Journal of Petroleum Science and Engineering 196 (2021) 107961

Contents lists available at ScienceDirect

Journal of Petroleum Science and Engineering

journal homepage: http://www.elsevier.com/locate/petrol

Accounting for model errors of rock physics models in 4D seismic history matching problems: A perspective of machine learning

Xiaodong Luo *, Rolf J. Lorentzen, Tuhin Bhakta

Norwegian Research Centre (NORCE), Norway

ARTICLE INFO

Keywords:
Seismic history matching
Iterative ensemble smoother
Model errors
Residual model
Machine learning

ABSTRACT

Model errors are ubiquitous in practical history matching problems. A common approach in the literature to accounting for model errors is to treat them as random variables following certain presumed distributions. While such a treatment renders algorithmic convenience, its underpinning assumptions are often invalid. In this work, we adopt an alternative approach, and treat model-error characterization as a functional approximation problem, which can be solved using a generic machine learning method. We then integrate the proposed model-error characterization approach into an ensemble-based history matching framework, and show that, with very minor modifications, existing ensemble-based history matching algorithms can be readily deployed to solve the history matching problem in the presence of model errors.

To demonstrate the efficacy of the integrated history matching framework, we apply it to account for potential model errors of a rock physics model in 4D seismic history matching applied to the full Norne benchmark case. The numerical results indicate that the proposed model-error characterization approach helps improve the qualities of estimated reservoir models, and leads to more accurate forecasts of production data. This suggests that accounting for model errors from a perspective of machine learning serves as a viable way to deal with model imperfection in practical history matching problems.

1. Introduction

Quantitative analyses of real-world phenomena often involve using certain numerical models. Due to the limitations of knowledge, capacity and/or available resources, certain simplifying assumptions or strategies are inevitably adopted in order to make such numerical models effective in practice. This situation is largely reflected by the aphorism "all models are wrong, but some are useful" (Box et al., 2011).

For geophysical data assimilation problems, such as reservoir characterization through computer-assisted history matching, a typical workflow consists of a forward simulator, a collection of measured data (called observations hereafter), and a data assimilation algorithm that aims to update model variables (state and/or parameters) to match the observations to a good extent. As a numerical model, the forward simulator will, more or less, suffer from certain model errors. Consequently, model errors will be propagated through the forward simulator and the inversion (data assimilation) algorithm, and will have an undesired impact on the estimated model variables.

As such, model errors are ubiquitous in geophysical data assimilation problems, whereas at the time being, how to properly handle model errors appears to remain as an open topic in data assimilation community. The practical challenges in dealing with model errors stem from factors like the complexities in quantitatively analyzing and characterizing the sources of model errors, the dependence of model errors on (uncertain) model variables, and consequently the tangled effects of both model errors and model variables on assimilation algorithms (Luo, 2019).

Therefore, to account for model errors in data assimilation problems, it is necessary to adopt certain strategies that take into account these noticed challenges. In this regard, perhaps the most commonly adopted strategy so far is to assume that model errors are independent of model variables, and treat model errors as random variables that follow certain presumed (typically Gaussian) distributions. By doing so, the impacts of model variables and model errors on the simulated data are disentangled, and the effect of model errors is taken into account by incorporating statistics (e.g., mean and covariance matrix) of model errors into the assimilation algorithm (Dee, 1995; Williamson et al., 2015; Evensen, 2019; Alfonzo and Oliver, 2020; Rammay et al., 2019; Lu and

* Corresponding author.
E-mail address: xluo@norceresearch.no (X. Luo).

https://doi.org/10.1016/j.petrol.2020.107961
Received 26 March 2020; Received in revised form 14 September 2020; Accepted 17 September 2020
Available online 7 October 2020

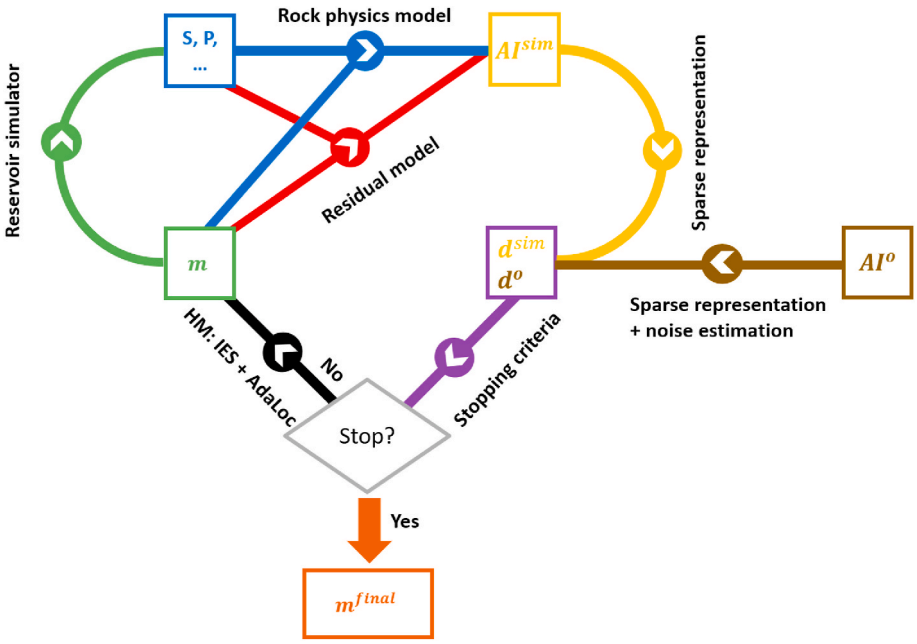


Fig. 1. A schematic outline of the 4D seismic history matching (SHM) framework used in the current study. For ease of reference, we adopt a color scheme for different procedures (represented by edges with arrows) and associated quantities (represented by rectangular boxes). More information regarding the 4D SHM framework is provided in the text. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Chen, 2020; He et al., 2020; Neto et al., 2020). While enjoying theoretical simplicity and algorithmic convenience, the underlying assumptions (e.g., model errors being independent of model variables, Gaussian model errors) with respect to this type of model-error treatment may be invalid in practice (Dee, 1995).

There are also investigations dedicated to studying the effects of model errors on history matching, but without assuming the independence between model errors and model variables, see, for example, (Stephen, 2007; Köpke et al., 2018; Luo, 2019). In practice, though, taking into account the dependence of model errors on model variables may nullify the conventional Gaussianity assumption adopted for the purpose of characterizing model errors.[1] In effect, this might then make the resulting history matching workflow more complicated, unless additional sources of information are assumed. For instance, Stephen (2007) investigated scale and process dependent model errors in a synthetic seismic history matching problem. To quantify the effect of model errors, the author assumed that the geo-model is accurate at some fine scale, whereas model errors are mainly resulted from the differences between the fine and coarse grid geo-models. In a similar setting, with model errors also stemming from the differences between some fine and coarse grid geo-models, Köpke et al. (2018) developed a local basis approach to characterizing model errors, but with an additional assumption that data residuals (defined as the difference between observed and simulated data) due to model errors and model variables are orthogonal.

In a recent study, Luo (2019) investigated an alternative approach to accounting for model errors in a class of data assimilation problems. The basic idea there is to treat model-error correction (MEC) as a data-driven functional optimization problem, which is then solved from a perspective of machine learning. Recent years have witnessed the wide applications of various machine learning technologies, e.g., supervised (Lee et al., 2019; Luo, 2019; Ma and Leung, 2020) or unsupervised (Shams et al., 2020; Mosser et al., 2017; Canchumuni et al., 2019) learning methods, to many problems in geosciences, see, for example, Tahmasebi et al. (2020) for a recent review. To our knowledge, though, using machine learning methods to tackle the issue of model errors appears to be a topic that is rarely touched.

By considering model-error as some unknown functional that relates model variables to data residuals, one can explicitly take into account the dependence of model errors on model variables and describe the effects of model errors and model variables on the assimilation algorithm, while without the need to presume a statistical distribution for model errors. In addition, this MEC approach can be integrated into ensemble-based history matching frameworks in such a way that only very minor modifications of the ensemble-based history matching algorithm are required. On the other hand, though, certain caution also has to be exercised regarding the capacity of this MEC approach. Indeed, as a data-driven approach, it may not be able to correctly identify the sources or origins of model errors, which is an issue that, in the authors' opinion, goes beyond the capacities of contemporary MEC approaches in most of the practical problems.

The goal of the current work is to demonstrate the applicability of the MEC approach in Luo (2019) to account for potential model errors from a rock physics model (RPM) used to history-match 4D seismic data in a full Norne field case study (Lorentzen et al., 2020). As the theoretical development of the data-driven MEC approach was carried out in the previous work Luo (2019), here we focus more on the practical aspects and document the details of the real field case study, while leaving out certain investigations, e.g., a comparison of the conventional MEC approach and the data-driven one used here, which were conducted in the previous work Luo (2019). In terms of novelty/contribution of the current work, to our knowledge, it is the first time that a data-driven MEC approach is applied to a real field case study and clearly demonstrates the potential of improving the history-matching performance.

For ease of comprehension, in the sequel we start from briefly introducing the 4D seismic history matching framework adopted in the current work, and explain how model errors are handled in ensemble-based history matching from a perspective of machine learning. Afterwards, we demonstrate the performance of the strengthened 4D seismic history matching framework (with the MEC approach being incorporated) through a case study in the full Norne field. We conclude the whole work with some technical discussions and thoughts for potential

[1] As in practice, model variables are often assumed to follow certain Gaussian distributions, while model errors may be nonlinearly related to model variables.

future work.

2. The 4D seismic history matching framework

With substantially larger spatial coverages, 4D seismic data provide valuable additional information that is complementary to that from the conventional core sample, well log and production data. For this reason, in the past two decades, there is increasing interest in incorporating 4D seismic data for quantitative reservoir characterization through computer-assisted history matching, for example, see Abadpour et al. (2013); Dadashpour et al. (2008); Emerick (2016); Gosselin et al. (2003); Ketineni et al. (2018); Lorentzen et al. (2019); Luo et al. (2017, 2018a); Obidegwu et al. (2017); Soares et al., 2020; Taha et al. (2019); Trani et al. (2012); Zhang and Leeuwenburgh (2017); Skjervheim et al. (2007).

