

BAYESIAN INVERSION OF CONVENTIONAL ELECTRIC LOGS

BOGDAN MIHAI NICULESCU

*University of Bucharest, Faculty of Geology and Geophysics, Department of Geophysics
6, Traian Vuia Street, 020956 Bucharest, Romania
e-mail: bogdan.niculescu@b.astral.ro, bogdan.niculescu@k.ro*

Modélisation inverse Bayésienne des diagraphies électriques conventionnelles. L'article présente l'élaboration d'un algorithme probabilistique Bayésien pour la modélisation inverse des données de résistivité apparente enregistrées à l'aide des dispositifs conventionnels (normaux et latéraux) de carottage électrique et l'implémentation de l'algorithme par un logiciel d'inversion efficace et rapide. Les modèles géoélectriques utilisés sont des milieux multistratifiés avec les interfaces de séparation planes-parallèles, leurs paramètres étant les profondeurs des interfaces et les résistivités des couches homogènes et isotropes. La formulation Bayésienne du problème inverse permet l'incorporation de l'incertitude associée aux paramètres initiaux du modèle ou aux données géophysiques enregistrées, conduisant à une estimation optimale et la plus probable du modèle d'interprétation. Les essais effectués avec l'algorithme proposé ont montré sa capacité d'optimiser un modèle géoélectrique initial et de produire le meilleur ajustement des données, tout en considérant le bruit géologique et les informations disponibles *a priori* sur les données ou les paramètres du modèle. Le domaine d'applicabilité de l'algorithme couvre non seulement la modélisation inverse des diagraphies électriques, mais aussi l'interprétation automatique de tous les types de données géophysiques, par une modification convenable du module de modélisation directe. Les procédés classiques d'inversion peuvent être obtenus comme solutions particulières de cette méthode probabilistique générale.

Key words: apparent resistivity, Bayesian, borehole geophysics, electric log, inversion, probability density.

1. INTRODUCTION

The classical automatic interpretation procedures of geophysical data which are based on inversion require a fitting of the measured values by means of a theoretical dataset, using the least-squares criterion. For surface and borehole apparent resistivity data inversion, frequently used were gradient, quasi-Newton or ridge regression methods (Levenberg, 1944; Marquardt, 1963; Inman, 1975; Rijo *et al.*, 1977; Petrick *et al.*, 1977; Pelton *et al.*, 1978; Hoversten and Morrison, 1982; Yang and Ward, 1984; Whitman *et al.*, 1989; Loke and Barker, 1996) as well as methods employing the generalized inverse matrix (Lanczos, 1961; Inman *et al.*, 1973; Jupp and Vozoff, 1975; Pous *et al.*, 1985). Due to the non-linearity of the

inverse problem, the geophysical interpretation model results from an iterative minimization procedure.

The main problem that apparent resistivity data inverse modeling has to deal with is the non-uniqueness of the solution caused by the existence of multiple equivalent geoelectric models. Also, it is possible that the obtained solution, depending on the selection of the initial geoelectric model, may not correspond to the geological reality or modify the *a priori* known parameters and constraints included in the initial model. The disadvantage of such an approach is related to convergence problems that may arise during the inverse modeling. Also, the likelihood of errors affecting the *a priori* information is ignored, even if this kind of information usually has a low degree of confidence.

From the probabilities theory perspective, the mathematical bases for using *a priori* information in the inverse modeling were elaborated by Goltsman (1971), Tarantola and Valette (1982), Jackson and Matsuura (1985) and Tarantola (1987), comprehensive overviews of this method being given by Mosegaard and Tarantola (2002, 2005). Up to now, the probabilistic inversion techniques have been applied for the interpretation of seismic data (Duijndam, 1988) and surface apparent resistivity surveys (Andersen *et al.*, 2003; Malinverno and Torres-Verdin, 2000; Pous *et al.*, 1987). The application of Bayes' theorem (1763) allows combining the information provided by the observational data with other types of information, geological or not, related to the parameters of the interpretation model or the constraints which may be applied to these parameters. The probabilistic treatment of observational data, *a priori* information and constraints, within the inverse modeling, leads to obtaining a unique solution in agreement with all the available information. Furthermore, the probabilistic inversion of geophysical data is an efficient technique for avoiding a non-convergence of the iterative process required to determine an optimum solution. Such problems may appear, for example, when the parameters of the initial interpretation model are badly chosen or when the observational data are not compatible with the imposed model (data measuring errors, interpretation models different from the real geological structure, etc.).

The probabilistic inverse modeling has a maximum degree of generality and may be applied for the interpretation of any kind of geophysical data. Classical inversion methods represent only particular cases of this general algorithm and are obtained by imposing particular probability distributions for the measured data or the model parameters.

2. DESCRIPTION OF THE PROPOSED METHOD

2.1 PROBABILISTIC INVERSION WITH *A PRIORI* INFORMATION

Let us consider a random vector \mathbf{y} whose components y_1, \dots, y_N are random variables representing discrete observational data, measured in the x_1, \dots, x_N positions of a non-random vector \mathbf{x} . The geophysical interpretation model is

determined by the components of a random vector $\mathbf{p} = p_1, \dots, p_M$, containing the unknown values of model parameters. Forward modeling relates to the determination of a theoretical set of values f_1, \dots, f_N representing the effect of a particular interpretation model (a formal, non-linear, functional f), in the x_1, \dots, x_N positions of a vector \mathbf{x} and for given fixed values of the model parameters. Inverse modeling tries to determine the components of the parameters vector \mathbf{p} , using the observational data \mathbf{y} .

For simplification purposes, it may be assumed that the $f(\mathbf{p}, \mathbf{x})$ response of an electric log only depends on the unknown parameters \mathbf{p} of the geoelectric model (for a multilayered model, these parameters might be the true resistivities and the depths of separation interfaces) and, also, that the recorded apparent resistivity curve \mathbf{y} was sampled with a constant depth step.

Disregarding the dependency upon the vector $\mathbf{x} = x_1, x_2, \dots, x_N$ which stores the N sampling points depths, it should be generally considered that

$$\mathbf{y} \equiv \mathbf{f}(\mathbf{p}). \quad (1)$$

As \mathbf{y} and \mathbf{p} are variable vectors, a probability density P can be associated to each of them. The Bayes theorem may be written

$$P(\mathbf{p} | \mathbf{y}) = \frac{P(\mathbf{y} | \mathbf{p}) P(\mathbf{p})}{P(\mathbf{y})}, \quad (2)$$

where: $P(\mathbf{p})$ – probability density for the unknown parameters \mathbf{p} of the geoelectric model (the *a priori* probability density), reflecting the degree of knowledge for these parameters in the absence of observational data; $P(\mathbf{y})$ – probability density for the observational data \mathbf{y} ; $P(\mathbf{y} | \mathbf{p})$ – conditional probability density of vector \mathbf{y} with respect to vector \mathbf{p} (the *likelihood function*), describing the theoretical link between the observational data and the geoelectric model parameters; $P(\mathbf{p} | \mathbf{y})$ – conditional probability density of vector \mathbf{p} with respect to vector \mathbf{y} (the *a posteriori* probability density), *i.e.*, the probability density for the unknown model parameters after the acquisition of observational data.

