

Notice that because the joint objective function is a logarithm of the posterior PDF, the factor structure in (3.10) transforms straightforwardly to the addition of objective function terms, each one accounting for the corresponding data likelihood, conditional probability, or prior probability.

The posterior model covariance is the inverse of the Hessian of the objective function, evaluated at the MAP configuration. It can be calculated by inverting the Hessian, but inverting the Curvature matrix (i.e., the product of the prior model covariance matrix and the Hessian of the objective function) and multiplying by the prior covariance is commonly a more stable procedure,

$$\mathbf{C}_m^{\text{posterior}} = \left(\mathbf{C}_m^{\text{prior}} \mathbf{Hess}[S(\mathbf{m}_{\text{MAP}})] \right)^{-1} \mathbf{C}_m^{\text{prior}}. \quad (3.19)$$

To work out the linear equations to search the MAP configuration, the data likelihood, the conditional PDFs, and priors need to be explicitly formulated. I will use a simple setting of a two-layered model and one data layer, as shown in Figure 3.7, to illustrate the formulation of the linear system of equations (3.18). It corresponds to the case of inverting pre-stack seismic data for joint estimation

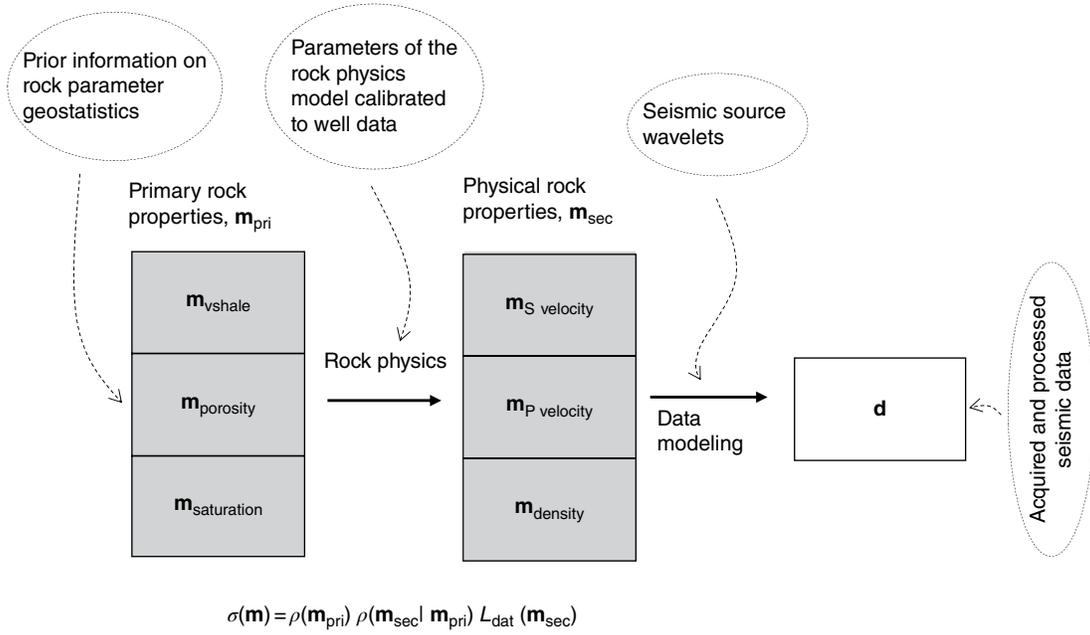


Figure 3.7 Inference network for a hierarchical two-layered model structure, illustrated with the case of a siliciclastic sedimentary basin description based on seismic data. Symbols are the same as in previous figures. The corresponding structure of the posterior PDF is shown at the bottom.

