

Characterization of the Measurement Error in Time-Lapse Seismic Data and Production Data with an EM Algorithm

Y. Zhao, G. Li and A.C. Reynolds

The University of Tulsa - U.S.A.

e-mail: yong-zhao@utulsa.edu - gaoming-li@utulsa.edu - reynolds@utulsa.edu

Résumé — Caractérisation des erreurs de mesure des données sismiques 4D et des données de production par l'algorithme EM — La caractérisation des erreurs de mesure est cruciale pour l'utilisation de l'approche de Bayes afin de conditionner les modèles de réservoir aux données dynamiques, c'est-à-dire les données sismiques 4D ainsi que les données de production, par l'*history matching* automatique. Dans la littérature, les erreurs de mesure pour chaque type de données sont généralement estimées en appliquant la technique de lissage sur tout le domaine des données engendrant, de ce fait, un sur-lissage des données (en particulier aux alentours des points pour lesquels les données réelles changent de façon drastique), ainsi qu'une surestimation des erreurs de mesure. Cet article présente une nouvelle procédure pour l'estimation de l'erreur de mesure. La méthode développée ici, est basée sur l'algorithme EM (*Expectation-Maximization*) modifié, combiné à un *fit* polynomial mouvant. Cette méthode fournit une estimation de la moyenne et de la covariance des erreurs de mesure. La procédure évite le lissage sur les discontinuités. L'algorithme est appliqué aussi bien aux données sismiques 4D synthétiques qu'aux données de champs ainsi qu'aux données de production. Les résultats sont comparés à ceux obtenus avec des algorithmes de lissage plus standard à fenêtre mouvante. Concernant l'exemple de données synthétiques, la procédure basée sur l'algorithme EM produit des résultats supérieurs à ceux obtenus par des méthodes basées sur une sorte de moyenne mouvante. En ce qui concerne les données de champs, l'EM semble aussi donner un résultat raisonnable.

Abstract — Characterization of the Measurement Error in Time-Lapse Seismic Data and Production Data with an EM Algorithm — The characterization of measurement error is important if one uses a Bayesian approach to condition reservoir models to dynamic data, e.g., time-lapse seismic and production data, by automatic history matching. In the literature, the measurement error for each data type is usually estimated by some smoothing technique in the whole data domain, which often over-smoothes the data (particularly around points where the underlying true data changes sharply) and results in over estimation of the measurement error. This paper presents a new procedure for measurement error estimation. The method is based on a modified EM (*Expectation-Maximization*) algorithm combined with a moving polynomial fit and provides an estimate of the mean and covariance of measurement errors. The procedure avoids smoothing over discontinuities. The algorithm is applied to both synthetic and field time lapse seismic data as well as production data. The results are compared with more standard moving window smoothing algorithms. For the synthetic example, the EM-based process yields results superior to standard smoothing procedures based on some type of moving average. For the field data, EM also appears to give a reasonable result.

INTRODUCTION

In order to characterize reservoir description, make performance predictions, track the movement of injected fluids, characterize the uncertainty in reservoir descriptions and performance predictions and optimize production, it is desirable to integrate all available data including static data (core, well logs) and dynamic production data (wellbore pressure, WOR, GOR or phase production rates) and 4D seismic data. A popular approach to characterize uncertainty involves generating multiple reservoir models (realizations) using randomized maximum likelihood (RML) (Oliver *et al.* 1996, Kitanidis 1995, Reynolds *et al.* 1999). An individual realization is constructed by minimizing an objective function, which is composed of a prior model mismatch term and a data mismatch term. The data mismatch term typically involves dynamic production and/or seismic data. The relative weights of different types of data in the objective function are determined by the covariance matrix of the measurement error. The balance among different types of data affects the suite of realizations obtained, so having a reasonable estimate of the measurement error is important. Estimation of the measurement error is also required if a set of realizations (ensembles) is generated with the ensemble Kalman filter (Evensen 1994, Naevdal *et al.* 2002, Naevdal *et al.* 2003, Evensen 2003).

The underlying assumption is that the measurement error for a particular data type (*e.g.* wellbore pressure versus time measured with a pressure gauge) has much higher frequency than the true underlying signal. Thus, we can smooth the data to estimate the true signal. Subtracting the estimated true signal from the measured data gives an estimate of the measurement error. Note this estimated measurement error can include components of processing errors, *e.g.*, in 4D seismic data, provided the processing errors are of much higher frequency than the underlying true signal. Throughout, our focus is on the estimation of measurement error in 4D seismic impedance change data, which is derived from time-lapse seismic surveys, and production WOR data.

Aanonsen *et al.* (2003) applied a moving average with equal weights to estimate measurement error on both production and seismic data with some success. The moving average method can introduce bias on the estimation when the underlying true signal is not a linear function of time or space. The bias becomes worse when the true signal contains discontinuities, because the simple moving average smoothes out discontinuities and thus some of the true signal is contained in the estimated measurement error. Intuitively, to accurately estimate measurement error in this case, it is necessary to find these discontinuities or group the data into regions to exclude the discontinuities and implement smoothing algorithm within each region.

In this study, we assume the data can be grouped so that data in a group can be approximated with a Gaussian distribution. The data in the whole reservoir is therefore repre-

sented by a Gaussian mixture model (GMM). The mean of each Gaussian model reflects the average value of the signal in that group and the variance reflects the variation of the underlying signal. As measurements in a specific group of data will be similar, a group should never contain points on the opposite side of a “discontinuity” such as a flood front. The task for grouping is now simplified to the problem of finding a Gaussian mixture model to describe the observed data. The Expectation Maximization (EM) algorithm (Hartley 1958, Dempster *et al.* 1977, Meng 1993, Meng 1994, Kung *et al.* 2004) has long been used for estimating the Gaussian Mixture Model parameters based on observation data. The EM algorithm is an iterative procedure and provides an approximation of the maximum likelihood estimate of the Gaussian model parameters using the observed data as well as soft membership information that gives the probability that a certain datum belongs to a certain group (Gaussian).

Since we need to implement smoothing within each group, it is desirable that data in a group are spatially continuous. Moreover, except for sharp changes in the data due, for example, to movements of fronts or abrupt changes in a well’s operating conditions, we expect that two data measured at nearby locations or at slightly different times will be similar in value. However, the traditional EM algorithm considers only the measured values of the data while ignoring the spatial continuity. This can result in groups such that a group contains data of similar value where the spatial coordinates of the data in the group represents many small disconnected regions in space. Although this is not always a significant problem, it can make smoothing more difficult. Allard and Guillot (1999) applied an approximation to the Classification EM algorithm (CEM, Celeux and Govaert 1991) to group irregularly spaced data and recover the spatial correlation of measurements in each groups. This algorithm ensures that the measurements in each group are spatially continuous, however, it requires a prior knowledge of the correct number of groups. In this paper, we propose an EM algorithm with spatial constraints to enhance the spatial continuity of each group. In our implementation of the spatial EM algorithm, grouping quality coefficients are applied to automatically delete some spatially scattered groups in order to find a parsimonious grouping of the measurements. After grouping, smoothing is done with a moving window quadratic fit within each individual group. The smoothed data represents the estimate of the underlying true signal and the difference of the smoothed data and original observed data is the estimate of measurement error.

In the process of revising this paper, we learned of geo-statistical filtering methods (factorial Kriging, Coleou *et al.* 2002 and Abreu *et al.* 2005), which directly separates the non-repeatable processing errors from measurements. The algorithm has been applied to remove process errors in 4D seismic surveys to enhance the 4D signature. The basic

idea of this algorithm is to decompose the variogram for each seismic survey into two parts: the common variogram shared by the two surveys and the residual variograms for the non-repeatable errors of each survey. Then the common part of the measurements from two surveys representing the true signal and the spatially independent residuals from each survey are separated using co-kriging. The algorithm seems to work well in the examples presented in the literature, but we have not yet applied this algorithm to estimate measurement errors.

