Please fill in the name of the event you are preparing this manuscript for.	2020 International Petroleum Technology Conference
Please fill in your 5-digit IPTC manuscript number.	20IPTC-(19856)-MS
Please fill in your manuscript title.	Efficient Joint 4D Seismic and Electromagnetic History Matching Using an Ensemble-based Approach

Please fill in your author name(s) and company affiliation.

Given Name	Surname	Company
Yanhui	Zhang	King Abdullah University of Science and Technology
Ibrahim	Hoteit	King Abdullah University of Science and Technology

This template is provided to give authors a basic shell for preparing your manuscript for submittal to an IPTC meeting or event. Styles have been included (Head1, Head2, Para, FigCaption, etc) to give you an idea of how your finalized paper will look before it is published by IPTC. All manuscripts submitted to IPTC will be extracted from this template and tagged into an XML format; IPTC's standardized styles and fonts will be used when laying out the final manuscript. Links will be added to your manuscript for references, tables, and equations. Figures and tables should be placed directly after the first paragraph they are mentioned in. The technical content of your paper WILL NOT be changed. Please start your manuscript below.

Abstract

This study aims at enhancing reservoir characterization through simultaneous history matching of time-lapse seismic and electromagnetic (EM) data. We propose an efficient ensemble-based history-matching approach by exploiting the complementary nature of these data types.

The developed workflow consists of two main steps. Firstly, water saturation and porosity as the rock cross-properties relating to both seismic velocities and formation conductivity are estimated by joint inversion of seismic and EM data using an iterative ensemble smoother with an adaptive non-distance-based Kalman gain localization. Prescribed rock-physics models are used to quantify the corresponding relationship of rock properties to seismic velocities and formation conductivity, based on which synthetic EM and full-waveform seismic data are simulated. Secondly, instead of integrating the inverted saturation field directly, front positions are identified from the inverted saturation field and used as the observed data for history matching. Model parameters are conditioned to the observed fronts using the iterative ensemble smoother with a feature-oriented distance parameterization.

The approach is verified with synthetic data sets. Results show that joint inversion of seismic and EM data leads to better estimates of rock properties when compared with those from the separate inversion. Moreover, by integrating the extracted fronts from the improved estimation of the saturation field, which synthesizes the information from both types of data, the proposed approach significantly reduces the redundancy and number of data while still capturing the essential information. The novelty of the proposed approach consists in combining the feature-based history matching with ensemble-based geophysical inversion to achieve an efficient joint integration of time-lapse seismic and EM data.

Introduction

Joint inversion of multiple sources of geophysical data has been actively studied in geophysical community due to its great potential to enhance the fidelity of inverted models (Dell'Aversana et al., 2011; Moorkamp et al., 2011; Liang et al., 2016; Sun and Li, 2016; Colombo and Rovetta, 2018). The core idea of joint inversion is to exploit the complementary nature between different types of geophysical data sets and combine the information into one single inversion scheme for better recovery of properties or structures within the reservoir. It is well known that seismic and electromagnetic (EM) data are complementary to each other and increasingly used for joint inversion (Abubakar et al., 2012; Gao et al., 2012). Seismic measurements provide effective information for reservoir structure while electromagnetic measurements are sensitive to fluid content (distinguishing between oil and water).

As discussed in Zhang and Hoteit (2019), the way to assimilate geophysical measurements like seismic and EM data is not unique. There are three different levels at which the extracted information from the raw data can be utilized for updating reservoir models. Zhang and Hoteit (2019) examined the integration of crosswell EM data from the first two levels, which correspond to the conditioning on raw EM observations (after necessary data processing) and interpreted conductivity through geophysical inversion, respectively. They showed that for the feature-based integration approach its performance somehow depends on the quality of inversion results and is susceptible to observation bias and propagated errors introduced by the inversion. In contrast, the third level, at which reservoir cross-properties (Dell'Aversana et al., 2011) such as porosity and saturation linking different geophysical attributes are inverted, provides an opportunity for joint inversion. Enhanced performance is therefore expected by combining the feature-based integration approach with joint inversion, which motivates the study we present here.

In the remainder of this paper, we first give an overview of the developed feature-oriented ensemble history-matching workflow, then explain the major components of the methodology in detail, followed by numerical experiments in which we examine the performance of the proposed method, and demonstrate the benefit to condition reservoir models on both seismic and EM data.

Feature-oriented ensemble history-matching workflow

This study aims at enhancing reservoir characterization through simultaneous history matching of time-lapse seismic and electromagnetic (EM) data. By exploiting the complementary nature of these data types, we introduce a feature-oriented ensemble history-matching workflow to iteratively update reservoir models as shown in Fig. 1. The developed workflow consists of two main steps. Firstly, water saturation and porosity as the rock cross-properties relating to both seismic velocities and formation conductivity are estimated by joint inversion of seismic and EM data (i.e., red loop). Secondly, the inverted saturation fields are then assimilated to update related reservoir properties such as permeability and porosity (i.e., blue loop).

The workflow starts with generating the initial ensemble of reservoir models through geostatistical modeling based on prior information. Proper sampling of the initial ensemble characterizing the prior uncertainty is essential to ensure the performance of ensemble-based data-assimilation methods, in the sense that the posterior solution is constrained within the subspace of the initial ensemble of model parameters (Aanonsen, 2009; Oliver and Chen, 2009). Reservoir porosity and permeability are considered as the uncertain model parameters for history matching, although the extension to include other types of model parameters is straightforward under the developed history-matching workflow.