of the isotropic elastic medium parameters, and primary rock parameters in a siliciclastic sedimentary medium, describing the total porosity, the shale factor and pore fluid phase (water–hydrocarbon) fraction. The formulation is the same for any posterior PDF with a similar structure. For describing the isotropic medium, various combinations of three parameters (elastic moduli, impedances, mass density, seismic velocities) can be selected; a common choice is the mass density and the seismic P and S velocities. To be explicit, I will describe the medium by specifying primary, $\mathbf{m}_{\text{pri}} = \{\mathbf{m}_{\text{vshale}}, \mathbf{m}_{\text{porosity}}, \mathbf{m}_{\text{saturation}}\}$, and secondary, $\mathbf{m}_{\text{sec}} = \{\mathbf{m}_{\text{P velocity}}, \mathbf{m}_{\text{S velocity}}, \mathbf{m}_{\text{density}}\}$, model parameters, the joint model space being $\mathbf{m} = \{\mathbf{m}_{\text{pri}}, \mathbf{m}_{\text{sec}}\}$. Commonly, these parameters are specified at each point over a 3D grid.

The seismic data depends explicitly on the elastic medium configuration and parameters associated with the seismic survey experiment (source function and geometry). The seismic observations in this problem could be travel times (tomography problems), reflection amplitudes (reflectivity inversion), or the full wave field. In any of these cases, we can formulate a forward modeling of the data, $\mathbf{g}(\mathbf{m}_{\text{sec}})$, based on the seismic wave mechanical theory, such that

$$\mathbf{d}_{\text{obs}} = \mathbf{g}(\mathbf{m}_{\text{sec}}) + \Delta \mathbf{d}, \quad (3.20)$$

where \mathbf{d}_{obs} is the observed data, and $\Delta \mathbf{d}$ are the deviations between the observed and the modeled data.

As already explained, rock physics is the natural link between the primary and secondary model parameters. After calibrating an appropriate rock physics model for the elastic parameters using local data if available, $\mathbf{f}(\mathbf{m}_{\text{pri}})$, we have

$$\mathbf{m}_{\text{sec}} = \mathbf{f}(\mathbf{m}_{\text{pri}}) + \Delta \mathbf{m}_{\text{sec}}, \quad (3.21)$$

where \mathbf{m}_{sec} are the true medium elastic parameters and $\Delta \mathbf{m}_{\text{sec}}$ are the deviations between the true and modeled elastic parameters.

The posterior PDF in this problem according to what has been previously explained is

$$\sigma(\mathbf{m}) = \rho(\mathbf{m}_{\text{pri}}) \rho(\mathbf{m}_{\text{sec}} | \mathbf{m}_{\text{pri}}) L_{\text{dat}}(\mathbf{m}_{\text{sec}}). \quad (3.22)$$

I will model the three factors in the posterior formulation by multivariate Gaussian functions. The first PDF describes the prior information on the primary model parameters, which could be defined with

$$\rho(\mathbf{m}_{\text{pri}}) = c_1 \exp \left[-\frac{1}{2} (\mathbf{m}_{\text{pri}} - \mathbf{m}_{\text{pri}}^{\text{prior}})^T \mathbf{C}_{\text{pri}}^{-1} (\mathbf{m}_{\text{pri}} - \mathbf{m}_{\text{pri}}^{\text{prior}}) \right], \quad (3.23)$$

with $\mathbf{m}_{\text{pri}}^{\text{prior}}$ being the expected prior primary parameters and \mathbf{C}_{pri} the prior covariance matrix; commonly, $\mathbf{m}_{\text{pri}}^{\text{prior}}$ is spatially dependent and previously modeled according to the geological stratification and formation horizon information. The likelihood function and the conditional rock

physics PDF are formulated by modeling the probability of the deviations $\Delta \mathbf{m}_{\text{sec}}$ and $\Delta \mathbf{d}$ in (3.20) and (3.21),

$$\begin{aligned} \rho(\mathbf{m}_{\text{sec}} | \mathbf{m}_{\text{prim}}) \\ = c_2 \exp \left[-\frac{1}{2} (\mathbf{m}_{\text{sec}} - \mathbf{f}(\mathbf{m}_{\text{prim}}))^T \mathbf{C}_{\text{seclprim}}^{-1} (\mathbf{m}_{\text{sec}} - \mathbf{f}(\mathbf{m}_{\text{prim}})) \right], \end{aligned} \quad (3.24)$$

