Please fill in the name of the event you are preparing this manuscript for.	SPE Reservoir Simulation Conference
Please fill in your 6-digit SPE manuscript number.	SPE-193808-MS
Please fill in your manuscript title.	Efficient Assimilation of Crosswell Electromagnetic Data using Ensemble-based History-matching Framework

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 SPE 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 SPE. All manuscripts submitted to SPE will be extracted from this template and tagged into an XML format; SPE'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

An ensemble-based history-matching framework is proposed to enhance the characterization of petroleum reservoirs through the assimilation of crosswell electromagnetic (EM) data. As one of advanced technologies in reservoir surveillance, crosswell EM tomography can provide a cross-sectional conductivity map and hence saturation profile at an interwell scale by exploiting the sharp contrast in conductivity between hydrocarbons and saline water. Incorporating this new information into reservoir simulation in combination with other available observations is therefore expected to enhance the forecasting capability of reservoir models and to lead to better quantification of uncertainty.

The proposed approach applies ensemble-based data-assimilation methods to build a robust and flexible framework under which various sources of available measurements can be readily integrated. Because the assimilation of crosswell EM data can be implemented in different ways (e.g., components of EM fields or inverted conductivity), a comparative study is conducted. The first approach integrates crosswell EM data in its original form which entails establishing a forward model simulating observed EM responses. In this work, the forward model is based on Archie's law that provides a link between fluid properties and formation conductivity, and Maxwell's equations that describe how EM fields behave given the spatial distribution of conductivity. Alternatively, formation conductivity can be used for history matching, which is obtained from the original EM data through inversion using an adjoint gradient-based optimization method. Because the inverted conductivity is usually of high dimension and very noisy, an image-oriented distance parameterization utilizing fluid front information is applied aiming to assimilate the conductivity field efficiently and robustly. Numerical experiments for different test cases with increasing complexity are carried out to examine the performance of the proposed integration schemes and potential of crosswell EM data for improving the estimation of relevant model parameters. The results demonstrate the efficiency of the developed history-matching workflow and added value of crosswell EM data in enhancing the characterization of reservoir models and reliability of model forecasts.

1 Introduction

Ensemble-based data-assimilation methods have witnessed a rapid development for solving reservoir history matching problems since the ensemble Kalman filter (EnKF) was first introduced to petroleum science (Aanonsen et al., 2009; Oliver and Chen, 2011). This is due to the efficient and flexible framework that these methods provide under which large number of model variables can be estimated conditional to various types of data. Besides, they are readily adapted to any reservoir simulator and easy to code and maintain in the sense that neither a tangent nor an adjoint model is required.

Many techniques have therefore been developed to overcome the inherent disadvantages of ensemble-based methods including the assumptions of linearity and Gaussianity that limit their performance in practical applications (Dovera and Della Rossa, 2011; Phale and Oliver, 2011; Hoteit et al., 2012; Luo et al., 2012; Chen and Oliver, 2013; Emerick and Reynolds, 2013; Zhang et al., 2014; Oliver and Chen, 2018). On the other hand, it is beneficial to incorporate as many data as possible to build reliable reservoir simulation models with enhanced prediction skills. Production data (water, oil and gas rates and pressures) have been commonly used in history matching. However, their information content is usually limited to accurately estimate large number of reservoir parameters (Oliver and Chen, 2011). In contrast, time-lapse seismic data possess much higher spatial (lateral) resolution and can potentially provide rich information about dynamic changes in the reservoir far from wells locations. Although the quantitative use of time-lapse seismic data for history matching still remains a challenging task, significant work has been conducted to overcome this difficulty (Skjervheim et al., 2007; Feng and Mannseth, 2010; Trani et al., 2013; Luo et al., 2017; Zhang and Leeuwenburgh, 2017; Oliver and Alfonzo, 2018). Meanwhile, a growing number of real-field history matching applications using ensemble-based assimilation methods have been presented showing satisfactory results in the literature (Haugen et al., 2008; Zhang and Oliver, 2011; Chen and Oliver, 2014; Abadpour et al., 2018).

While seismic techniques have been the method of choice for exploration and reservoir monitoring, EM techniques have shown their value in providing new information to overcome some of the shortcomings of seismic techniques (Al-Ali et al., 2009). As one of these techniques, crosswell EM technique was developed (Wilt et al., 1995; Spies and Habashy, 1995), reaching deep into the reservoir to produce a cross-sectional display of conductivity distribution between wells (Abubakar et al., 2008; Zhdanov and Yoshioka, 2003; Newman and Alumbaugh, 2000) by exploiting the strong conductivity contrast between hydrocarbons and salt water. The results from field studies presented by Marsala et al. (2017, 2008) showed the ability of crosswell EM tomography to deliver useful interwell conductivity, and hence saturation, mapping even at widely-spaced wells.

There has been a growing interest in enhancing reservoir characterization through history matching using EM observations. Katterbauer et al. (2016) presented a history matching study using EM data attribute (conductivity, assumed to be known) to estimate the components of a compositional reservoir model with the EnKF, in which the uncertainty in Archie's parameters and the variance of observation error was also considered. In addition to EM conductivity attribute, Katterbauer et al. (2015) also incorporated seismic, gravimetry and InSAR data for history matching using the EnKF to study the synergy effect that could result from combining them. Matching enhancements were observed when all data sets were integrated, and the most influential impact from EM data was suggested. Similarly, Liang et al. (2016) developed a deterministic joint inversion approach incorporating EM (components of magnetic fields), seismic, and production data to map relevant reservoir properties and monitor water flooding.

In this paper, our focus is on the investigation of the performance of different approaches to integrate crosswell EM data into the history-matching process and the effect of combined use of crosswell EM and well production data on the characterization of reservoir models. In the remainder of this paper, we first give an overview of the ensemble-based history matching workflow and propose two schemes for the integration of crosswell EM data. We then explain the major components of each scheme in detail,

followed by a couple of numerical examples in which the performance of the proposed schemes is examined, and positive synergy effect by joint assimilation of production and EM data is demonstrated.

2 Ensemble-based history matching framework

The general practice to model a petroleum reservoir consists of two necessary steps, which are known as geological modeling and history matching. First, geoscientists build geological models that are conditioned to all of the static data such as geology, geophysics, petrophysics, seismic data, and geostatistical information. These models provide a geologically realistic representation of the petroleum reservoir, but they rarely honor the reservoir's dynamic behavior. Then these models are passed to petroleum engineers who incorporate the dynamic data such as well rates, pressures, and 4D seismic surveys to improve the reservoir characterization and further reduce the uncertainty.

Ensemble-based data-assimilation methods provide a flexible framework under which any uncertain model parameters can be considered and various types of measurements can be readily incorporated (Chen and Oliver, 2014; Katterbauer et al., 2015). The flowchart shown in Fig. 1 presents such an ensemble-based history-matching workflow that couples the two processes mentioned above and forms a closed loop to iteratively update reservoir models. Here we use the term “*history matching*” in a broad sense to describe the whole process of reservoir characterization.

Fig. 1 Ensemble-based history-matching workflow with different assimilation schemes of EM data. Three integration schemes of crosswell EM data are indicated by different colors. Shared steps (after flow simulation) between different schemes are indicated by the combination of the colors in the arrow.

The workflow starts with generating the initial ensemble of reservoir models through geological modeling based on available static data. Generally, the initial ensemble represents all our prior knowledge about the reservoir and approximates the probability density function (pdf) of the prior distribution of uncertain model parameters. Sampling the initial ensemble plays a vital role in 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). Therefore, it is crucial for the success of these methods to ensure that the initial ensemble of model parameters properly characterizes the prior uncertainty.