Step 1: ensemble-based joint inversion. The initial ensemble of saturation field at the seismic and EM survey time is sampled by running a reservoir flow simulator with the generated ensemble of porosity and permeability, although it can be generated with any suitable sampling method. To simulate the seismic and EM responses, we need to implement rock physics modeling first needed to build the relationship of rock cross-properties (i.e., porosity and saturation) to seismic velocities and formation conductivity. With the inputs from the established rock-physics models, predicted full-waveform seismic and EM data are then obtained from seismic and EM forward simulators. Conditioning on the observed seismic and EM

measurements, we calibrate the prior ensemble of porosity and saturation using an iterated ensemble smoother (Chen and Oliver, 2013).

Figure 1—Feature-oriented ensemble history-matching workflow for assimilation of time-lapse seismic and EM data.

Step 2: feature-oriented ensemble update. The goal of this step is to incorporate the information contained in the inverted saturation fields into the porosity and permeability fields. Instead of integrating the inverted saturation field directly, front positions are identified from the inverted saturation field and used as the observed data for history matching. The iterated ensemble smoother combined with a feature-oriented distance parameterization method (Zhang and Leeuwenburgh, 2017) is applied to condition the prior permeability ensemble and updated porosity ensemble on the observed fronts.

When time-lapse seismic and EM data are assimilated sequentially, the two steps can be coupled in different ways depending on how the prior ensemble of saturation field is generated. For instance, we can do the first step only at first to integrate seismic and EM data sequentially in time by sampling the prior saturation ensemble at each survey time from the forward flow simulation using the initial ensemble of porosity and permeability fields. After that, we implement the second step to assimilate the inverted time-lapse saturation fields. As an alternative, the two steps can be implemented alternately at each survey time by sampling the realizations of saturation field based on the updated ensemble of porosity and permeability fields. Conceptually, the second approach seems better in the sense that more information from the data is incorporated into the prior ensemble. However, this may not practically be true partly because the variability of the updated model variables is often over-reduced due to the limited ensemble size that is normally used. In this paper, we choose the first approach to sample the prior saturation field, but we also alternate the two steps at each survey time so that the inverted saturation information can be incorporated timely into the porosity and permeability fields. In the following, we present the details of the methodology, focusing on the major components of the developed history-matching workflow.

Rock physics modeling

Petro-elastic model. The reservoir fluid considered in this paper consists of two phases, oil, and water. Gassmann's equations (Gassmann, 1951) are used to link seismic velocities to reservoir properties, i.e., porosity ϕ and saturation. The compressional wave (P-wave) velocity V_p in fluid-saturated rocks is expressed as follows:

$$V_p = \sqrt{\frac{K_s + \frac{4}{3}\mu_s}{\rho_s}}, \quad (1)$$

where

$$K_s = (1 - \beta)K_m + \beta^2 \left(\frac{\beta - \phi}{K_m} + \frac{\phi}{K_f} \right)^{-1}, \quad (2)$$

$$K_f = \left(\frac{S_w}{K_w} + \frac{S_o}{K_o} \right)^{-1}, \quad (3)$$

$$\rho_s = \phi(S_w\rho_w + S_o\rho_o) + (1 - \phi)\rho_m, \quad (4)$$

$$\beta = \min\left(\frac{\phi}{\phi_c}, 1\right). \quad (5)$$

In the above equations, K_s , μ_s , and ρ_s are saturated-rock bulk modulus, shear modulus, and bulk density, respectively; K_m , K_o , and K_w are the bulk moduli of rock matrix, oil, and water, respectively; ρ_m , ρ_o , and ρ_w are the densities of rock matrix, oil, and water, respectively; S_o and S_w are oil and water saturations, respectively; β is the Biot coefficient which is computed by the critical porosity model of Nur (1992) (i.e., Eq. 5), in which $\min(\cdot)$ denotes the minimum operator; and ϕ_c is the critical porosity above which solids become suspensions. Here, we use the acoustic approximation so that the shear modulus is neglected.

Petro-electric model. Archie's law (Archie, 1942) is used to quantify the relationship of formation conductivity to relevant reservoir properties such as saturation and porosity, which is expressed as follows:

$$\sigma = a^{-1}\sigma_w\phi^m S_w^n, \quad (6)$$

where σ and σ_w are the formation and the brine conductivity, respectively, and a , m , and n are Archie's parameters whose values normally depend on the rock's cementation, wettability, compaction, pore structure and so on.

Seismic and EM modeling

Seismic forward model. Full-waveform seismic data are simulated by solving the Helmholtz wave equation in an isotropic and inhomogeneous medium. An acoustic approximation is used as follows:

$$\nabla \cdot \left[\left(\frac{1}{\rho(\mathbf{r})} \right) \nabla u(\mathbf{r}) \right] + \frac{\omega^2}{\lambda(\mathbf{r})} = -i\omega q(\mathbf{r}_s), \quad (7)$$

where u is the pressure field; q is the monopole source; ω is the angular frequency and $i^2 = -1$; λ is the Lamé coefficient related to seismic velocity through the relation $V_p = \sqrt{\lambda/\rho}$; $\mathbf{r}$ denotes spatial positions; $\mathbf{r}_s$ denotes source positions, and ∇ is the spatial differential operator with respect to $\mathbf{r}$. The above equation is solved using the frequency-domain finite-difference approach with a perfectly matched layer absorbing

boundary condition (Operto et al., 2007; Silva and Herrmann, 2019). Seismic measurements in crosswell setting are defined as a vector of pressure field responses, generated by the sources at positions $\mathbf{r}_s$ and observed at receiver locations $\mathbf{r}_r$ along boreholes, denoted by

$$\mathbf{d}_{obs,\omega_j}^{Seis} = \begin{bmatrix} \mathbf{Re}\left(u_{\omega_j}(\mathbf{r}_r)\right) \\ \mathbf{Im}\left(u_{\omega_j}(\mathbf{r}_r)\right) \end{bmatrix}, \quad \text{for } j = 1, 2, \dots, N_{\omega}^{Seis}, \quad (8)$$

where $\mathbf{Re}(\cdot)$ and $\mathbf{Im}(\cdot)$ denote the operators taking the real and imaginary parts of the argument, respectively. N_{ω}^{Seis} is the number of frequencies that are selected for full-waveform inversion.