with $\mathbf{C}_{\text{seclprim}}$ being the covariance matrix of the rock physics model deviations $\Delta \mathbf{m}_{\text{sec}}$. Similarly, the data likelihood function is

$$L_{\text{dat}}(\mathbf{m}_{\text{elas}}) = \exp \left[-\frac{1}{2} (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}_{\text{sec}}))^T \mathbf{C}_{\text{dat}}^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}_{\text{sec}})) \right], \quad (3.25)$$

with \mathbf{C}_{dat} being the data covariance, encompassing data observational and modeling uncertainties. By adding exponents of the three modeled factors of the posterior PDF, the full objective functions has each of the three information components,

$$\begin{aligned} S(\mathbf{m}) = & \frac{1}{2} \left[\underbrace{(\mathbf{m}_{\text{prim}} - \mathbf{m}_{\text{prim}}^{\text{prior}})^T \mathbf{C}_{\text{prim}}^{-1} (\mathbf{m}_{\text{prim}} - \mathbf{m}_{\text{prim}}^{\text{prior}})}_{\text{Prior information term}} \right] \\ & + \frac{1}{2} \left[\underbrace{(\mathbf{m}_{\text{sec}} - \mathbf{f}(\mathbf{m}_{\text{prim}}))^T \mathbf{C}_{\text{seclprim}}^{-1} (\mathbf{m}_{\text{sec}} - \mathbf{f}(\mathbf{m}_{\text{prim}}))}_{\text{Rock physics term}} \right] \\ & + \frac{1}{2} \left[\underbrace{(\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}_{\text{sec}}))^T \mathbf{C}_{\text{dat}}^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}_{\text{sec}}))}_{\text{Geophysical term}} \right]. \end{aligned} \quad (3.26)$$

Expressions $\mathbf{g}(\mathbf{m}_{\text{sec}})$ and $\mathbf{f}(\mathbf{m}_{\text{prim}})$ are in general nonlinear, and hence the search of the model update needs successive iterations in the application of the Gauss–Newton’s method described in expression (3.15). The first and second derivatives of the objective function, as well as algebraic simplifications, required to calculate the model update are detailed in the work by *Bosch* [2004]. The resulting model update for the secondary model parameters satisfies the linear system of equations,

$$\mathbf{A} \Delta \mathbf{m}_{\text{sec}} = \mathbf{b}, \quad (3.27)$$

with the left-hand side being

$$\mathbf{A} = \mathbf{I} + (\mathbf{C}_{\text{seclprim}} + \mathbf{F} \mathbf{C}_{\text{prim}} \mathbf{F}^T) \mathbf{G}^T \mathbf{C}_{\text{dat}}^{-1} \mathbf{G}. \quad (3.28)$$

Above, matrices \mathbf{G} and \mathbf{F} are the Jacobian matrices of $\mathbf{g}(\mathbf{m}_{\text{sec}})$ and $\mathbf{f}(\mathbf{m}_{\text{prim}})$ correspondingly, and \mathbf{I} is the identity matrix. The right-hand side of (27) is

$$\begin{aligned} \mathbf{b} = & \mathbf{f}(\mathbf{m}_{\text{prim}}) - \mathbf{m}_{\text{sec}} + \mathbf{F}(\mathbf{m}_{\text{prim}}^{\text{prior}} - \mathbf{m}_{\text{prim}}) \\ & + (\mathbf{C}_{\text{seclprim}} + \mathbf{F} \mathbf{C}_{\text{prim}} \mathbf{F}^T) \mathbf{G}^T \mathbf{C}_{\text{dat}}^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}_{\text{sec}})). \end{aligned} \quad (3.29)$$

After the model update for the secondary parameters is calculated, the model update for the primary parameters is obtained by [*Bosch*, 2004]

$$\begin{aligned} \Delta \mathbf{m}_{\text{prim}} = & \mathbf{m}_{\text{prim}}^{\text{prior}} - \mathbf{m}_{\text{prim}} \\ & + \mathbf{C}_{\text{prim}} \mathbf{F}^T \mathbf{G}^T \mathbf{C}_{\text{dat}}^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}_{\text{sec}}) - \mathbf{G} \Delta \mathbf{m}_{\text{sec}}). \end{aligned} \quad (3.30)$$

Once (3.27) and (3.30) are calculated, the model parameters are jointly updated and a new evaluation of (3.27)–(3.30) can be obtained by iterating towards proximity of the MAP configuration. Commonly, convergence monitoring the data residual evolution until substantial reduction and stabilization.