Simulating observed dynamic data via forward modeling with the generated ensemble of model parameters as part of its inputs is the next step, also known as the forecast step in filtering theory, which

is also the most time-consuming part in the whole workflow, typically requiring hundreds of forward runs. However, the computational overhead can be significantly reduced by executing these runs in parallel, depending on available computing resources. Among the forward models, reservoir flow simulator predicting reservoir state variables and fluid flow dynamics due to well production activities, should be the most fundamental one, based on which the links to other types of measurements (e.g., time-lapse seismic and EM data) are usually established. For the case of crosswell EM data, a rock physics model is first needed to build direct or indirect connections between the model parameters of interest and formation conductivity. Archie's law takes such a role in this work (Archie, 1942) empirically relating porosity, brine conductivity and brine saturation to formation conductivity. Then a forward EM simulator needs to be implemented to model EM responses to conductivity changes by solving Maxwell's equations (Stratton, 2007).

The information contained in different observations is finally assimilated to update the prior ensemble of model parameters through the innovations (i.e., the difference between actual and simulated data) at the analysis step. Similar to time-lapse seismic data or many other types of geophysical data, the way to assimilate crosswell EM data is not unique. There are basically three different levels at which the extracted information from crosswell EM surveys can be utilized for updating reservoir models. At one extreme, as shown by scheme 1 of Fig. 1, the raw EM data (e.g., components of electromagnetic fields) are used for history matching by going through the full forward modeling procedure. In this case, establishing a solid foundation underpinning the forward modeling would be the key, although it is usually difficult and time-consuming due to the absence of a universally established rock physics theory and a general lack of experimental field data. At the other extreme, as shown by scheme 3 of Fig. 1, a direct comparison of water saturation changes is pursued through the process of EM inversion which is, however, notoriously known as an ill-posed inverse problem with non-unique solutions. An intermediate level lying in between is the comparison of electrical conductivity depicted by scheme 2 of Fig. 1, which can be seen as a compromise between the full-modeling scheme 1 and the full-inversion scheme 3. It is generally not clear which approach is the best. In this paper, we focus on schemes 1 and 2 and examine their performances for history matching in conjunction with production data. The methodology used in scheme 2 as will be shown later also applies to scheme 3. When the same methodology is used for both schemes, the quality of corresponding inversion results plays a dominant role in their assimilation performances. Because Archie's law states a log-linear relationship between saturation and conductivity, it is expected that schemes 2 and 3 would exhibit comparable performances in such context. However, it is conceivable that the saturation profile in scheme 3 can be improved by incorporating more types of data through the joint inversion, for example, of seismic and EM data. An investigation in this aspect will be conducted in the future. The major components of the proposed workflow are described in more detail in the following sections.

3 Iterative ensemble smoother

The standard EnKF has been shown to provide acceptable history matches and uncertainty estimations in many reservoir applications (Aanonsen et al., 2009; Oliver and Chen, 2011). Even though the EnKF was originally proposed as an alternative to the extended KF for applications with nonlinear dynamical systems, the update step in the EnKF is still linear. For highly nonlinear problems, however, the EnKF may yield unphysical updates of model parameters and poor data matches (Zafari and Reynolds, 2007). In such cases, iterative EnKF methods are usually used in order to overcome the high nonlinearity and improve the quality of data match (Gu and Oliver, 2007; Li and Reynolds, 2009). As an alternative ensemble-based assimilation method, the ensemble smoother (ES) was first applied for petroleum history matching by Skjervheim et al. (2011), which mainly motivated by the computational inefficiency of EnKF caused by the frequent restart of the reservoir simulator at every update step and convergence issues associated with updates to the state variables. In a typical application of the ES, all available data are

assimilated simultaneously with a single analysis step taking the similar form to the EnKF, in which only the model parameters are updated so that simulator restarts are avoided. However, the ES in its original form is not as effective as the EnKF especially when the nonlinearity of the concerned problem is strong and when the changes to the model parameters are dramatic (Chen and Oliver, 2012). Many iterative methods based on the ES have been proposed (Chen and Oliver, 2013; Emerick and Reynolds, 2013; Iglesias, 2015; Luo et al., 2015), most of which are achieved through the way of minimizing a stochastic objective function, or equivalently maximizing the posterior probability of each realization. In general, most of these methods lead comparable performances with appropriate settings of corresponding tuning parameters and sample from the posterior distribution (Evensen, 2018).

3.1 LM-EnRML

In this paper, we choose to use an iterative ensemble smoother named LM-EnRML (Chen and Oliver, 2013), in which two different versions of the algorithm are proposed. At the ℓ th iteration, the update to the i th realization of model parameters $\mathbf{m}_i^\ell \in \mathcal{R}^{N_m}$ is written as, in its full form

$$\delta \mathbf{m}_i^{\ell, \text{full}} = \delta \mathbf{m}_i^{\ell, \text{approx}} + \Delta \mathbf{M}^\ell \left[(1 + \lambda_\ell) \mathbf{I}_{N_d} + \Delta \mathbf{D}^\ell \Delta \mathbf{D}^{\ell T} \right]^{-1} \Delta \mathbf{M}^{\ell T} (\Delta \mathbf{M}^0 \Delta \mathbf{M}^{0T})^{-1} (\mathbf{m}_i^0 - \mathbf{m}_i^\ell), \quad (1)$$

and in its approximate form

$$\delta \mathbf{m}_i^{\ell, \text{approx}} = \underbrace{\Delta \mathbf{M}^\ell \Delta \mathbf{D}^{\ell T} \left[(1 + \lambda_\ell) \mathbf{I}_{N_d} + \Delta \mathbf{D}^\ell \Delta \mathbf{D}^{\ell T} \right]^{-1}}_{\text{Kalman gain}} \overbrace{\mathbf{C}_D^{-\frac{1}{2}} \left(\mathbf{d}_{\text{obs}, i} - \mathbf{g}(\mathbf{m}_i^\ell) \right)}^{\text{innovation}}, \quad (2)$$

where $\mathbf{I}_{N_d} \in \mathcal{R}^{N_d \times N_d}$ is an identity matrix; the superscript 0 denotes the initial state before iteration; $\mathbf{C}_D \in \mathcal{R}^{N_d \times N_d}$ is the covariance matrix of data noise, which is assumed to be diagonal in the examples considered in this paper; $\mathbf{d}_{\text{obs}, i} \in \mathcal{R}^{N_d}$ is a vector of perturbed data sampled from a multivariate Gaussian distribution $\mathcal{N}(\mathbf{d}_{\text{obs}}, \mathbf{C}_D)$ and $\mathbf{d}_{\text{obs}} \in \mathcal{R}^{N_d}$ 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 $\Delta \mathbf{D}^\ell \in \mathcal{R}^{N_d \times N_e}$ and $\Delta \mathbf{M}^\ell \in \mathcal{R}^{N_m \times N_e}$ are the normalized deviation ensemble of simulated data and the deviation ensemble of model parameters, respectively, denoted as

$$\begin{aligned} \Delta \mathbf{D}^\ell &= \frac{1}{\sqrt{N_e - 1}} \mathbf{C}_D^{-\frac{1}{2}} \mathbf{D}^\ell \left(\mathbf{I}_{N_e} - \frac{1}{N_e} \mathbf{1}_{N_e} \mathbf{1}_{N_e}^T \right), \\ \Delta \mathbf{M}^\ell &= \frac{1}{\sqrt{N_e - 1}} \mathbf{M}^\ell \left(\mathbf{I}_{N_e} - \frac{1}{N_e} \mathbf{1}_{N_e} \mathbf{1}_{N_e}^T \right), \end{aligned} \quad (3)$$

where $\mathbf{I}_{N_e} \in \mathcal{R}^{N_e \times N_e}$ is an identity matrix and $\mathbf{1}_{N_e} \in \mathcal{R}^{N_e}$ is a column vector with all elements equal to one. Each column in $\mathbf{M}^\ell \in \mathcal{R}^{N_m \times N_e}$ and $\mathbf{D}^\ell \in \mathcal{R}^{N_d \times N_e}$ respectively represents one realization of model parameters $\mathbf{m}_i^\ell$ and corresponding simulated data $\mathbf{g}(\mathbf{m}_i^\ell)$. The full updating equation (1) can be seen as a synthesis of the updates from the data mismatch term and from the model mismatch term of the objective function defined in (Chen and Oliver, 2013). The approximate form of LM-EnRML (2) ignores the update from the model mismatch term, so that its formulation is very similar to the ES-MDA (Emerick and Reynolds, 2013), except for the choice of the tuning parameter and the resampling of the perturbed observations.