Fig. 1 presents a schematic outline of the 4D SHM framework used in the current study, which consists of a few procedures adopted for specific purposes. For ease of reference, we adopt a color scheme to represent different procedures and quantities that are fed into or generated by the procedures. In the sequel, we provide short descriptions of these procedures and the associated quantities.

2.1. Forward seismic simulations

Forward seismic simulations aim to generate simulated acoustic impedance (AI) data AI^{sim} (in orange) based on input reservoir models $\mathbf{m}$ (in green). In Fig. 1, this procedure involves three components, namely, the reservoir simulator (in green), the rock physics model (in blue), and the residual model (in red) introduced to account for potential model errors from the RPM. The reservoir simulator takes a reservoir model as the input, and outputs profiles of fluid saturation $\mathbf{S}$ and pressure $\mathbf{P}$ (both in blue). The simulated saturation and pressure profiles, together with some of the reservoir model variables, such as porosity (and sometimes net-to-gross ratio, as in Lorentzen et al., 2020, 2019), are then taken as the inputs to both the RPM and the residual model. The RPM is often field-dependent and varies from case to case, whereas in the current work the residual model is built from a perspective of machine learning as in Luo (2019). More information regarding the RPM and the residual model will be provided in the case-study part later.

In Fig. 1, AI is taken as the attribute in SHM. The same framework can also be extended to other types of attributes, such as amplitude-versus-angle data, as was done in the previous work (Luo et al., 2018a, 2017; Soares et al., 2020).

2.2. Sparse data representation and noise level estimation

The sparse data representation procedure aims to reduce the size of observations in SHM, while retaining as much information as possible. The main idea behind sparse representation is to apply a transform to the data, and represent the original data in another domain, in terms of certain "dominant" coefficients whose number is substantially smaller than the original data size. It is these "dominant" coefficients that will be used as the observations for model updates later on, through a certain history matching algorithm. Therefore, data-size reduction through sparse representation helps mitigate the issue of high-demanding requirements on computational resources (e.g., computer memory) in big data assimilation problems.

In Fig. 1, the sparse representation procedure is applied to both the simulated seismic data AI^{sim} (in orange) and the observed ones AI^{o} (in coffee color). Since in this work we revisit the real field case study in Lorentzen et al. (2020), for consistency we adopt the discrete wavelet transforms (DWT) for sparse data representation. Consequently, we apply a wavelet-based denoising algorithm to estimate noise levels in the observed data AI^{o}. Readers are referred to Lorentzen et al. (2020) for more information in this regard.

2.3. Model updates

The SHM framework in Fig. 1 entails an iterative process. Given sparse representations of both the simulated and observed data, we first calculate the corresponding data mismatch values and then use them to decide whether the iterative process should be stopped or not (in grey), according to certain preset stopping criteria (in purple) (the definition of data mismatch and the concrete stopping criteria will be provided later in the case-study part). If the history matching process terminates, then one obtains a final ensemble of estimated reservoir models $\mathbf{m}^{final}$ (in brown). Otherwise, a certain history matching algorithm (in black) will be used to further update the reservoir models, which will then be used as the inputs to the reservoir simulator to start a new iteration.

In the current work, the iterative ensemble smoother (IES) from Luo et al. (2015), called the regularized Levenberg-Marquardt algorithm for a minimum average cost (RLM-MAC) problem, is adopted, although other algorithms, e.g., the ensemble smoother with multiple data assimilation (ES-MDA) (Emerick and Reynolds, 2012), an approximate form of the ensemble-based randomized maximum likelihood (EnRML) algorithm (Chen and Oliver, 2013), or an ensemble subspace implementation of the EnRML (Raanes et al., 2019), may also be employed. In applications of ensemble-based methods to real-world problems, especially those with relatively big datasets but relatively small ensemble sizes, ensemble collapse is a commonly encountered problem. To mitigate this issue, localization is typically adopted. The conventional localization schemes are often based on distances between the physical locations of both observations and reservoir model variables (Chen and Oliver, 2010; Raniolo et al., 2013; Emerick and Reynolds, 2011). In our SHM framework, however, the effective observations (e.g., wavelet coefficients) resulting from sparse data representation do not have associated physical locations, and this makes it challenging to apply distance-based localization. As a result, we adopt a correlation-based adaptive localization scheme instead, which helps overcome the aforementioned problem, while achieving a few additional benefits (Luo et al., 2018b, 2019). More details of the implementation of correlation-based adaptive localization will also be provided in the case-study part later.

3. Accounting for model errors in history matching problems: a perspective of machine learning

Suppose that in a generic history matching problem, we have a collection $\mathbf{d}^{o}$ of observations, which is produced through the following observation system

$$\mathbf{d}^{o} = \mathbf{f}(\mathbf{m}^{tr}) + \boldsymbol{\varepsilon}. \tag{1}$$

In Eq. (1), from right to left, ε stands for some additive observation noise which is assumed to follow a certain Gaussian distribution with zero mean and covariance matrix $\mathbf{C_d}$; $\mathbf{m}^{tr}$ for the true reservoir model (truth); and $\mathbf{f}$ for the perfect forward simulator without any model error.

Due to the presence of model errors, one ends up with using an imperfect simulator, denoted by $\mathbf{g}$ here, for forward simulations. In this case, given an input reservoir model $\mathbf{m}$, the simulated observations $\mathbf{d}^{sim}$ are generated by using $\mathbf{g}$ to map $\mathbf{m}$ onto the observation space, i.e.,

$$\mathbf{d}^{sim} = \mathbf{g}(\mathbf{m}). \tag{2}$$

We note that $\mathbf{m}$ in Eq. (2) may have a different dimensionality from that of $\mathbf{m}^{tr}$ in Eq. (1), if one considers the fact that in practice, the uncertain parameters in a reservoir model may not be able to completely explain the observations $\mathbf{d}^{o}$. In other words, the model variables in $\mathbf{m}$ may just constitute an incomplete list of possible factors that have resulted in the observed data.

During history matching, we adopt an IES to iteratively update an ensemble of reservoir models until certain stopping criteria are met (the concrete stopping criteria in the field case study will be specified later in

§4.3). Suppose that, at the ith iteration step, we have an ensemble of N_e reservoir models, denoted by $\mathscr{M}^i \equiv \{\mathbf{m}_j^i\}_{j=1}^{N_e}$ hereafter, as the background (i.e., the prior), where j is the index of ensemble members. Here our objective is to update $\mathscr{M}^i$ to a new ensemble $\mathscr{M}^{i+1} \equiv \{\mathbf{m}_j^{i+1}\}_{j=1}^{N_e}$, which serves as the background ensemble for the next iteration step if none of the stopping criteria is met, or as the final ensemble of the history matching process otherwise.

3.1. Model updates without accounting for model errors

Without taking into account the presence of potential model errors, a new ensemble $\mathscr{M}^{i+1}$ is obtained by solving the following minimum-average-cost (MAC) problem (Luo et al., 2015):

$$\underset{\{\mathbf{m}_j^{i+1}\}_{j=1}^{N_e}}{\operatorname{argmin}} \frac{1}{N_e}\sum_{j=1}^{N_e}\left\{\left(\mathbf{d}^o-\mathbf{g}\left(\mathbf{m}_j^{i+1}\right)\right)^T\mathbf{C}_\mathbf{d}^{-1}\left(\mathbf{d}^o-\mathbf{g}\left(\mathbf{m}_j^{i+1}\right)\right)+\boldsymbol{\gamma}^\mathbf{i}\left(\mathbf{m}_\mathbf{j}^{\mathbf{i+1}}-\mathbf{m}_\mathbf{j}^\mathbf{i}\right)^\mathbf{T}\left(\mathbf{C}_\mathbf{m}^\mathbf{i}\right)^{-\mathbf{1}}\left(\mathbf{m}_\mathbf{j}^{\mathbf{i+1}}-\mathbf{m}_\mathbf{j}^\mathbf{i}\right)\right\}, \tag{3}$$

where $\mathbf{C}_\mathbf{m}^i$ is the sample covariance matrix induced by the ensemble $\mathscr{M}^i$, in the form of $\mathbf{C}_\mathbf{m}^i = \mathbf{S}_m^i(\mathbf{S}_m^i)^T$, whereas the square root matrix $\mathbf{S}_m^i$ will be defined later (cf. Eq. (7)); and γ^i is a parameter that is adaptive to the iteration process, following certain preset rules (Luo et al., 2015).

The main idea in Eq. (3) is to minimize the average of an ensemble of cost functions. Each cost function consists of two parts, namely, the data mismatch term (the first term counting from the left side) which measures the distance between simulated and observed data, and the regularization term (the second one) which is introduced to mitigate the ill-posedness in the history matching problem, and also to prevent the IES from overfitting the observations.

Through a linearization-based approximation strategy, Eq. (3) is solved as follows:

$$\mathbf{m}_j^{i+1} = \mathbf{m}_j^i + \mathbf{K}^i\left(\mathbf{d}^o-\mathbf{g}\left(\mathbf{m}_j^i\right)\right),\ j=1,2,\cdots,N_e; \tag{4}$$

$$\mathbf{K}^i \equiv \mathbf{S}_m^i\left(\mathbf{S}_d^i\right)^T\left(\mathbf{S}_d^i\left(\mathbf{S}_d^i\right)^T+\gamma^i\mathbf{C}_d\right)^{-1}; \tag{5}$$

$$\overline{\mathbf{m}}^i \equiv \frac{1}{N_e}\sum_{j=1}^{Ne} N_e\mathbf{m}_j^i; \tag{6}$$

$$\mathbf{S}_m^i \equiv \frac{1}{\sqrt{N_e-1}}\left[\mathbf{m}_1^i-\overline{\mathbf{m}}^i,\mathbf{m}_2^i-\overline{\mathbf{m}}^i,\cdots,\mathbf{m}_{N_e}^i-\overline{\mathbf{m}}^i\right]; \tag{7}$$

$$\mathbf{S}_d^i \equiv \frac{1}{\sqrt{N_e-1}}\left[\mathbf{g}\left(\mathbf{m}_1^i\right)-\mathbf{g}\left(\overline{\mathbf{m}}^i\right),\mathbf{g}\left(\mathbf{m}_2^i\right)-\mathbf{g}\left(\overline{\mathbf{m}}^i\right),\cdots,\mathbf{g}\left(\mathbf{m}_{N_e}^i\right)-\mathbf{g}\left(\overline{\mathbf{m}}^i\right)\right]. \tag{8}$$

On the other hand, a practical implementation of the IES deviates the theoretical form presented in Eqs. (4) and (5), in order to improve the numerical stability and history matching performance of the algorithm. The deviations stem from the following two changes: (1) a truncated singular value decomposition (TSVD) conducted to approximate the square root matrix $\mathbf{S}_d^i$ in Eq. (8); and (2) localization applied to a certain Kalman-gain-like matrix.