3.6. DISCUSSION

Once the corresponding posterior probability density is defined according to the models involved, the solution of the inverse problem relies in generating object outcomes that summarize the posterior information—that is, the combination of the various data likelihoods, conditionals, and prior densities. Two major approaches have been described here: (1) sampling object configurations in proportion to the posterior PDF and (2) searching maximum posterior probability object configurations. A comparison between the two options is a common subject of discussion.

A first point to mention is that the two approaches do not provide the same description of the posterior information. The sampling approach provides a complete description of the posterior PDF. Theoretically, with a large enough set of samples of the model parameters, all marginals, conditionals, and the joint posterior PDFs can be approximated by the sample statistics with arbitrarily small deviations. Also, expected values, standard deviations and frequency histograms of the modeled object parameters can be computed. Hence, the sampling approach produces a full solution of the inverse problem. The optimization approach, on the other hand, searches for the mode (the local MAP configurations) that is the closest to a starting search point. It is essentially local, as the model parameter space can be always divided in a series of local mode sectors. In the neighborhood of the mode the posterior covariance matrix can be calculated to describe the uncertainties and posterior correlations around the mode.

Another known limitation of the MAP configuration is that it is usually smooth in physical (3D plus time) space, because all the spatially distributed parameters align with their posterior expected values. Realizations produced from the sampling process better represent the true spatial variability of the parameters, and they are often more useful when a representation of the spatial heterogeneity is relevant. A typical case is for modeling

the fluid flow in a permeable medium, where the driving flow locations correspond to the large permeability channels and not to the average permeability [Dubrule, 2003]. The MAP configuration commonly underestimates by large the fluid flow. This limitation can partly be overcome by producing posterior simulations centered at the MAP with superposed deviations generated according to the posterior covariance matrix.

The computational costs involved in the optimization and the sampling approaches are highly dependent on the specific case and objectives of the inference; a comparison requires a case-by-case analysis. In general, the number of iterations involved in the search of the MAP is much smaller than the number of model parameters (a few iterations), whereas sampling chains require a length (number of realizations) of several times the number of model parameters. However, the operations required in the computation of a single iteration for the optimization approach are much larger than the computations involved in generating one sampling step, and they increase faster with the number of model parameters. Direct matrix methods solving linear systems as expression (3.17) typically require operations in the order between the square and the cube of the number of model parameters. Nevertheless, methods that take advantage of the sparsity and/or structure of the system (3.17), or approximate iterative solvers such as the conjugate gradients method, decrease this dependency, commonly reaching performances with operations orders beneath the square of the number of model parameters. Such numerical methods are required for the solution of large to very large inverse problems with the optimization approach.

Another issue of discussion refers to the various spaces associated with the inference: model parameters, modeled object, modeled observations, and data parameters. Implicit applications link the model parameter space to the model object space, and they also link the modeled observation space to the data space. These applications are not commonly explicit in the general formulation of the inference problems, but need to be accounted to model the involved functions and PDFs.

The formulation presented herein relates data and model parameters with basis on a combination of multiple modeling processes, described by conditional probability densities and likelihood functions. The basic knowledge to establish these links across model components and data (physical laws, geostatistical relationships) is not given primarily on the model parameter space but on physical space (3D plus time plus modeled matter). For this reason, it is useful to comment on the difference between the space of model parameters, which here is given by random variables supporting a physical object model, and the space of physical modeled objects. A set of object mode-

ling rules, sometimes identified as *parameterization* rules, are required to transform a model parameter outcome to a modeled object configuration in physical 3D space and in some cases time. These rules commonly involve the physical identification of the parameters and the construction of their outcome in space accordingly. Sometimes, they are straightforward, like in the case of assigning property values to a Euclidean three-dimensional grid, but in other cases can be more elaborated. Examples are parameters being coefficients of polynomials defining geological body surfaces, or when the model elements are defined over curved coordinate systems.