Similar to the standard LM algorithm, adjusting λ influences not only the step size but also the search direction. By increasing the value of λ , the search direction gets closer to the steepest descent direction with a decreasing step size, so the updates are small and the convergence rate is slow. Conversely, reducing the value of λ makes the search direction less likely to be descent and imposes large corrections on the model variables, but it may accelerate the convergence rate if the current iterate is not far from the solution. Therefore, the typical tuning strategy is that λ starts with a relatively large value and “smartly” decreases as the minimization gets more stable with iterations. In this paper, the starting value of λ is chosen of the same order of magnitude of the data mismatch and then reduced by a constant factor with each successful iteration. The stopping criteria used for terminating the iteration process include: exceeding a prescribed maximum number of iterations; reaching a prescribed tolerance factor of minimum reduction of data mismatch; or reaching a prescribed tolerance factor of minimal changes to model parameters.

3.2 Kalman gain localization

One fundamental issue for ensemble-based data-assimilation methods relates to the fact that the number of ensemble members is usually quite small compared to the dimension of the model space. Theoretically, increasing the ensemble size could solve much of the difficulty, however, it would sacrifice the computational efficiency of ensemble-based methods, which is valuable in practical applications. Therefore, the goal is always to use as small as an ensemble as possible. It has been well recognized that two problems associated with a limited ensemble size become prominent: (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 EM measurements, the number of data could become very large, which necessitates the use of some type of regularization to overcome the mentioned issues.

Localization is one of the most common regularization techniques used in ensemble-based history-matching applications (Oliver and Chen, 2011). In general, localization methods fall into two categories: distance-based (Chen and Oliver, 2010; Emerick and Reynolds, 2011) and non-distance-based (Zhang and Oliver, 2010; Luo et al., 2018) localization methods. In distance-based localization, one assumes that any covariance beyond a certain distance is purely a result of sampling errors. The reduction of spurious correlations is commonly realized by multiplying the ensemble estimate of error covariance matrices, or Kalman gain matrix, element-wise by a localization matrix, the elements of which are typically formed by a certain type of taper function with some predefined parameters. The selection of taper function is mainly based on an ad hoc basis. Chen and Oliver (2017) compared the performance of Kalman gain localization to two forms of local analysis within the iterative ensemble smoothers. They concluded that different localization methods can lead to equivalent results under iterations if the localization range is chosen appropriately, although the corresponding convergence rates may vary. On the other hand, there exist model variables and/or observations for which a distance interpretation is not sensible, such as the depth of the oil-water contact. In such situations, non-distance-based localization methods are more general and straightforward to use.

In view of the spatial dependence of both production and crosswell EM data on the considered model parameters in this study, we choose to apply the distance-based localization to the Kalman gain matrix of the approximate form of LM-EnRML (2). The localized Kalman gain is given by

$$\mathbf{K}^{\text{loc}} = \boldsymbol{\rho}_{\text{md}} \circ \left\{ \Delta \mathbf{M}^{\ell} \Delta \mathbf{D}^{\ell \text{T}} \left[(1 + \lambda_{\ell}) \mathbf{I}_{N_d} + \Delta \mathbf{D}^{\ell} \Delta \mathbf{D}^{\ell \text{T}} \right]^{-1} \right\}, \quad (4)$$

where $\boldsymbol{\rho}_{\text{md}}$ is the localization matrix and $\circ$ denotes the Schur product or element-wise multiplication. To construct $\boldsymbol{\rho}_{\text{md}}$, we use the Gaspari-Cohn function (Gaspari and Cohn, 1999) in which the only free parameter is the range of a correlation determining the radius of localization. As revealed in (Chen and

Oliver, 2010; Emerick and Reynolds, 2011), the choice of localization range should be a joint consideration of prior model covariance, data sensitivity and ensemble size. In most cases, it needs to balance the harm done by spurious correlations due to limited ensemble against the harm done by excluding true correlations.

4 Direct integration of crosswell EM data

The forward crosswell EM modeling consists of a rock physics model linking rock and fluid properties to formation conductivity, and an EM simulator to predict observed EM responses by solving Maxwell's equations in a crosswell configuration.

4.1 Rock physics model

There are many empirical formulae proposed for quantifying the relationship of formation conductivity to other relevant reservoir properties such as saturation, salt concentration, porosity, temperature, or etc. Here, we choose to use Archie's law (Archie, 1942)

$$\sigma = a^{-1} \sigma_w \phi^m S_w^n, \quad (5)$$

where σ and σ_w are the formation and the brine conductivity, respectively; ϕ is the porosity; S_w is the brine saturation; 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. As outlined in Hamada et al. (2013), the estimates of Archie's parameters can be heterogeneous and quite uncertain in some applications, especially in the case of carbonate reservoirs. Although these uncertain parameters can be readily incorporated into the history-matching framework (Katterbauer et al., 2016), we neglect this uncertainty in this study and assume that Archie's parameters are homogeneous and already determined from core or log analysis. In Eq. 5, the brine conductivity σ_w is computed from an empirical relationship (Dresser, 1982),

$$\sigma_w = \left[\left(0.0123 + \frac{3647.5}{C_w^{0.955}} \right) \frac{82}{1.8T + 39} \right]^{-1}, \quad (6)$$

where C_w is the salt concentration (in ppm) and T is the formation temperature (in °C). Note that the rock physics model is built on the fluid flow simulator from which the distributions of S_w and C_w are obtained.

4.2 EM simulator

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 rearrangement of Maxwell's equations, the second-order electromagnetic equations can be obtained as below:

$$i\omega\mu(\sigma(\mathbf{r}) - i\omega\epsilon)\mathbf{E}(\mathbf{r}) - \nabla \times \nabla \times \mathbf{E}(\mathbf{r}) = -i\omega\mu\mathbf{J}(\mathbf{r}_s) - \nabla \times \mathbf{M}(\mathbf{r}_s), \quad (7)$$

where $\mathbf{r}$ denotes spatial positions; $\mathbf{r}_s$ denotes source positions; ω is the angular frequency and $i^2 = -1$; 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 equations (7) in which only $\mathbf{E}$ is the unknown are solved with a 3D multigrid solver described in (Mulder, 2006; Plessix and Mulder, 2008). When the electric field $\mathbf{E}$ is known, the magnetic field $\mathbf{H}$ can be obtained from

$$\mathbf{H}(\mathbf{r}) = \frac{1}{i\omega\mu} \nabla \times \mathbf{E}(\mathbf{r}). \quad (8)$$

Here, we define the crosswell EM measurements as a vector of magnetic field responses, generated by magnetic current source at positions $\mathbf{r}_s$ and observed at receiver locations $\mathbf{r}_r$ along boreholes, denoted by

$$\mathbf{d}_{\text{obs}}^h = \begin{bmatrix} \text{Re}(\mathbf{S}(\mathbf{r}_r, \mathbf{r})\mathbf{H}(\mathbf{r})) \\ \text{Im}(\mathbf{S}(\mathbf{r}_r, \mathbf{r})\mathbf{H}(\mathbf{r})) \end{bmatrix}, \quad (9)$$

where $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ denote the operators taking the real and imaginary parts of the argument, respectively, and $\mathbf{S}(\cdot)$ is the sampling operator that maps the field to the receiver positions $\mathbf{r}_r$. Note that $\mathbf{d}_{\text{obs}}^h$ depends not only on the source-receiver positions but also on frequencies, which means that $\mathbf{d}_{\text{obs}}^h$ contains all data induced at each frequency when a survey is conducted at multiple frequencies.