EM forward model. The EM response is governed by Maxwell's equations. In crosswell setting, an EM survey is normally conducted at only a few low frequencies (more often single frequency), so the problem is usually formulated and solved in the frequency domain. After the rearrangement of Maxwell's equations, the second-order electromagnetic equations can be obtained as below:

$$i\omega\mu(\sigma(\mathbf{r}) - i\omega\varepsilon)\mathbf{E}(\mathbf{r}) - \nabla \times \nabla \times \mathbf{E}(\mathbf{r}) = -i\omega\mu\mathbf{J}(\mathbf{r}_s) - \nabla \times \mathbf{M}(\mathbf{r}_s), \quad (9)$$

where μ is the magnetic permeability; ε is the dielectric permittivity; $\nabla \times (\cdot)$ represents the curl operator; $\mathbf{E}$ is the electric field; and $\mathbf{J}$ and $\mathbf{M}$ denote the electric and magnetic current sources, respectively. Perfectly electrically conducting boundary conditions (i.e., $\mathbf{n} \times \mathbf{E} = 0$) is imposed, where $\mathbf{n}$ is the outward normal on the boundary of the domain. With the input of $\sigma(\mathbf{r})$ provided by the rock physics model, the resulting system of Eq. 9 is solved with a 3D multigrid solver described in (Mulder, 2006; Plessix and Mulder, 2008). Once the electric field $\mathbf{E}$ is known, the magnetic field $\mathbf{H}$ can be obtained from $\mathbf{H}(\mathbf{r}) = (i\omega\mu)^{-1}\nabla \times \mathbf{E}(\mathbf{r})$. The crosswell EM measurements are defined as a vector of magnetic field responses, generated by a magnetic current source at positions $\mathbf{r}_s$ and observed at receiver locations $\mathbf{r}_r$ along boreholes, denoted by

$$\mathbf{d}_{obs,\omega_j}^{EM} = \begin{bmatrix} \mathbf{Re}\left(\mathbf{H}_{\omega_j}(\mathbf{r}_r)\right) \\ \mathbf{Im}\left(\mathbf{H}_{\omega_j}(\mathbf{r}_r)\right) \end{bmatrix}, \quad \text{for } j = 1, 2, \dots, N_{\omega}^{EM}, \quad (10)$$

where the subscript ω_j indicates that the EM measurements are frequency-dependent and N_{ω}^{EM} is the number of frequencies at which a crosswell EM survey is conducted.

Iterated ensemble smoother

For weakly nonlinear and approximately Gaussian problems, the standard EnKF or ES usually works well. However, iterative versions of EnKF or ES would be necessary if the nonlinearity is strong. In this paper, we choose to use an iterative ensemble smoother named LM-EnRML (Chen and Oliver, 2013), although there are a number of suitable iterative algorithms based on ES available (Emerick and Reynolds, 2013; Iglesias, 2015; Luo et al., 2015). In the approximate form of LM-ENRML, the update to the i th realization of model parameters $\mathbf{m}_i^{\ell} \in \mathcal{R}^{N_m}$ at the ℓ th iteration is written as,

$$\delta\mathbf{m}_i^{\ell} = \Delta\mathbf{M}^{\ell}\Delta\mathbf{D}^{\ell T}[(1 + \lambda_{\ell})\mathbf{I}_{N_d} + \Delta\mathbf{D}^{\ell}\Delta\mathbf{D}^{\ell T}]^{-1}\mathbf{C}_D^{-\frac{1}{2}}(\mathbf{d}_{obs,i} - \mathbf{g}(\mathbf{m}_i^{\ell})), \quad (11)$$

where $\mathbf{I}_{N_d} \in \mathcal{R}^{N_d \times N_d}$ is an identity matrix; $\mathbf{C}_D \in \mathcal{R}^{N_d \times N_d}$ is the covariance matrix of data noise, which is assumed to be diagonal in the examples considered in this paper; $\mathbf{d}_{obs,i} \in \mathcal{R}^{N_d}$ is a vector of perturbed

data sampled from a multivariate Gaussian distribution $\mathcal{N}(\mathbf{d}_{obs}, \mathbf{C}_D)$ and $\mathbf{d}_{obs} \in \mathbb{R}^{N_d}$ is the vector of observed data; $\mathbf{g}(\cdot)$ denotes the operator mapping the model parameters to predicted data; λ_ℓ is the Levenberg-Marquardt (LM) tuning parameter at the ℓ th iteration; N_d is the number of data; and N_m is the number of model parameters. The matrices $\Delta \mathbf{D}^\ell \in \mathbb{R}^{N_d \times N_e}$ and $\Delta \mathbf{M}^\ell \in \mathbb{R}^{N_m \times N_e}$ are the normalized deviation ensemble of simulated data and the deviation ensemble of model parameters, respectively, denoted as

$$\begin{aligned}\Delta \mathbf{D}^\ell &= \frac{1}{\sqrt{N_e - 1}} \mathbf{C}_D^{-\frac{1}{2}} \mathbf{D}^\ell \left(\mathbf{I}_{N_e} - \frac{1}{N_e} \mathbf{1}_{N_e} \mathbf{1}_{N_e}^T \right), \\ \Delta \mathbf{M}^\ell &= \frac{1}{\sqrt{N_e - 1}} \mathbf{M}^\ell \left(\mathbf{I}_{N_e} - \frac{1}{N_e} \mathbf{1}_{N_e} \mathbf{1}_{N_e}^T \right),\end{aligned}\tag{12}$$

where $\mathbf{I}_{N_e} \in \mathbb{R}^{N_e \times N_e}$ is an identity matrix; $\mathbf{1}_{N_e} \in \mathbb{R}^{N_e}$ is a column vector with all elements equal to one; and N_e is the ensemble size. Each column in $\mathbf{M}^\ell \in \mathbb{R}^{N_m \times N_e}$ and $\mathbf{D}^\ell \in \mathbb{R}^{N_d \times N_e}$ respectively represents one realization of model parameters m_i^ℓ and corresponding simulated data $\mathbf{g}(\mathbf{m}_i^\ell)$.