A realization of an object model configuration can be regarded as the process that combines (1) drawing an outcome of the random variables (in model parameter space) according to the correspondent PDF and (2) passing these parameters through the object modeling engine to end up with a configuration of the object in the modeled physical (3D and time) space. The modeled object configuration is the result of this realization process. Once the object modeling rules have been established and behave as a bijective function, each outcome of the model parameters is associated with a correspondent outcome of the modeled object configuration, and vice versa. An outcome of the model parameters implicitly indicates a realization of a modeled object configuration; hence they are sometimes treated as the same entity.

A similar distinction applies for the observation of experiments made on the studied object (geophysical surveys, well-log measurements) and the data, conceived as parameters that describe the observations. The data also requires a series of configuration rules to have physical meaning (in 3D space, time, and observation nature). Also, commonly the data used in the inverse problems involves processing from raw (lower level) field or instrumental data. The understanding of the differences between model parameters, modeled object, experimental observation, and data is useful for the formulation of the relations between data and model parameters and the complete description of the related uncertainties. In particular, uncertainties should involve the various modeling processes that are present.

The issue of modeling parameter probability densities is closely related with the object modeling transform, as different parameterizations produce different PDFs for the same state of information on the modeled object. In particular, homogeneous (i.e., null-information) parameter probability densities are straightforwardly related to the parameterization [see Tarantola, 2005]. The object and observational modeling transforms are implicit components in the corresponding conditionals and likelihoods, as well as in the geophysical (3.20) and petrophysical (3.21) functions used to model the objective function terms.

3.7. CONCLUSIONS

Inference problems in Earth sciences involve the integration of multiple types of knowledge, observations, and information, which ultimately is done by the expert and the scientific community at large by continuous processes of partial analysis and synthesis. To support these processes, methods for quantitative inference in complex models and multiple data are in progress. The inference formulation is done via the definition of probability densities over parameter spaces that model the object or phenomenon to be described. To model the posterior state of information, after a set of multiple observations have been included, the data likelihood functions can be factorized across surveys and observational methods, assuming independence of the observed data uncertainties.

To couple components of the object model that are responsible for diverse observations, the knowledge about inner relationships across the object model parameters need to be used as part of the prior information, entering as conditional probability densities between components of the object model. The identification of relevant dependencies and acceptable independencies across model components is an important issue for the pertinence of the model and the reliability of the inference. The presentation of these dependencies via direct acyclic graphs is useful, allowing a straightforward formulation of the posterior PDF.

Once the posterior PDF is modeled, the generation of the object posterior configurations can follow two lines: drawing multiple realizations from the posterior PDF (sampling approach) or searching maximum posterior PDF configurations and their local covariance, a procedure that commonly has only local validity depending on the modes of the PDF.

The theoretical capabilities of these methods are unlimited, in a mathematical sense, depending for their application on the computational capacities and the ability to reliably describe the object/phenomenon laws and inner relationships across their components.

ACKNOWLEDGMENTS

The author acknowledges the Universidad Central de Venezuela, the Institute de Physique du Globe de Paris, and the University of Cambridge, which housed different periods of the author's work in the subjects developed in the chapter. Special acknowledgment is due to Albert Tarantola, professor and friend, who promoted the interest in inverse problems and their probabilistic treatment in professional and academic Earth science community.

