Analyse bayésienne par chaîne de Markov de cartographie pôle-pôle. Aspects théoriques et exemples d'application

Jean-Jacques Schott[1], Myriam Schmutz[2], Eric Frugier[1], Vincent Girard[3]

[1] *Ecole et Observatoire des Sciences de la Terre, 5 rue René Descartes, 67084 Strasbourg cedex, France*
[2] *C.E.R.E.G., 3 rue de l'Argonne, 67083 Strasbourg cedex, France*
[3] *E.N.S.P.S., Boulevard Brandt, 67400 Illkirch, France*

Abstract

Bayesian formulation of the inverse problem in electrical methods leads to the estimation of the posterior probability distribution of the unknown parameters, given a prior distribution and the set of measurements. In our method, the posterior distribution is explored with the help of an appropriate Markov chain. This is easily done in 1D modelling but is much more tedious in higher dimensional cases. As a preliminary step before true 2D or 3D bayesian analysis, we have implemented a method based upon 1D modelling. The successive soundings carried out on a given profile are inverted simultaneously, the inversions being linked together by smoothing factors which control the horizontal as well as the vertical variation of the resistivity. Thus, we obtain approximate 2D smooth models, each block of constant resistivity being characterized by a marginal posterior probability density which shows the resolution we can expect. This analysis has been applied to two very different contexts :

- in the frame of a hydrogeological study undertaken by the L.T.H.E., Grenoble, we have carried out measurements in sand and gravel formations draining groundwater ;
- the structure of a landslide located in marl formations (Barcelonnette, Alpes de Haute Provence, France) has been investigated.

Introduction

Nous présentons une application de l'analyse bayésienne à l'inversion de panneaux de mesures électriques pôle-pôle, dans l'approximation où le problème direct est résolu dans un cadre 1D. Rappelons que la formulation bayésienne du problème inverse s'écrit :

$$P_{pd}(\mathbf{X}=\mathbf{r}\,/\,\mathbf{Y}=\mathbf{d})=\frac{P(\mathbf{Y}=\mathbf{d}\,/\,\mathbf{X}=\mathbf{r})P_{ad}(\mathbf{X}=\mathbf{r})}{\sum_{r\ \Omega}P(\mathbf{Y}=\mathbf{d}\,/\,\mathbf{X}=\mathbf{r})P_{ad}(\mathbf{X}=\mathbf{r})} \qquad (1)$$

où **X** est le vecteur des paramètres, prenant ses valeurs dans un domaine Ω de ^{N}, **Y** le vecteur des données, prenant ses valeurs dans un domaine de ^{D}, P_{ad} la probabilité a priori sur **X**, P(**Y** = **d / X** = **r**) la probabilité conditionnelle sur les observables **Y** pour un modèle donné et P_{pd} (**X** = **r / Y** = **r**) la probabilité a posteriori sur les paramètres sachant les observables.

Dans notre cas, **X** = $(X_1, X_2, ..., X_N)^T$, **Y** = $(Y_1, Y_2, ..., Y_D)^T$. X_i est une variable aléatoire discrète prenant ses valeurs dans l'ensemble E = $\{\rho_j\}$, $j = 1,2,..., N_E$ et **Y** prend des valeurs de la forme $(d_1, d_2,...,d_D)$. Avec ces notations, (1) s'écrit explicitement :

$$P_{pd}(\mathbf{X}=\mathbf{r})\,/\,\mathbf{Y}=\mathbf{d})=\frac{P(\mathbf{Y}=\mathbf{d}\,/\,\mathbf{X}=\mathbf{r})\,P_{ad}(\mathbf{X}=\mathbf{r})}{\sum_{j_1,\dots,j_N}P(\mathbf{Y}=\mathbf{d}\,/\,\mathbf{X}=\mathbf{r})P_{ad}(\mathbf{X}=\mathbf{r})} \qquad (2)$$

Classiquement, la loi de probabilité des observables, pour un modèle fixé, s'écrit sous une forme gaussienne (Sen et Stoffa, 1996) :

$$P(\mathbf{Y}=\mathbf{d}_{obs}\,/\,\mathbf{X}) = C\exp\left(-\frac{1}{2\sigma^2}(\mathbf{d}_{obs}-F(\mathbf{X}))^2\right) \qquad (3)$$

où F(**X**) est la fonctionnelle du problème direct, permettant de calculer la résistivité apparente pour un dispositif donné. Nous utilisons la méthode du filtrage linéaire en utilisant l'algorithme proposé par Anderson (1979).