In this work, the TSVD is conducted following the previous work Luo et al. (2019). On the other hand, we adopt correlation-based adaptive localization, since in practice it tends to be more flexible than distance-based localization, as elaborated in Luo et al. (2019). For more information of how correlation-based adaptive localization is conducted, readers are referred to Luo et al. (2018b, 2019); Lorentzen et al. (2019).

3.2. Model updates taking into account model errors

To take into account model errors in history matching problems, a natural choice is to provide a certain mechanism that, to some extent, compensates for model imperfection. If there are certain prior knowledge, experience and/or expertise that help identify potential sources of errors, then they should be exercised for model improvements. Our starting point here is that all such attempts for model improvements have been tried. Under this assumption, we consider a data-driven approach for possible further model improvements.

One way of providing a compensation is to introduce a residual functional $\mathbf{r}$ to the imperfect simulator $\mathbf{g}$ in Eq. (2), such that we obtain a modified forward simulator

$$\mathbf{d}^{sim} = \mathbf{g}(\mathbf{m}) + \mathbf{r}(\mathscr{T}(\mathbf{m}^g)), \tag{9}$$

where $\mathscr{T}$ represents a certain transform operator that maps a set of variables $\mathbf{m}^g$ to some other quantities (e.g., saturation or pressure fields through reservoir simulations, in case that $\mathbf{m}^g$ is identical to the reservoir model $\mathbf{m}$). Reflecting the previously mentioned difference between $\mathbf{m}^{tr}$ in Eq. (1) and $\mathbf{m}$ in Eq. (2), $\mathbf{m}^g$ in Eq. (9) may also not be the same as $\mathbf{m}$ in general. In fact, one of the main challenges in practice is how to efficiently identify a proper set of variables $\mathbf{m}^g$ that help explain the differences between observed and simulated data.[2] The approach to be presented below is not proposed to address this issue. Instead, we focus on how to find a proper functional form $\mathbf{r}$, for a given transform operator $\mathscr{T}$ and a chosen list of uncertain variables $\mathbf{m}^g$. For ease of discussion, in the sequel we simply let $\mathbf{m}^g = \mathbf{m}$, while the derivation below can also be extended to the case $\mathbf{m}^g \neq \mathbf{m}$. Also, without loss of generality, we further simplify the situation and let $\mathscr{T}$ be an identity operator (otherwise we can consider the composition functional of $\mathscr{T}$ and $\mathbf{r}$). As a result, we end up with the following simplified forward simulator

$$\mathbf{d}^{sim} = \mathbf{g}(\mathbf{m}) + \mathbf{r}(\mathbf{m}). \tag{10}$$

To take into account model errors in history matching problems, we then replace the original forward simulator in Eq. (2) by that in Eq. (10). However, note that the functional form of $\mathbf{r}$ is not known in advance. As a result, during history matching we need to simultaneously optimize both the reservoir model variables $\mathbf{m}$ and the functional $\mathbf{r}$. As discussed in Luo (2019), this optimization problem becomes intractable if without any constraint on the functional form of $\mathbf{r}$ (as this leads to an optimization problem in an infinite-dimensional space). For this reason, we assume that the residual functional $\mathbf{r}$ belongs to a certain Reproducing Kernel Hilbert Space, such that $\mathbf{r}$ can be approximated by a set of corresponding kernel functions, in the form of

$$\widehat{\mathbf{r}}(\mathbf{m};\boldsymbol{\theta}) = \sum_{k=1}^{Ncp} N_{cp}c_k\,\mathbf{K}_k(\mathbf{m}_k^{cp},\mathbf{m};\boldsymbol{\beta}_k), \tag{11}$$

[2] In principle, one can of course conduct intense search in a large enough space through the trial-and-error approach. This, however, will become computationally inefficient.

Table 1
Parameterization of the reservoir models.

Parameter	Krg	Krw	NTG	MULTFLT	MULTREGT	MULTZ	OWC	PERMX	PORO
Number	4	4	44,927	53	3	13,309	5	44,927	44,927

$$\boldsymbol{\theta} \equiv \left[c_1, c_2, \cdots, c_{N_{cp}}, \boldsymbol{\beta}_1^T, \boldsymbol{\beta}_2^T, \cdots, \boldsymbol{\beta}_{N_{cp}}^T\right]^T, \tag{12}$$

where N_{cp} is the total number of center points (cp), denoted by $\mathbf{m}_k^{cp}$, which are chosen (and then fixed) in advance; c_k is the weight coefficient associated with the kth kernel function $\mathbf{K}_k$ (which is also pre-chosen and fixed); $\boldsymbol{\beta}_k$, as a (column) vector in general, represents a set of intrinsic parameters of $\mathbf{K}_k$; and $\boldsymbol{\theta}$ is the set of total parameters resulting from kernel-based functional approximation. Details of implementing Eq. (11) in the field case study will be provided later.

We have a few remarks related to the issue of optimizing the residual functional $\mathbf{r}$. First, for the purpose of approximating $\mathbf{r}$ itself, one is not restricted to adopt the kernel based method, as in Eq. (11). Instead, there are many options, e.g., generic parametrized learning models such as various shallow or deep neural networks, which one can exploit, as long as such models serve the purpose of functional approximation well. However, we note that, in this context, the problem of optimizing the residual functional is different from a supervised learning problem, as discussed in Luo (2019). For the latter problem, typically one has many instances of input-output pairs to learn the functional relations. While for the former problem considered here, it is more in line with the typical setting of a data assimilation problem, in which there is only one instance of realized observation at a given spatial and temporal location. In addition, unlike the supervised learning problem, we have no access to the true model (truth) that generates the observations. This means that, the observed residual with respect to $\mathbf{m}$, in terms of $\mathbf{d}^o - \mathbf{g}(\mathbf{m})$, would depend on both the model $\mathbf{m}$ and the truth $\mathbf{m}^{tr}$, rather than only on $\mathbf{m}$. Since the truth $\mathbf{m}^{tr}$ is unknown, a part of the information will miss for the purpose of estimating $\mathbf{r}$. To mitigate this problem, a possible strategy is to also take the observation $\mathbf{d}^o$ as a part of the inputs, as $\mathbf{d}^o$ is implicitly related to $\mathbf{m}^{tr}$. In this case, however, we will also bring the noise term $\boldsymbol{\varepsilon}$ into the estimation of $\mathbf{r}$.

Second, by approximating $\mathbf{r}$ by a suitable parametrized learning model, one converts the problem of optimizing the residual functional into a problem of optimizing a set of parameters (e.g., c_k and $\boldsymbol{\beta}_k$ in Eq. (11)) associated with the chosen learning model. This conversion ensures very minor modifications of the IES introduced previously for the purpose of accounting for model errors in history matching problems, as will become evident soon. Third, in general, the kernel-based learning model in Eq. (11) will be able to accommodate a mixture of different types of kernel functions, which is the reason that we adopt a subscript k for each kernel function $\mathbf{K}_k$ therein. In addition, the intrinsic kernel parameters $\boldsymbol{\beta}_k$ will be estimated in the process of history matching, instead of being manually set a priori.

In consistency with ensemble-based history matching, let us define $\boldsymbol{\Theta}^{i+1} \equiv \{\boldsymbol{\theta}_j^{i+1}\}_{j=1}^{N_e}$ as an ensemble of learning-model parameters at the $(i+1)$ th iteration step, then to take into account the presence of model errors, the MAC problem in Eq. (3) is modified as follows:

$$\underset{\left\{\widehat{\mathbf{m}}_j^{i+1}\right\}_{j=1}^{N_e}}{\operatorname{argmin}} \frac{1}{N_e} \sum_{j=1}^{N_e} \left\{ \left(\mathbf{d}^o - \widehat{\mathbf{g}}\left(\widehat{\mathbf{m}}_j^{i+1}\right)\right)^T \mathbf{C}_{\mathbf{d}}^{-1} \left(\mathbf{d}^o - \widehat{\mathbf{g}}\left(\widehat{\mathbf{m}}_j^{i+1}\right)\right) + \gamma^i \left(\widehat{\mathbf{m}}_j^{i+1} - \widehat{\mathbf{m}}_j^i\right)^T \left(\mathbf{C}_{\widehat{\mathbf{m}}}^i\right)^{-1} \left(\widehat{\mathbf{m}}_j^{i+1} - \widehat{\mathbf{m}}_j^i\right) \right\}, \tag{13}$$

$$\widehat{\mathbf{m}}_j^{i+1} \equiv \left[\left(\mathbf{m}_j^{i+1}\right)^T, \left(\boldsymbol{\theta}_j^{i+1}\right)^T\right]^T, \tag{14}$$

$$\widehat{\mathbf{g}}\left(\widehat{\mathbf{m}}_j^{i+1}\right) \equiv \mathbf{g}\left(\mathbf{m}_j^{i+1}\right) + \widehat{\mathbf{r}}\left(\mathbf{m}_j^{i+1}; \boldsymbol{\theta}_j^{i+1}\right), \tag{15}$$

where $\widehat{\mathbf{m}}_j^{i+1}$ is an augmented vector consisting of both reservoir model variables $\mathbf{m}_j^{i+1}$ and learning-model parameters $\boldsymbol{\theta}_j^{i+1}$; $\widehat{\mathbf{g}}$ is the modified forward simulator, which takes as the output the sum of the simulated data from the original imperfect forward simulator $\mathbf{g}$ and a residual term calculated by $\widehat{\mathbf{r}}$; and $\mathbf{C}_{\widehat{\mathbf{m}}}^i$ stands for the sample covariance matrix derived from the augmented ensemble $\widehat{\mathscr{M}}^{i+1} \equiv \{\widehat{\mathbf{m}}_j^{i+1}\}_{j=1}^{N_e}$.