1 NON-SEQUENTIAL GAUSSIAN MIXTURE MODEL AND NON-SEQUENTIAL EM ALGORITHM

The measurements studied in most EM applications are non-sequential *i.e.*, the measurements are not spatially or temporally related. However, the measurements discussed in this paper (4D seismic data and production data) are spatially or temporally measured. Except for sharp changes in the underlying signal, we want to constrain the grouping of data to make it more likely that data measured at spatial or time points close together will be in the same group. For convenience, in our discussion, we will refer to both spatial and temporal constraints as a spatial constraint. We denote the spatially measured data as

$$D = \{d_{a,i}, d_{s,i}, i = 1, N\} \quad (1)$$

Here, $d_{a,i}$ and $d_{s,i}$, respectively, are the value and spatial coordinates (or time coordinate) of the i -th measurement. We assume these measurements can be divided into M groups, and the measured values in each group can be modeled by a Gaussian, *i.e.*, the measurements are sampled from a Gaussian mixture model. As in this section, we ignore the spatial coordinate(s) of the data, we temporarily define D as

$$D = \{d_{a,i}, i = 1, N\} \quad (2)$$

as a set of N realizations from a Gaussian mixture model.

We assign each Gaussian model a label (indicator), *i.e.* c_j is the label for the j -th Gaussian model and define

$$C = \{c_j, j = 1, M\} \quad (3)$$

A membership indicator can be defined for each datum to indicate the group it belongs to, *i.e.*, we let

$$Z = \{z_i, i = 1, N\} \quad (4)$$

where z_i can take M discrete values in C ($z_i \in C$). Thus, if $z_i = c_j$, the i -th datum ($d_{a,i}, d_{s,i}$) belongs to the j -th group. Although each measurement must be a sample from only one of these groups, the set of Gaussian parameters (Θ) for all the groups are to be estimated by trying to maximize

an appropriate log-likelihood function discussed later. The parameters for the Gaussian mixture model are given by

$$\Theta = \{(\pi_j, \mu_j, \sigma_j^2), j = 1, M\} \quad (5)$$

where μ_j and σ_j^2 , respectively, denote the mean and variance of the j -th Gaussian, and π_j is the probability (mixing proportion) of the j -th Gaussian model. We must have

$$\sum_{j=1}^M \pi_j = 1 \quad (6)$$

Note because of Equation 6, we can delete one of the proportions, say π_M , from the set of parameters to be estimated.

The maximization step is only feasible if we have approximate values for the membership indicator, *i.e.*, the z_i 's. More specifically, we define an $N \times M$ membership matrix H with the entry in the i -th row and j -th column defined by

$$h_i^j(\Theta) = P(z_i = c_j | d_{a,i}, \Theta) \quad (7)$$

where P denotes a probability. Note that $h_i^j(\Theta)$ represents the probability that the i -th measurement belongs to the j -th group when $d_{a,i}$ and the Gaussian model, Θ , are given. Here, we only consider the case that we group by the measured value of the data, $d_{a,i}$; spatial information will be considered later. We let δ_i^j be a random indicator variable which is equal to 1 when the i -th datum belongs to the j -th group and zero otherwise, *i.e.*,

$$\delta_i^j = \begin{cases} 1 & i\text{-th datum is from } j\text{-th group} \\ 0 & i\text{-th datum is not from } j\text{-th group} \end{cases} \quad (8)$$

As shown in Hastie *et al.* (2001) and Kung *et al.* (2004), Equation 7 can be written as

$$h_i^j(\Theta) = P(\delta_i^j = 1 | d_{a,i}, \Theta) = E[\delta_i^j] \quad (9)$$

where E denotes expectation. The last equation provides the reason why the calculation of h_i^j is called the expectation step in the EM algorithm.

Given only a known set of measurements D , the membership Z and the parameters Θ of Gaussian mixture model are both unknown. In order to estimate Z , Θ has to be known, and vice versa. Thus our problem is a ‘‘missing data problem’’. The EM algorithm can be used to solve this missing data problem.

If we consider only the measured values of data, the probability density function (pdf) for a Gaussian mixture model is given by the following equation:

$$P(d_{\text{obs}} | \Theta) = \sum_{j=1}^M \pi_j P(d_{\text{obs}} | \mu_j, \sigma_j) \quad (10)$$

This Gaussian mixture model is composed of M Gaussian models and $P(d_{\text{obs}} | \mu_j, \sigma_j)$ is the pdf of the j -th Gaussian with

mean μ_j and standard deviation σ_j for the random variable d_{obs} . Note the set of $d_{a,i}$'s present realizations (samples) or d_{obs} .

These data, can be divided into groups according to which Gaussian model they are sampled from. For example, when a datum $d_{a,i}$ is sampled from the j -th Gaussian model, we say that the datum $d_{a,i}$ belongs to the j -th group so $z_i = c_j$.

Since π_j represents the probability of the j -th Gaussian model, when a datum is sampled from the Gaussian mixture model Θ , the probability that this datum is sampled from the j -th model is π_j . Therefore, we have:

$$\pi_j = P(\delta_i^j = 1|\Theta), \quad j = 1, M \quad (11)$$

With a given model $\Theta = \{(\pi_j, \mu_j, \sigma_j), j = 1, M\}$, the probability density function for the i -th datum is

$$\begin{aligned} P(d_{a,i}|\Theta) &= \sum_{j=1}^M P(d_{a,i}, \delta_i^j = 1|\Theta) \\ &= \sum_{j=1}^M P(\delta_i^j = 1|\Theta)P(d_{a,i}|\delta_i^j = 1, \Theta) \end{aligned} \quad (12)$$

Following Kung *et al.* (2004), the log-likelihood function to be maximized for non-sequential measurements is

$$\begin{aligned} L(\Theta|D) &= \ln \left[P(D|\Theta) \right] \\ &= \sum_{i=1}^N \ln \left[\sum_{j=1}^M \pi_j P(d_{a,i}|\delta_i^j = 1, \Theta_j) \right] \end{aligned} \quad (13)$$

Equation (13) is the log-likelihood of model parameters Θ given measurement D . Note in this section, D only contains the value part of the measurements. The EM algorithm starts from an initial guess Θ_0 of the model Θ , and gradually increases the log likelihood function by updating the model through an expectation step and a maximization step.

Suppose the model at the n -th iteration is Θ_n , in the expectation step, the expression for each entry of the membership matrix $(h_i^j)_n$ defined by Equation 7 can be derived by using Bayes theorem and Equation 11 as follows:

$$\begin{aligned} (h_i^j)_n &= \frac{P(\delta_i^j = 1|d_{a,i}, \Theta_n)}{P(\delta_i^j = 1, d_{a,i}|\Theta_n)} \\ &= \frac{P(d_{a,i}|\Theta_n)}{P(d_{a,i}|\delta_i^j = 1, \Theta_n)P(\delta_i^j = 1|\Theta_n)} \\ &= \frac{P(d_{a,i}|\Theta_n)}{\pi_j P(d_{a,i}|\delta_i^j = 1, \Theta_n)} \\ &= \frac{1}{\sum_{k=1}^M \pi_k P(d_{a,i}|\delta_i^k = 1, \Theta_n)} \end{aligned} \quad (14)$$

In the maximization step, the $(n+1)$ -th model Θ_{n+1} is proposed by maximizing the Q function, *i.e.*,

$$\Theta_{n+1} = \text{Argmax}_{\Theta} Q(\Theta|\Theta_n) \quad (15)$$

where from Kung *et al.* (2004), the Q function can be defined and expanded as

$$\begin{aligned} Q(\Theta|\Theta_n) &= \sum_{i=1}^N \sum_{j=1}^M (h_i^j)_n \ln [P(d_{a,i}, \delta_i^j = 1|\Theta)] \\ &= \sum_{i=1}^N \sum_{j=1}^M (h_i^j)_n \left\{ \ln \left[\frac{\pi_j}{\sigma_j \sqrt{2\pi}} \exp \left(- \frac{(d_{a,i} - \mu_j)^2}{2(\sigma_j)^2} \right) \right] \right\} \end{aligned} \quad (16)$$

The above Q function is the expectation of the log-likelihood function (Equation 13) with respect to the group indicator Z given the estimate of Θ_n (Kung *et al.* 2004). With a known model for Θ_n , any new model Θ_{n+1} that satisfies $Q(\Theta_{n+1}|\Theta_n) \geq Q(\Theta_n|\Theta_n)$ also gives a higher value of the log-likelihood function, *i.e.* $L(\Theta_{n+1}|D) \geq L(\Theta_n|D)$. Thus, given Θ_n and associated entries of the membership matrix, we maximize the Q function to find Θ_{n+1} in order to increase the log-likelihood.