La difficulté, dans cette approche, est de calculer, sur toutes les valeurs possibles des paramètres, la somme figurant au dénominateur de la formule (2). Nous l'avons résolue par l'élaboration d'une chaîne de Markov qui possède la loi a posteriori sur **X** comme loi d'équilibre. Par ailleurs, nous nous intéressons à des modèles à grand nombre de couches, qu'on peut considérer comme la discrétisation d'une variation continue de la résistivité avec la profondeur.

Estimation de la loi a posteriori dans le cas 1D

Si la chaîne de Markov est irréductible et ergodique, la loi d'équilibre est aussi la loi invariante, ce qui se traduit par la relation :

$$\lim_{n\to\infty} P(\mathbf{X}_n=\mathbf{r}\,/\,\mathbf{X}_0=\mathbf{s}) = \lim_{n\to\infty} P_{tr}^{n}(\mathbf{s},\mathbf{r}) = P_{pd}(\mathbf{r}) \qquad (4)$$

où P_{tr} est la loi de transition de la chaîne. Nous définissons la probabilité de transition par la propriété :

$$P_i(\mathbf{s},\mathbf{r}) = P_{pd}(\mathbf{X}=\mathbf{r}\,/\,f_i(\mathbf{X})=f_i(\mathbf{r})) \qquad \text{quand } f_i(\mathbf{r}) = f_i(\mathbf{s})$$

$$P_i(\mathbf{s},\mathbf{r}) = 0 \qquad \text{quand } f_i(\mathbf{r}) \neq f_i(\mathbf{s})$$

où f_i (**r**) = $\left(\rho_{j_1},\ldots\rho_{j_{i-1}},\rho_{j_{i+1}},\ldots,\rho_{j_N}\right)$ correspond à une transition au cours de laquelle seule la résistivité de la couche i est susceptible de varier.

Posant $q_{k_i} = P_{pd}(\mathbf{X}=\mathbf{r} \,/\, f_i(\mathbf{X}) = f_i(\mathbf{r})$, q_{k_i} est donné par la relation :

$$q_{k_i} = \frac{P_{pd}(\mathbf{X} = (\rho_{j_1},\ldots,\rho_{j_{i-1}},\rho_{k_i},\rho_{j_{i+1}},\ldots,\rho_{j_N}))}{\sum_{k_i} P_{pd}(\mathbf{X} = (\rho_{j_1},\ldots,\rho_{j_N}))} \tag{5}$$

La comparaison des équations (2) and (5) montre que le calcul de q_{k_i} est beaucoup plus aisé que celui de P_{pd}. Les transitions de probabilité élémentaires P_i sont recomposées pour construire la probabilité de transition complète de la chaîne.

Rappelons que les modèles employés sont constitués d'un grand nombre de couches (supérieur en général au nombre de données). Dans ces conditions, et compte tenu, de plus, des lois d'équivalence et de suppression, la solution est sous-déterminée, ce qui se traduit par un comportement instable de la chaîne de Markov.

Grandis (1994), Menvielle et Roussignol (1995) ont montré que le processus pouvait être stabilisé si l'on introduit une condition a priori de lissage. L'effet de cette contrainte est de modifier la loi a priori par l'intermédiaire d'un facteur de la forme :

$$\exp\left(-\lambda\left(\mathrm{Log}\frac{\rho_{k_i}}{\rho_{j_{i-1}}}\right)^2 - \lambda\left(\mathrm{Log}\frac{\rho_{j_{i+1}}}{\rho_{k_i}}\right)^2\right) \tag{6}$$

où λ est un facteur de lissage qu'il faut choisir de manière à atteindre un bon compromis entre la finesse de résolution de la courbe ρ = f(z) et le caractère plausible des variations observées d'une couche à l'autre.