5 Alternative integration of inverted conductivity

As an alternative to the direct integration approach, crosswell EM measurements can be first inverted to conductivity data and then assimilated to calibrate reservoir models through the history-matching process. To demonstrate the flexibility of the developed workflow, we choose to solve the inverse problem by gradient-based optimization, which is typically used in the literature and real applications. Ensemble-based methods can be directly applied for the inversion, which seems more consistent with the current framework and may give a better quantification of uncertainty (Tveit et al., 2015).

5.1 Conductivity inversion of crosswell EM data

Given the crosswell EM measurements $\mathbf{d}_{\text{obs}}^h \in \mathbb{R}^{N_h}$, we pose the inverse problem of finding the conductivity σ as a minimization problem with the objective function

$$O_T(\sigma) = \frac{1}{2} \left(\mathbf{d}_{\text{obs}}^h - \mathbf{s}(\sigma) \right)^T \mathbf{C}_H^{-1} \left(\mathbf{d}_{\text{obs}}^h - \mathbf{s}(\sigma) \right) + O_R(\sigma), \quad (10)$$

where $\mathbf{s}(\cdot)$ represents the forward EM simulator as described in the previous section, $\mathbf{C}_H \in \mathbb{R}^{N_h \times N_h}$ is a diagonal matrix whose elements are estimates of the variance of the data noise, and O_R is a regularization function. There are many options for the regularization function to restrain the solution space of the ill-posed inverse problem. Here, we adopt the minimum support regularization (Portniaguine and Zhdanov, 1999) that preserves the feature of high conductivity contrast, of the form

$$O_R = \int_D \alpha(\mathbf{r}) \frac{|\sigma(\mathbf{r}) - \sigma^{\text{ref}}(\mathbf{r})|^2}{\beta^2 + |\sigma(\mathbf{r}) - \sigma^{\text{ref}}(\mathbf{r})|^2}, \quad (11)$$

where D denotes the inversion domain, σ^{ref} is a reference model, α is the regularization factor which is set by trial and error, and β is a positive number that controls the degree of smoothing on σ . The estimate

of β requires prior knowledge on σ , for instance, from core or log analysis. It is easy to see that the minimum support regularization imposes more suppression on smaller changes to σ^{ref} , which is equal to the initial guess of σ in this paper. A quasi-Newton method known as L-BFGS algorithm (Byrd et al., 1995) is used to solve the minimization problem, in which the gradient is computed with the adjoint method described in Plessix and Mulder (2008).

5.2 Image-oriented distance parameterization

Instead of integrating the inverted conductivity directly, we employ a distance parameterization method (Leeuwenburgh and Arts, 2014; Zhang and Leeuwenburgh, 2017) that utilizes the shape information of the extracted features from the attribute of interest. Compared to the straight integration of inverted conductivity, the advantages of the image-oriented 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 conductivity. 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.

In the case considered here, the feature is related to the shape of the spatial changes of conductivity induced by saturation changes and sharp conductivity contrast between water and hydrocarbons. Let A and B represent two shapes extracted from the inverted and simulated conductivity, 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 $\mathbf{I}_B$. Then, the innovation (defined in Eq. 2) in the image-oriented distance parameterization is computed by the local Hausdorff distance (Baudrier et al., 2008):

$$\text{LHD}(A, B) = \mathbf{I}_A \circ \mathbf{D}_B + \mathbf{I}_B \circ \mathbf{D}_A, \quad (12)$$

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 a nearest point on the shape in the image. In essence, the LHD 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).

Based on Eq. 12, three parameterization approaches were developed in Zhang and Leeuwenburgh (2017). Here, we use the approach termed as PLHDC

$$\text{PLHDC}(A_c, B_c) = \mathbf{I}_{A_c} \circ \mathbf{D}_{B_c}, \quad (13)$$

considering the facts that the reservoir condition and well configuration are relatively simple in a typical crosswell EM survey and PLHDC is particularly effective in reducing the number of data. In Eq. 13, the subscript c denotes that the shape is represented by a contour (boundary or front). Therefore, PLHDC is actually a partial LHD measuring the distance only from the simulated front to the “observed” front. Because nonzero distance values exist only on the “observed” front, the number of data reduces from the number of grid points on the map to the number of grid points on the “observed” front.

6 Numerical experiments

To examine the potential of crosswell EM data for improving reservoir characterization and uncertainty quantification with the proposed workflow, two test cases with increasing complexities are presented. In the first test case, a relatively simple two-dimensional (2D) channel model is used and a comparison is conducted between the two assimilation methods. The second test case considers a more realistic three-

dimensional (3D) reservoir model with a focus on the joint assimilation of production and crosswell EM data.

6.1 2D channel model

The reference model (used to generate observations) shown in Fig. 2a is a two-dimensional binary reservoir model with channel-like features oriented in the east-west direction. It covers an area of $225 \times 225 \text{ m}^2$ with a thickness of 20 m, with a simulation grid of dimensions 45×45 . The reference model is randomly drawn from the realizations generated by a multipoint-based geostatistical algorithm (Strebelle, 2002). There are two facies types in the true model. The permeability in the channel facies (sand) is 1000 md, and the permeability in the background facies (shale) is 50 md. The porosity is assumed to be uniform and equal to 0.2 everywhere. The fluid system consists of two immiscible phases (oil and brine water) with a connate water saturation of 0.2, residual oil saturation of 0.2, initial formation pressure of 310 bar, and initial salt concentration of 160 kg/sm^3 . There are two horizontal wells, one producer (shown as black solid circles) and one injector (shown as white crosses), in the simulation model with an interwell separation of 125 m. The producer is under bottom-hole pressure (BHP) control at 138 bar, and the injector is on a rate control at $200 \text{ sm}^3/\text{day}$. The concentration of salt in the injected water is 50 kg/sm^3 . The formation temperature is assumed constant at 90°C . For the rock physics model, Archie's parameters are assumed to be known with $m = 2.2$ and $n = 2.3$. The model variables updated with LM-EnRML are the log-permeability of each gridblock, and the initial ensemble consists of 100 unconditional realizations (not conditioned to the information at well locations). Flow simulations are performed with the ECLIPSE (Schlumberger, 2014) reservoir simulator.

Fig. 2 The true log-permeability field (a) from which synthetic historical data are generated. True conductivity fields (b-d) for the crosswell EM surveys conducted at days 60, 120, and 180, respectively. Trajectories of the horizontal injector are shown as white crosses and the horizontal producer are shown as solid black circles.

For the crosswell EM setup, transmitters and receivers are placed in the boreholes of the injector and the producer, respectively. There are 29 transmitters and 29 receivers that are uniformly distributed from the heel to the toe of each horizontal well with a spacing of 5 m. 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. The resistivity of the reservoir formation before waterflooding is $30 \Omega\text{m}$, and the resistivity of the surrounding formation is assumed to be known with a value of $10 \Omega\text{m}$. The EM model has the dimensions of $64 \times 64 \times 16$, with uniform grid size of 5 m within the domain of interest (the reservoir area, the same lateral resolution as the reservoir flow model), and with variable spacing away from the uniform region for computational efficiency. Time-lapse EM surveys are conducted at days 60, 120, and 180, respectively. The synthetic data are generated by solving the forward flow and EM problem

with the reference model, and corrupted with 2% Gaussian random white noise. The number of data for each survey amounts to 1682 ($= 29 \times 29 \times 2$). Only EM data are considered in this example. The true conductivity fields at the survey times are illustrated in Figs. 2b-d. As water moves from the injector to the producer, the high-permeable channel structures are gradually illuminated through the variation of conductivity profile, which is detected by the crosswell EM surveys. This information embedded in the time-lapse EM data is what we expect to extract through the process of history matching.