A comparison between the MAC problems in Eqs. (3) and (13) reveals that they are essentially the same type of optimization problem. As a result, the IES algorithm (including the adopted TSVD and localization techniques for performance improvement) introduced in the preceding sub-section can be readily applied to solve the MAC problem in Eq. (13), but with some necessary minor modifications as indicated in Eqs. (14) and (15), namely, the augmentation of reservoir model variables and learning-model parameters, and the modification of the "effective" forward simulator, respectively.

4. Application to 4D SHM in the full norne field case study: case-study settings

Norne is an offshore oil field located in the Norwegian Sea. In the current study, we use a numerical reservoir model for the full Norne field from the website http://www.ipt.ntnu.no/~norne. This numerical model consists of four segments: C, D, E and G, and has five primary formations (from top to bottom): Garn, Not, Ile, Tofte and Tilje. The dimension of the reservoir model is $46 \times 112 \times 22$. Therefore, the total number of reservoir gridblocks is 113,344, among which 44,927 gridblocks are active.

History matching real production and/or seismic data of the Norne field through a certain ensemble based method has been carried out in some recent work, e.g., Chen and Oliver (2014); Evensen and Eikrem (2018); Lorentzen et al. (2020); Lu and Chen (2020); Luo et al. (2019). Especially, Lu and Chen (2020) considered to account for potential model errors through assimilating production data. In the current work, we focus on accounting for model errors from a perspective of machine

learning. We adopt the kernel-based learning approach, which can be considered as an analogy to a neural network with a single hidden layer, and is trained by the IES. We use seismic data as the "training dataset", and adopt production data to cross-validate the qualities of reservoir models obtained from 4D SHM.

This section mainly aims to explain the case-study settings in the field case study. Following the outline of the 4D SHM framework in Fig. 1, we document the details of setting up different components within the following three procedures, namely, forward seismic simulations, data processing (sparse representation and noise estimation) and model updates through history matching.

4.1. The forward seismic simulator

As indicated in Fig. 1 and discussed in the corresponding text, the forward seismic simulator contains a forward reservoir simulator and a rock physics model. In addition, a residual model is also included to account for potential model errors in forward seismic simulations.

4.1.1. The forward reservoir simulator and reservoir model parametrization

A forward reservoir simulator takes a numerical reservoir model as the input to simulate certain dynamical fields (e.g., fluid saturation and pressure) at some specified time instances. In our case study, we use the ECLIPSE© black oil simulator, which was executed under the control mode of reservoir volume for production wells, and the mode of fluid (water alternating gas) rates (RATE) for injection wells.

The parametrization of the numerical reservoir model, in terms of uncertain parameters to be estimated in history matching, is summarized in Table 1. Specifically, in history matching, the uncertain parameters include porosity (PORO) Φ, x-dimensional permeability (PERMX, in the scale of natural logarithm)[3] and net-to-gross (NTG) ratio distributed over active reservoir gridblocks. In addition, we also estimate 13,309 vertical transmissibility multipliers (MULTZ, in the scale of $\log_{10}$) distributed over active reservoir gridblocks in six layers (namely, layer 1, 8, 11, 12, 15 and 18), 53 fault transmissibility multipliers (MULTFLT, in the scale of $\log_{10}$), 8 multipliers for scaled end-point water (Krw) and gas (Krg) relative permeabilities of four geological formations (with the inactive Not formation being excluded), 3 transmissibility multiplier between regions (MULTREGT, from region number 17 to 13, 18 to 14, and 19 to 15, respectively, in the scale of $\log_{10}$), and 5 oil-water-contact (OWC) depth values in equilibration (EQUIL) data specification. Under this setting, the total number of parameters to estimate is 148,159. An initial ensemble of 100 reservoir models with this parametrization setting can be found from GitHub (Lorentzen, 2018), which is the one used in the 4D SHM study below. We note that the same initial ensemble was previously used in both synthetic and real field case studies, see Lorentzen et al. (2019, 2020).

4.1.2. The rock physics model

A rock physics model (RPM) (also referred to as petro-elastic model) aims to link properties within reservoir and seismic domains. They are used to simulate various seismic attributes, such as P-wave velocity (v_p), S-wave velocity (v_s), density (ρ), acoustic impedance (z_p), shear impedance (z_s) etc., using both dynamic and static reservoir model variables like water saturation (S_w), pore-pressure (P), porosity (Φ) and so on, as the inputs.

In the current work, we use acoustic impedances (AI) z_p distributed over active reservoir gridblocks as the seismic attribute in 4D SHM, whereas the corresponding RPM is outlined in Appendix A. Putting together Eqs. (A.1) – (A.7) therein, the effective RPM is summarized as

Table 2
Ranges of the first five inputs of the residual model.

Input	Φ	NTG	S_w	S_g	P (MPa)
Range	$[0.001, 0.4]$	$[0.001, 1]$	$[0.001, 1]$	$[0.001, 1]$	$[1, 65.22]$

$$z_p = \mathrm{RPM}\left(\Phi, NTG, S_w, S_g, P, \rho_w, \rho_g, \rho_o\right), \tag{16}$$

which depends on both static reservoir model variables Φ (porosity) and NTG to be estimated through SHM, and dynamical variables like fluid saturations S_w and S_g, pore pressure P and fluid densities ρ_w, ρ_g and ρ_o that are all generated by the reservoir simulator (ECLIPSE). Eq. (16) does not contain oil saturation S_o, as it can be calculated through the constraint $S_o = 1 - S_w - S_g$. On the other hand, all the other quantities that are involved in Eqs. (A.1) – (A.7) but do not appear in Eq. (16) (e.g., K_{quartz} and G_{quartz}), have preset and fixed values. We treat these quantities as constants of the RPM and thus do not include them in Eq. (16). We note that the dynamical variables such as S_w, S_g and P are obtained from reservoir simulations, and thus (implicitly) depend on static reservoir model variables like porosity, NTG, permeability and so on. As such, the RPM in Eq. (16) can be considered as a function that essentially describes the relations between static reservoir model variables and the AI data.

4.1.3. The residual model

Complementary to the RPM in Eq. (16), we introduce a residual model $\widehat{r}$ to each active reservoir gridblock. For simplicity, $\widehat{r}$ is set to be the same for all active reservoir gridblocks at all seismic survey time instances. Following Eq. (30) of Luo (2019), the residual model $\widehat{r}$ on the ℓ -th gridblock consists of a set of Gaussian kernels, defined as follows (also cf. Eq. (11) of the current work)[4]:

$$\widehat{r}(\mathbf{V}_\ell; \boldsymbol{\theta}) = \sum_{k=1}^{Ncp} N_{cp} c_k \exp\Big\{ -\langle \boldsymbol{\beta}_k^2, (\mathbf{V}_\ell - \mathbf{V}_k^{cp})^2 \rangle \Big\}, \ \ell = 1, 2, \cdots, N_{agb}, \tag{17}$$

$$\boldsymbol{\theta} \equiv \left[c_1, c_2, \cdots, c_{N_{cp}}, \boldsymbol{\beta}_1^T, \boldsymbol{\beta}_2^T, \cdots, \boldsymbol{\beta}_{N_{cp}}^T \right]^T, \tag{18}$$

$$\langle \boldsymbol{\beta}_k^2, (\mathbf{V}_\ell - \mathbf{V}_k^{cp})^2 \rangle \equiv \frac{1}{2m_v} \sum_{s=1}^{mv} m_v \beta_{k,s}^2 \left(V_{\ell,s} - V_{k,s}^{cp} \right)^2, \tag{19}$$

where c_k, $\boldsymbol{\beta}_k$ and $\mathbf{V}_k^{cp}$ are weight parameter, scale parameter and center point associated with the k-th Gaussian kernel function, respectively; N_{agb} ($= 44927$) is the number of active gridblocks; and N_{cp} is the number of kernel functions (or center points), which is set to 20,000 through a trial-and-error procedure, by examining its impact on the data mismatch of forecast production data. The scale parameter $\boldsymbol{\beta}_k$ is in the same dimension (denoted by m_v) as $\mathbf{V}_k^{cp}$, meaning that for each element of $\mathbf{V}_\ell$ and $\mathbf{V}_k^{cp}$, denoted by $V_{\ell,s}$ and $V_{k,s}^{cp}$ respectively in Eq. (19), there is an associated scale coefficient $\beta_{k,s}$. In 4D SHM, the weight and scale parameters will be simultaneously estimated together with the aforementioned reservoir model variables (cf. Table 1).

In Eq. (17), the vector $\mathbf{V}_\ell$ contains a set of properties as the inputs into the residual model. In the current work,

$$\mathbf{V}_\ell = \left[\Phi_\ell, NTG_\ell, S_{w,\ell}, S_{g,\ell}, P_\ell, z_{p,\ell}^o \right]^T, \tag{20}$$

where Φ_ℓ, NTG_ℓ, $S_{w,\ell}$, $S_{g,\ell}$, P_ℓ and $z_{p,\ell}^o$ represent porosity, NTG, water and gas saturations, pressure and observed AI data, respectively, on the

[3] We did not take permeability along y (PERMY) and z (PERMZ) directions as uncertain parameters, because in the reservoir models, the values of PERMY and PERMZ are proportional to those of PERMX.

[4] For notational convenience and distinction, here we have used a different notation, $\mathbf{V}$ rather than $\mathbf{m}$, to represent a set of model parameters associated with each active reservoir gridblock.