REFERENCES

- Alpak, F., C. Torres-Vedin, and T. Habashy (2008), Estimation of in-situ petrophysical properties from wireline formation tester and induction logging measurements: A joint inversion approach, *J. Petroleum Sci. Eng.*, *63*, 1–17.
- Bhattacharjya, D., and Mukerji, T. (2006), Using influence diagrams to analyze decisions in 4D seismic reservoir monitoring, *The Leading Edge*, *25*, 1236–1239.
- Bosch, M. (1999), Lithologic tomography: From plural geophysical data to lithology estimation, *J. Geophys. Res.*, *104*, 749–766.
- Bosch, M. (2004), The optimization approach to lithological tomography: Combining seismic data and petrophysics for porosity prediction, *Geophysics*, *69*, 1272–1282, doi: 10.1190/1.1801944.
- Bosch, M., C. Barnes, and K. Mosegaard (2000), Multi-step samplers for improving efficiency in probabilistic geophysical inference, in *Methods and Applications of Inversion*, Springer, New York.
- Bosch, M., L. Cara, J. Rodrigues, A. Navarro, and M. Díaz (2007), A Monte Carlo approach to the joint estimation of reservoir and elastic parameters from seismic amplitudes, *Geophysics*, *72*(6), O29–O39, doi: 10.1190/1.2783766.
- Bosch, M., C. Carvajal, J. Rodrigues, A. Torres, M. Aldana, and J. Sierra (2009), Petrophysical seismic inversion conditioned to well-log data: Methods and application to a gas reservoir, *Geophysics*, *74*(2), O1–O15, doi: 10.1190/1.3043796.
- Bosch, M., Guillen, A., and Ledru, P. (2001), Lithologic tomography: An application to geophysical data from the Cadomian belt of northern Brittany, France, *Tectonophysics*, *331*, 197–227.
- Bosch, M., R. Meza, R. Jimenez, and A. Honig (2004), Joint gravity and magnetic inversion in 3D using Monte Carlo methods, *Geophysics*, *71*, G153–G156, doi:10.1190/1.2209952.
- Bosch, M., T. Mukerji, and E. Gonzalez (2010), Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: A review, *Geophysics*, *75*, A165–A176.
- Bourne, J. H. (1993), Use of magnetic susceptibility, density and modal mineral data as a guide to composition of granitic plutons, *Math. Geol.*, *25*, 357–375.
- Buland, A., and O. Kolbjomsen (2012), Bayesian inversion of CSEM and magnetotelluric data: *Geophysics*, *77*, E33–E42.
- Caers J., S. Srinivasan, and A. Journel (2000), Geostatistical quantification of geological information for a fluvial-type North Sea reservoir, *SPE Reservoir Eval. Eng.*, *3*, 457–467.
- Chen, J., and G. M. Hoversten (2012), Joint inversion of marine seismic AVA and CSEM data using statistical rock-physics models and Markov random fields, *Geophysics*, *77*, R65–R80.
- Chiles, J.-P., and P. Delfiner (2009), *Geostatistics: Modeling Spatial Uncertainty*, John Wiley & Sons, Hoboken, NJ.
- Connolly, J. A. D. (2005), Computation of phase equilibria by linear programming tool for geodynamic modeling and an application to subduction zone decarbonation, *Earth Planet. Sci. Lett.*, *236*, 524.
- Dadashpour, M., D. Echeverria-Ciaurri, J. Kleppe, and M. Landro (2009), Porosity and permeability estimation by