6.1.1 Scheme 1: integration of crosswell EM data

The approximate form of LM-EnRML (Eq. 2) with Kalman gain localization (Eq. 4) is used for history matching. The construction of localization matrix $\mathbf{p}_{md}$ is typically done in an ad hoc way, which needs to take into account the correlation scale of model parameters and data sensitivity. It is difficult to compute the sensitivity of crosswell EM data to model parameters directly from a chain of forward models including reservoir flow, rock physics, and crosswell EM. But it is possible to obtain a rough estimation indirectly from the sensitivity of the crosswell measurement to conductivity changes. Spies and Habashy (1995) studied the sensitivity patterns for low-frequency crosswell EM with a focus on the operating frequency, strike length of scatterers, source-receiver angle and different components of the magnetic field. Here, we compute the sensitivity of the crosswell measurement to conductivity changes for the specific crosswell configuration considered in this section by an adjoint method (Abubakar et al., 2008).

Fig. 3 Sensitivities (in dB) for real (a and c) and imaginary (b and d) components of the magnetic field for the considered crosswell configuration with two different source-receiver angles. Small values are truncated to zero for better display. Taper functions (e and f) of two different source-receiver pairs corresponding to two columns in the localization matrix.

The sensitivities for real and imaginary components of the magnetic field are shown in Figs. 3a-d. The sensitivity is clearly strongest in the immediate vicinity of the source and receiver and decreases away

from the wells. The sensitivity in the region between the wells has a sign opposite to the one outside the wells. As shown in Spies and Habashy (1995), under a certain frequency, the sensitivity region becomes more elongated along the axis joining the source and receiver as the operating frequency increases. The sensitivity analysis can be generalized for other crosswell setting by use of the electromagnetic scaling law (Sinclair, 1948; Spies and Habashy, 1995), which states that field patterns and ratios are identical if the product $\sigma L^2 f$ is held constant, where L is the distance between the source and receiver and f is the frequency. Because L and f are fixed in a crosswell EM survey, the sensitivity mainly depends on the value of σ . Analogous to the effect of frequency, the sensitivity range shrinks with increasing conductivity. In the case of time-lapse crosswell EM, instead of using different taper functions for data obtained at different times, we choose to apply the same localization matrix to all data sets, however, with a localization range large enough to disregard the time dependence of data sensitivity. For instance, the sensitivity range estimated with the background conductivity before waterflooding as shown in Fig. 3a-d. Combining the correlation feature of the permeability field, an anisotropic taper function based on Gaspari-Cohn function (Emerick and Reynolds, 2011) is used with two critical localization ranges equal to 180 m and 90 m, respectively. The same localization function is used for both real and imaginary parts of the EM data. Figs. 3e-f show the taper functions of two different source-receiver pairs corresponding to two columns in the localization matrix.

The measurements from time-lapse crosswell EM surveys can be either treated independently or used in the form of the difference of time-lapse data (if without repeatability issues). Here, we choose the former way to assimilate the time-lapse EM data based on the following considerations: (1) The computational cost is much higher for the difference data when assimilated in sequential mode because it requires to run the forward EM model at two different survey times for each realization to obtain the ensemble of predicted difference data. (2) The data noise is increased in the difference data which tends to blur the signal contained in the original data.

Fig. 4 Comparison of data mismatch for the initial ensemble (labeled by “Prior”) and the final updated ensemble (labeled by “Post”) at three crosswell EM surveys (days 60, 120, and 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.

Fig. 4 compares the data mismatch for the initial ensemble and the final updated ensemble at three crosswell EM surveys which are assimilated sequentially in time. The data mismatch $S_{d,i}$ for each model realization is defined by

$$S_{d,i} = (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}_i))^T \mathbf{C}_D^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{g}(\mathbf{m}_i)). \quad (14)$$

Clearly, the initial data mismatch is reduced significantly and the final updated ensemble provides a good data match for all three crosswell EM surveys.

Fig. 5 shows the evolution of the ensemble mean and standard deviation during the sequential assimilation of time-lapse crosswell EM data. As the variability keeps reducing which originates from the surrounding area of the injector spreading out to the producer, the high-permeable channel structures in the reference model are gradually reproduced in the updated ensemble mean. One log-permeability realization is presented in Fig. 6 showing the continuous refinement during the process of history matching.

Fig. 5 The means (first row) and standard deviations (second row) of the initial ensemble and the updated ensembles after each assimilation step. Each plot is shown in the log-transformed permeability field.

Fig. 6 The evolution of one log-permeability realization during the sequential assimilation of time-lapse crosswell EM data. Each plot is shown in the log-transformed permeability field.

6.1.2 Scheme 2: integration of inverted conductivity

This alternative scheme utilizes the information of extracted features from inverted conductivity field to calibrate the reservoir model. The inversion is implemented with the gradient optimization algorithm as described in Section 5.1. It is well understood that a good starting model is essential in leading the inversion to converge to a useful solution. To study the impact of the starting model on the outcome of the inversion, we test two different starting models in this example, a uniform model with background conductivity as shown in Fig. 7a and a model equal to the mean of the initial ensemble of predicted conductivity at the first survey time as shown in Fig. 7c. Obviously, the second model contains more prior information about the spatial distribution of conductivity than the first one. For the regularization term as shown in Eq. 11, the parameter β is set to 0.04 S/m based on the estimation of the conductivity contrast between hydrocarbon and brine. The reference model σ^{ref} is equal to the starting model. The regularization parameter α is determined by

$$\alpha(r) = \tau s_{\sigma}^{-1}(r) \quad (15)$$

where s_{σ} is the standard deviation of the initial ensemble of predicted conductivity at corresponding survey time, and τ is a scalar constant over the entire inversion domain of interest whose value is resolved by trial and error. A small value of 0.001 is enforced for where the standard deviation is zero. By taking into account the variation information, the idea is to guide the inversion to focus more on the uncertain area (with relatively large standard deviation so that the corresponding regularization parameter is small).

The inverted conductivity fields with two different starting models at the first survey time are shown in Figs. 7b and 7d, respectively. In both cases, the value of the objective function is reduced to a comparable level. Even though similar patterns resembling the ones in the true conductivity field (Fig. 2b) are reproduced in both cases, there is a clear difference on the magnitudes of the inverted conductivity. The case with more informative starting model definitely stays closer to the truth. In the meanwhile, the inversion results also motivate the use of feature-based approach to assimilate the inverted conductivity, in the sense that compared with the amplitude of each individual grid cell the feature seems to carry the most essential information with less uncertainty.

Fig. 7 Two different starting models of conductivity (**a** and **c**) and the corresponding inverted conductivity fields (**b** and **d**). Note that the color scale in plot **b** is different from the others for better display of reproduced features.

We proceed with the result of Fig. 7d in the following inversion and infer the conductivity fields from the other two crosswell EM surveys. During the sequential inversion process, previous inverted result is used as the starting model for the next inversion. In doing so, the idea is to incorporate more information into the starting model so that it is better constrained. To reduce inversion artifacts and extract coherent

features, we apply a lateral smoothing to the inverted conductivity fields as shown in Figs. 8d-f. Apparently, the inverted conductivity fields fail to capture some features observed in the reference model and exhibit some bias on reproduced features. For comparison, the smoothed true conductivity fields as shown in Figs. 8a-c are also used for history matching, which can be seen as an ideal case of scheme 2 representing the best performance it can potentially achieve. The observed front contour (boundary) positions highlighted by the black dots in Fig. 8 are derived by applying a threshold value of 0.1 to the smoothed fields. Uncorrelated distance measurement errors are assumed with a standard deviation of 5 m (one grid cell length). Note that the number of data (equal to the number of points on the front contour) for each case considered here is less than 200, which is drastically reduced compared with the data size used in scheme 1. Therefore, the distance parameterization approach somehow relieves the necessity for localization.