The probabilistic inverse modeling is carried out by maximizing the *a posteriori* probability density $P(\mathbf{p} | \mathbf{y})$ and selecting the corresponding set of parameters \mathbf{p} , as the ones which optimally, and most probably, explain the recorded apparent resistivity data:

$$P(\mathbf{p} | \mathbf{y}) \text{ maximum} \Rightarrow \mathbf{p}. \quad (3)$$

The probability density used within the elaborated inversion algorithm is the normal (Gaussian) distribution of a random variable p , given by

$$P(p, \mu_p, \sigma_p) = \frac{1}{\sigma_p \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{p - \mu_p}{\sigma_p} \right)^2 \right\}, \quad (4)$$

where μ_p represents the mean (most probable) value of the random variable p and σ_p its standard deviation. The square of this value, $D(p) = \sigma_p^2$, stands for the variance (dispersion) of variable p .

For M independent random variables p , characterized by the probability densities

$$P(p_1, \mu_{p_1}, \sigma_{p_1}) = \frac{1}{\sigma_{p_1} \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{p_1 - \mu_{p_1}}{\sigma_{p_1}} \right)^2 \right\}$$

$$P(p_2, \mu_{p_2}, \sigma_{p_2}) = \frac{1}{\sigma_{p_2} \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{p_2 - \mu_{p_2}}{\sigma_{p_2}} \right)^2 \right\} \quad (5)$$

$$\vdots$$

$$\vdots$$

$$\vdots$$

$$P(p_M, \mu_{p_M}, \sigma_{p_M}) = \frac{1}{\sigma_{p_M} \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{p_M - \mu_{p_M}}{\sigma_{p_M}} \right)^2 \right\},$$

their simultaneous probability density is

$$P(\mathbf{p}) = P(p_1, p_2, \dots, p_M) = P(p_1) P(p_2) \dots P(p_M). \quad (6)$$

After the corresponding calculations, one obtains

$$P(p_1, p_2, \dots, p_M) = \frac{(2\pi)^{-M/2}}{\prod_{i=1}^M \sigma_{p_i}} \exp \left\{ -\frac{1}{2} \left[\frac{(p_1 - \mu_{p_1})^2}{\sigma_{p_1}^2} + \frac{(p_2 - \mu_{p_2})^2}{\sigma_{p_2}^2} + \dots + \frac{(p_M - \mu_{p_M})^2}{\sigma_{p_M}^2} \right] \right\} \quad (7)$$

Taking into consideration the difference vector

$$(\mathbf{p} - \boldsymbol{\mu}_p) = \begin{bmatrix} p_1 - \mu_{p_1} \\ p_2 - \mu_{p_2} \\ \vdots \\ \vdots \\ p_M - \mu_{p_M} \end{bmatrix}, \quad (8)$$

the covariance matrix \mathbf{C}_p for the geoelectric model parameters,

$$\mathbf{C}_p = \begin{bmatrix} \sigma_{p_1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{p_2}^2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \sigma_{p_M}^2 \end{bmatrix}, \quad (9)$$

and the inverse of this matrix ($\mathbf{C}_p^{-1} = \mathbf{W}_p =$ parameters weighting matrix), the probability density for the unknown parameters of the geoelectric model, described by equation 7, becomes

$$P(\mathbf{p}) = P(p_1, p_2, \dots, p_M) = \frac{(2\pi)^{-M/2}}{\prod_{i=1}^M \sigma_{p_i}} \exp\left\{-\frac{1}{2}(\mathbf{p} - \boldsymbol{\mu}_p)^T \mathbf{C}_p^{-1} (\mathbf{p} - \boldsymbol{\mu}_p)\right\} \quad (10)$$

or, in a symbolic notation,

$$P(\mathbf{p}) = \text{constant} \exp\left\{-\frac{1}{2}(\mathbf{p} - \boldsymbol{\mu}_p)^T \mathbf{C}_p^{-1} (\mathbf{p} - \boldsymbol{\mu}_p)\right\}. \quad (11)$$

The vector of discrete apparent resistivity data \mathbf{y} may be written as

$$\mathbf{y} = \mathbf{f}(\mathbf{p}) + \mathbf{n}, \quad (12)$$

where $\mathbf{f}(\mathbf{p})$ is the vector of theoretical values corresponding to the adopted geoelectric model and $\mathbf{n} = n_1, n_2, \dots, n_N$ represents the noise which may affect the observational data and/or the theoretical estimation errors. Considering that the

errors \mathbf{n} are independent of $\mathbf{f}(\mathbf{p})$ and characterized by a probability density P_n , the conditional probability density $P(\mathbf{y}|\mathbf{p})$, or the likelihood function, is

$$P(\mathbf{y}|\mathbf{p}) = P_n(\mathbf{y} - \mathbf{f}(\mathbf{p})). \quad (13)$$

If the errors associated with the measured apparent resistivity values have zero mean and a covariance matrix

$$\mathbf{C}_y = \begin{bmatrix} \sigma_{y_1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{y_2}^2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \sigma_{y_N}^2 \end{bmatrix}, \quad (14)$$

the probability density of the errors vector may be symbolically expressed as

$$P(\mathbf{n}) = \text{constant} \exp\left\{-\frac{1}{2} \mathbf{n}^T \mathbf{C}_y^{-1} \mathbf{n}\right\}, \quad (15)$$

where the inverse $\mathbf{C}_y^{-1} = \mathbf{W}_y$ may be regarded as a data weighting matrix. Taking into account expression 13, the likelihood function becomes

$$P(\mathbf{y} | \mathbf{p}) = \text{constant} \exp\left\{-\frac{1}{2} (\mathbf{y} - \mathbf{f}(\mathbf{p}))^T \mathbf{C}_y^{-1} (\mathbf{y} - \mathbf{f}(\mathbf{p}))\right\}. \quad (16)$$

Therefore, the numerator of expression 2 takes the form

$$(17)$$

$$P(\mathbf{y} | \mathbf{p})P(\mathbf{p}) = \text{constant} \exp\left\{-\frac{1}{2} \left[(\mathbf{p} - \boldsymbol{\mu}_p)^T \mathbf{C}_p^{-1} (\mathbf{p} - \boldsymbol{\mu}_p) + (\mathbf{y} - \mathbf{f}(\mathbf{p}))^T \mathbf{C}_y^{-1} (\mathbf{y} - \mathbf{f}(\mathbf{p})) \right]\right\}$$

Maximizing the *a posteriori* probability density $P(\mathbf{p}|\mathbf{y})$ reduces to maximizing the expression 17, which leads to minimizing its exponent. Consequently, the objective function to be minimized may be written in matrix notation as

$$E(\mathbf{p}) = (\mathbf{p} - \boldsymbol{\mu}_p)^T \mathbf{C}_p^{-1} (\mathbf{p} - \boldsymbol{\mu}_p) + (\mathbf{y} - \mathbf{f}(\mathbf{p}))^T \mathbf{C}_y^{-1} (\mathbf{y} - \mathbf{f}(\mathbf{p})) \rightarrow \text{minimum} \quad (18)$$

and, after the necessary substitutions,

$$E(\mathbf{p}) = \sum_{i=1}^M \frac{(p_i - \mu_{p_i})^2}{\sigma_{p_i}^2} + \sum_{i=1}^N \frac{(y_i - f_i(\mathbf{p}))^2}{\sigma_{y_i}^2} \rightarrow \text{minimum}. \quad (19)$$