It has been well recognized that two problems become prominent in a standard implementation of EnKF with a limited ensemble size (Aanonsen et al., 2009b): (1) spurious correlations arising in the ensemble approximation of the covariance matrix impose changes to model or state variables in regions that should not be updated, (2) limited number of degrees of freedom constrains the ability to assimilate large amounts of independent data. In the case of crosswell seismic and EM measurements, the number of data could become very large, and the redundancy of the measurements could be high due to the dense and repeated spatial coverage of the interwell region, which necessitates the use of some regularization to overcome the mentioned issues.

In the previous study (Zhang and Hoteit, 2019), a distance-based Kalman gain localization was applied to integrate crosswell EM data, which was shown to be a viable option if the localization matrix was determined appropriately. However, to choose a proper localization matrix is not an easy task that usually involves a joint consideration of prior ensemble covariance, data sensitivity and ensemble size. It becomes particularly difficult for the full-waveform seismic data given their complex sensitivity patterns with respect to related reservoir properties. To circumvent these difficulties, we adopt a non-distance-based Kalman gain localization method introduced by Zhang and Oliver (2010). They applied the bootstrap method to assess the confidence level of each element in the Kalman gain matrix so that the real and the spurious correlations in the Kalman gain matrix can be discriminated. The localized Kalman gain is given by

$$\mathbf{K}^{loc} = \mathbf{L} \circ \underbrace{\left\{ \Delta \mathbf{M}^\ell \Delta \mathbf{D}^{\ell T} \left[(1 + \lambda_\ell) \mathbf{I}_{N_d} + \Delta \mathbf{D}^\ell \Delta \mathbf{D}^{\ell T} \right]^{-1} \right\}}_{\text{Kalman gain } \mathbf{K}},\tag{13}$$

where $\mathbf{L}$ is the localization matrix and $\circ$ denotes the Schur product or element-wise multiplication. The element L_{ij} named confidence factor in the localization matrix is given by

$$L_{ij} = \frac{1}{1 + \mathbf{R}_{ij}^2 (1 + 1/\gamma^2)}, \quad \text{for } i = 1, 2, \dots, N_m, j = 1, 2, \dots, N_d,\tag{14}$$

$$\mathbf{R}_{ij}^2 = \frac{\frac{1}{N_b} \sum_{l=1}^{N_b} (\mathbf{K}_{ij}^l - \mathbf{K}_{ij})^2}{\mathbf{K}_{ij}^2},\tag{15}$$

where $\mathbf{K}_{ij}^l$ represents the element from the l th bootstrapped sample of the Kalman gain matrix and N_b is

the number of bootstrapped samples. The weighting factor γ is introduced to regulate the estimate of $\mathbf{L}$ to avoid taking negative values and adjust the degree of screening to spurious correlations. The value of $\mathbf{L}_{ij}$ reflects the accuracy of the correlations present in the Kalman gain matrix. Small value of $\mathbf{L}_{ij}$ indicates an unreliable correlation of the corresponding model variable with the data which thus should be eliminated or suppressed. By decreasing the value of γ , the confidence factor $\mathbf{L}_{ij}$ will get close to 0 especially for the ones with large variance, so that the Kalman gain is increasingly screened. Because $\mathbf{L}_{ij}$ is never equal to 0, the suppressed spurious correlations will be gradually added back to the model update with iteration, which allows for the selection of a lower value of γ compared with the case without iteration. On the other hand, reducing the value of γ will slow down the convergence rate so that more iterations are required. In this paper, the value of γ is determined by experimentation.

Distance parameterization of saturation fronts

A distance parameterization method (Leeuwenburgh and Arts, 2014; Zhang and Leeuwenburgh, 2017) is employed to facilitate the assimilation of the inverted saturation field from seismic and EM data. The distance parameterization utilizes the shape information of the extracted features from the attribute of interest. Compared with the direct integration of the inverted saturation field, the advantages of the distance parameterization approach are at least twofold. On the one hand, the information contained in the shape or outline of a feature (if can be extracted coherently) is generally more reliable and informative than the one contained in the amplitude of individual grid cell of inverted saturation. On the other hand, higher computational efficiency can possibly be achieved since the dimensions of the feature space are typically much lower than the original data space.