- integration of production and time-lapse near and far offset seismic data: *J. Geophys. Eng.*, *6*, 325–344, doi: 10.1088/1742-2132/6/4/001.
- Deutsch, C. R., and A. G. Journel (1992), *GSLIB, Geostatistical Software Library and User's Guide*, Oxford University Press, New York, 340 pages.
- Deustch, C. V., and L. Wang (1996), Hierarchical object based stochastic modeling of fluvial reservoirs, *Math. Geol.*, *28*, 857–880.
- Doetsch, J., N. Linde, I. Coscia, S. Greenhalgh, and A. Green (2010), Zonation for 3D aquifer characterization based on joint inversion of mutimethod crosshole geophysical data, *Geophysics*, G53–G64.
- Dubrule, O. (2003), *Geostatistics for Seismic Data Integration in Earth Models*, SEG, New York.
- Eidsvik J., P. Avseth, H. More, T. Mukerji, and G. Mavko (2004), Stochastic reservoir characterization using prestack seismic data, *Geophysics*, *69*, 978–993.
- Gallardo, L., and M. Meju (2004), Joint two-dimensional DC resistivity and seismic travel-time inversion with cross-gradients constraints, *J. Geophys. Res.: Solid Earth*, *109*, B03311.
- Gallardo, L., M. A. Perez, and E. Gomez (2003), A versatile algorithm for joint 3D inversion of gravity and magnetic data, *Geophysics*, *68*, 949–959.
- Gelfand, A. E., and A. F. M. Smith (1990), Sampling-based approaches to calculating marginal densities, *J. Am. Stat. Assoc.*, *85*, 398–409.
- Geyer, C. J. (1992), Practical Markov chain Monte Carlo, *Stat. Sci.*, *7*, 473–551.
- González, E., T. Mukerji, and G. Mavko (2008), Seismic inversion combining rock physics and multiple-point geostatistics, *Geophysics*, *73*, R11–R21.
- Grana, D. (2014), Probabilistic approach to rock physics modeling, *Geophysics*, *79*, D123–D143.
- Grana, D., and E. Della Rossa (2010), Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion, *Geophysics*, *75*, O21–O37.
- Grana, D., T. Mukerji, J. Dvorkin, and G. Mavko (2012), Stochastic inversion of facies from seismic data based on sequential simulations and probability perturbation method, *Geophysics*, *77*, M53–M72.
- Griffiths, T. C. Kemp, and J. Tenenbaum (2008), Bayesian models of cognition, in *The Cambridge Handbook of Computational Psychology*, R. Sun, ed., Cambridge University Press, New York.
- Guillen, A., P. Calcagno, G. Courrioux, A. Joly, and P. Ledru (2007), Geological modelling from field data and geological knowledge, Part II—Modeling validation using gravity and magnetic data inversion, *Phys. Earth Planetary Interiors*, *171*, doi:10.1016/j.pepi.2008.06.014.
- Haber, E., and D. Oldenburg (1997), Joint inversion: A structural approach, *Inverse Problems*, *13*, 63–77.
- Hacker, B. R., G. A. Abers, and S. M. Peacock (2003), Subduction factory 1. Theoretical mineralogy, densities, seismic wave speeds and H₂O contents, *J. Geophys. Res.*, *108*, doi:10.1029/2001JB001127.
- Hilterman, F. J. (2001), *Seismic Amplitude Interpretation*, SEG, New York.
- Holden, L., R. Hauge, O. Skare, A. Skorstad (1998), Modeling of fluvial reservoirs with object models, *Math. Geol.*, *30*, 473–496.
- Huang, X. (2001), *Integrating Time-Lapse Seismic with Production Data: A Tool for Reservoir Engineering*, The Leading Edge, New York.
- Khan, A., and J. A. D. Connolly (2008), Constraining the composition and thermal state of Mars from inversion of geophysical data, *J. Geophys. Res.*, *113*, E7003.
- Khan, A., J. A. D. Connolly, J. MacLennan, and K. Mosegaard (2006), Joint inversion of seismic and gravity data for lunar composition and thermal state, *Geophys. J. Int.*, *168*, 243–258.
- Khan, A., J. A. D. Connolly, and S. R. Taylor (2008), Inversion of seismic and geodetic data for the major element chemistry and temperature of the Earth's mantle, *J. Geophys. Res.*, *113*, B9308.
- Larsen A., M. Ulvmoen, H. Omre, and A. Buland (2006), Bayesian lithology/fluid prediction and simulation on the basis of a Markov-chain prior model, *Geophysics*, *71*, R69–R78.
- Linde, N., J. Chen, M. Kowalsky, and S. Hubbard (2006a), Hydrogeophysical parameter estimation approaches for field scale characterization, In *Applied Hydrogeophysics*, H. Verreken, A. Binley, G. Cassiani, A. Revil, and K. Titov, eds., Springer, New York.
- Linde, N., A. Binley, A. Tryggvason, L. Pedersen, and A. Revil (2006b), Improved hydrogeophysical characterization using joint inversion of cross-hole electrical resistance and ground-penetrating radar traveltimes data, *Water Resources Res.*, *42*, W12404.
- Lines, L., A. Schultz, and S. Treitel (1988), Cooperative inversion of geophysical data, *Geophysics*, *53*, 8–20.
- Liu, S. J. (1998), Metropolized independent sampling with comparison to rejection sampling and importance sampling, *Stat. Comput.*, *6*, 1113–1119.
- Martinelli, G., Eidsvik, J., Sinding-Larsen, R., Rekstad, S., and Mukerji, T. (2013), Building Bayesian networks from basin-modelling scenarios for improved geological decision making, *Petroleum Geosci.*, *19*(3), 289–304.
- Mavko, G., and T. Mukerji (1998), A rock physics strategy for quantifying uncertainty in common hydrocarbon indicators, *Geophysics*, *63*, 1997–2008.
- Mavko, G., T. Mukerji, and J. Dvorkin (2003), *The Rock Physics Handbook*, Cambridge University Press, New York.
- Mezghani, M., A. Fornel, V. Langlais, and N. Lucet (2004), *History Matching and Quantitative Use of 4D Seismic Data for an Improved Reservoir Characterization*, Society of Petroleum Engineers, Houston, TX, SPE-90420-MS, doi: 10.2118/90420.
- Mosegaard, K. (2011), Quest for consistency, symmetry and simplicity—The legacy of Albert Tarantola, *Geophysics*, *76*, W51–W61.
- Mosegaard K., and A. Tarantola (1995), Monte Carlo sampling of solutions to inverse problems, *J. Geophys. Res.*, *100*, 12431–12447.
- Mosegaard K., and A. Tarantola (2002), Probabilistic approach to inverse problems, in *International Handbook of Earthquake*