The objective function represented by expression 19 is linearized using a Taylor series expansion around an initial estimate \mathbf{p}^0 of the unknown parameters vector, retaining the first and second order terms of the expansion. After computation of the partial derivatives of the objective function with respect to the parameters p of the geoelectric model and cancellation of the derivatives, the minimization equation is obtained as

$$\mathbf{H} \Delta \mathbf{p} = -\mathbf{g}, \quad (20)$$

where \mathbf{H} is the Hessian matrix of second order partial derivatives of the objective function with respect to model's parameters, \mathbf{g} is the gradient vector containing the first order derivatives of the objective function and $\Delta \mathbf{p}$ is the parameters update vector, representing the solution of the inverse problem.

The elements of the gradient vector are

$$g_j = \left(\frac{\partial E(\mathbf{p})}{\partial p_j} \right)_0 ; j = 1, \dots, M, \quad (21)$$

which leads to

$$g_j = \frac{\partial}{\partial p_j} \left[\sum_{i=1}^M \frac{(p_i - \mu_{p_i})^2}{\sigma_{p_i}^2} + \sum_{i=1}^N \frac{(y_i - f_i(\mathbf{p}))^2}{\sigma_{y_i}^2} \right]_{p_j=p_j^0}. \quad (22)$$

Carrying out the corresponding calculations, the following equations are obtained:

$$\begin{aligned} \left(\frac{\partial E(\mathbf{p})}{\partial p_1} \right)_0 &= -2 \left\{ \sum_{i=1}^N \left[\frac{(y_i - f_i(\mathbf{p}))}{\sigma_{y_i}^2} \left(\frac{\partial f_i(\mathbf{p})}{\partial p_1} \right)_0 \right] - \frac{p_1^0 - \mu_{p_1}}{\sigma_{p_1}^2} \right\} \\ \left(\frac{\partial E(\mathbf{p})}{\partial p_2} \right)_0 &= -2 \left\{ \sum_{i=1}^N \left[\frac{(y_i - f_i(\mathbf{p}))}{\sigma_{y_i}^2} \left(\frac{\partial f_i(\mathbf{p})}{\partial p_2} \right)_0 \right] - \frac{p_2^0 - \mu_{p_2}}{\sigma_{p_2}^2} \right\} \end{aligned} \quad (23)$$

$$\left(\frac{\partial E(\mathbf{p})}{\partial p_M} \right)_0 = -2 \left\{ \sum_{i=1}^N \left[\frac{(y_i - f_i(\mathbf{p})) \left(\frac{\partial f_i(\mathbf{p})}{\partial p_M} \right)_0}{\sigma_{y_i}^2} \right] - \frac{p_M^0 - \mu_{p_M}}{\sigma_{p_M}^2} \right\}.$$

Taking into consideration the Jacobian matrix containing the first order derivatives of the formal functional relationship $f(\mathbf{p})$ (the forward model, describing the theoretical response of the geoelectric model) with respect to the unknown parameters p ,

$$\mathbf{J} = \begin{bmatrix} \left(\frac{\partial f_1(\mathbf{p})}{\partial p_1} \right)_0 & \left(\frac{\partial f_1(\mathbf{p})}{\partial p_2} \right)_0 & \dots & \left(\frac{\partial f_1(\mathbf{p})}{\partial p_M} \right)_0 \\ \left(\frac{\partial f_2(\mathbf{p})}{\partial p_1} \right)_0 & \left(\frac{\partial f_2(\mathbf{p})}{\partial p_2} \right)_0 & \dots & \left(\frac{\partial f_2(\mathbf{p})}{\partial p_M} \right)_0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \left(\frac{\partial f_N(\mathbf{p})}{\partial p_1} \right)_0 & \left(\frac{\partial f_N(\mathbf{p})}{\partial p_2} \right)_0 & \dots & \left(\frac{\partial f_N(\mathbf{p})}{\partial p_M} \right)_0 \end{bmatrix}, \quad (24)$$

and defining the vector of differences between the mean values μ_p of the model's unknown parameters and their estimated values \mathbf{p}^0 ,

$$\mathbf{dp} = (\mu_p - \mathbf{p}^0) = \begin{bmatrix} \mu_{p_1} - p_1^0 \\ \mu_{p_2} - p_2^0 \\ \cdot \\ \cdot \\ \cdot \\ \mu_{p_M} - p_M^0 \end{bmatrix}, \quad (25)$$

the gradient vector takes the form

$$\mathbf{g} = -2 \left[\mathbf{J}^T \mathbf{C}_y^{-1} (\mathbf{y} - \mathbf{f}(\mathbf{p})) + \mathbf{C}_p^{-1} (\boldsymbol{\mu}_p - \mathbf{p}^0) \right] \quad (26)$$

or

$$\mathbf{g} = -2 \left[\mathbf{J}^T \mathbf{W}_y \Delta \mathbf{y} + \mathbf{W}_p \mathbf{d} \mathbf{p} \right], \quad (27)$$

where the covariance matrices \mathbf{C}_p and \mathbf{C}_y are given by the expressions 9 and 14 and $\Delta \mathbf{y}$ is the residual vector storing the differences between the measured apparent resistivity values and the theoretical ones, computed via forward modeling for a current geoelectric model:

$$\Delta \mathbf{y} = \begin{bmatrix} y_1 - f_1(\mathbf{p}) \\ y_2 - f_2(\mathbf{p}) \\ \vdots \\ y_N - f_N(\mathbf{p}) \end{bmatrix}; f_j(\mathbf{p}) = f(\mathbf{p}, x_j). \quad (28)$$

The elements of the Hessian matrix are

$$H_{ij} = \left(\frac{\partial^2 E(\mathbf{p})}{\partial p_i \partial p_j} \right)_0; i = 1, \dots, M; j = 1, \dots, M, \quad (29)$$

which may be written as

$$H_{ij} = \frac{\partial}{\partial p_j} \left\{ \frac{\partial}{\partial p_i} \left[\sum_{k=1}^M \frac{(p_k - \mu_{p_k})^2}{\sigma_{p_k}^2} + \sum_{i=1}^N \frac{(y_i - f_i(\mathbf{p}))^2}{\sigma_{y_i}^2} \right]_{p_i=p_i^0} \right\}_{p_j=p_j^0}. \quad (30)$$