Let A and B represent the water fronts (i.e., contours) extracted from the inverted and the simulated saturation fields, respectively, with their corresponding binary images (e.g., with 0 indicating the background and with 1 identifying the shape) denoted by the matrices $\mathbf{I}_A$ and $\mathbf{I}_B$. Three parameterization schemes were developed in Zhang and Leeuwenburgh (2017). Here, we use the scheme termed as LHDC,

$$\text{LHDC}(A, B) = \mathbf{I}_A \circ \mathbf{D}_B + \mathbf{I}_B \circ \mathbf{D}_A, \quad (16)$$

where $\mathbf{D}_A$ and $\mathbf{D}_B$ are two distance maps for shapes A and B , respectively, which are computed by using a fast marching method (Zhang and Leeuwenburgh, 2016). The value at each point of a distance map measures the distance from this point to the nearest point on the shape in the image. Additionally, signed distance fields are used such that distance information is available also for cells within the flooded area. Positive distance values ahead of the front and negative distance values behind the front are assigned. In essence, LHDC quantifies the differences with two directed distance maps in complementary directions, $\mathbf{I}_A \circ \mathbf{D}_B$ (distance from B to A) and $\mathbf{I}_B \circ \mathbf{D}_A$ (distance from A to B). Figure 2 shows a stepwise diagram illustrating the computation of the LHDC as a measure of local dissimilarity for two shapes. After applying LHDC, the observed fronts correspond to zero distance map (i.e., $\mathbf{d}_{obs} = \text{LHDC}(A, A) = \mathbf{0}$) and the simulated fronts correspond to the dissimilarity distance map (i.e., $\mathbf{d}_{sim,j} = \text{LHDC}(A, B_j)$), which are subsequently integrated by the localized LM-EnRML as in Eqs. 11 and 13.

Application

We examine the performance of the developed history-matching workflow using a two-dimensional (2D) channelized reservoir model with a crosswell configuration for EM and seismic surveys.

Model setup. The reference porosity and log-transformed permeability fields (used to generate observations) are shown in Fig. 3 displaying channel-like features oriented in the east-west direction. The reservoir model has the dimensions of $45 \times 45 \times 1$ with a uniform grid size of 5 m in the x - and y -directions, and 20 m in the y -direction. There are two facies types in the true model. The realizations of facies model are first generated by a multipoint-based geostatistical algorithm (Strebelle, 2002), based on

which the realizations of porosity and log-permeability are generated by the sequential Gaussian simulation algorithm using an exponential variogram model with an isotropic range of 8 gridblocks. The reference model is randomly drawn from the generated realizations. As can be seen in Fig. 3, the channel facies (sand) is characterized by high permeability and porosity while the background facies (shale) exhibits opposite features.

Figure 2—Schematic diagram of LHDC (Eq. 16) for shape comparison based on contour.

Figure 3—The true porosity and log-permeability fields from which synthetic historical data are generated. Trajectories of the horizontal injector are shown as white crosses and the horizontal producer are shown as solid black circles.

The fluid system consists of two immiscible phases (oil and brine) with a connate water saturation of 0.2, residual oil saturation of 0.2, and initial formation pressure of 310 bar. There are two horizontal wells, one producer (shown as solid black circles) and one injector (shown as white crosses), with an interwell separation of 125 m. The producer is under bottom-hole pressure (BHP) control at 138 bar, and the injector is on rate control at 200 sm^3/day . Table 1 summarizes the input parameters of the rock-physics models used in this experiment.

Measurement setup. For the crosswell configuration of both seismic and EM surveys, transmitters and receivers are placed in the boreholes of the producer and the injector, respectively. There are 15 transmitters and 15 receivers that are uniformly distributed from the heel to the toe of each horizontal well.

In the crosswell EM surveys, the transmitters are axial magnetic dipoles, and the measured data are the horizontal components of the magnetic fields. The frequency of operation is 500 Hz. In the crosswell seismic surveys, monopole sources are used, and the receivers measure scalar pressure fields. We choose to use the data that correspond to operating frequencies of 20, 60 and 100 Hz for the inversion. The EM model has the dimensions of $64 \times 64 \times 16$, with a uniform grid size of 5 m within the reservoir, which is the domain of interest. It has the same lateral resolution as the reservoir flow model. Away from the domain of interest, a variable spacing is chosen for computational efficiency so that perfectly conducting boundary conditions can be satisfied. The seismic model has the dimensions of $45 \times 45 \times 7$ with the same grid size as the EM model. All three model grids are superposed within the domain of interest. Time-lapse crosswell seismic and EM surveys are conducted at days 60 and 180, respectively. The true saturation fields at these survey times are shown in Fig. 4. The synthetic seismic and EM data are corrupted with 5% Gaussian random white noise. The total number of seismic and EM data for each survey amounts to 1800 ($= 15 \times 15 \times 2 \times 3^{\text{Seis}} + 15 \times 15 \times 2^{\text{EM}}$). For multifrequency seismic data, a data-weighting scheme introduced by Hu et al. (2009) is used to restrain the high-frequency component from dominating the history-matching process and introducing undesirable artifacts to the estimated reservoir properties.

Table 1—Input parameters of the rock-physics models.

Petro-elastic model		Petro-electric model	
ϕ_c	0.4	a	1
ρ_o	885 kg/m ³	m	1.5
ρ_w	1050 kg/m ³	n	2
ρ_m	2640 kg/m ³	σ_w	4.5 S/m
K_m	35 GPa		
K_w	2.85 GPa		
K_o	0.71 GPa

Figure 4—The true saturation fields for the crosswell seismic and EM surveys conducted at days 60 and 180, respectively.

History matching. The initial ensemble consists of 100 realizations that are not conditioned on well data. For the LM-EnRML, the starting value of the LM tuning parameter λ is chosen on the same order of magnitude as $\sum_{j=1}^{N_e} S_d(\mathbf{m}_j^0) / (2N_e)$, where $\mathbf{m}^0$ is the initial model variables and the data mismatch S_d is defined by

$$S_d(\mathbf{m}_j) = \left(\mathbf{d}_{obs} - \mathbf{g}(\mathbf{m}_j) \right)^T \mathbf{C}_D^{-1} \left(\mathbf{d}_{obs} - \mathbf{g}(\mathbf{m}_j) \right). \quad (17)$$

We follow the tuning strategy and stopping criteria used in Chen and Oliver (2013). At each iteration, if both the ensemble mean and the standard deviation of $\{S_d(\mathbf{m}_j)\}_{j=1}^{N_e}$ are reduced, the model update is accepted and the value of λ is divided by 10; if only the mean of the data mismatch is reduced, the model update is accepted but the value of λ is unchanged; otherwise, if the mean of the data mismatch is increased, the model update is rejected, and the value of λ is multiplied by 10. The iterative updating stops when the number of iterations exceeds 15, or the reduction of the mean of $\{S_d(\mathbf{m}_j)\}_{j=1}^{N_e}$ at two consecutive iterations is less than 1%. The weighting factor $\gamma = 0.3$ is used for the bootstrap-based Kalman gain localization.