The maximizer of the Q function satisfies the following three equations:

$$\frac{\partial Q(\Theta|\Theta_n)}{\partial \mu_k} = 0, \quad k = 1, M \quad (17)$$

$$\frac{\partial Q(\Theta|\Theta_n)}{\partial \sigma_k} = 0, \quad k = 1, M \quad (18)$$

$$\frac{\partial}{\partial \pi_k} \left[Q(\Theta|\Theta_n) + \lambda \left(\sum_{l=1}^M \pi_l - 1 \right) \right] = 0, \quad k = 1, M \quad (19)$$

Note that in the last equation, a Lagrangian multiplier λ is applied. From the preceding three equations, we have the following solution for the maximization step:

$$\begin{aligned} \mu_{j,n+1} &= \frac{\sum_{i=1}^N (h_i^j)_n d_{a,i}}{\sum_{i=1}^N (h_i^j)_n} \\ \sigma_{j,n+1}^2 &= \frac{\sum_{i=1}^N (h_i^j)_n (d_{a,i} - \mu_{j,n})^2}{\sum_{i=1}^N (h_i^j)_n} \\ \pi_{j,n+1} &= \frac{1}{N} \sum_{i=1}^N (h_i^j)_n \end{aligned} \quad (20)$$

These three equations are used to update the model in the maximization step to obtain Θ_{n+1} using the membership matrix H_n , which is the output of the expectation step. In order to group the data using the spatial information, the spatial EM is considered next, after which the step by step procedure will be described.

2 SPATIAL GAUSSIAN MIXTURE MODEL AND SPATIAL EM ALGORITHM

The traditional EM algorithm groups data according only to their values, while ignoring the spatial relationship between them. When this algorithm is applied to spatially correlated data, it gathers all the data with similar values together

regardless of their locations. As noted in more detail earlier, this can create groups which are highly discontinuous spatially, which can make smoothing within groups tenuous. Moreover, in most cases we expect data measured at close locations to have similar values. To overcome this disadvantage in the traditional EM algorithm, two strategies have been proposed to construct a spatially-constrained EM in the literatures. One strategy is to smooth the membership matrix h_i^j (Diplaros *et al.* 2004) group by group (for different j 's). The basic idea of smoothing h_i^j is that when two data points are spatially close to each other, they tend to fall into the same group, which means they should have a similar probability (h_i^j) for being in any particular group (or are likely to be samples from the same Gaussian model). In the spatially-constrained EM algorithm used by Diplaros *et al.* (2004), a smoothing step is added between the E-step and the M-step. The other strategy consists of adding a spatial penalty term to the log-likelihood function (Neighborhood EM algorithm, Ambroise and Govaert 1995). Since the penalty term of the spatial neighborhood EM algorithm does not contain the Gaussian mixture model parameters, Θ , defined in the previous section, the maximization step of this algorithm is equivalent to the traditional EM algorithm, but the membership matrix is modified to include the spatial information. This is equivalent to smoothing the membership matrix using the spatial information implicitly. Our method for imposing a spatial constraint is somewhat similar to the Neighborhood EM algorithm, but the way we incorporate the spatial information is different. Moreover, our final algorithm incorporates a method to delete groups so we do not need to know the number of groups *a priori*.

In order to include spatial information in the Gaussian mixture model, we need to define a spatial PDF for the coordinates of each measurement, because we wish the measurements in a specific group to be close in the measured value, and also wish a group to be spatially continuous. Suppose the model parameters can be rewritten as

$$\Theta = \{\Theta_{a,j}, \Theta_{s,j}, j = 1, M\}$$

where $\Theta_{a,j} = \{\pi_j, \mu_j, \sigma_j^2\}$, and $\Theta_{s,j}$ represents the model parameters of the j -th spatial PDF:

$$\Theta_{s,j} = \{P_j, r_0, j = 1, M\} \quad (21)$$

Here, $P = \{P_1, P_2, \dots, P_M\}$ are the M partitions of measurements associated to the group indicator Z , *i.e.* $P_j = \{i | z_i = c_j\}$. Here, r_0 is a distance weighting factor used to construct spatial PDF. In this paper, r_0 is fixed as we have been unable to construct a stable algorithm for estimating it directly.

Suppose P_j contains the single entry i_0 , then we define the j -th spatial Gaussian PDF by

$$P(d_{s,i} | \delta_i^j = 1, \Theta_{s,j}) = \frac{1}{2\pi r_0^2} \exp \left[-\frac{\|d_{s,i} - d_{s,i_0}\|^2}{2r_0^2} \right] \quad (22)$$

For a general case, the spatial PDF is defined by

$$P(d_{s,i} | \delta_i^j = 1, \Theta_{s,j}) = \frac{1}{N_j} \sum_{k \in P_j} \left\{ \frac{1}{2\pi r_0^2} \exp \left[-\frac{\|d_{s,i} - d_{s,k}\|^2}{2r_0^2} \right] \right\} \quad (23)$$

where N_j is the elements in the set P_j .

In Figure 1, the annular region (gray colored) represents the spatial coordinates of one group of data. The inside and outside of this annular are the spatial coordinates of two other groups. When the spatial PDF of Equation 23 is evaluated at point A ($d_{s,i}$ is at A), and the annular region shaded gray is the j -th model, only the measurements inside both the circle (radius of $3r_0$) and gray-shaded region will contribute significantly to the spatial PDF evaluated at $d_{s,i}$.

With the definition of the value and spatial PDFs, and assuming the spatial coordinate and the value of the measurements are independent, the spatial PDF for the random vector (d_a, d_s) evaluated at $(d_{a,i}, d_{s,i})$ is defined by

$$P(d_{a,i}, d_{s,i} | \Theta) = \sum_{j=1}^M \pi_j P(d_{a,i} | \delta_i^j = 1, \Theta_{a,j}) P(d_{s,i} | \delta_i^j = 1, \Theta_{s,j}) \quad (24)$$

To integrate the spatial constraint, the membership matrix H at the n -th iteration is redefined in accordance with Equations 14 and 24 by

$$(h_i^j)_n = P(\delta_i^j = 1 | d_i, \Theta_n) = \frac{(\pi_j)_n P(d_{a,i} | \delta_i^j = 1, \Theta_{a,j,n}) P(d_{s,i} | \delta_i^j = 1, \Theta_{s,j,n})}{\sum_{k=1}^M (\pi_k)_n P(d_{a,i} | \delta_i^k = 1, \Theta_{a,k,n}) P(d_{s,i} | \delta_i^k = 1, \Theta_{s,k,n})} \quad (25)$$

Here $\Theta_{a,j,n}$ ($(\pi_j)_n, (\mu_j)_n, (\sigma_j^2)_n$) and $\Theta_{s,j,n}$ ($(P_j)_n$) represent the current estimates of model parameters.