After the necessary calculations, the following equations are derived:

$$\left(\frac{\partial^2 E(\mathbf{p})}{\partial p_i \partial p_i} \right)_0 = 2 \left[\sum_{k=1}^N \frac{\left(\frac{\partial f_k(\mathbf{p})}{\partial p_i} \right)_0 \left(\frac{\partial f_k(\mathbf{p})}{\partial p_i} \right)_0}{\sigma_{y_k}^2} + \frac{1}{\sigma_{p_i}^2} \right]; \text{ for } j = i \quad (31)$$

$$\left(\frac{\partial^2 E(\mathbf{p})}{\partial p_i \partial p_j} \right)_0 = 2 \sum_{k=1}^N \frac{\left(\frac{\partial f_k(\mathbf{p})}{\partial p_i} \right)_0 \left(\frac{\partial f_k(\mathbf{p})}{\partial p_j} \right)_0}{\sigma_{y_k}^2}; \text{ for } j \neq i. \quad (32)$$

Considering the expressions of the matrices \mathbf{C}_p , \mathbf{C}_y and \mathbf{J} from 9, 14 and 24, the Hessian matrix takes the form

$$\mathbf{H} = 2[\mathbf{J}^T \mathbf{C}_y^{-1} \mathbf{J} + \mathbf{C}_p^{-1}] \quad (33)$$

or

$$\mathbf{H} = 2[\mathbf{J}^T \mathbf{W}_y \mathbf{J} + \mathbf{W}_p]. \quad (34)$$

The initial optimization equation, after substituting the expressions 26 and 33 or 27 and 34, becomes

$$[\mathbf{J}^T \mathbf{C}_y^{-1} \mathbf{J} + \mathbf{C}_p^{-1}] \Delta \mathbf{p} = [\mathbf{J}^T \mathbf{C}_y^{-1} (\mathbf{y} - \mathbf{f}(\mathbf{p})) + \mathbf{C}_p^{-1} (\boldsymbol{\mu}_p - \mathbf{p}^0)] \quad (35)$$

or

$$[\mathbf{J}^T \mathbf{W}_y \mathbf{J} + \mathbf{W}_p] \Delta \mathbf{p} = [\mathbf{J}^T \mathbf{W}_y \Delta \mathbf{y} + \mathbf{W}_p \mathbf{d}\mathbf{p}]. \quad (36)$$

Finally, the update vector for the unknown parameters of the geoelectric model may be determined from the matrix equations

$$\Delta \mathbf{p} = [\mathbf{J}^T \mathbf{C}_y^{-1} \mathbf{J} + \mathbf{C}_p^{-1}]^{-1} [\mathbf{J}^T \mathbf{C}_y^{-1} (\mathbf{y} - \mathbf{f}(\mathbf{p})) + \mathbf{C}_p^{-1} (\boldsymbol{\mu}_p - \mathbf{p}^0)] \quad (37)$$

or

$$\Delta \mathbf{p} = [\mathbf{J}^T \mathbf{W}_y \mathbf{J} + \mathbf{W}_p]^{-1} [\mathbf{J}^T \mathbf{W}_y \Delta \mathbf{y} + \mathbf{W}_p \mathbf{d}\mathbf{p}]. \quad (38)$$

If the mean of the probability distribution for the geoelectric model parameters ($\boldsymbol{\mu}_p$) coincides with their estimated values (\mathbf{p}^0), the update vector turns into

$$\Delta \mathbf{p} = [\mathbf{J}^T \mathbf{C}_y^{-1} \mathbf{J} + \mathbf{C}_p^{-1}]^{-1} \mathbf{J}^T \mathbf{C}_y^{-1} (\mathbf{y} - \mathbf{f}(\mathbf{p})) \quad (39)$$

or

$$\Delta \mathbf{p} = [\mathbf{J}^T \mathbf{W}_y \mathbf{J} + \mathbf{W}_p]^{-1} \mathbf{J}^T \mathbf{W}_y \Delta \mathbf{y}. \quad (40)$$

It should be noted that the probabilistic inverse modeling of apparent resistivity data requires the computation of matrices \mathbf{J} and \mathbf{J}^T and the selection of covariance matrices \mathbf{C}_p and \mathbf{C}_y , whose inverses $\mathbf{C}_p^{-1} = \mathbf{W}_p$ and $\mathbf{C}_y^{-1} = \mathbf{W}_y$ function as weighting matrices during the inversion. Due to the linearization of the objective function $E(\mathbf{p})$ through a truncated Taylor series, equations 39 or 40 are solved iteratively. In each iteration the global fitting error of the observational data is minimized, the parameters update vector $\Delta\mathbf{p}$ is computed and a new set of parameters \mathbf{p} is determined by correcting the previous ones, *i.e.*, $\mathbf{p} = \mathbf{p}^0 + \Delta\mathbf{p}$. The iterative process which leads to the determination of an optimal geoelectric model is continued until the global fitting error reaches a minimum or falls below a user-imposed tolerance.

The $\Delta\mathbf{p}$ solution from equations 39 and 40 represents a very general *maximum a posteriori estimation* (MAP), which may take particular forms depending on the proper selection of the matrices \mathbf{C}_p and \mathbf{C}_y or the corresponding weighting matrices \mathbf{W}_p and \mathbf{W}_y . If \mathbf{C}_p and \mathbf{C}_y are diagonal, with $\mathbf{C}_y = \sigma_y^2 \mathbf{I}$ and $\mathbf{C}_p = \sigma_p^2 \mathbf{I}$ (where \mathbf{I} denotes the identity matrix) and $\sigma_p = \infty$, a standard least-squares minimization algorithm is obtained,

$$\Delta\mathbf{p} = (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T (\mathbf{y} - \mathbf{f}(\mathbf{p})), \quad (41)$$

the infinite variance of the geoelectric model's parameters meaning the absence of *a priori* information about them.

For any \mathbf{C}_y matrix, if $\mathbf{C}_p = \sigma_p^2 \mathbf{I}$ with $\sigma_p = \infty$, a *maximum likelihood estimator* (MLE) is derived as

$$\Delta\mathbf{p} = (\mathbf{J}^T \mathbf{C}_y^{-1} \mathbf{J})^{-1} \mathbf{J}^T \mathbf{C}_y^{-1} (\mathbf{y} - \mathbf{f}(\mathbf{p})). \quad (42)$$

In this case, the *a priori* information about the model parameters is also missing, but the distribution of errors which affect the measured values is known. If \mathbf{C}_y matrix is diagonal, the MLE estimator reduces to a weighted least-squares algorithm.

It is important to emphasize that the optimal geoelectric model \mathbf{p} resulted from the inversion of apparent resistivity data is not unique, but largely depends on the estimation of an initial, sufficiently correct, model \mathbf{p}^0 . The fundamental ambiguity of inverse geophysical modeling also arises in the case of borehole apparent resistivity logs inversion, so the selection of an erroneous initial model may lead to convergence problems and, also, to a final solution which is substantially different with respect to the real geoelectric model. The advantage of a Bayesian probabilistic approach, using parameters and data covariance matrices along with a suitable inversion algorithm, is represented by the possibility to determine the "most probable" geoelectric model, if adequate probability distributions are selected for the observational data and the model parameters.