Fig. 8 Smoothed conductivity fields at three crosswell EM survey times from the reference model (a-c) and from the inversion (d-f), respectively. The interpreted front contour positions used for history matching are highlighted with black dots. The plots in each row are arranged in ascending order of survey time from left to right.

The full form of LM-EnRML as in Eq. 1 without localization is used to history match the front measurements extracted from the inverted conductivity. The measurements are assimilated sequentially in time. Fig. 9 shows the mean and standard deviation of the updated ensemble at each assimilation step using front contours from the true conductivity fields. The result is quite similar to the one shown in Fig. 5 of scheme 1, except that the standard deviations are slightly lower around the boundary area of the reservoir than the counterparts in scheme 1. This means that under ideal conditions both schemes give comparable performances.

Fig. 10 shows the mean and standard deviation of the updated ensemble at each assimilation step using front contours from the inverted conductivity fields. In contrast to the ideal case, the updated ensemble mean reflects the main features that appear in the reference model and the standard deviation is reduced to a similar degree. However, as expected, degraded performance occurs due to the propagated uncertainty

introduced by the inversion, for example, the loss of sinuosity in the upper channel and missing of branches in the lower channel. Therefore, a good estimation or treatment of observation error and bias is particularly important for the feature-based integration scheme.

Fig. 9 The means (first row) and standard deviations (second row) of the updated ensembles after each assimilation step with the front contours from the true conductivity fields. Each plot is shown in the log-transformed permeability field.

6.1.3 Discussion

For the purpose of demonstration, the conductivity inversion presented here is not performed in an optimal way, which is also beyond the focus of this study. The real geophysical inversion often involves a delicate process in which various sources of relevant information (e.g., well logs and core analysis) can be incorporated and many advanced optimization and regularization algorithms are available. However, no matter how the inversion is implemented, it is generally inevitable to introduce biases and propagated errors into the inferred solution. On the other hand, model errors are ignored from the forward modeling of crosswell EM observations, which in practical applications, however, may result from many causes such as unmodeled physical process, insufficient identification of necessary model parameters, numerical discretization or approximation errors, and so on. In addition, the errors in EM observations have been assumed to be independent, which is a strong assumption given the fact that for time-lapse EM data the error is not only spatially but also temporally correlated. Taking these factors into account, the real history-matching problem comes down to how to calibrate imperfect models to biased observations. Without accounting for model error and observation bias properly incomplete results can be obtained. Oliver and Alfonzo (2018) proposed a systematic approach to deal with this kind of problem, in which an iterative procedure involving model criticism and improvement is introduced following the process of model creation and calibration.

Fig. 10 The means (first row) and standard deviations (second row) of the updated ensembles after each assimilation step with the front contours from the inverted conductivity fields. Each plot is shown in the log-transformed permeability field.

6.2 3D Egg model

In this test case, we consider a two-phase flow problem in a 3D water flood using an open synthetic reservoir dataset named Egg Model published by Jansen et al. (2014), to further examine the behavior of the proposed history-matching workflow for joint integration of production and crosswell EM data. As shown in Fig. 11, the egg model is a channelized oil reservoir having $60 \times 60 \times 7$ gridblocks with grid cell size of 8 m in the x- and y-directions and 4 m in the z-direction, of which 18553 grid cells are active resulting in an egg-shaped model. The dataset contains an ensemble of 100 unconditional permeability realizations in addition to the reference permeability field shown in Fig. 11a. All these permeability fields are hand-drawn with the assistance of a computer drawing program, aiming to mimic typical meandering river patterns featured by intertwined high-permeable channels in a low-permeable background. There is a strong correlation in the vertical direction among the seven layers. Based on the existing permeability fields, we generate the same number of porosity fields by assuming a log-linear relationship between these properties. The generated reference porosity field sharing the same geological patterns as the reference permeability field is shown in Fig. 11b.

There are 12 vertical wells completed in all layers, including 8 water injectors and 4 oil producers. The injectors are controlled by constant injection rate of $49.5 \text{ m}^3/\text{day}$ and the producers are controlled by constant bottom-hole pressure of 395 bar. The injected water has a salt concentration of $50 \text{ kg}/\text{m}^3$. The initial salt concentration in the water phase is $160 \text{ kg}/\text{m}^3$, and the formation temperature is assumed constant at 90°C . For the rock physics model, we assume that Archie's parameters m and n are both equal to 2. In the crosswell EM surveys, transmitters are placed in the injector I4 located in the middle of the reservoir, and receivers are placed in the remaining 11 wells. The distribution of transmitters and receivers covers an interval that extends above and below the reservoir by a distance, with a uniform spacing of 4 m within the reservoir. The data are collected with 13 transmitters and 13×11 receivers. The transmitters

are vertical magnetic dipoles and the measured data are the vertical components of the magnetic fields. The frequency of operation is 500 Hz. The conductivity of the surrounding formation is assumed to be known with a value of 0.1 S/m.

Fig. 11 Reference log-permeability (a) and porosity fields (b) which are used to generate synthetic data. Injecting and producing wells are marked in blue and black colors, respectively.

The simulation time is 97 months (30 days/month) composed of a 36-month history-matching period and a 41-month forecasting period. Observations for the production data which consist of water and oil production rates and BHPs of injectors are collected every month, and for crosswell EM data are gathered every 12 months (accumulating to 3 EM surveys). The standard deviation of data noise is 5 sm^3/day for oil production rate, 8 sm^3/day for water production rate, 1 bar for BHP, and 2% error rate for EM data. To analyze the synergy effect by joint use of production and EM data, we conduct three experiments in which the first one uses only production data (with $N_d = 576$ data points), the second one uses only EM data (with $N_d = 11154$ data points), and the last one uses both types of data (with $N_d = 11730$ data points). Model parameters that are updated during the history matching process include permeability and porosity at each grid cell which amount to 37106 parameters in total. The approximate form of LM-EnRML (Eq. 2) with Kalman gain localization (Eq. 4) is used for history matching, in which the EM data is assimilated by scheme 1. The Gaspari-Cohn taper function is used for localization with the range equal to 296 m (isotropic) for production data as shown in Fig. 12a and the ranges equal to 296 and 148 m (anisotropic) for EM data as shown in Fig. 12b. No localization is applied in the vertical direction for both types of data, considering the fact that all the wells are completed in every layer of the reservoir and the vertical changes of EM responses are relatively small within the limited interval of the reservoir.

Fig. 12 Taper functions for well production data (a) and crosswell EM data (b) corresponding to two columns in the localization matrix.

Fig. 13 compares the match and forecast to production data between the initial ensemble and final updated ensembles from designed experiments. Although the prediction from the initial ensemble (in gray) exhibits high variability and poor match to data, the ensemble spreads for all updated ensembles are significantly reduced and aligned closer to the reference solution (in red). The best data match and forecast (in green) are obtained when both production and EM data are assimilated. With the use of EM data alone, the prediction at each well location (in yellow) is improved but mostly biased. As to the match to EM data, consistent result is observed as shown in Fig. 14 with the maximum reduction of data mismatch for the combined use of both types of data. Incorporation of production data alone only slightly improves the match to EM data.

Fig. 13 Production data match and forecast at all wells for the initial ensemble (gray) and final updated ensembles with production data only (blue), EM data only (yellow), and both (green). The red line corresponds to the data from the reference model. The history matching and forecasting periods are separated by the vertical magenta dashed line.