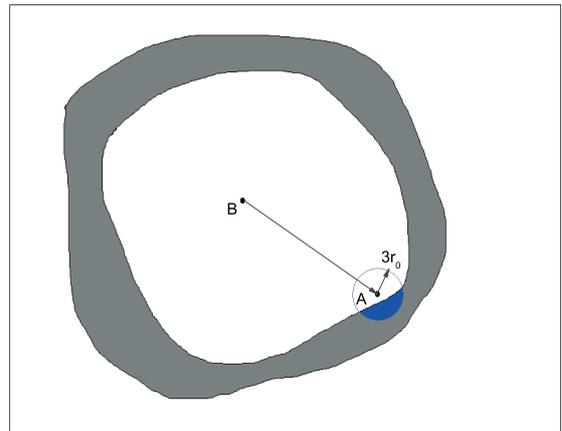


Figure 1

The example of annular region to show the spatial PDF.

From Equation 23, the spatial PDF defined at the current model of Θ_n can be written as:

$$P(d_{s,i}|\delta_i^j = 1, \Theta_{s,j,n}) = \frac{1}{(N_j)_n} \sum_{k \in (P_j)_n} \left\{ \frac{1}{2\pi r_0^2} \exp \left[-\frac{\|d_{s,i} - d_{s,k}\|^2}{2r_0^2} \right] \right\} \quad (26)$$

With the PDF of Equation 24, the Q function is defined by

$$\begin{aligned} Q(\Theta|\Theta_n) &= \sum_{i=1}^N \sum_{j=1}^M \{ (h_i^j)_n \ln[\pi_j(\Theta) P(d_{a,i}|\delta_i^j = 1, \Theta_{a,j}) \\ &\quad P(d_{s,i}|\delta_i^j = 1, \Theta_{s,j})] \} \\ &= \sum_{i=1}^N \sum_{j=1}^M \{ (h_i^j)_n \{ \ln[P(d_{a,i}|\delta_i^j = 1, \Theta_{a,j})] + \\ &\quad \ln[P(d_{s,i}|\delta_i^j = 1, \Theta_{s,j})] + \ln[\pi_j] \} \} \end{aligned} \quad (27)$$

If we define:

$$Q_a(\Theta|\Theta_n) = \sum_{i=1}^N \sum_{j=1}^M \{ (h_i^j)_n \{ \ln[P(d_{a,i}|\delta_i^j = 1, \Theta_{a,j})] + \ln[\pi_j] \} \} \quad (28)$$

$$Q_s(\Theta|\Theta_n) = \sum_{i=1}^N \sum_{j=1}^M \{ (h_i^j)_n \{ \ln[P(d_{s,i}|\delta_i^j = 1, \Theta_{s,j})] \} \} \quad (29)$$

then

$$Q(\Theta|\Theta_n) = Q_a(\Theta|\Theta_n) + Q_s(\Theta|\Theta_n) \quad (30)$$

In Equation 30, the division of the two parts of the Q function are according to their relationship with the maximization step. As discussed in detail in Section 5, in the optimization step we maximize only $Q_a(\Theta|\Theta_n)$ to find the new values of $\{\pi_{j,n+1}, \mu_{j,n+1}, \sigma_{j,n+1}\}$, $j = 1, 2, \dots, M$ for the next iteration using Equation 20. $Q_s(\Theta|\Theta_n)$ is held constant based on the partition determined from the $(h_i^j)_n$'s at the old iteration. Because of this, we can not prove that the log-likelihood function is non-decreasing from iteration to iteration, but we have tested the algorithm on many synthetic examples and have consistently obtained good results.

3 SPATIAL EM ALGORITHM WITH GROUP QUALITY COEFFICIENT

The above spatial EM algorithm does not account for the uncertainty in the number of groups. With this algorithm we find that the number of groups tends to be conserved. If we know the number of groups, the algorithm is similar to the Allard and Guillot (1999) clustering EM algorithm. From iteration to iteration in most cases, we need to estimate the number of the groups in the observations. Richardson and Green (1996) use a reversible jump Markov chain Monte Carlo (McMC) method to find the optimal number of

groups. Because the McMC method is very computationally demanding, we use another approach to find the number of the groups.

We start from a fairly large number of groups, and apply additional multipliers in the membership matrix. The membership matrix is modified to

$$h_i^j(\Theta_n) = \frac{(\pi_j)_n (F_j)_n (A_i^j)_n (S_i^j)_n}{\sum_{j'=1}^M (\pi_{j'})_n (F_{j'})_n (A_i^{j'})_n (S_i^{j'})_n} \quad (31)$$

where

$$\begin{aligned} (A_i^j)_n &= P(d_{a,i}|\delta_i^j = 1, \Theta_{a,j,n}) \\ &= \frac{1}{\sqrt{2\pi(\sigma_j^2)_n}} \exp \left[-\frac{(d_{a,i} - (\mu_j)_n)^2}{2(\sigma_j^2)_n} \right] \end{aligned} \quad (32)$$

$$\begin{aligned} (S_i^j)_n &= P(d_{s,i}|\delta_i^j = 1, \Theta_{s,j,n}) \\ &= \frac{1}{(N_j)_n} \sum_{k \in (P_j)_n} \left\{ \frac{1}{2\pi r_0^2} \exp \left[-\frac{\|d_{s,i} - d_{s,k}\|^2}{2r_0^2} \right] \right\} \end{aligned} \quad (33)$$

$(F_j)_n$ is a variable used to indicate the quality of the j -th group at the n -th iteration. It is used to cause groups of low quality to be absorbed by groups of higher quality so that the number of groups will be gradually decreased until convergence. $(F_j)_n$ is defined based on a grouping score matrix $(E_n)_{M \times N}$ at the n -th iteration:

$$(E_i^j)_n = \frac{(N_j)_n (S_i^j)_n}{\sum_{j'=1}^M (N_{j'})_n (S_i^{j'})_n} \quad (34)$$

Note that $(E_i^j)_n$ is a number between 0 and 1. From Equation 33, Equation 34 can be rewritten as

$$(E_i^j)_n = \frac{\sum_{k \in (P_j)_n} \left\{ \exp \left[-\frac{\|d_{s,i} - d_{s,k}\|^2}{2r_0^2} \right] \right\}}{\sum_{k \in (P_n)} \left\{ \exp \left[-\frac{\|d_{s,i} - d_{s,k}\|^2}{2r_0^2} \right] \right\}} \quad (35)$$

If the i -th measurement is partitioned to the j -th group, there are two types of situations:

- Most points within a distance of $5r_0$ from the i -th measurement fall into the j -th group. In this case, Equation 35 will give a value close to 1. This type of measurement exists inside the spatially continuous regions of the j -th group.
- Most of the points within a distance of $5r_0$ from the i -th measurement do not fall into the j -th group. In this case, Equation 35 will yield a value much smaller than 1. This type of measurement exists near the boundary of the j -th group.

We view groups which have a large percentage of their data satisfying the first situation to be of higher quality than groups which have a large percentage of their data satisfying the second situation. Based on this, the group quality

coefficient for the j -th group is defined as the average of the grouping scores of the measurements in the j -th group, *i.e.*, we define the group quality coefficients by

$$(F_j)_n = \frac{a}{(N_j)_n} \sum_{k \in (P_j)_n} (E_k^j)_n \quad (36)$$

for $j = 1, 2, \dots, M$, where

$$a = \left(\sum_{j=1}^M \left[(\pi_j)_n \frac{1}{(N_j)_n} \sum_{k \in (P_j)_n} (E_k^j)_n \right] \right)^{-1} \quad (37)$$

is a normalizing constant which ensures that

$$\sum_{j=1}^M (\pi_j)_n (F_j)_n = 1 \quad (38)$$

The preceding equation ensures that

$$P(d_a, d_s | \Theta_n) = \sum_{j=1}^M (\pi_j)_n (F_j)_n P(d_a | \delta_i^j = 1, \Theta_{a,j,n}) P(d_s | \delta_i^j = 1, \Theta_{s,j,n}) \quad (39)$$

defines a conditional PDF for the random vector (d_a, d_s) .