To analyze the synergy effect by the joint use of seismic and EM data, we conduct three experiments in which the first one uses only EM data, the second one uses only EM data, and the last one uses both types of data. The time-lapse seismic and EM data are assimilated sequentially in time in all cases.

Results: day 60 – step 1. Figure 5 shows the distribution of the data mismatch with iteration for separate and joint assimilation of seismic and EM data at day 60. The initial data mismatch for EM and seismic data is on similar order of magnitude. In all cases, the data mismatch is reduced consistently along with the iteration and reaches an acceptable level.

Figure 5—Boxplots of data mismatch as functions of iteration step for separate and joint assimilation of seismic and EM data at day 60. The bounds of the box are 25% and 75% quantiles, the whiskers are the extremes, the line in the box is the median, the dot is mean, and the pluses are outliers.

Figure 6 compares the root-mean-square error (RMSE) of the updated ensembles of porosity and water saturation from the three experiments, in which the RMSE is defined as

$$RMSE = \sqrt{\frac{1}{N_m} \sum_{j=1}^{N_m} (\mathbf{m}_{i,j} - \mathbf{m}_{t,j})^2}, \quad (18)$$

where the subscripts i , j , and t represent realization number, element index, and reference model property (porosity, permeability), respectively. The behavior of the change of RMSE along with the iteration somehow reflects the characteristics of the information content of each data type. In the EM only case, a larger and more consistent reduction of RMSE for saturation is observed compared with the one for porosity, the final RMSE of which goes up slightly. On the contrary, in the seismic only case, the reduction of RMSE of porosity is stronger and more stable than the one for saturation. The complementary effect is seen when both types of data are used, but seismic data plays a dominant role in this specific case, partly because EM data are in single frequency while seismic data contain multiple frequencies.

Figure 6—Boxplots of RMSEs of porosity and water saturation as functions of iteration step for separate and joint assimilation of seismic and EM data at day 60. The bounds of the box are 25% and 75% quantiles, the whiskers are the extremes, the line in the box is the median, the dot is mean, and the pluses are outliers.

Figure 7 shows the means and standard deviations of the porosity field for the initial ensemble and the updated ensembles of the three cases. The initial mean exhibits a rather homogenous profile with large standard deviation over the entire field. In the case that only EM data is integrated, some channel features observed in the reference porosity field appear in the updated mean field, and reduced variability concentrates in the neighborhood of well locations, particularly around the injector. In contrast, a significantly refined channel profile in better agreement with the reference field is observed for the cases that seismic data is incorporated. The standard deviation is reduced prominently across the whole domain, particularly in the interwell region. In contrast, the difference in the mean fields of saturation for the three cases shown in Fig. 8 is less obvious, even though the difference in the standard deviation fields is still dramatic. In general, the mean saturation fields in all cases reproduce the main features observed in the reference saturation field as shown in Fig. 4a.

Results: day 60 – step 2. We proceed with the second step to assimilate the inverted saturation information from the joint use of seismic and data, as shown in Fig 8e. The determination and extraction of saturation fronts from the inverted saturation fields is summarized in Fig. 9. To reduce inversion artifacts and extract coherent features, we apply a lateral smoothing to the inverted mean saturation field (see Fig. 9a). A threshold value is necessary in order to derive the observed front positions. Here, we determine the threshold value by identifying the unstable area of the contour map of the smoothed mean saturation field, as shown in Fig. 9b highlighted by the black rectangle. In these areas, saturation front changes dramatically so that biases can be introduced when the threshold value is too small or important information can be lost when the threshold value is too large. Therefore, we take the value in between to avoid extremes. A threshold value of 0.3 is used to derive the observed front positions. Uncorrelated distance measurement errors are assumed with a standard deviation of 5 m (one grid cell length), which

is roughly estimated from the spread of the extracted fronts from the updated ensemble of saturation fields as shown in Fig. 9c.

Figure 7—The means and standard deviations (STD) of porosity field for the initial ensemble (a and b) and the updated ensembles (c through h) after separate and joint assimilation of seismic and EM data at day 60.

Figure 8—The means and standard deviations (STDs) of water saturation field for the initial ensemble (a and b) and the updated ensembles (c through h) after separate and joint assimilation of seismic and EM data at day 60.

Figure 9—Determination and extraction of saturation fronts from the inverted saturation field by joint assimilation of seismic and EM data at day 60. The black rectangle in (b) highlights the unstable changing area of saturation contours and the red isoline corresponding to 0.3 indicates the observed front selected for history matching. The corresponding extracted front from the reference saturation field in Fig. 4a is marked with green circle in (c). The grayscale in (c) indicates the count of occurrence of the water fronts at a location extracted from the inverted ensemble of saturation fields. The scale is truncated for better display.

The initial ensemble of permeability and the updated ensemble of porosity from the first step are then conditioned on the observed fronts using the feature-oriented assimilation approach. Figure 10 shows the consistent reduction of data mismatch and the RMSEs of porosity and permeability along with the iteration, even though the RMSE of porosity only decreases slightly.