La probabilié de transition élémentaire s'écrit alors :

➔ *pour i différent de 1 et N*

$$q_{k_i} = \frac{\exp\left(\frac{-d^2}{2}\right)\exp\left(-\lambda\left(\mathrm{Log}\frac{\rho_{k_i}}{\rho_{j_{i-1}}}\right)^2 - \lambda\left(\mathrm{Log}\frac{\rho_{j_{i+1}}}{\rho_{k_i}}\right)^2\right)}{\sum_{k_i}\exp\left(\frac{-d^2}{2}\right)\exp\left(-\lambda\left(\mathrm{Log}\frac{\rho_{k_i}}{\rho_{j_{i-1}}}\right)^2\right)\exp\left(-\lambda\left(\mathrm{Log}\frac{\rho_{j_{i+1}}}{\rho_{k_i}}\right)^2\right)} \tag{7a}$$

➔ *pour i = 1*

$$q_{k_1} = \frac{w(k_1)\exp\left(\frac{-d^2}{2}\right)\exp\left(-\lambda\left(\mathrm{Log}\frac{\rho_{j_2}}{\rho_{k_1}}\right)^2\right)}{\sum_{k_1} w(k_1)\exp\left(\frac{-d^2}{2}\right)\exp\left(-\lambda\left(\mathrm{Log}\frac{\rho_{j_2}}{\rho_{k_1}}\right)^2\right)} \tag{7b}$$

➔ *pour i = N*

$$q_{k_N} = \frac{v(k_N)\exp\left(\frac{-d^2}{2}\right)\exp\left(-\lambda\left(\mathrm{Log}\frac{\rho_{k_N}}{\rho_{j_{N-1}}}\right)^2\right)}{\sum_{k_N} v(k_N)\exp\left(\frac{-d^2}{2}\right)\exp\left(-\lambda\left(\mathrm{Log}\frac{\rho_{k_N}}{\rho_{j_{N-1}}}\right)^2\right)} \qquad (7c)$$

où **w** et **v** sont les vecteurs propres, à gauche et à droite respectivement, de la matrice $M(i,j) = \exp\left[-\lambda\,\mathrm{Log}\left(\frac{\rho_i}{\rho_j}\right)^2\right]$, associés à la valeur propre maximale.

La chaîne de Markov ainsi créée permet de calculer aisément les lois marginales a posteriori des résistivités, ainsi que les lois jointes deux à deux.

Extension à une analyse à deux dimensions

L'extension à deux dimensions est très simple : il suffit d'introduire une contrainte de lissage à la fois dans le sens horizontal et dans le sens vertical. Le facteur donné par la formule (6) doit être modifié de la manière suivante :

$$\exp\left(-\lambda_1\left(\mathrm{Log}\frac{\rho_{k_i,l}}{\rho_{j_{i-1},l}}\right)^2 - \lambda_1\left(\mathrm{Log}\frac{\rho_{j_{i+1},l}}{\rho_{k_i,l}}\right)^2 - \lambda_2\left(\mathrm{Log}\frac{\rho_{k_i,l+1}}{\rho_{j_{i-1},l}}\right)^2 - \lambda_2\left(\mathrm{Log}\frac{\rho_{j_{i+1},l}}{\rho_{k,l+1}}\right)^2\right)$$

où l et 1+1 désignent 2 sondages successifs.

Le balayage est effectué, à chaque itération, sur l'ensemble des paramètres, donc des résistivités des cellules rectangulaires résultant du découpage horizontal en grand nombre de couches et du découpage vertical lié à l'espacement des points de mesure (c'est à dire, avec le dispositif pôle-pôle, à l'espacement des électrodes d'injection).