4 SPATIAL EM ALGORITHM WITH GROUP QUALITY COEFFICIENT, IMPLEMENTATION

Spatial EM algorithm starts from an initial grouping and iteratively adjusts the model parameters in each iteration until the grouping stops changing and the proportions stop changing. Here, we discuss steps of the spatial EM algorithm using the group quality coefficients.

4.1 Initialization

The initialization step is used to construct the first membership matrix H_1 . To do so, we simply divide the data into a fairly large number of groups. This can be done either by value, by spatial location or randomly. Since we have partitioned the data into M initial groups, we know $(Z)_1$, the set of initial values for the z_i 's as well as the initial partition, $(P)_1$. Initially, we set $\{(F_j)_1 = 1.0, j = 1, M\}$. All the other parameters are constructed from the initial values of the z_i 's or the set of $(P_j)_1$'s, $j = 1, 2, \dots, M$. For $j = 1, 2, \dots, M$, we let $(N_j)_1$ denote the number of elements in the set $(P_j)_1$. In particular, $\Theta_{a,j,1}$ can be calculated as

$$\begin{aligned} (\pi_j)_1 &= \frac{(N_j)_1}{N} \\ (\mu_j)_1 &= \frac{1}{(N_j)_1} \sum_{k \in (P_j)_1} d_{a,k} \\ (\sigma_j^2)_1 &= \frac{1}{(N_j)_1} \sum_{k \in (P_j)_1} (d_{a,k} - (\mu_j)_1)^2 \end{aligned} \quad (40)$$

With the initial guess of the partition $\{(P_j)_1, j = 1, 2, \dots, M\}$, spatial PDF term can be calculated as:

$$(S_i^j)_1 = \frac{1}{(N_j)_1} \sum_{k \in (P_j)_1} \left\{ \frac{1}{2\pi r_0^2} \exp \left[-\frac{\|d_{s,i} - d_{s,k}\|^2}{2r_0^2} \right] \right\} \quad (41)$$

4.2 Membership Matrix Update (E-Step)

With the above definitions, we can start the expectation step for the first iteration by calculating the initial membership matrix H_1 .

In the expectation step, the probabilities of each datum belonging to each group are evaluated based on the current model, which is the membership matrix shown in Equation 31. Here, $H_n = \{(h_i^j)_n\}$ is evaluated at:

$$\{(\pi_j)_n, (\mu_j)_n, (\sigma_j)_n, (S_i^j)_n, (F_j)_n, j = 1, M, i = 1, N\}$$

which are updated in the maximization step of the previous iteration. For H_1 , set $n = 1$ in the preceding equation.

4.3 Model Parameter Update (M-Step)

In the maximization step, the model parameters are updated using the membership matrix evaluated in the E-step. Here we calculate the updated parameter $\Theta_{a,n+1}$ using Equation 20.

Next, the partitions are updated from the current membership matrix H_n . The new group indicators $(Z)_{n+1}$ can be updated in two ways:

- MAP estimate: $(z_i)_{n+1}$ is calculated as

$$(z_i)_{n+1} = \ell$$

where

$$\ell = \text{Argmax}_j [(h_i^j)_n].$$

- Stochastic estimate: $(z_i)_{n+1}$ can be evaluated stochastically by sampling the cumulative distribution function for the i -th row of the current membership matrix $\{(h_i^j)_n, j = 1, 2, \dots, M\}$.

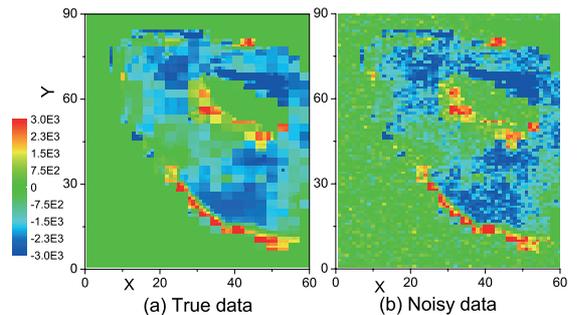


Figure 2

Noisy 4D acoustic impedance change from PUNQ3-D.

In the examples presented here, we use stochastic grouping because computational experiments indicate that it is more robust than grouping based on the MAP estimate of the indicator variable. Given $(Z)_{n+1}$, find the new partition $(P_j)_{n+1}$, from which we can evaluate $\{(F_j)_{n+1}$ and $(S_i^j)_{n+1}, j = 1, M\}$ using Equations 26, 34 and 36. Up to now, all the parameters that are necessary for calculating the new membership matrix H_{n+1} are evaluated. Then we go to the E-step of iteration $n + 1$.

After the model parameter update, we delete some groups with very small group probabilities, if no data tends to fall into them. To do this, the groups are sorted by their proportions, so that $\pi_1 \geq \pi_2 \geq \dots \geq \pi_M$. We delete groups $j' = j_0, j_0 + 1, \dots, M$, if group j_0 satisfies the following two criteria:

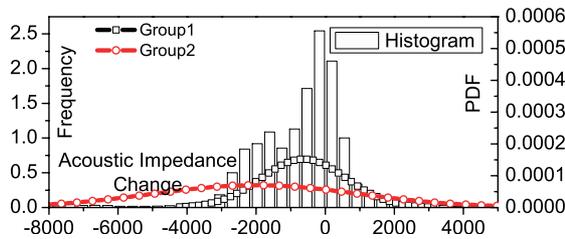
- For $j' = j_0, j_0 + 1, \dots, M, (z_i)_{n+1} \neq C_{j'}$ for any i ;
- $\pi_{j_0} \leq \epsilon_g$, where ϵ_g is a small value set by the user (*i.e.* 0.0001).

Under these conditions, groups j_0 to M have a negligible effect on the Q function and are not expected to increase in size. Thus, we delete these groups and set $M = j_0 - 1$.

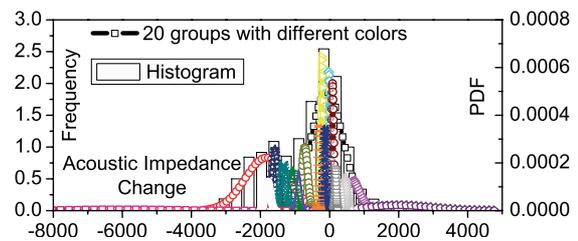
4.4 Stopping Criteria

We use the following two conditions for the criteria to terminate the iterative process:

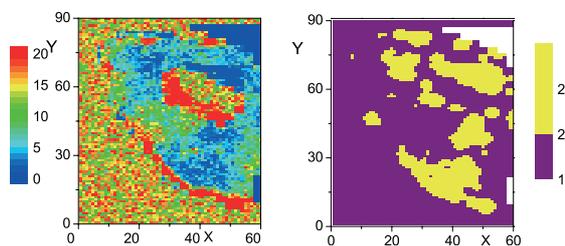
- No group is deleted in the current iteration;
- $\frac{\|\pi_j^{n+1} - \pi_j^n\|}{\|\pi_j^n\|} < \epsilon_1$ for $j = 1, 2, \dots, M$; normally, we use $\epsilon_1 = 10^{-4}$.