Figure 10—Boxplots of data mismatch and the RMSEs of permeability and porosity as functions of iteration step after the assimilation of saturation fronts at day 60. The bounds of the box are 25% and 75% quantiles, the whiskers are the extremes, the line in the box is the median, the dot is mean, and the pluses are outliers.

Figure 11 shows the means and standard deviations of the permeability field for the initial ensemble and the updated ensemble. The standard deviation is significantly reduced especially in the surrounding area of the injector from where the water front originates. The essential information carried by the observed front is illustrated by the upper high-permeable channel structure in the reference model reproduced in the updated mean field. Figure 12 shows that the predicted ensemble of saturation field from the updated ensemble of porosity and permeability provides a good match to the observed front with high variation concentrating around the boundary area.

Figure 11—The means and standard deviations (STDs) of the initial ensemble and the updated ensemble of the log-permeability field after the assimilation of saturation fronts at day 60.

Figure 12—The mean and standard deviation (STD) of the predicted ensemble of saturation field at day 60 by rerunning the simulation with the updated porosity and permeability fields.

Results: day 180 – step 1. We proceed to the next survey time with the updated ensemble of porosity at the last time step and the prior ensemble of saturation (mean field shown in Fig. 15a) based on the initial ensemble of permeability and porosity. Note that we do not implement the second step at day 60 for the EM only and seismic only cases, considering the fact that the porosity ensemble is only calibrated slightly at this step as shown in Fig. 10c. The distribution of the data mismatch along with the iteration for separate and joint assimilation of seismic and EM data is shown in Fig. 13.

Figure 13—Boxplots of data mismatch as functions of iteration step for separate and joint assimilation of seismic and EM data at day 180. The bounds of the box are 25% and 75% quantiles, the whiskers are the extremes, the line in the box is the median, the dot is mean, and the pluses are outliers.

Figure 14 shows that the RMSEs of porosity in all cases are hardly reduced, which reflects the dynamic nature of time-lapse seismic and EM data. The means of the updated ensembles of saturation fields for the three cases are shown in Fig. 15. The main difference comes from the area where a small spike of water front is induced by a short thin channel as seen in the reference model.

Figure 14—Boxplots of the RMSEs of porosity and saturation as functions of iteration step for separate and joint assimilation of seismic and EM data at day 180. The bounds of the box are 25% and 75% quantiles, the whiskers are the extremes, the line in the box is the median, the dot is mean, and the pluses are outliers.

Figure 15—The means of the initial ensemble and the updated ensembles of saturation fields for separate and joint assimilation of seismic and EM data at day 180.

Results: day 180 – step 2. Figure 16 compares the extracted fronts in the same way as before from the three cases, which shows a positive synergy effect by the joint use of seismic and EM. Figure 17 shows the means and standard deviations of the final ensembles of porosity and permeability, and the corresponding predicted ensemble of water saturation. The improvement in the permeability field is distinct with more channel features reproduced.

In the experiments considered above, porosity and permeability fields are assumed to be highly correlated. However, in the developed workflow, the calibration of these model parameters are separated. As we can see in Fig. 17, the separate updating procedure would result in a loss of cross-correlation that is captured by the initial ensemble. One way to retain the correlation is by applying a standard ensemble update to the permeability field before the second step using the updated ensemble of porosity from the first step as the data. As shown in Fig. 18, this cross-correlation can be retained when the prior permeability and porosity fields are highly correlated. On the other hand, spurious correlations and biases can be introduced if the cross-correlation is not represented properly by the prior ensemble. On such occasions, the separate updating procedure would be a better choice.

Figure 16—Comparison of extracted saturation fronts from the inverted saturation field by separate and joint assimilation of seismic and EM data at day 180. The black rectangle in the second column highlights the unstable changing area of saturation contours and the red isolines corresponding to the extracted front selected for history matching. The corresponding extracted front from the reference saturation field in Fig. 4b is marked with green circle (last column). The grayscale in the last column indicates the count of occurrence of the water front at a location. The scale is truncated for better display.

Figure 17—The means and standard deviations (STDs) of the updated ensembles of porosity and permeability fields, and the corresponding predicted ensemble of water saturation field after the assimilation of saturation fronts at day 180.

Figure 18—The mean and standard deviation (STD) of the updated ensemble of permeability obtained by being conditioned on the updated ensemble of porosity field from the joint assimilation of seismic and EM data at day 60.

Conclusions

In this paper, we introduce a feature-oriented ensemble history-matching workflow focusing on the joint assimilation of time-lapse seismic and EM data. The developed workflow divides the history-matching process into two steps and separates reservoir cross-properties such as porosity and saturation that link

both rock elastic and electric properties from the other model variables. The first step involves a joint inversion of seismic and EM data, from which the uncertain cross-properties are updated using the ensemble-based data assimilation methods. The number of data for seismic and EM measurements is typically large, which causes difficulty for ensemble-based methods in which a small ensemble size is usually employed. To deal with that, we adopt an adaptive Kalman gain localization method based on bootstrap sampling. In the second step, the inverted saturation information is integrated to update related model variables by use of an image-oriented distance parameterization combined with the ensemble-based assimilation method.

We examine the performance of the developed history-matching workflow using a synthetic channelized reservoir model with a crosswell configuration for seismic and EM surveys. The experimental results reveal the complementary nature between seismic and EM data and show a positive synergy effect on the characterization of model variables by jointly assimilating both types of data. It demonstrates that the developed workflow provides a novel and effective way to incorporate the information from multiple sources of geophysical data into reservoir models.