ℓ -th reservoir gridblock. Likewise, $\mathbf{V}_k^{cp}$ are defined in a similar way:

$$\mathbf{V}_k^{cp} = \left[\Phi_k^{cp}, NTG_k^{cp}, S_{w,k}^{cp}, S_{g,k}^{cp}, P_k^{cp}, z_{p,k}^{o,cp}\right]^T. \tag{21}$$

The first five inputs of $\mathbf{V}_\ell$, namely Φ_ℓ, NTG_ℓ, $S_{w,\ell}$, $S_{g,\ell}$ and P_ℓ, are adopted in light of the RPM in Eq. (16). Our experience indicates that the fluid densities, namely ρ_w, ρ_g and ρ_o, have relatively narrow ranges with respect to different ensemble members, thus we do not include them in Eq. (20) to reduce the dimension of the input $\mathbf{V}_\ell$ (which in turn helps to save computer memory, as will be discussed later). On the other hand, we choose to add the observed AI data $z_{p,\ell}^o$ into $\mathbf{V}_\ell$, as we expect that $z_{p,\ell}^o$ contain certain information of the underlying RPM that led to the observed data. In addition, this also allows us to include hard data (if any) into the residual model, as will be explained soon.

The vectors $\mathbf{V}_k^{cp}$ in Eq. (17) represent a set of N_{cp} center points which are pre-chosen and fixed throughout the field case study. Note that $\mathbf{V}_k^{cp}$ are not required to be affiliated with any reservoir gridblock (or any physical location), therefore the subscript ℓ is not present therein. If there are hard data (e.g., core or well log data), then they can be used as some of the center points, in the form of, e.g., $[\Phi^{hard}, NTG^{hard}, S_w^{hard}, S_g^{hard}, P^{hard}, z_p^{hard}]^T$. In the case study, we do not use any hard data due to their availability (e.g., well log data contain part, but not all, of the required information). Instead, to generate $\mathbf{V}_k^{cp}$, we first specify the ranges of the first five inputs, Φ, NTG, S_w, S_g and P (in the unit of MPa), of the residual model, which are summarized in Table 2. Based on these ranges, we then adopt Latin hypercube sampling (McKay et al., 1979) to generate N_{cp} samples of $[\Phi_k^{cp}, NTG_k^{cp}, S_{w,k}^{cp}, S_{g,k}^{cp}, P_k^{cp}]^T$.

Given the above sample points, the next step is to specify the associated AI data $z_{p,k}^{o,cp}$. In general, we do not have precise information to find the proper values of $z_{p,k}^{o,cp}$ (perhaps except for those of hard data). As a result, some empirical approach will have to be adopted here. In this study, we carry out the allocations of $z_{p,k}^{o,cp}$ to $[\Phi_k^{cp}, NTG_k^{cp}, S_{w,k}^{cp}, S_{g,k}^{cp}, P_k^{cp}]^T$ in the following way: With the initial ensemble of reservoir models, we run reservoir simulations and collect the profiles of simulated pressure, water and gas saturations at all seismic survey time instances. From these, we compute mean fields of Φ, NTG, S_w, S_g and P that are distributed over active reservoir gridblocks.[5] Let

$$\mathbf{U}_\ell^{mean} \equiv \left[\Phi_\ell^{mean}, NTG_\ell^{mean}, S_{w,\ell}^{mean}, S_{g,\ell}^{mean}, P_\ell^{mean}\right]^T, \ell = 1, 2, \cdots, N_{agb}, \tag{22}$$

be a vector containing the values of the mean fields of Φ, NTG, S_w, S_g and P distributed on the ℓ -th (active) reservoir gridblock, then we treat $\{\mathbf{U}_\ell^{mean}\}_{\ell=1}^{N_{agb}}$ as if it were the "true" inputs of the RPM that led to the observed AI data $\{z_{p,\ell}^o\}_{\ell=1}^{N_{agb}}$. In other words, we allocate $z_{p,\ell}^o$ to $\mathbf{U}_\ell^{mean}$, and then create a set of mean-field vectors $\{\mathbf{V}_\ell^{mean}\}_{\ell=1}^{N_{agb}}$, with

$$\mathbf{V}_\ell^{mean} \equiv \left[\left(\mathbf{U}_\ell^{mean}\right)^T, z_{p,\ell}^o\right]^T. \tag{23}$$

Analogous to Eq. (22), let us define

$$\mathbf{U}_k^{cp} \equiv \left[\Phi_k^{cp}, NTG_k^{cp}, S_{w,k}^{cp}, S_{g,k}^{cp}, P_k^{cp}\right]^T, k = 1, 2, \cdots, N_{cp} \tag{24}$$

for the k-th center point $\mathbf{V}_k^{cp}$. To obtain $z_{p,k}^{o,cp}$ that is associated with $\mathbf{U}_k^{cp}$, we then search K_{nn} vectors from the set $\{\mathbf{U}_\ell^{mean}\}_{\ell=1}^{N_{agb}}$ that are the closest ones to $\mathbf{U}_k^{cp}$. Here let us denote the set of indices of the K_{nn} nearest neighbours of $\mathbf{U}_k^{cp}$ by

$$\mathscr{S}_k^{cp}(K_{nn}) \equiv \left\{\ell_s^k : \mathbf{U}_{\ell_s^k}^{mean} \text{ are the } K_{nn} \text{ nearest neighbours of } \mathbf{U}_k^{cp}\right\}, k = 1, \cdots, N_{cp},$$

then $z_{p,k}^{o,cp}$ is calculated as the average of the observed AI data $z_{p,\ell_s^k}^o$ that are associated with the vectors $\mathbf{U}_{\ell_s^k}^{mean}$ in $\mathscr{S}_k^{cp}(K_{nn})$, i.e.,

$$z_{p,k}^{o,cp} = \frac{1}{K_{nn}} \sum_{s=1}^{Knn} K_{nn} z_{p,\ell_s^k}^o, \ \ell_s^k \in \mathscr{S}_k^{cp}(K_{nn}). \tag{25}$$

In the current work, we set $K_{nn} = 10$. In general, $z_{p,k}^{o,cp}$ calculated through Eq. (25) may not be consistent with the corresponding $\mathbf{U}_k^{cp}$ in Eq. (24). The impact of this inconsistency, however, can be somewhat mitigated by adjusting the associated scale parameter in Eq. (19).

After choosing $\mathbf{V}_\ell$ and $\mathbf{V}_k^{cp}$, we are able to determine the size of the residual model, in terms of the number of elements of the parameter vector $\boldsymbol{\theta}$ in Eq. (18). Under this setting, the residual-model size is then equal to $(1 + m_v) \times N_{cp}$. In the current case study, $m_v = 6$ and $N_{cp} = 20000$, therefore the residual-model size is 140,000.

To integrate the residual model into ensemble-based 4D SHM, we need to create an initial ensemble of the residual-model parameters (cf Eq. (13) – (15)), which is denoted by $\boldsymbol{\Theta}^0 \equiv \{\boldsymbol{\theta}_j^0\}_{j=1}^{N_e}$ hereafter. The initialization is carried out as follows.

At the first step, we initialize the scale parameters $\beta_{k,s}$ associated with each element of the center point $\mathbf{V}_k^{cp}$ (cf. Eq. (19)). For Gaussian kernels, these scale parameters can be interpreted as the inverse of standard deviations (STD) of certain Gaussian distributions. As such, we use the STDs of the mean property fields (cf. Eq. (22)) or observed AI data to create the initial ensemble. Take the 6th element, $z_{p,k}^{o,cp}$, of $\mathbf{V}_k^{cp}$ as an example, we first compute the STD of observed AI data $z_{p,\ell}^o$ distributed on active gridblocks over all survey time instances, denoted by σ_{z_p} hereafter. Let $\beta_{k,6,j}^0$ be the scale parameter associated with $z_{p,k}^{o,cp}$ in the j-th ensemble member $\boldsymbol{\theta}_j^0$, then $\beta_{k,6,j}^0$ is given by

$$\beta_{k,6,j}^0 = e^\xi / \sigma_{z_p}, \quad \text{for} \ \ k = 1, 2, \cdots, N_{cp}; \ j = 1, 2, \cdots, N_e, \tag{26}$$

where ξ is a random number drawn from the normal distribution with zero mean and unit variance for $N_{cp} \times N_e$ times. Scale parameters $\beta_{k,s,j}^0$ ($s = 1, 2, \cdots, 5$) associated with the others element of the center point $\mathbf{V}_k^{cp}$ are computed in a way similar to that in Eq. (26), but with the STD values therein being replaced by those of the respective mean property fields in Eq. (22).

Given the values of $\beta_{k,s,j}^0$, at the second step we aim to compute the weight parameter, $c_{k,j}^0$, of the k-th kernel function in the j-th ensemble member $\boldsymbol{\theta}_j^0$. At this stage, we want the residual model in Eq. (17) to be able to partially explain the residuals (differences) between the observed AI data and the simulated AI data (through Eq. (16)) with respect to the initial ensemble of reservoir models. For this reason, we do not adopt the random sampling strategy to initialize $c_{k,j}^0$. Instead, we use some anchor points for the purpose of initialization. More specifically, given the residual fields (distributed on active reservoir gridblocks over all seismic surveys), we randomly select N_e points, denoted by $\tilde{r}_j^{o,ir}$ ($j = 1, 2, \cdots, N_e$), where the superscripts o and ir stand for "observed" and "initial residual", respectively. In light of the residual model in Eq. (17), for each ensemble-member index j, we want to find a weight vector

$$\mathbf{c}_j^0 \equiv \left[c_{1,j}^0, c_{2,j}^0, \cdots, c_{N_{cp},j}^0\right]^T, \tag{27}$$

which approximately solves the following equation

$$\tilde{r}_j^{o,ir} = \left(\mathbf{c}_j^0\right)^T \mathbf{K}_j^{vec}, \tag{28}$$

[5] Note that the mean fields of Φ and NTG are over the ensemble members since they are static variables, whereas those of S_w, S_g and P are over both the ensemble members and the seismic surveys.