One of the main components of the elaborated inversion software is a forward modeling algorithm (Niculescu, 2002, 2006) used to compute the theoretical apparent resistivity response vector $\mathbf{f}(\mathbf{p}, \mathbf{x})$ of a particular logging device, for the selected geoelectric interpretation model. The algorithm finds a solution of the Laplace equation for the electric field's potential in cylindrical coordinates, by constructing a fundamental system of boundary conditions (potential's continuity across the separation interfaces and continuity of the current's density normal component, in addition to particular conditions for the potential near the current source and at infinity). The system is solved via a fast Gaussian elimination routine, in order to determine a set of kernel coefficients $a(\lambda)$ and $b(\lambda)$ which appear in the general integral expression of the potential, then the potentials of the measuring electrodes, along with the corresponding apparent resistivity, are computed using a first degree (trapeze) or second degree (Simpson) numerical quadrature procedure.

During the iterative inversion process, the elements of the Jacobian matrix \mathbf{J} (the sensitivities of the forward modeling algorithm $f(\mathbf{p}, \mathbf{x})$ with respect to the geoelectric model parameters) are computed through numerical differentiation, in each iteration and for each unknown parameter p the theoretical response of the model being evaluated twice, by imposing a sufficiently small \pm variation for the respective parameter around its current value. The amplitude of parameters numerical variations, expressed as percents of their current values, is selected by the user, smaller variation bounds meaning a more precise evaluation of the partial derivatives. After computing the elements of the Hessian matrix \mathbf{H} , the gradient vector \mathbf{g} and the misfit vector $\Delta\mathbf{y}$, matrix \mathbf{H} is inverted and the parameters update vector $\Delta\mathbf{p}$ is determined, gradually optimizing the initial interpretation model \mathbf{p}^0 . Inversion's convergence is tested in each iteration, the process ending after minimizing the root mean square data fitting error or reaching a user-imposed number of iterations.

The uncertainties of the geoelectric model's parameters as well as the apparent resistivity data uncertainties (the elements of \mathbf{C}_p and \mathbf{C}_y matrices) are provided as constant or variable standard deviations of the corresponding quantities, leading to different optimization modes for the inversion algorithm, which may function as a standard least-squares or ridge regression procedure, a maximum likelihood one or as a very general Bayesian method. The selection of a specific inversion mode strongly influences the results and allows the determination of realistic geoelectric models, taking into consideration the noise which may affect the recorded data and/or the variable degree of knowledge of the model parameters.

3. CASE STUDIES AND RESULTS

The applicability of the borehole apparent resistivity inversion software will be illustrated by means of several case studies which use a theoretical geoelectric model and synthetic datasets representing its response. The model comprises 3 separation interfaces and 4 layers, with the first and last layer infinitely extended,

the corresponding parameters (depths z_i of the separation interfaces and true resistivities ρ_i) being mentioned in Table 1.

Table 1

Layered model used for testing the inversion algorithm

h_1 (m)	h_2 (m)	h_3 (m)	ρ_1 (Ωm)	ρ_2 (Ωm)	ρ_3 (Ωm)	ρ_4 (Ωm)
2	3	5	2	10	5	8

The theoretical response of this model, with a conventional vertical extent of 7 m, was determined through forward modeling for an ideal normal logging device with $AM = 0.2$ m length, using a 0.2 m depth sampling step and a 10^{-6} precision for the Simpson numerical quadrature involved in the electric field's potentials and apparent resistivity computations. The resulted apparent resistivity curve (Fig. 2) was considered observational data and used as input for the inversion algorithm.

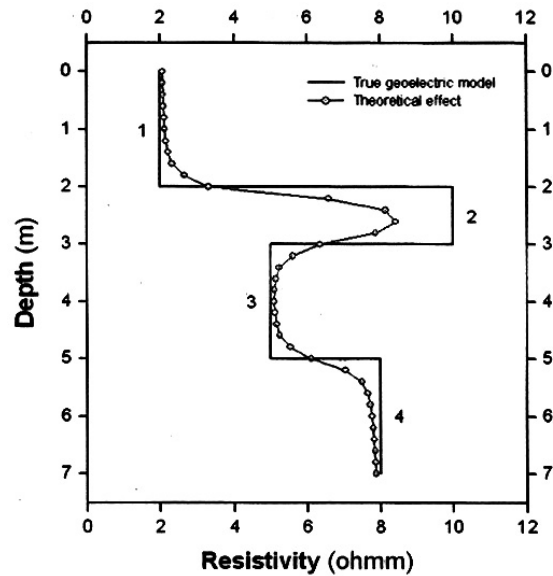


Fig. 2 – Configuration of the true geoelectric model used for testing the probabilistic inversion algorithm and its theoretical effect corresponding to an ideal normal device $AM = 0.2$ m.

If some of the geoelectric model parameters are already known, from geological data or the interpretation of electric logs, this may be taken into account in order to avoid the modification of parameters values during the inversion. In the first case, it was assumed that the separation interfaces depths may be determined by analyzing the apparent resistivity curves and, consequently, this *a priori*

information was incorporated, the inversion algorithm being used to determine only the true resistivities. The adopted initial interpretation model has correct depth values and erroneous resistivity estimates, the chosen standard deviation for all z_i parameters was $\sigma_p = 0.001$, while for the ρ_i parameters $\sigma_p = 1$ standard deviations were used. The covariance matrix of the observational data was chosen as diagonal, $C_y = \sigma_y^2 \mathbf{I}$, with $\sigma_y = 0.001$, the inversion method being a MAP one.

Table 2-a presents the results of the first 5 inversion iterations and Table 2-b shows the estimation errors for the parameters of the initial and final geoelectric models. The values in square brackets from Table 2-a – iteration "0" are referring to the parameters which are not allowed to change during model optimization, e_{RMS} is the root mean square data fitting error and e_{MR} is the mean relative data fitting error.

Table 2-a

Results of the probabilistic inversion of apparent resistivity data corresponding to an ideal AM = 0.2 m normal device (depths of the separation interfaces are assumed known *a priori*)

Iteration	Geoelectric model's parameters							e_{RMS} (Ωm)	e_{MR} (%)
	h_1 (m)	h_2 (m)	h_3 (m)	ρ_1 (Ωm)	ρ_2 (Ωm)	ρ_3 (Ωm)	ρ_4 (Ωm)		
0	[2]	[3]	[5]	1	9	4	9	0.985	24.598
1	2.000	3.000	5.000	1.934	9.650	4.880	8.198	0.145	2.586
2	2.000	3.000	5.000	1.954	9.672	4.911	8.207	0.135	2.138
3	2.000	3.000	5.000	1.954	9.672	4.911	8.207	0.135	2.139
4	2.000	3.000	5.000	1.954	9.672	4.911	8.207	0.135	2.139
5	2.000	3.000	5.000	1.954	9.672	4.911	8.207	0.135	2.139

Table 2-b

Parameters estimation errors corresponding to the initial and final models from Table 2-a

Parameter	h_1	h_2	h_3	ρ_1	ρ_2	ρ_3	ρ_4
Initial estimation error (%)	0.000	0.000	0.000	50.000	10.000	20.000	12.500
Final estimation error (%)	0.000	0.000	0.000	2.312	3.279	1.787	2.593

The results of the optimization process show that the obtained models are progressively matching the true geoelectric model from Table 1, interfaces depths being unaltered in each iteration. The initial mean relative data fitting error of about 25% was reduced to 2% in iteration 2, inversion's convergence being very fast and stable around the optimal solution. It is possible to further reduce the final estimation errors for the model resistivities, by using larger variation bounds (*i.e.*, larger σ_p standard deviations) for these parameters. Figs. 3-a and 3-b present the initial geoelectric model, the final model resulted from inversion and the theoretical effects of the models, in comparison with the observational data.