- and *Engineering Seismology*, W. Lee, H. Kanamori, P. Jennings, and C. Kisslinger, eds., Academic Press, New York.
- Mukerji, T., A. Jorstad, P. Avseth, and J. R. Granli (2001), Mapping lithofacies and pore–fluid probabilities in a North Sea reservoir: Seismic inversions and statistical rock physics, *Geophysics*, 66, 988–1001.
- Pearl, J. (1986), Fusion, propagation and structuring in belief networks, *Artifi. Intell.*, 29, 241–288.
- Pearl, J. (1994), Causal diagrams for empirical research, *Biometrika*, 82, 669–688.
- Rimstad K., P. Avseth, and H. Omre (2012), Hierarchical Bayesian lithology/fluid prediction: A North Sea case study, *Geophysics*, 77, B69–B85.
- Smith, A. F., and G. O. Roberts (1993), Bayesian computations via the Gibbs sampler and related Markov chain Monte Carlo methods, *J. Royal Stat. Soc.*, 55, 3–23.
- Spikes, K., T. Mukerji, and G. Mavko (2007), Probabilistic seismic inversion on rock-physics models, *Geophysics*, R87–R97.
- Strebelle, S. (2002), Conditional simulation of complex geological structures using multiple-point statistics, *Math. Geol.*, 34, 1–22.
- Suman, A., and T. Mukerji (2013), Sensitivity study of rock-physics parameters for modeling time-lapse seismic response of Norne field, *Geophysics*, 78, D511–D523.
- Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, Philadelphia.
- Thulasiraman, K., and M. Swamy (1992), *Graphs: Theory and Algorithms*, John Wiley & Sons, New York.
- Tiberi, C., M. Diament, J. Deverchere, C. Petit-Mariani, V. Mikhailov, S. Tikhostsky, and U. Achauer (2003), Deep structure of the Baikal rift zone by joint inversion of gravity and seismology, *J. Geophys. Res. Solid Earth*, 108, 2133.
- Tierney, L. (1994), Markov-chains for exploring posterior distributions, *Ann. Stat.*, 22, 1702–1762.
- Torres-Verdin, C., A. Revil, M. Oristaglio and T. Mukerji (2012), Multiphysics borehole geophysical measurements, formation evaluation, petrophysics and rock physics—Introduction, *Geophysics*, 77, WA1–WA2.
- Ulvmoen M., H. Omre, and A. Buland (2010), Improved resolution in Bayesian lithology/fluid inversion from prestack seismic data and well observations, Part 2—Real case study, *Geophysics*, 75, B73–B82.