Fig. 14 Distribution of data mismatch for the initial ensemble (labeled by “Ini”) and the final updated ensembles (labeled by “Prod” for production data only case, “EM” for EM data only case, and “Both” for both production and EM data case) at three crosswell EM survey times (labeled by “T1”, “T2”, and “T3”). 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.

Figs. 15 shows the means and standard deviations of the permeability field in layer 3 (similar results seen in the other layers) for the initial ensemble and final updated ensembles. The initial mean exhibits a rather homogenous profile with large standard deviation over the entire field. In the case when only production data is integrated, some channel features observed in the reference permeability field vaguely appear in the updated mean field, and reduced variability concentrates in the neighborhood of well locations. In contrast, refined channel profile in better agreement with the reference field is observed when EM data is incorporated. The standard deviation is reduced prominently across the whole domain, particularly in the interwell regions connecting the injector I4 to other surrounding wells. This significant reduction of variability is partly due to the large number of low-noise EM measurements and relatively loose estimation on the localization range, however, it also partly indicates the high information content and great spatial sensitivity of the EM data. The mean field updated with the joint use of both types of data appears to display the greatest similarity to the reference permeability field among the three cases. This is further confirmed by comparing the root-mean-square error (RMSE) of the final updated ensembles of permeability as shown in Fig. 16, in which the RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{1}{N_m} \sum_{j=1}^{N_m} (\mathbf{m}_{i,j} - \mathbf{m}_{t,j})^2}, \quad (16)$$

where the subscripts i , j , and t represent realization number, element index, and truth model property, respectively. Similar results presented here are observed for the porosity field as well.

Fig. 15 The means (a-d) and standard deviations (e-h) of log-permeability field in layer 3 for the initial ensemble (first column) and final updated ensembles with production data only (second column), EM data only (third column), and both (last column).

Fig. 16 Boxplot of the RMSE of log-permeability for the initial ensemble (labeled by “Ini”) and the final updated ensembles labeled by “Prod” for production data only case, “EM” for EM data only case, and “Both” for both production and EM data case. 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.

Fig. 17 shows the mean conductivity fields resulting from the final updated ensemble (third row) using both production and EM data at three survey times. The initial mean fields (second row) are very homogeneous and smooth, in which the high conductivity regions originating from injecting wells expands over time due to water saturation changes. However, after history matching the updated mean fields exhibit high heterogeneity and clear resemblance to the reference conductivity fields (first row), evidencing the efficiency of the proposed workflow for integration of EM data.

Fig. 17 Reference conductivity fields (first row) and the means of the conductivity field in layer 3 from the initial ensemble (second row) and final updated ensemble with joint use of production and EM data, respectively, at three EM survey times.

7 Conclusions

We proposed two integration schemes for crosswell EM data using ensemble-based data-assimilation methods. The first scheme takes a direct approach to incorporate raw EM observations (after necessary data processing) into the history matching framework. To achieve that, a forward model simulating observed EM responses need to be established, which requires a good understanding of the underlying physics and a solid theoretical and/or experimental foundation to support it. In the second scheme, the original EM observations are inverted to formation. A feature-based distance parameterization method is used aiming to assimilate the inverted conductivity efficiently and robustly.

We first examined the performance of each scheme for history matching with a simple channel model. The experimental result demonstrated that both schemes are effective in converting the information embedded in the data into valid calibration of model variables, but also exposed some potential issues of each scheme in practical applications. Because the direct integration approach normally requires many forward runs for simulating EM responses, an efficient EM solver is necessary, particularly when a crosswell survey is conducted at multiple frequencies and has dense source locations. Besides, the number of raw EM observations is typically large, which causes a difficulty for ensemble-based history-matching methods in which a small ensemble is usually used. It was shown in the example that distance-based localization is a viable option to overcome this difficulty, meanwhile, reduces the effect of sampling errors. For the feature-based approach, the essential information of observations (inverted conductivity) is retained and the number of data is drastically reduced, which relieves the necessity for localization; however, its performance somehow depends on the quality of inversion results and is susceptible to

observation bias and propagated errors introduced by the inversion. In such a case, it is important to have a good estimation of observation errors or to introduce a systematic measure to take these factors into account. We further investigated the joint assimilation of production and EM data for history matching using the proposed workflow with a more realistic egg model. The results revealed the complementary nature between the two types of data and showed a positive synergy effect on the data match and characterization of model variables when both types of data were used.

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 which formed the basis for our inversion.

References

- Aanonsen, S.I., Nævdal, G., Oliver, D.S., Reynolds, A.C. and Vallès, B. 2009. The ensemble Kalman filter in reservoir engineering--a review. *SPE Journal*, 14(03): 393–412.
- Abadpour, A., Adejare, M., Chugunova, T., Mathieu, H., Haller, N. and others. 2018. Integrated geo-modeling and ensemble history matching of complex fractured carbonate and deep offshore turbidite fields, generation of several geologically coherent solutions using ensemble methods. In *Abu Dhabi International Petroleum Exhibition & Conference*.
- Abubakar, A., Habashy, T.M., Druskin, V.L., Knizhnerman, L. and Alumbaugh, D.L. 2008. 2.5D forward and inverse modeling for interpreting low-frequency electromagnetic measurements. *Geophysics*, 73(4): F165–F177.
- Al-Ali, Z., H. Al-Buali, M., AlRuwaili, S., Ma, S.M., Marsala, A.F., Hulme, C. and Wilt, M. 2009. Looking deep into the reservoir. *Oilfield Review*, 21(2): 38–47.
- Archie, G.E. 1942. The electrical resistivity log as an aid in determining some reservoir characteristics. *Transactions of the AIME*, 146(01): 54–62.
- Baudrier, É., Nicolier, F., Millon, G. and Ruan, S. 2008. Binary-image comparison with local-dissimilarity quantification. *Pattern Recognition*, 41(5): 1461–1478.
- Byrd, R.H., Lu, P., Nocedal, J. and Zhu, C. 1995. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5): 1190–1208.
- Chen, Y. and Oliver, D.S. 2010. Cross-covariances and localization for EnKF in multiphase flow data assimilation. *Computational Geosciences*, 14(4): 579–601.
- Chen, Y. and Oliver, D.S. 2012. Ensemble randomized maximum likelihood method as an iterative ensemble smoother. *Mathematical Geosciences*, 44(1): 1–26.
- Chen, Y. and Oliver, D.S. 2014. History matching of the norne full-field model with an iterative ensemble smoother. *SPE Reservoir Evaluation & Engineering*, 17(02): 244–256.
- 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.
- Chen, Y. and Oliver, D.S. 2017. Localization and regularization for iterative ensemble smoothers. *Computational Geosciences*, 21(1): 13–30.
- Dovera, L. and Della Rossa, E. 2011. Multimodal ensemble Kalman filtering using Gaussian mixture models. *Computational Geosciences*, 15(2): 307–323.
- Dresser, A. 1982. *Well logging and interpretation techniques. The course for home study*. Dresser Atlas.
- Emerick, A. and Reynolds, A. 2011. Combining sensitivities and prior information for covariance localization in the ensemble Kalman filter for petroleum reservoir applications. *Computational Geosciences*, 15(2): 251–269.
- Emerick, A.A. and Reynolds, A.C. 2013. Ensemble smoother with multiple data assimilation. *Computers & Geosciences*, 55(0): 3–15.
- Evensen, G. 2018. Analysis of iterative ensemble smoothers for solving inverse problems. *Computational Geosciences*, 22(3): 885–908.
- Feng, T. and Mannseth, T. 2010. Impact of time-lapse seismic data for permeability estimation. *Computational Geosciences*, 14(4): 705–719.
- Gaspari, G. and Cohn, S.E. 1999. Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125(554): 723–757.
- Gu, Y. and Oliver, D.S. 2007. An iterative ensemble Kalman filter for multiphase fluid flow data assimilation. *SPE Journal*, 12(04): 438–446.
- Hamada, G.M., Almajed, A.A., Okasha, T.M. and Algahe, A.A. 2013. Uncertainty analysis of Archie’s parameters