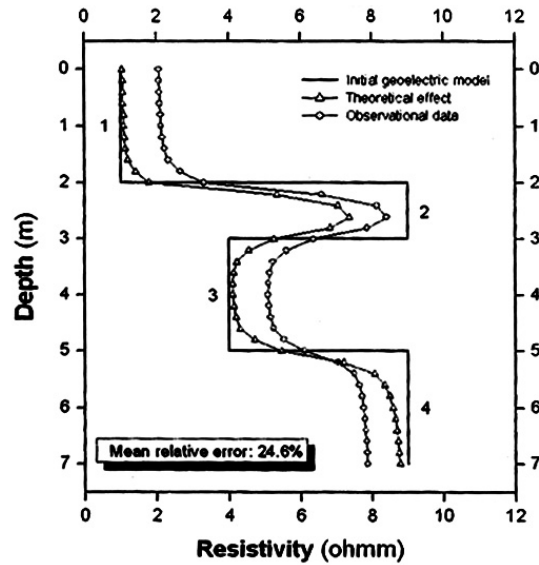


Fig. 3-a – Estimated geolectric model (depths of the separation interfaces are assumed known *a priori*), its theoretical effect and the observational data corresponding to an ideal normal device $AM = 0.2$ m. The initial mean relative data fitting error is 24.6%.

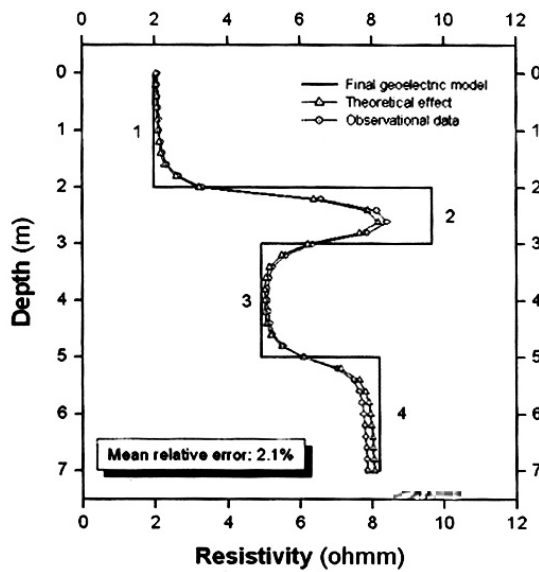


Fig. 3-b – Optimal geolectric model resulted from the probabilistic inversion, its theoretical effect and the observational data corresponding to an ideal normal device $AM = 0.2$ m. The final mean relative data fitting error is 2.1%.

The probabilistic inversion algorithm was also tested in the hypothesis of available *a priori* information about some of the model resistivities. By analyzing the theoretical response of the normal AM = 0.2 m device, it is easily seen that the true resistivities for layers 1, 3 and 4 can be estimated from the dataset itself. Consequently, an initial interpretation model was selected for which these resistivity values were assumed known, then the model was optimized without altering them. Standard deviations of $\sigma_p = 0.001$ were used for the resistivities of layers 1, 3 and 4, the rest of the model parameters having assigned $\sigma_p = 1$ standard deviations. The covariance matrix of the observational data was chosen as diagonal, $C_y = \sigma_y^2 \mathbf{I}$, with $\sigma_y = 0.001$, the inversion operating in MAP mode.

Table 3-a presents the first 8 iterations of the optimization procedure. The values in square brackets, in iteration "0", are referring to the parameters which are assumed known and not allowed to change. It may be observed that the inversion algorithm has improved the initial interpretation model, the mean relative data fitting error e_{MR} being reduced from about 15% to 1% in 5 iterations. The convergence rate was a bit slower with respect to the previous case, the model parameters and the associated estimation errors becoming stable in iteration 5. Further iterations did not refine the model, but the estimation errors can be reduced even more by specifying larger variation bounds for the unknown parameters. The resistivities of layers 1, 3 and 4 were practically unaltered, due to the very small *a priori* variances associated to them. On the other hand, it should be emphasized that the algorithm uses Gaussian probability densities, not uniform ones, so it is possible that the values of *a priori* known parameters might be slightly changed during the inversion.

Table 3-b shows the initial and final estimation errors for the geoelectric model parameters, the final model being substantially improved. The negligible altering of some parameters for which correct starting values were estimated (depth of interface 2 and resistivity of layer 4), without imposing hard constraints, is a specific effect of any automatic interpretation algorithm which tries to find a global optimal solution. Figs. 4-a and 4-b present the initial and final geoelectric models, as well as their theoretical effects in comparison with the observational data.

A final example of automatic interpretation by means of probabilistic inverse modeling concerns the processing of apparent resistivity logs affected by geological noise. For this purpose, the observational data represented by the normal AM = 0.2 m device response were contaminated with a pseudorandom noise of $\pm 10\%$ amplitude. Fig. 5-a shows the estimated geoelectric model, its theoretical effect and one of pseudorandom datasets. Diagonal covariance matrices were used for processing, $C_y = \sigma_y^2 \mathbf{I}$ and $C_p = \sigma_p^2 \mathbf{I}$, with $\sigma_y = 0.1$ and $\sigma_p = 1$. In this case, the probabilistic inversion algorithm behaves like a ridge regression one (Levenberg, 1944; Marquardt, 1963), the ratio $k = \sigma_y / \sigma_p$ measuring the relative importance of the information provided by the data in comparison with the *a priori* information about the model parameters.

Table 3-a

Results of the probabilistic inversion of apparent resistivity data corresponding to an ideal AM = 0.2 m normal device (some of the model resistivities are assumed known *a priori*)

Iteration	Geoelectric model's parameters							e_{RMS} (Ωm)	e_{MR} (%)
	h_1 (m)	h_2 (m)	h_3 (m)	ρ_1 (Ωm)	ρ_2 (Ωm)	ρ_3 (Ωm)	ρ_4 (Ωm)		
0	1.5	3	5.5	[2]	8.5	[5]	[8]	1.129	14.910
1	1.579	3.280	5.149	2.000	7.104	5.000	8.079	0.825	10.046
2	1.992	3.117	5.073	2.000	8.119	5.000	8.081	0.381	2.520
3	1.948	3.086	5.018	2.000	9.089	5.000	8.081	0.155	1.420
4	1.954	3.068	5.013	2.000	9.175	5.000	8.081	0.133	1.218
5	1.955	3.048	5.012	2.000	9.211	5.000	8.081	0.125	1.126
6	1.954	3.054	5.012	2.000	9.196	5.000	8.081	0.127	1.155
7	1.955	3.051	5.012	2.000	9.204	5.000	8.081	0.126	1.138
8	1.955	3.052	5.012	2.000	9.202	5.000	8.081	0.126	1.143

Table 3-b

Parameters estimation errors corresponding to the initial and final models from Table 3-a