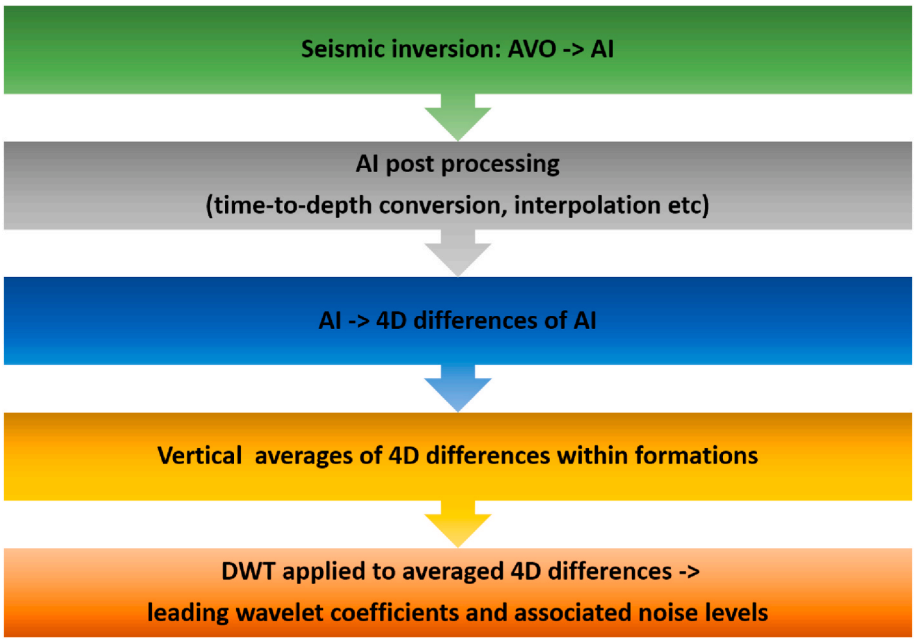


Fig. 2. Procedures involved in seismic data processing.

$$\mathbf{K}_j^{vec} \equiv \left[\exp\left\{-\left\langle\left(\boldsymbol{\beta}_{1,j}^0\right)^2, \left(\mathbf{V}\left(\tilde{r}_j^{o,ir}\right) - \mathbf{V}_1^{cp}\right)^2\right\rangle\right\}, \exp\left\{-\left\langle\left(\boldsymbol{\beta}_{2,j}^0\right)^2, \left(\mathbf{V}\left(\tilde{r}_j^{o,ir}\right) - \mathbf{V}_2^{cp}\right)^2\right\rangle\right\}, \cdots, \exp\left\{-\left\langle\left(\boldsymbol{\beta}_{N_{cp},j}^0\right)^2, \left(\mathbf{V}\left(\tilde{r}_j^{o,ir}\right) - \mathbf{V}_{N_{cp}}^{cp}\right)^2\right\rangle\right\}\right]^T \tag{29}$$

$$\boldsymbol{\beta}_{k,j}^0 \equiv \left[\beta_{k,1,j}^0, \beta_{k,2,j}^0, \cdots, \beta_{k,6,j}^0\right]^T, \tag{30}$$

where $\mathbf{V}(\tilde{r}_j^{o,ir})$ represent an input vector of the residual model Eq. (17). The static elements (Φ and NTG) of $\mathbf{V}(\tilde{r}_j^{o,ir})$ correspond to those in the j-th reservoir model of the initial ensemble, on the same reservoir gridblock as $\tilde{r}_j^{o,ir}$. Likewise, the dynamical variables (S_w, S_g and P) of $\mathbf{V}(\tilde{r}_j^{o,ir})$ correspond to the respective saturation or pressure profile generated by the j-th reservoir model of the initial ensemble, on the same reservoir gridblock and at the same seismic survey time instance as $\tilde{r}_j^{o,ir}$. The observed AI data (z_p^o) contained in $\mathbf{V}(\tilde{r}_j^{o,ir})$ are identified in a similar way.

Eq. (28) is approximately solved in terms of

$$\mathbf{c}_j^0 = \tilde{r}_j^{o,ir} \times \mathbf{K}_j^{vec} \Big/ \left(b + \left(\mathbf{K}_j^{vec}\right)^T \mathbf{K}_j^{vec}\right), \tag{31}$$

where b is a relatively small constant introduced to avoid the potential issue of division-by-zero, and is set to 0.01 throughout this work.

At the last step of computing the initial ensemble $\boldsymbol{\Theta}^0$ of residual-model parameters, we take into account the fact that the scale parameters $\boldsymbol{\beta}_{k,j}^0$ from the first step is obtained through random sampling. As a result, we choose to update the ensemble $\boldsymbol{\Theta}^0$ so that the residual model Eq. (17) can better match the field of initial residuals distributed over active reservoir gridblocks over all seismic surveys, given the vector fields $\mathbf{V}$ with respect to the initial ensemble of reservoir models. Essentially, this boils down to a parameter estimation problem (similar to history matching), and can thus be solved by ensemble-based data assimilation. For this reason, we choose to use the IES introduced previously to update $\boldsymbol{\Theta}^0$ (without localization, and with the observation error covariance matrix $\mathbf{C}_d$ being an identity matrix). To avoid overfitting, only one iteration is conducted. After the update, the obtained ensemble $\boldsymbol{\Theta}^0$ is deployed for 4D SHM later on.

4.2. Data processing for history matching

The 4D seismic datasets used in the current study are the same as those in Lorentzen et al. (2020). Here we only outline the essential procedures involved in seismic data processing (see Fig. 2). For more information, readers are referred to Lorentzen et al. (2020).

The Norne open dataset (http://www.ipt.ntnu.no/~norne) contains amplitude-versus-offset (AVO) data from four seismic surveys in 2001 (base), 2003 (monitor #1), 2004 (monitor #2) and 2006 (monitor #3), respectively. In the field case study, however, we converted the AVA data into AI data at each survey time, through a procedure described in Lorentzen et al. (2020). The actual data used in SHM are the 4D differences of AI data between monitor and base surveys. As a result, there are three AI-difference datasets.

In the Norne field model, the thickness of each individual layer is around 5 m on average, and is below the typical vertical resolution of seismic data, which is in the order of 10–20 m (Souza et al., 2019). To limit the sub-scale variations of the AI data that are not captured with the relatively low seismic vertical resolution, the AI-difference data are averaged within each reservoir formation (Garn, Ile, Tilje, and Tofte). This formation-based averaging strategy does not necessarily lead to any theoretical advantages, but is chosen largely from an engineering perspective to mitigate the impact of insufficient vertical resolution of seismic data (Lorentzen et al., 2020).

DWT are then applied to these averaged AI-difference data, while a set of leading wavelet coefficients is selected for sparse data representation in the wavelet domain, following the work of Luo et al. (2017,

2018a); Lorentzen et al. (2019). The number of original AI datasets (from four seismic surveys) on active reservoir gridblocks is $4 \times 44927 = 179708$, after the processing procedures (4D difference, formation average and DWT), the total number of leading wavelet coefficient is reduced to 24,232.

Moreover, as a by-product of the sparse representation procedure, one can also estimate the noise levels associated with the leading wavelet coefficients through a wavelet-based denoising method (Luo and Bhakta, 2017). The information of noise levels will be subsequently exploited by the history matching algorithm.

Since history matching typically exploits the differences between observed and simulated data for reservoir model updates, when a sparse data representation procedure is applied to the observed data, the same procedure also needs to be applied to the simulated data in order for the comparison between observed and simulated data to make sense. As a result, DWT are also applied to the simulated averaged AI-difference data generated by the RPM and the subsequent procedures of 4D difference and formation average. For the purpose of comparison, we select wavelet coefficients of the simulated averaged AI-difference data that are located in the same positions as the leading wavelet coefficients of observed averaged AI-difference data.

In relation to the focus of the current study, when we use the RPM to generate simulated AI data, we consider two situations in which the previously described residual model is or is not added to the original RPM. These two scenarios correspond to the choices of whether to account for model errors from RPM or not.

The Norne open dataset also contains a set of production data from 1997 to 2006. For our purpose here, production data are not used for history matching. Instead, they are adopted to cross-validate the qualities of estimated reservoir models from SHM, in terms of averaged data mismatch (to be defined later) between the observed production data and the forecast production data from the estimated reservoir models. Following the previous work (Chen and Oliver, 2014; Lorentzen et al., 2020; Luo et al., 2019), the production data used for cross-validation include oil, water and gas production rates from production wells, denoted by WOPR, WWPR and WGPR, respectively. The total number of these production data is 5038. In order to compute data mismatch values, the noise levels associated with WOPR, WWPR and WGPR are set to 100 Sm^3/d, 200 Sm^3/d and 20,000 Sm^3/d, respectively.

4.3. Model updates through the IES

RLM-MAC proposed by Luo et al. (2015) is adopted as the history matching algorithm in the field case study. To improve numerical stability, TSVD is conducted in the implementation of the IES (see, e.g., Iterative-Ensemble-Smoother-master, 2019), in such a way that the sum of the retained leading singular values is no less than 95% of the sum of all singular values.

Given the relatively large amount of observations (24,232 data points) in SHM, we choose to project all observations onto a sub-space spanned by the leading singular vectors obtained from TSVD, following Lorentzen et al. (2020).

Due to the projection, the "effective" observations do not possess clear physical locations. As such, we adopt correlation-based adaptive localization in history matching, using the hard thresholding strategy originally proposed in Luo et al. (2018b).