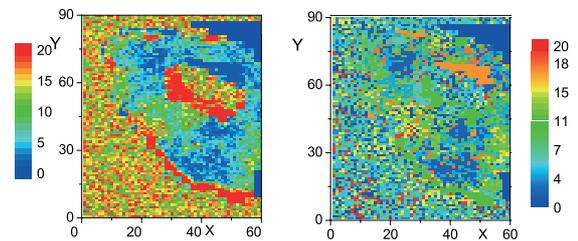
(a) Histogram and final PDFs



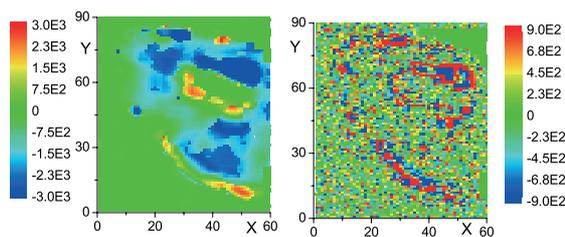
(a) Histogram and final PDFs



(b) Initial groups (left) and final groups (right)



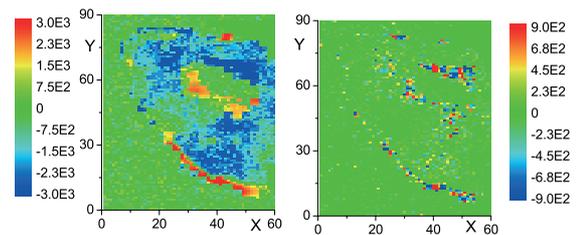
(b) Initial groups (left) and final groups (right)



(c) Estimated true signal (left) and estimated measurement error (right)

Figure 3

EM grouping of synthetic data, membership matrix smoothing.



(c) Estimated true signal (left) and estimated measurement error (right)

Figure 4

EM grouping of synthetic data, no F_j used, initially sorted by measured values into 20 groups, 20 final groups.

5 EXAMPLES

Synthetic 4D acoustic impedance change is used to illustrate the utility of the improved spatial EM algorithm for grouping spatially measured data. The algorithm is also applied to a field 4D acoustic impedance change data. In the synthetic case, two types of initializations are discussed:

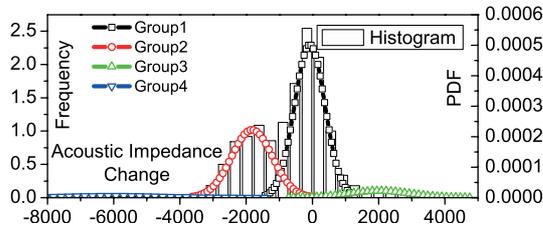
- random initialization, in which the initial grouping of measurements is generated randomly;
- value initialization, in which the measurements are evenly divided into a number of groups by their measured values.

In both cases, the initial number of groups is specified. Results obtained with F_j (improved spatial EM), without F_j ($F_j = 1$) and the Diplaros *et al.* (2004) algorithm are also compared in the synthetic case.

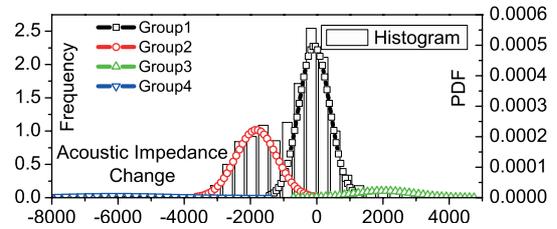
For all cases, we use $r_0 = 2.0$ (spatially with correlation range of about 5 gridblocks), which is a reasonable value and works in all synthetic cases that we have ever investigated. For the synthetic example considered here, we have tried different values for r_0 , and virtually identical results are obtained for the $r_0 = 2$ and $r_0 = 3$ cases, but the results were significantly less accurate with $r_0 = 5$. In order to separate the true signal and measurement error, quadratic fitting is applied within each group using a 11×11 moving window for 2-D cases and a moving window of length 21 for the 1-D WOR case.

5.1 Synthetic Case

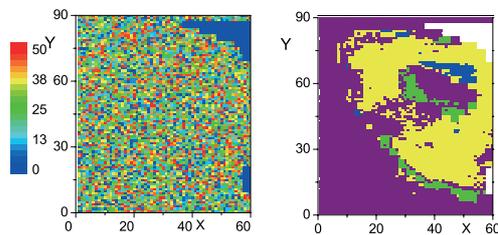
The synthetic 4D acoustic impedance change data is simulated from the PUNQ-3D model. Detailed information on



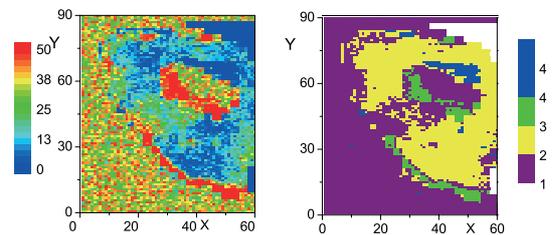
(a) Histogram and final PDFs



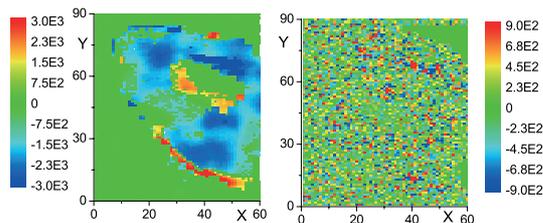
(a) Histogram and final PDFs



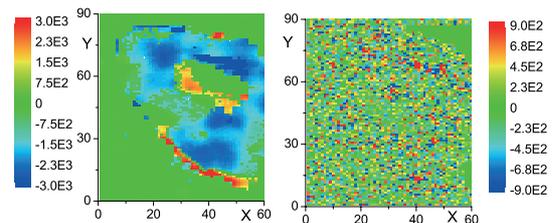
(b) Initial groups (left) and final groups (right)



(b) Initial groups (left) and final groups (right)



(c) Estimated true signal (left) and estimated measurement error (right)



(c) Estimated true signal (left) and estimated measurement error (right)

Figure 5

EM grouping of synthetic data, F_j is used, random initial 50 groups, 4 final groups.

Figure 6

EM grouping of synthetic data, F_j is used, initially sorted by measured values into 50 groups, 4 final groups.

this reservoir can be found in Barker *et al.* (2001), Floris *et al.* (2001), and Gao and Reynolds (2006). The reservoir is bounded with two faults, one is located at the east side and another at the south side. At the north and west sides, there is a strong aquifer. As in Gao and Reynolds (2006), we use a numerical aquifer. At gridblocks within the aquifer, we set water saturation equal to 1.0 and set $\phi = 0.95$. In the example considered here, a $60 \times 90 \times 5$ simulation grid is used. With the simulated pressures and saturations, the same rock physics models used by Dong and Oliver (2003) are applied to determine the acoustic impedance change. The resulting synthetic 4D acoustic impedance change data are shown in Figure 2(a), which shows a comparatively smooth signal. The right panel of this plot is the noisy data after uncorrelated noise is added. The group colored green consists mainly of the gridblocks within the aquifer plus the gridblocks in the original gas cap. The gridblocks in the original gas cap are those fairly near the top and upper right which are surrounded by cells colored blue red or yellow. The blue cells correspond to the oil column at the end of the second seismic survey. The red and yellow gridblocks “below” the original gas cap represent oil influx into the gas cap and those colored the darkest blue “above” the original gas cap represent gas influx into the original oil zone. The band of yellow and red gridblocks near the “lower boundary” of the blue region correspond to water influx into the oil column.

Three methods are considered:

- The Diplaros *et al.* (2004) algorithm is applied. The results are shown in Figure 3. In each EM iteration, we apply a 5×5 Gaussian smoother to the membership matrix as in Diplaros *et al.* (2004). Starting from 20 initial groups (sorted by values) we finally get 2 groups, whereas, based on the physics, we should expect to have at least four groups. The yellow colored group in the right panel of Figure 3(b) actually includes the influx

regions, but it includes both high values and low values as can be seen in the estimated PDF in Figure 3(a). The estimated measurement error as shown in the right panel of Figure 3(c) contains some of the structure of the true signal which indicates that we have over smoothed the data. The estimated measurement error has a variance of 5.6×10^5 and a correlation length of 3, which is not very accurate as the true variance is 1.8×10^5 and the true noise is uncorrelated.