Parameter	h_1	h_2	h_3	ρ_1	ρ_2	ρ_3	ρ_4
Initial estimation error (%)	25.000	0.000	10.000	0.000	15.000	0.000	0.000
Final estimation error (%)	2.251	1.602	0.239	0.000	7.890	0.000	1.010

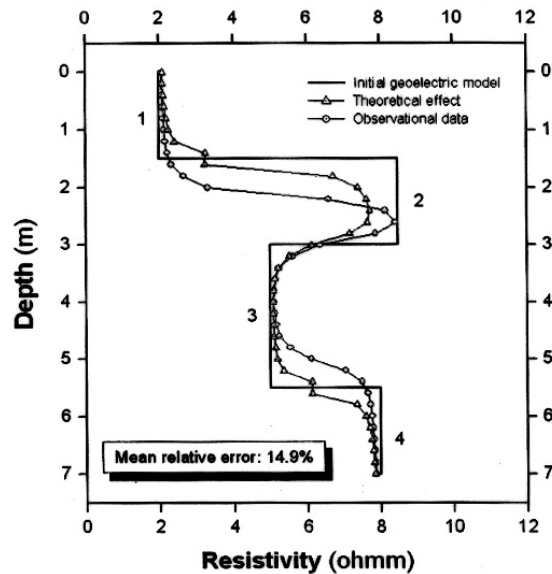


Fig. 4-a – Estimated geoelectric model (some of the model resistivities are assumed known *a priori*), its theoretical effect and the observational data corresponding to an ideal normal device AM = 0.2 m. The initial mean relative data fitting error is 14.9%.

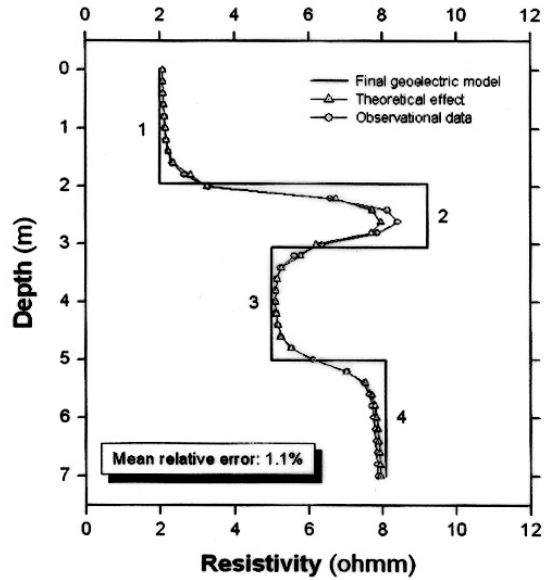


Fig. 4-b – Optimal geoelectric model resulted from the probabilistic inversion, its theoretical effect and the observational data corresponding to an ideal normal device $AM = 0.2$ m. The final mean relative data fitting error is 1.1%.

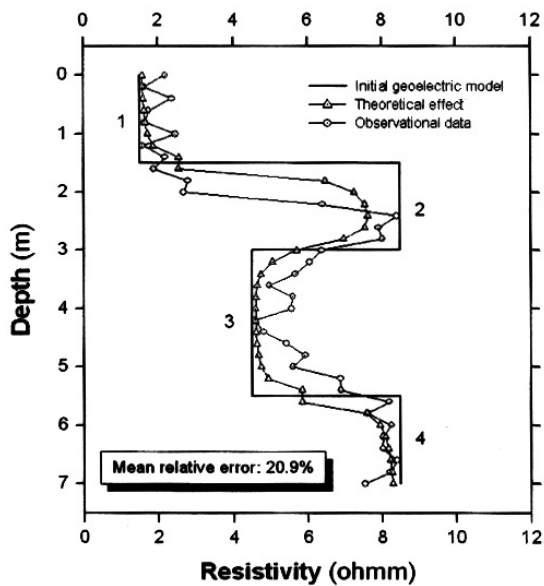


Fig. 5-a – Estimated geoelectric model, its theoretical effect and the observational data ($\pm 10\%$ pseudorandom noise added) corresponding to an ideal normal device $AM = 0.2$ m. The initial mean relative data fitting error is 20.9%.

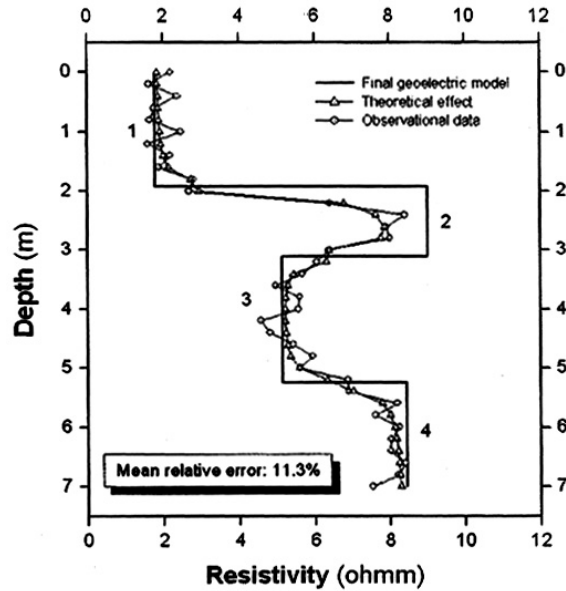


Fig. 5-b – Optimal geoelectric model resulted from the probabilistic inversion, its theoretical effect and the noisy observational data corresponding to an ideal normal device $AM = 0.2$ m. The final mean relative data fitting error is 11.3%.

Table 4-a illustrates the optimization of the initial geoelectric model, described by iteration "0". Due to the noise, the final data fitting errors were larger than in the previous cases, the initial mean relative error e_{MR} of about 21% being reduced to 11% in 3 iterations, when convergence was obtained. Table 4-b presents the parameters estimation errors for the initial and final models, with respect to the true parameters from Table 1. It may be observed that the inversion has considerably optimized the initial model, justifying the use of such probabilistic algorithms for the interpretation of real electric logs affected by the inherent geological noise. Fig. 5-b presents the optimal geoelectric model resulted from inversion and the noisy data fitting by the model theoretical response.

The C_y and C_p matrices used in the probabilistic inversion algorithm are, in fact, covariance matrices, their diagonal elements storing the variances (dispersions) of the probability densities assigned to observational data and, respectively, the model parameters. The diagonal character of these matrices is justified by the lack of any correlation between different data samples or between the model parameters.