The threshold values in correlation-based adaptive localization are computed in the following way: For reservoir model variables, such as PERMX, PORO and NTG in Table 1, which have spatial distributions (hence unambiguous physical locations), we first compute the correlation fields between these model variables and the projected data. By treating these correlation fields as images, we apply a wavelet-based image denoising method (Luo and Bhakta, 2017) to calculate the noise levels of these correlation fields, and then compute estimated threshold values through the universal rule (Donoho and Johnstone, 1994). To avoid possible underestimations of the threshold values, we further specify a lower bound for them, in the form of $3/\sqrt{N_e}$ (based on an asymptomatic result discussed in Luo and Bhakta, 2020), whereas the final threshold values are the maximum ones between the estimated threshold values and the lower bound. On the other hand, for reservoir model variables without any spatial distributions and clear physical locations (e.g., Krg and Krw as in Table 1), we simply use $1/\sqrt{N_e}$ as the threshold value, following the choice in Lorentzen et al. (2020, 2019); Luo et al. (2019).[6]

Table 3
Data mismatch of simulated seismic data with respect to the final ensembles obtained in four sets of history-matching studies, in terms of mean ± STD. For the cases with MEC, two sets of data mismatch values are presented, in the form of A/B, which correspond to the cases that residuals are included (A)/not included (B) in the course of calculating data mismatch, respectively.

	With MEC	No MEC
$\eta^0 = 1$	$(2.8552 \pm 2.1247) \times 10^7$ / $(1.2457 \pm 0.5492) \times 10^7$	$(0.7052 \pm 0.2445) \times 10^7$
$\eta^0 = 10$	$(2.0788 \pm 1.2039) \times 10^7$ / $(1.2773 \pm 0.5655) \times 10^7$	$(1.5284 \pm 0.7247) \times 10^7$

Table 4
Data mismatch of forecast production data with respect to the final ensembles obtained in four sets of history-matching studies, in terms of mean ± STD.

	With MEC	No MEC
$\eta^0 = 1$	$(4.9841 \pm 0.7099) \times 10^5$	$(5.8287 \pm 0.6741) \times 10^5$
$\eta^0 = 10$	$(5.2984 \pm 0.9342) \times 10^5$	$(5.3273 \pm 0.9182) \times 10^5$

When the kernel-based residual model is introduced to account for potential model errors in the RPM, we have an additional set of residual-model parameters (140,000 per realization), which are associated with individual kernel functions and do not possess physical locations. For the purpose of conducting localization for these parameters, we adopt correlation-based adaptive localization based on the hard thresholding strategy again. In this case, we choose to use $3/\sqrt{N_e}$ as the threshold value.

The configuration of the IES in the current study largely follows that in Luo et al. (2019). Specifically, based on the basic update formulae Eqs. (4) and (5), one needs to specify how the scalar γ^i therein evolves along the iteration process. Here we adopts the following rule:

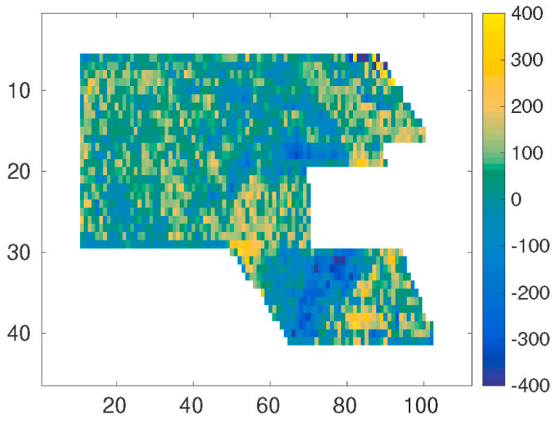
$$\gamma^i = \eta^i \, \mathrm{trace}\left(\tilde{\mathbf{S}}_d^i \left(\tilde{\mathbf{S}}_d^i\right)^T\right) / \mathrm{trace}(\mathbf{I}), \tag{32}$$

where trace represents the operator that gets the trace of a relevant matrix, $\tilde{\mathbf{S}}_d^i$ is a square root matrix obtained by applying the TSVD to $\mathbf{S}_d^i$ in Eq. (5), and $\mathbf{I}$ is the identity matrix with a suitable dimension. Readers are referred to Luo et al. (2015, 2019) for the rationale behind the parameter rule in Eq. (32). The scalar factor η^i itself also varies with the iteration, in the form of:

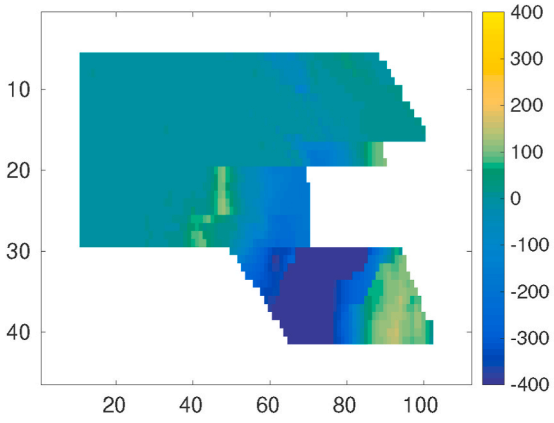
$$\eta^{i+1} = \beta \, \eta^i, \quad \text{with } 0 < \beta < 1 \, . \tag{33}$$

Given a limited number of iteration steps, the choice of the initial value η^0 has a substantial impact on the history matching performance when model errors are not accounted for, as will be demonstrated later.

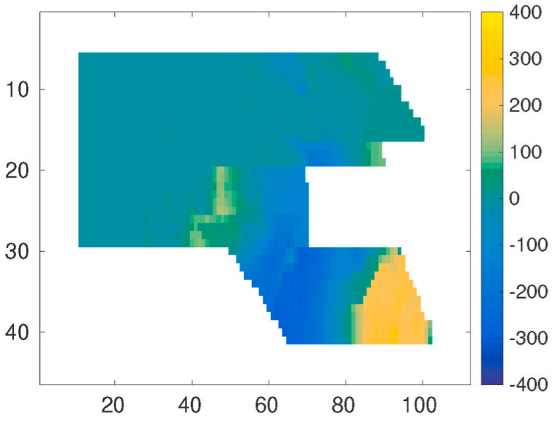
[6] This choice was made in light of the observation that sample correlations between this kind of reservoir model variables and projected data tend to be weaker, in comparison to sample correlations with respect to reservoir model variables with physical locations. We thus decided to lower the threshold values to allow the latter type of reservoir model variables to have more impact on history matching in our previous work.



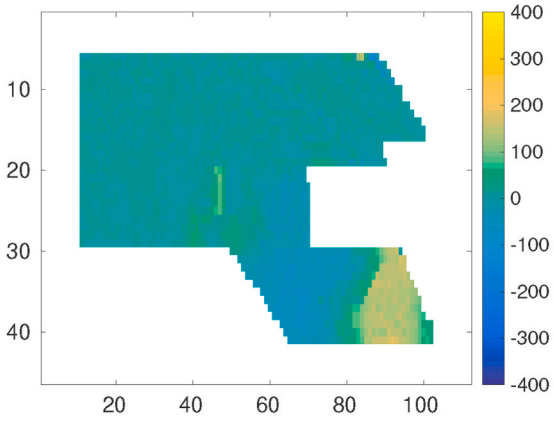
(a) Inverted AI difference (2004-2001,Garn)



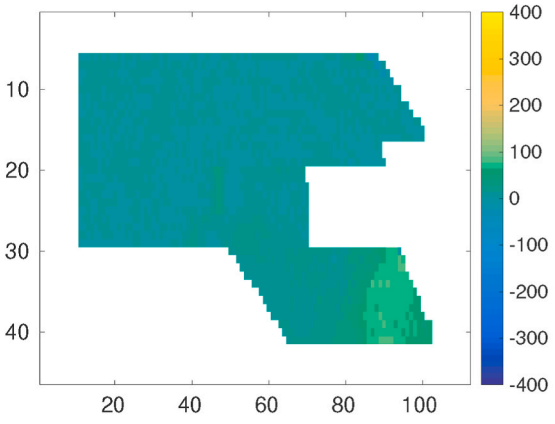
(b) Corresponding simulated AI diffrence (initial)



(c) Corresponding simulated AI diffrence (final)



(d) Corresponding residual diffrence (initial)



(e) Corresponding residual diffrence (final)

Fig. 3. Inverted or simulated seismic data, in the unit of (m/s) (g/cc). (a) Seismic data inverted from the field AVO datasets. Here the seismic data map is the vertical average (within the Garn formation) of the time-difference (2004–2001) of acoustic impedance (AI) data; (b) Map of corresponding seismic data (averaged over ensemble members), generated by the original rock physics model (RPM) using the initial ensemble of reservoir models; (c) As in (b), but with respect to the final ensemble of reservoir models in the case of MEC; (d) Map of corresponding residual data (averaged over ensemble members), generated by the residual model using the initial ensemble of reservoir models and kernel parameters; (e) As in (d), but with respect to the final ensemble of reservoir models and kernel parameters in the case of MEC.