- Twenty initial groups from value initialization are applied in the spatial EM algorithm, and the grouping quality coefficients F_j 's are not used. The results are shown in Figure 4. Starting from 20 groups, we still have 20 groups at convergence. From Figure 4(a), we see that the component PDFs for the Gaussian mixture model fit the histogram very well, but the spatial continuity of each group is poor as shown in the right panel of Figure 4(b), and do not have any physical meaning. The separation of the smoothed signal and measurement error is also relatively poor as shown in Figure 4(c): some measurement error remains in the smoothed data, while some structure of the true signal remains in the estimated measurement error. The estimated measurement error has a correlation length of 2 and variance of 0.45×10^5 . The variance of the true measurement error is uncorrelated with the variance of 1.8×10^5 , which indicate that the variance of the measurement error is underestimated by a factor of 4.0 and the correlation length is overestimated. These results indicate that the EM algorithm without F_j tends to maintain the initial number of groups, and the estimation of measurement error can be poor if the initial number of groups is inappropriate.
- Random initialization (50 groups) and value initialization (50 groups) are both used in the spatial EM algorithm with the grouping quality coefficient. Both of these initializations generated 4 final groups, and the results are

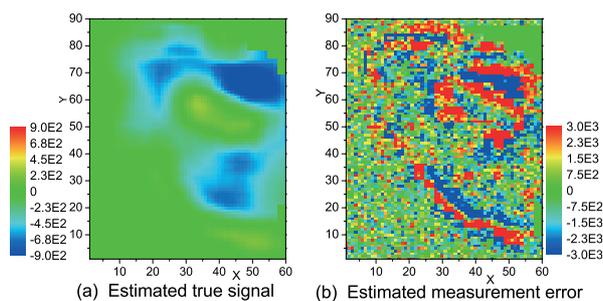


Figure 7

Direct moving average without using EM groups (11×11).

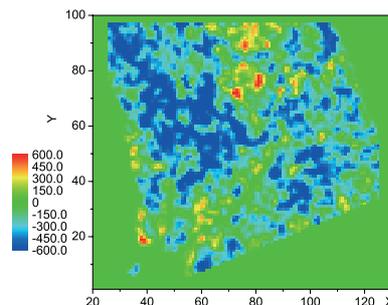


Figure 8

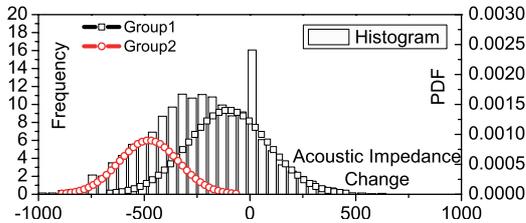
Field 4-D acoustic impedance change.

very close to each other as shown in Figures 5 and 6. The spatial distributions of these groups are close to each other, and are consistent with the physical regions:

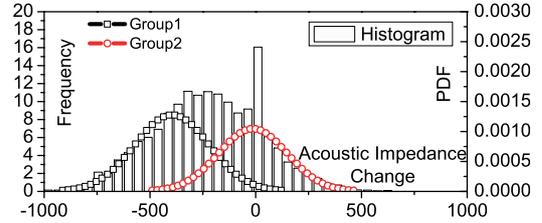
- the wine colored gridblocks correspond mainly to the water region (the aquifer) and the region occupied by the initial gas cap;
- the yellow colored gridblocks correspond to the oil region,
- the green colored gridblocks correspond to water influx into the oil region and oil influx into gas region,
- the blue colored gridblocks correspond to the gas influx into the oil column.

For both initializations, the estimated measurement errors are uncorrelated spatially with the variance of 2.1×10^5 , which is close to the true variance of 1.8×10^5 .

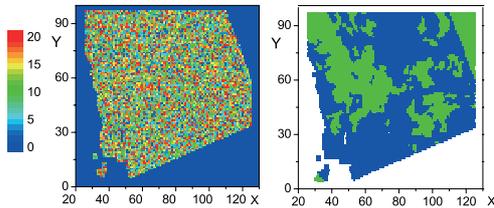
These cases show that the improved spatial EM algorithm can successfully resolve the number of groups, and increase the robustness of the estimation, because two distinctly types of initializations gave the same results. Also note that Figures 4, 5 and 6 illustrate that histogram of the data can be well fit with the Gaussian mixture model based on the groups obtained from the EM algorithm. This indicates the assumption that measurements can be represented by a Gaussian mixture model is valid. Only uncorrelated noise is used in this example, but the algorithm also works for correlated noise with a short correlation length Zhao *et al.* (2006). As the correlation length becomes longer, the estimation of the measurement error using this algorithm becomes less robust, since measurement error with a long correlation length is a smooth component, which violates the assumption that only the underlying true signal is smooth.



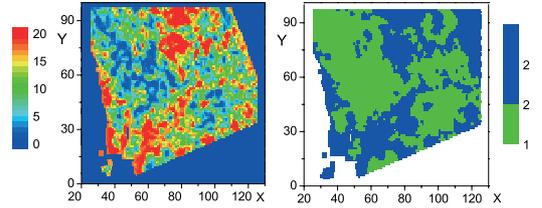
(a) Histogram and final PDFs



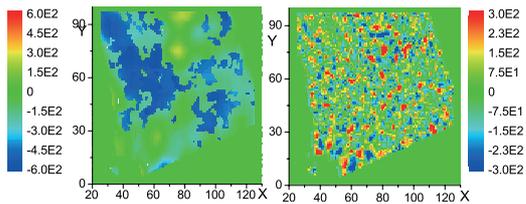
(a) Histogram and final PDFs



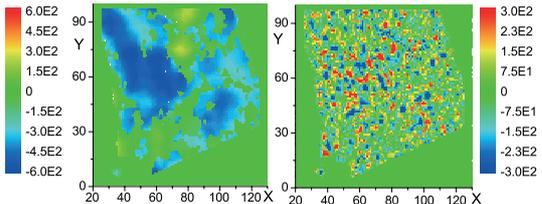
(b) Initial groups (left) and final groups (right)



(b) Initial groups (left) and final groups (right)



(c) Estimated true signal (left) and estimated measurement error (right)



(c) Estimated true signal (left) and estimated measurement error (right)

Figure 9

EM grouping of field 4-D acoustic impedance change, F_j is used, random initial 20 groups, 2 final groups (and a tiny group).

Figure 10

EM grouping of field 4-D acoustic impedance change, F_j is used, initially sorted by measured values into 20 groups, 2 final groups.

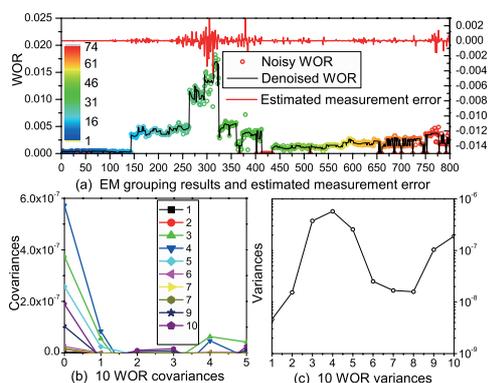


Figure 11

EM grouping of field WOR data, F_j is used, spatially initiated into 100 groups, 74 final groups.

Figure 7 shows the results when the standard moving window smoothing is used for measurement error estimation. In this case, a window size of 11×11 is applied and the arithmetic average in the window represents the estimated true signal. Figure 7(a) shows the estimated true signal with this method and Figure 7(b) gives the estimated measurement error. The estimated true signal is too smooth compared to the actual true signal (Fig. 2a). The estimated measurement error exhibits much of the structure of the true signal, which is a result of smoothing over the discontinuities. This example shows that grouping the data with the EM algorithm before smoothing can avoid smoothing over discontinuities.