Table 4-a

Results of the probabilistic inversion of apparent resistivity data corresponding to an ideal AM = 0.2 m normal device ($\pm 10\%$ pseudorandom noise added to data)

Iteration	Goelectric model's parameters							e_{RMS} (Ωm)	e_{MR} (%)
	h_1 (m)	h_2 (m)	h_3 (m)	ρ_1 (Ωm)	ρ_2 (Ωm)	ρ_3 (Ωm)	ρ_4 (Ωm)		
0	1.5	3	5.5	1.5	8.5	4.5	8.5	1.275	20.929
1	1.536	3.244	5.329	1.549	7.189	5.507	8.252	0.930	15.066
2	1.773	2.496	5.133	1.562	7.707	5.232	8.342	0.976	13.553
3	2.145	2.927	5.245	1.896	9.278	5.601	8.383	0.741	10.318
4	2.129	2.983	5.256	2.210	10.491	5.247	8.401	0.624	11.320
5	2.154	2.980	5.249	2.225	10.688	5.237	8.393	0.624	11.373
6	2.121	2.993	5.259	2.202	10.350	5.245	8.451	0.626	11.323
7	2.159	2.985	5.244	2.234	10.697	5.195	8.394	0.625	11.466
8	2.114	2.999	5.255	2.199	10.240	5.244	8.399	0.626	11.316

Table 4-b

Parameters estimation errors corresponding to the initial and final models from Table 4-a

Parameter	h_1	h_2	h_3	ρ_1	ρ_2	ρ_3	ρ_4
Initial estimation error (%)	25.000	0.000	10.000	25.000	15.000	10.000	6.250
Final estimation error (%)	5.717	0.029	5.099	9.981	2.401	4.875	4.997

4. CONCLUSIONS

An inverse modeling method based on a maximum generality Bayesian approach is proposed as a technique for the interpretation of conventional borehole apparent resistivity logs. The algorithm uses multilayered interpretation models with planar-parallel separation interfaces, being able to considerably improve an initial estimate of the true goelectric model by taking into account the available *a priori* information related to the model parameters or the measured data, as well as the random geological noise.

One of the probabilistic inversion's advantages is represented by the flexibility of selecting the initial goelectric model of the formations crossed by a borehole. In this respect, the inversion's convergence may be obtained through a convenient choice of the model parameters covariance matrix C_p , if no *a priori* information is available. During the optimization, if abnormal values are obtained for some parameters the process can be restarted after changing their corresponding standard deviations, which may lead to convergence and a final solution which is very close to the true goelectric model. Another advantage consists in the possibility of incorporating, in a probabilistic way, physical or dimensional

constraints regarding the model parameters, in order to avoid final interpretation models which are unrealistic from a physical or geological perspective.

The efficiency of the elaborated inverse modeling software has been demonstrated through several examples, but its application possibilities are much more extended and able to cover complex geological situations. Although the software was especially designed for the inversion of conventional borehole apparent resistivity logs, the mathematical and numerical algorithm may be used for the automatic interpretation of any kind of geophysical data, by suitable modification of the $f(\mathbf{p}, \mathbf{x})$ forward modeling component.

REFERENCES

- ANDERSEN, K.E., BROOKS, S.P., HANSEN, M.B. (2003), *Bayesian inversion of geoelectrical resistivity data*. J. Roy. Stat. Soc. B, **65**, Part 3, 619–642.
- DUIJNDAM, A.J. (1988), *Bayesian estimation in seismic inversion. Part I: Principles*. Geophys. Prospect., **36**, 878–898.
- DUIJNDAM, A.J. (1988), *Bayesian estimation in seismic inversion. Part II: Uncertainty analysis*. Geophys. Prospect., **36**, 899–918.
- GOLTSMAN, F.M. (1971), *Statistical interpretation models*. Nauka Publishing House, Moscow (in Russian).
- HOVERSTEN, G.M., MORRISON, H.F. (1982), *Comparison of five least-squares inversion techniques in resistivity sounding*. Geophys. Prospect., **30**, 688–715.
- INMAN, J.R., RYU, J., WARD, S.H. (1973), *Resistivity inversion*. Geophysics, **38**, 1088–1108.
- INMAN, J.R. (1975), *Resistivity inversion with ridge-regression*. Geophysics, **40**, 798–817.
- JACKSON, D.D., MATSUURA, M. (1985), *A Bayesian approach to nonlinear inversion*. J. Geophys. Res., **90**, 581–591.
- JUPP, D.L., VOZOFF, K. (1975), *Stable iterative methods for the inversion of geophysical data*. Geophys. J. Roy. Astr. S., **42**, 67–72.
- LEVENBERG, K. (1944), *A Method for the Solution of Certain Problems in Least Squares*. Q. Appl. Math., **2**, 164–168.
- MARQUARDT, D. (1963), *An Algorithm for Least-Squares Estimation of Nonlinear Parameters*. SIAM J. Appl. Math., **11**, 431–441.
- MOSEGAARD, K., TARANTOLA, A. (2002), *Probabilistic Approach to Inverse Problems*. In: W. Lee, H. Kanamori, P. Jennings and C. Kisslinger, eds., International Handbook of Earthquake & Engineering Seismology, Part A: Academic Press Inc., e-book.
- LANCZOS, C. (1961), *Linear Differential Operators*. Van Nostrand Co., London.
- LOKE, M.H., BARKER, R.D. (1996), *Rapid least-squares inversion of apparent resistivity pseudosections by a quasi-Newton method*. Geophys. Prospect., **44**, 131–152.
- MALINVERNO, A., TORRES-VERDIN, C. (2000), *Bayesian inversion of DC electrical measurements with uncertainties for reservoir monitoring*. Inverse Probl., **16**, 1343–1356.
- NICULESCU, B.M. (2006), *Forward electric modeling – Applications in borehole geophysics*. University of Bucharest Publishing House, Bucharest (in Romanian).
- NICULESCU, B.M. (2006), *Inverse electric modeling Applications in borehole geophysics*. University of Bucharest Publishing House, Bucharest (in Romanian).
- NICULESCU, B.M. (2002), *Contributions to the electric modeling of stratified media. Specific applications in borehole geophysics*. Ph.D. Thesis, University of Bucharest, Bucharest (in Romanian).

- PELTON, W.H., RIJO, L., SWIFT, C.M. (1978), *Inversion of two-dimensional resistivity and induced polarization data*. Geophysics, **43**, 788–803.
- PETRICK, W.R., PELTON, W.H., WARD, S.H. (1977), *Ridge regression applied to crustal resistivity sounding data from South Africa*. Geophysics, **42**, 955–1005.
- POUS, J., MARCUELLO, A., QUERALT, P. (1985), *Inversion with an alternative error function in resistivity measurements*. Geoexploration, **23**, 527–536.
- POUS, J., MARCUELLO, A., QUERALT, P. (1987), *Resistivity inversion with a priori information*. Geophys. Prospect., **35**, 590–603.
- RIJO, L., PELTON, W.H., FEITOSA, E.C., WARD, S.H. (1977), *Interpretation of apparent resistivity data from Apodi Valley, Brazil*. Geophysics, **42**, 811–822.
- TARANTOLA, A. (1987), *Inverse problem theory – methods for data fitting and model parameter estimation*. Elsevier Science Publishing Co., Amsterdam.
- TARANTOLA, A. (2005), *Inverse problem theory and model parameter estimation*. SIAM, Philadelphia, e-book.
- WHITMAN, W.W., TOWLE, G.H., KIM, J.H. (1989), *Inversion of normal and lateral well logs with borehole compensation*. The Log Analyst, **30**, 1–10.
- YANG, F.W., WARD, S.H. (1984), *Inversion of borehole normal resistivity logs*. Geophysics, **49**, 1541–1548.

Received: March 1, 2007

Accepted for publication: May 23, 2007