Exemples d'application

❖ *Site expérimental de la Côte-St-André (Isère)*

Ce site fait l'objet d'études hydrologiques réalisées par le Laboratoire d'Etude des Transferts en Hydrologie et Environnement et destinées à mettre en évidence les transferts de fluides et de solutés entre la surface du sol et la nappe phréatique située, à cet endroit, à 10 m de profondeur environ.

Le site est localisé dans la vallée de la Bièvre, qui est essentiellement un remplissage de sables et de graviers grossiers, avec des lentilles d'argile. Cette vallée est orientée E-W et bordée de dépôts morainiques.

Nous avons effectué une série de 5 profils de 104 m de longueur chacun, avec un sondage tous les 4 m, les profils étant eux-même espacées de 4 m. Les distances a entre l'électrode d'injection et l'électrode de mesure vont de 0,6 m à 68 m, et sont en progression géométrique de raison 1,6. Les mesures ont été réalisées avec le dispositif multiélectrodes de Iris Instruments comprenant 18 électrodes. Les espacements et le facteur de progression géométriques choisis permettent d'optimiser l'utilisation des électrodes.

La méthode d'analyse décrite plus haut se justifie relativement bien dans cet exemple car les variations latérales sont faibles, quoique présentes. La méthode de discrétisation de la résistivité avec la profondeur et la nécessité du lissage ne permettent pas de situer le toit de la nappe avec précision. Néanmoins, l'estimation qu'on peut en donner est en bon accord avec les mesures fournies par les piézomètres voisins.

❖ *Glissement de terrain de Sauze (Barcelonnette, Alpes de Haute Provence)*

Le glissement de terrain est situé dans les marnes callovo-oxfordiennes de la fenêtre de Barcelonnette. L'objectif de l'étude entreprise est de déterminer la structure du glissement de terrain et la topographie du paléorelief indispensable à la compréhension des modalités du glissement.

L'interprétation des mesures électriques est plus délicate dans cet exemple, en raison des faibles valeurs de résistivité, de leur faible contraste (elles se situent, pour l'essentiel, entre 20 et 50 Ω.m) et de la topographie très accidentée. Nous avons ici effectué 5 profils pôle-pôle de 60 m de longueur, espacés de 4 m. Sur chaque profil, il y a 25 sondages avec des espacements a en progression arithmétique de 2,5 à 12,5 m et un pas de 2,5 m, soit 5 mesures par sondage. L'utilisation de la méthode décrite plus haut est beaucoup moins justifiée dans ce cas en raison de la topographie et de l'importance des variations latérales. De plus, l'espacement maximum adopté est insuffisant. On obtient cependant une image électrique intéressante, corrélée avec les structures observées sur le terrain et fournissant une bonne première approximation pour une modélisation plus adaptée.

Conclusions

La méthode d'analyse bayésienne mise en oeuvre, bien qu'utilisant la solution du problème direct dans le cadre limité du modèle 1D, fournit, par l'intermédiaire des lois marginales a posteriori sur les paramètres, une indication quantitative précieuse sur la résolution que l'on peut espérer à partir de mesures de surface, en même temps qu'une première solution pour une modélisation plus élaborée. Les exemples proposés illustrent son intérêt dans deux situations de complexité très différente. L'étape suivante consistera à adopter comme solution du problème direct, celle d'une méthode numérique véritablement 2D ou 3D. La méthode des éléments finis, en particulier, est bien adaptée à la prise en compte des effets topographiques.

Bibliographie

Anderson W.L., 1979. Computer program. Numerical integration of related Hankel transforms of order 0 and 1 by adaptive digital filtering. Geophysics, 44, 1 287-1 305.

Grandis H., 1994. Imagerie électromagnétique bayésienne par la simulation d'une chaîne de Markov. Thèse, Univ. Paris VII, 250 pp.

Menvielle M. et Roussignol M., 1995. Imagerie électromagnétique bayésienne. C.N.F.G.G., rapport quadriennal 1991-1994, 177-190.

Sen M.K., Stoffa P.L., 1996. Bayesian inference, Gibb's sampler and uncertainty estimation in geophysical inversion. Geophysical Prospecting, 44, 313-350.