5.2 Field 4D Seismic Data Case

As is shown in Figure 8, a layer of 4D acoustic impedance change (AIC) is used as the data to be analyzed using the EM algorithm. The data consists of 7049 active grids. Both random and value initialization are used here. The results of random initialization (20 initial groups) are shown in Figure 9, and those of value initialization (20 initial groups) are shown in Figure 10. Both these cases generated 2 final groups. Figure 9(a) and Figure 10(a) compare the estimated PDF with the histogram of the data. Although the histogram is like a single Gaussian, the spatial EM separated it into two Gaussian using the spatial information. From the two different initializations, the sizes of the two final groups are different, but the estimated covariance are very close (variance of 1.9×10^4 for value initialization and 2.0×10^4 for random initialization). Moreover, the estimated correlation length is 3 in all directions. The estimated true signal (smoothed signal) has clear boundaries, and the estimated measurement error seems to be reasonable.

Although the estimated covariance for this example was isotropic, it is important to note that there is nothing inherent in the algorithm that assumes that measurement errors are isotropic. In fact, Zhao *et al.* (2006) show synthetic examples where the method yields a reasonable estimate of anisotropic covariance function. These authors have also shown that the method can be applied for some multivariate cases, *i.e.*, where the measurements include both time lapse impedance data and the change in Poisson's ratio.

5.3 Field WOR Data Case

Daily field WOR data of a single well is shown in Figure 11(a). Initially these data are evenly divided into 100 groups according to time. By applying the spatial EM algorithm with grouping quality coefficients, the number of the groups is decreased to 74. Different colors in Figure 11(a) refer to different groups. The edges (discontinuities) are naturally detected by the EM algorithm. The measurement error is estimated by applying quadratic fitting within each group, which avoids smoothing over sharp changes, and the window length used is 21 days. The estimated measurement error is shown in red at the top of this figure. The amplitude of the measurement error changes from early times to late times, *i.e.*, the measurement error is non-stationary. We pick 10 sections from the estimated measurements to calculate covariances separately. The first section is from point 1 to point 161, and the second is from point 81 to point 241 etc., which means each section contains 161 measurements and the sections overlap. Figure 11(b) and (c) shows the 10 covariances and temporally distributed variances in sequence.

CONCLUSIONS

The examples presented indicate that the improved EM algorithm, *i.e.*, the EM algorithm with a spatial constraint and grouping quality coefficients can efficiently group the time lapse seismic data into regions, which correspond to physical changes in the reservoir between two seismic surveys, and can also be applied to the production data. The estimate of covariance of the measurement error based on quadratic fitting within each group is superior to that obtained with a constant moving window average.

The improved EM algorithm can be used to determine an appropriate number of groups. The grouping quality coefficient is used to enhance spatial continuity within each group and eliminate low continuity groups. Results from the synthetic data presented here and in Zhao *et al.* (2006) indicate that the final grouping can be used to obtain a reliable characterization of measurement error.

With the improved EM algorithm, the estimated measurement error from different initializations are virtually identical for both the synthetic seismic data and the field seismic

data, although in the field case, the size of the two final groups are somewhat different for the two initializations. The results suggest that for the purpose of characterizing measurement error, the improved EM algorithm is robust.

ACKNOWLEDGEMENTS

This work was supported by the member companies of TUPREP and the U.S. Department of Energy under Award No. DE-FC26-04NT15517. However, any opinions, findings, conclusions or recommendations herein are those of the authors and do not necessarily reflect the views of the DOE.

REFERENCES

- Aanonsen S.I., Aavatsmark I., Barkve T., Cominelli A., Gonard R. Gosselin O., Kolasinski M., and Reme H. (2003) Effect of scale dependent data correlations in an integrated history matching loop combining production data and 4D seismic data (SPE 79665), in *Proceedings of the 2003 SPE Reservoir Simulation Symposium*.
- Abreu C., Lucet N., Nivlet P., and Royer J.J. (2005) Improving 4D seismic data interpretation using geostatistical filtering. *9th International Congress of Brazilian Geophysical Society, Expanded abstracts*.
- Allard D. and Guillot G. (1999) Clustering geostatistical data. *Geostatistics 2000 Cape Town*, 1, 49-63.
- Ambroise C. and Govaert G. (1995) Spatial clustering and the EM algorithm. <http://www.isip.msstate.edu/projects/speech/sup-port/help/bibliography/index.html>.
- Barker J.W., Cuypers M., and Holden L. (2001) Quantifying uncertainty in production forecasts: Another look at the PUNQ-S3 problem. *SPE J.*, 6, 433-441.
- Celeux G., and Govaert G. (1991) *A classification EM algorithm for clustering and two stochastic versions*. Rapports de Recherche.
- Coleou T.L., Hoerber H., and Lecerf D. (2002) Multivariate geostatistical filtering of time-lapse seismic data for an improved 4D signature. *72nd Annual International Meeting, SEG, Expanded Abstracts*, 1662-1665.
- Dempster A.P., Laird N.M., and Rubin D.B. (1977) Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc. B*, 39, 1-38.
- Diplaros A., Gevers T., and Vlassis N. (2004) Skin detection using the EM algorithm with spatial constraints. In *Systems, Man and Cybernetics, 2004 IEEE International Conference*, 4.
- Dong Y., and Oliver D. (2003) Automatic history matching of production data and seismic impedance change data. In *TUPREP Research Report 20*. The University of Tulsa.
- Evensen G. (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, 99, 143-162.
- Evensen G. (2003) The ensemble kalman filter: theoretical formulation and practical implementation. *Ocean Dynam.*, 53, 343-367.
- Floris F.J.T., Bush M.D., Cuypers M., Roggero F., and Syversveen A.-R. (2001) Methods for quantifying the uncertainty of production forecasts: A comparative study. *Petrol. Geosci.*, 7, S87-S96.
- Gao G., and Reynolds A.C. (2006) An improved implementation of the LBFSG algorithm for automatic history matching. *SPE J.*, 11, 5-17.
- Hartley H. (1958) Maximum likelihood from incomplete data. *Biometrics*, 14, 174-194.
- Hastie T., Tibshirani R., and Friedman J. (2001) *The Elements of Statistical Learning*, Springer-Verlag, New York.
- Kitanidis P.K. (1995) Quasi-linear geostatistical theory for inversion. *Water Resour. Res.*, 31, 2411-2419.
- Kung S.Y., Mak M.W., and Lin S.H. (2004) *Biometric Authentication: A Machine Learning Approach*. Prentice Hall.
- Meng X.-L. (1993) Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika*, 80, 267-278.
- Meng X.-L. (1994) On the rate of convergence of the ECM algorithm. *Ann. Stat.*, 22, 326-339.
- Naevdal G., Mannseth T., and Vefring E.H. (2002) Near-well reservoir monitoring through ensemble Kalman filter (SPE-75235), in *Proceeding of SPE/DOE Improved Oil Recovery Symposium*.
- Naevdal G., Johnsen L.M., Aanonsen S.I., and Vefring E.H. (2003) Reservoir monitoring and continuous model updating using ensemble Kalman filter (SPE-84372), in *2003 SPE Annual Technical Conference and Exhibition*.
- Oliver D.S., He N., and Reynolds A.C. (1996) Conditioning permeability fields to pressure data, in *European Conference for the Mathematics of Oil Recovery*, V, 1-11.
- Reynolds A.C., He N., and Oliver D.S. (1999) Reducing uncertainty in geostatistical description with well testing pressure data, in *Reservoir Characterization – Recent Advances*, (edited by Schatzinger R.A. and Jordan J.F.). American Association of Petroleum Geologists, 149-162.
- Richardson S., and Green P.J. (1997) On Bayesian analysis of mixtures with an unknown number of components. *J. R. Stat. Soc.*, 59, 731-792.
- Zhao Y., Li G., and Reynolds A.C. (2006) Characterizing measurement error with the EM algorithm. *TUPREP Report*, 14-147.

Final manuscript received in December 2006

Copyright © 2007 Institut français du pétrole

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than IFP must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee: Request permission from Documentation, Institut français du pétrole, fax. +33 1 47 52 70 78, or revueogst@ifp.fr.