- determination techniques in carbonate reservoirs. *Journal of Petroleum Exploration and Production Technology*, 3(1): 1–10.
- Haugen, V.E., Naevdal, G., Natvik, L.-J., Evensen, G., Berg, A. and Flornes, K. 2008. History matching using the ensemble Kalman filter on a north sea field case. *SPE Journal*, 13(4): 382–391.
- Hoteit, I., Luo, X. and Pham, D.-T. 2012. Particle Kalman filtering: a nonlinear bayesian framework for ensemble Kalman filters. *Monthly Weather Review*, 140(2): 528–542.
- Iglesias, M.A. 2015. *Iterative regularization for ensemble data assimilation in reservoir models*.
- Iglesias, M.A. and Dawson, C. 2013. The regularizing Levenberg-Marquardt scheme for history matching of petroleum reservoirs. *Computational Geosciences*, 17(6): 1033–1053.
- Jansen, J., Fonseca, R., Kahrobaei, S., Siraj, M., Essen, G. Van and Hof, P. Van Den. 2014. The egg model - a geological ensemble for reservoir simulation. *Geoscience Data Journal*, 1(2): 192–195.
- Katterbauer, K., Arango, S., Sun, S. and Hoteit, I. 2016. Enhanced characterization of reservoir hydrocarbon components using electromagnetic data attributes. *Journal of Petroleum Science and Engineering*, 140: 1–15.
- Katterbauer, K., Arango, S., Sun, S. and Hoteit, I. 2015. Multi-data reservoir history matching for enhanced reservoir forecasting and uncertainty quantification. *Journal of Petroleum Science and Engineering*, 128: 160–176.
- 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.
- Li, G.M. and Reynolds, A.C. 2009. Iterative ensemble Kalman filters for data assimilation. *SPE Journal*, 14(3): 496–505.
- 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., Bhakta, T., Jakobsen, M. and Nævdal, G. 2017. An ensemble 4D-seismic history-matching framework with sparse representation based on wavelet multiresolution analysis. *SPE Journal*, 22(03): 985–1010.
- Luo, X., Bhakta, T. and Nævdal, G. 2018. Correlation-based adaptive localization with applications to ensemble-based 4D-seismic history matching. *SPE Journal*, 23(02): 396–427.
- Luo, X., Hoteit, I., Duan, L. and Wang, W. 2012. Review of nonlinear Kalman, ensemble and particle filtering with application to the reservoir history matching problem. *Nonlinear Estimation and Applications to Industrial Systems Control*: 197–224.
- 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.
- Marsala, A.F., Lyngra, S., Ma, S.M., Alsaif, S.F., Aramco, S., Zhang, P., Ramadan, M. and Abdallah, W. 2017. Saturation mapping in the interwell reservoir volume: a new technology breakthrough. In *SPE Annual Technical Conference and Exhibition*.
- Marsala, A.F., Ruwaili, S., Ma, S.M., Ali, Z., Buali, M., Donadille, J.M., Crary, S. and Wilt, M. 2008. Crosswell electromagnetic tomography : from resistivity mapping to interwell fluid distribution. In *International Petroleum Technology Conference*.
- Mulder, W.A. 2006. A multigrid solver for 3D electromagnetic diffusion. *Geophysical Prospecting*, 54(5): 633–649.
- Newman, G.A. and Alumbaugh, D.L. 2000. Three-dimensional magnetotelluric inversion using non-linear conjugate gradients. *Geophysical Journal International*, 140(2): 410–424.
- Oliver, D.S. and Alfonzo, M. 2018. Calibration of imperfect models to biased observations. *Computational Geosciences*, 22(1): 145–161.
- Oliver, D.S. and Alfonzo, M.A. 2018. Seismic data assimilation with an imperfect model. In *15th European Conference on the Mathematics of Oil Recovery (ECMOR XV)*.
- Oliver, D.S. and Chen, Y. 2018. Data assimilation in truncated plurigaussian models: impact of the truncation map. *Mathematical Geosciences*, 50(8): 867–893.
- Oliver, D.S. and Chen, Y. 2009. Improved initial sampling for the ensemble Kalman filter. *Computational Geosciences*, 13(1): 13–27.
- Oliver, D.S. and Chen, Y. 2011. Recent progress on reservoir history matching: a review. *Computational Geosciences*, 15(1): 185–221.
- Phale, H.A. and Oliver, D.S. 2011. Data assimilation using the constrained ensemble Kalman filter. *SPE Journal*, 16(02): 331–342.
- Plessix, R.E. and Mulder, W.A. 2008. Resistivity imaging with controlled-source electromagnetic data: depth and data weighting. *Inverse Problems*, 24(3): 034012.
- Portniaguine, O. and Zhdanov, M.S. 1999. Focusing geophysical inversion images. *Geophysics*, 64(3): 874–887.
- Schlumberger. 2014. ECLIPSE industry reference reservoir simulator. <https://www.software.slb.com/products/eclipse>.
- Sinclair, G. 1948. Theory of models of electromagnetic systems. *Proceedings of the IRE*, 36(11): 1364–1370.
- Skjervheim, J.-A., Evensen, G., Aanonsen, S.I., Ruud, B.O. and Johansen, T.-A. 2007. Incorporating 4D seismic data in reservoir simulation models using ensemble Kalman filter. *SPE Journal*, 12(03): 282–292.
- Skjervheim, J., Evensen, G., Hove, J. and Vabø, J.G. 2011. An ensemble smoother for assisted history matching. In *SPE Reservoir Simulation Symposium. Society of Petroleum Engineers*.
- Spies, B.R. and Habashy, T.M. 1995. Sensitivity analysis of crosswell electromagnetics. *Geophysics*, 60(3): 834–845.

-
- Stratton, J.A. 2007. *Electromagnetic theory*. John Wiley & Sons.
- Strebel, S. 2002. Conditional simulation of complex geological structures using multiple-point statistics. *Math. Geol.*, 34(1): 1–21.
- Trani, M., Arts, R. and Leeuwenburgh, O. 2013. Seismic history matching of fluid fronts using the ensemble Kalman filter. *SPE Journal*, 18(01): 159–171.
- Tveit, S., Bakr, S.A., Lien, M. and Mannseth, T. 2015. Ensemble-based bayesian inversion of CSEM data for subsurface structure identification. *Geophysical Journal International*, 201(3): 1849–1867.
- Wilt, M.J., Alumbaugh, D.L., Morrison, H.F., Becker, A., Lee, K.H. and Deszcz-Pan, M. 1995. Crosswell electromagnetic tomography: System design considerations and field results. *Geophysics*, 60(3): 871–885.
- Zafari, M. and Reynolds, A.C. 2007. Assessing the uncertainty in reservoir description and performance predictions with the ensemble Kalman filter. *SPE Journal*, 12(3): 382–391.
- 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. 2011. History matching using the ensemble Kalman filter with multiscale parameterization: a field case study. *SPE Journal*, 16(2): 307–317.
- Zhang, Y. and Oliver, D.S. 2010. Improving the ensemble estimate of the Kalman gain by bootstrap sampling. *Mathematical Geosciences*, 42(3): 327–345.
- Zhang, Y., Oliver, D.S., Chen, Y. and Skaug, H.J. 2014. Data assimilation by use of the iterative ensemble smoother for 2D facies models. *SPE Journal*, 20(01): 169–185.
- Zhdanov, M.S. and Yoshioka, K. 2003. Cross-well electromagnetic imaging in three dimensions. *Exploration Geophysics*, 34(2): 34–40.