Acknowledgments

Support for the authors is provided by the research project “Efficient Integration of Electromagnetic Tomography into Reservoir History Matching” which is funded by Saudi Aramco. The authors also thank Femke C. Vossepoel, Wim Mulder, and Marwan Wirianto for providing the multigrid EM forward solver.

References

- Aanonsen, S.I., Nævdal, G., Oliver, D.S., Reynolds, A.C. and Vallès, B. 2009a. The ensemble Kalman filter in reservoir engineering--a review. *SPE Journal*, 14(03): 393–412.
- Aanonsen, S.I., Nævdal, G., Oliver, D.S., Reynolds, A.C. and Vallès, B. 2009b. The Ensemble Kalman Filter in Reservoir Engineering – a Review. *SPE Journal*, 14(03): 393–412.
- Abubakar, A., Gao, G., Habashy, T.M. and Liu, J. 2012. Joint inversion approaches for geophysical electromagnetic and elastic full-waveform data. *Inverse Problems*, 28(5).
- Archie, G.E. 1942. The electrical resistivity log as an aid in determining some reservoir characteristics. *Transactions of the AIME*, 146(01): 54–62.
- Chen, Y. and Oliver, D.S. 2013. Levenberg-Marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification. *Computational Geosciences*, 17(4): 689–703.
- Colombo, D. and Rovetta, D. 2018. Coupling strategies in multiparameter geophysical joint inversion. *Geophysical Journal International*, 215(2): 1171–1184.
- Dell’Aversana, P., Bernasconi, G., Miotti, F. and Rovetta, D. 2011. Joint inversion of rock properties from sonic, resistivity and density well-log measurements. *Geophysical Prospecting*, 59(6): 1144–1154.
- Emerick, A.A. and Reynolds, A.C. 2013. Ensemble smoother with multiple data assimilation. *Computers & Geosciences*, 55(0): 3–15.
- Gao, G., Abubakar, A., Habashy, T.M. and Pan, G. 2012. Three-dimensional joint petrophysical inversion of electromagnetic and seismic data. *Society of Exploration Geophysicists International Exposition and 82nd Annual Meeting 2012, SEG 2012*, 77(3): 579–584.
- Gassmann, F. 1951. Elastic waves through a packing of spheres. *Geophysics*, 16: 673–685.
- Hu, W., Abubakar, A. and Habashy, T.M. 2009. Simultaneous multifrequency inversion of full-waveform seismic data. *Geophysics*, 74(2).
- Iglesias, M.A. 2015. Iterative regularization for ensemble data assimilation in reservoir models. *Computational Geosciences*, 19(1): 177–212.
- Leeuwenburgh, O. and Arts, R. 2014. Distance parameterization for efficient seismic history matching with the ensemble Kalman Filter. *Computational Geosciences*, 18(3–4): 535–548.
- Liang, L., Abubakar, A. and Habashy, T.M. 2016. Reservoir property mapping and monitoring from joint

- inversion of time-lapse seismic, electromagnetic, and production data. *Geophysics*, 81(5): 73–84.
- Luo, X., Stordal, A.S., Lorentzen, R.J. and Nævdal, G. 2015. Iterative ensemble smoother as an approximate solution to a regularized minimum-average-cost problem : theory and applications. *SPE J.*, 20(05): 962–982.
- Moorkamp, M., Heincke, B., Jegen, M., Roberts, A.W. and Hobbs, R.W. 2011. A framework for 3-D joint inversion of MT, gravity and seismic refraction data. *Geophysical Journal International*, 184(1): 477–493.
- Mulder, W.A. 2006. A multigrid solver for 3D electromagnetic diffusion. *Geophysical Prospecting*, 54(5): 633–649.
- Nur, A. 1992. Critical porosity and the seismic velocities in rocks. *EOS Transactions AGU*, 73: 66.
- Oliver, D.S. and Chen, Y. 2009. Improved initial sampling for the ensemble Kalman filter. *Computational Geosciences*, 13(1): 13–27.
- Operto, S., Virieux, J., Amestoy, P., L’Excellent, J.-Y., Giraud, L. and Ali, H.B.H. 2007. 3D finite-difference frequency-domain modeling of visco-acoustic wave propagation using a massively parallel direct solver: A feasibility study. *Geophysics*, 72(5): SM195--SM211.
- Plessix, R.E. and Mulder, W.A. 2008. Resistivity imaging with controlled-source electromagnetic data: depth and data weighting. *Inverse Problems*, 24(3): 034012.
- Silva, C. Da and Herrmann, F. 2019. A Unified 2D/3D Large-Scale Software Environment for Nonlinear Inverse Problems. *ACM Transactions on Mathematical Software (TOMS)*, 45(1): 7.
- Strebelle, S. 2002. Conditional simulation of complex geological structures using multiple-point statistics. *Math. Geol.*, 34(1): 1–21.
- Sun, J. and Li, Y. 2016. Joint inversion of multiple geophysical data using guided fuzzy c-means clustering. *Geophysics*, 81(3): ID37–ID57.
- Zhang, Y. and Hoteit, I. 2019. Efficient Assimilation of Crosswell Electromagnetic Data Using Ensemble-Based History-Matching Framework. In Society of Petroleum Engineers.
- Zhang, Y. and Leeuwenburgh, O. 2016. Ensemble-based seismic history matching with distance parameterization for complex grids. In *ECMOR XV - 15th European Conference on the Mathematics of Oil Recovery*.
- Zhang, Y. and Leeuwenburgh, O. 2017. Image-oriented distance parameterization for ensemble-based seismic history matching. *Computational Geosciences*, 21(4): 713–731.
- Zhang, Y. and Oliver, D.S. 2010. Improving the ensemble estimate of the Kalman gain by bootstrap sampling. *Mathematical Geosciences*, 42(3): 327–345.