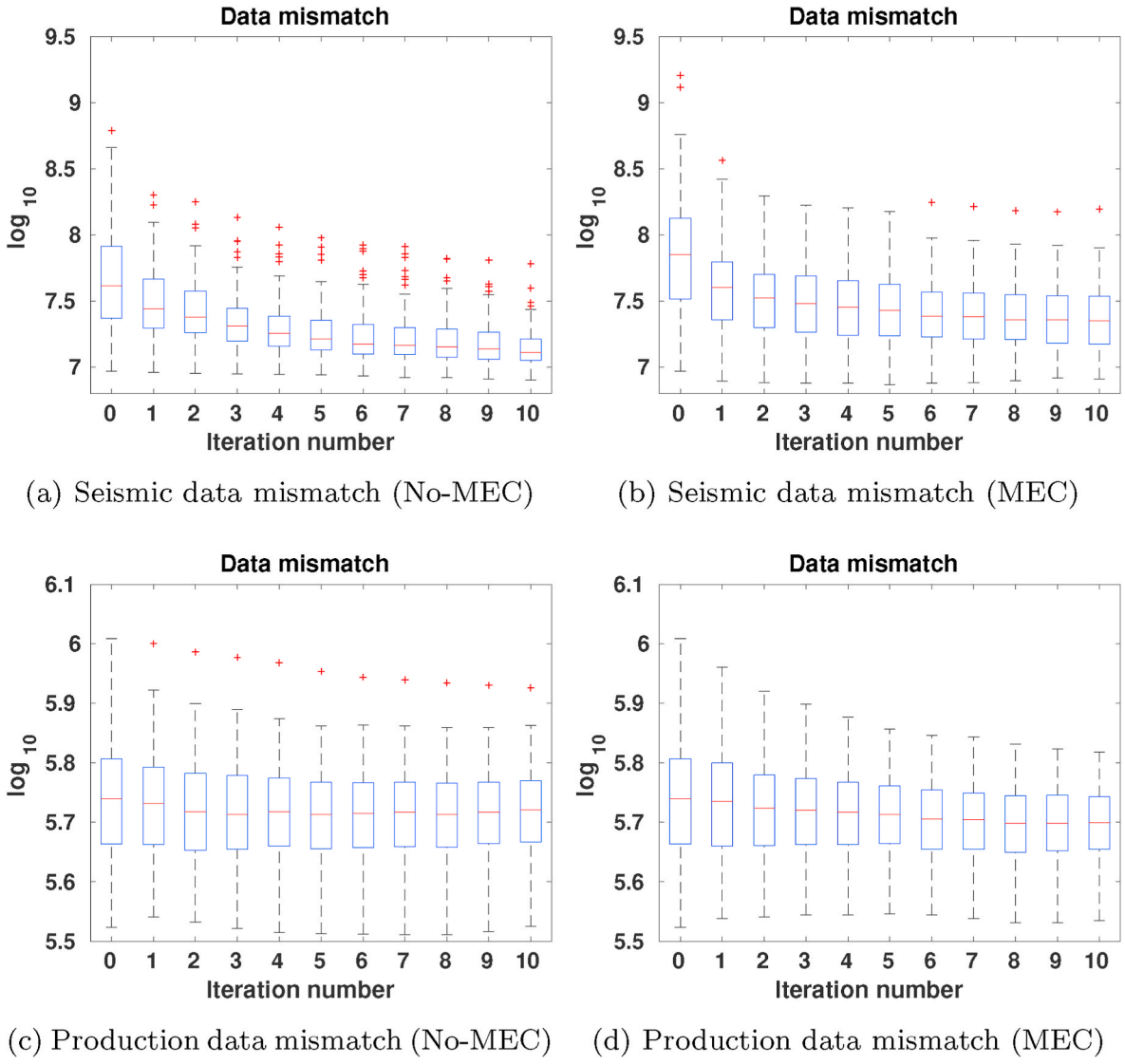


Fig. 4. Box plots of seismic (top) and production (bottom) data mismatch (in the scale of $\log_{10}$) with respect to the ensembles of reservoir models at different iteration steps. In the MEC case, the data mismatch values are calculated by combining the simulated seismic data from the original RPM and the outputs of the residual model.

The iteration process of IES is guided by the average data mismatch (ADM) with respect to the ensemble of reservoir models. Given an ensemble $\mathscr{M}^{i+1} = \{\mathbf{m}_j^{i+1}\}_{j=1}^{N_e}$ of reservoir models, the corresponding ADM $\Xi(\mathscr{M}^{i+1})$ is calculated as

$$\Xi(\mathscr{M}^{i+1}) \equiv \frac{1}{N_e}\sum_{j=1}^{N_e}\left(\mathbf{d}^o - \mathbf{g}^{eff}\left(\mathbf{m}_j^{i+1}\right)\right)^T \mathbf{C}_d^{-1}\left(\mathbf{d}^o - \mathbf{g}^{eff}\left(\mathbf{m}_j^{i+1}\right)\right) \tag{34}$$

where $\mathbf{g}^{eff}$ is the effective forward simulator, which is equal to the original physics-based forward simulator $\mathbf{g}$ if model errors are not taken into account, or the modified forward simulator $\widehat{\mathbf{g}} = \mathbf{g} + \widehat{\mathbf{r}}$ if a residual model $\widehat{\mathbf{r}}$ is introduced to account for potential model errors in $\mathbf{g}$ (cf. Eq. (15)).

Through model updates, if a new ensemble $\mathscr{M}^{i+1}$ leads to lower ADM than the current ensemble $\mathscr{M}^{i}$, then we accept $\mathscr{M}^{i+1}$ and use it as the background ensemble at the next iteration step to search another new ensemble that has a potential to further reduce the ADM. Otherwise, if $\Xi(\mathscr{M}^{i+1}) \geq \Xi(\mathscr{M}^{i})$, then an inner loop is adopted to try a reduced step size by replacing η^i by $\alpha\,\eta^i$ ($\alpha > 1$) until lower ADM is found, or the maximum number of the inner loop iteration is reached, see, for instance, the demonstration in (Iterative-Ensemble-Smoother-master, 2019). In this study, we set $\alpha = 2, \beta = 0.9$, and the maximum number of the inner loop iteration to be 5.

To run the IES, we also need to specify certain stopping criteria. In this work, we adopt the following three stopping criteria as in Luo et al. (2019): (1) the maximum number of (outer) iteration steps reaches 10; (2) the relative change of the average data mismatch in two consecutive outer iteration steps is less than 1%; and (3) the average data mismatch is less than $4p$, with p being the number of observations.

5. Numerical results

This section aims to investigate the impact of model errors on the performance of SHM. To this end, we conduct four sets of history-matching studies (cf. Tables 3 and 4), among which the differences lie in (a) whether a MEC mechanism (through the kernel based residual model) is introduced or not; and (b) the scalar factor η^0 (cf. Eq. (32) – (33)) is set to 1 or 10. Apart from these mentioned differences, all other settings are identical in the studiess.

Based on Eqs. (4), (5) and (32), it is clear that a larger value for η^0 implies a larger value for γ^0, hence a smaller step size for model updates.

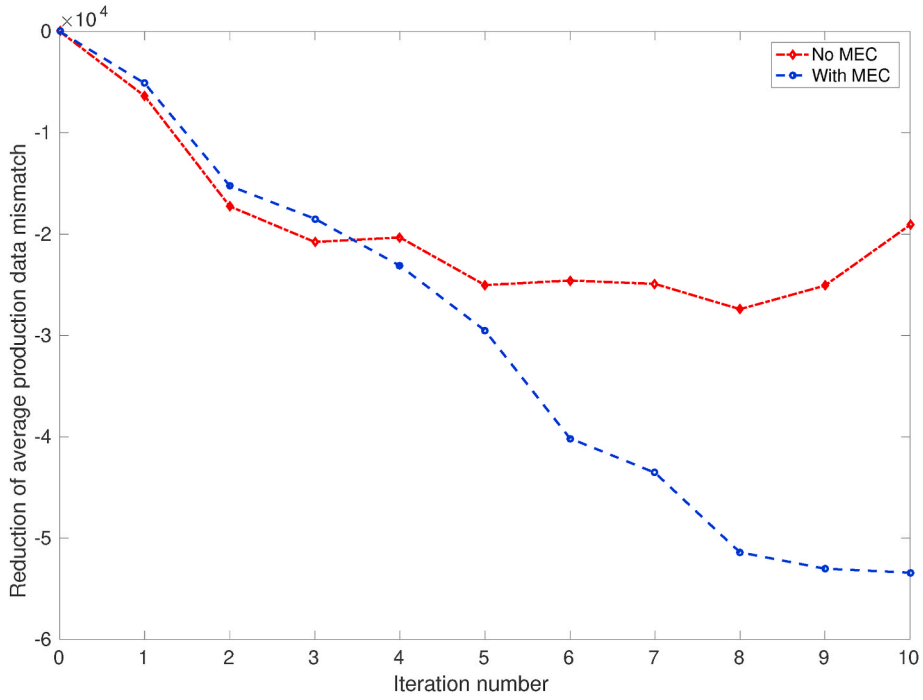


Fig. 5. Changes of average data mismatch (ADM) of production data at different iteration steps, with respect to ADM of production data from the initial ensemble, in cases of MEC (blue) and No-MEC (red). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

In a practical history matching problem, given the limited computational resources, the IES is often applied with a relatively small number of iteration steps (e.g., 10 steps or even less). As a result, the choice of η^0 has a substantial impact on the history matching performance, which is measured in terms of ADM of forecast against observed data in the current work.

Table 3 reports the data mismatch values, in terms of mean $\pm$ STD, with respect to seismic data from the final ensembles obtained in four sets of history-matching studies. For the cases with MEC, we show data mismatch values in two scenarios, which correspond to the cases that residuals are included (Scenario A)/not included (Scenario B) in the course of calculating data mismatch, respectively. Note that in history matching, it is the ADM values of Scenario A that are used to guide the iteration process, whereas the ADM values of Scenario B are calculated after history matching.

From Table 3, it is clear that, with residuals being included (Scenario A), the ADM values of seismic data in the cases with MEC are higher than those in the cases without MEC. After excluding residuals from the calculations (Scenario B), the ADM values in the cases with MEC become lower than those in the case without MEC at $\eta^0 = 10$, but still higher than those in the case without MEC at $\eta^0 = 1$. Meanwhile, using different η^0 values in the cases with MEC tends to result in less variations of ADM values, in contrast to the cases without MEC. These results suggest that, in the current field cast study, the residual model has a substantial impact on the performance of SHM; and in the cases with MEC, the history matching performance appears less sensitive to the choice of η^0.

For illustration, Fig. 3 shows a set of inverted or simulated seismic data. Fig. 3(a) plots the seismic data map with respect to the time-difference (2004–2001) of inverted AI data, averaged vertically within the Garn formation (also see our discussion on data processing in the preceding section). Fig. 3(b –c) present the corresponding simulated seismic data maps (excluding residuals) with respect to the initial and final ensembles of reservoir models, respectively, in the case of MEC, while Fig. 3(d –e) depict associated residual maps, with respect to the initial and final ensembles of reservoir models and kernel parameters, respectively. As a common procedure to generate the maps in Fig. 3(b – e), given an ensemble of 4D datasets (simulated AI or residual data), we first compute the average of the 4D datasets over ensemble members (called ensemble-mean hereafter), then compute the 4D differences of the ensemble-mean datasets. After that, for a given time-difference dataset, we further compute the averages over the layers within individual formations of the reservoir model, and plot the average values onto the active reservoir gridblocks of respective layers.

Comparing Fig. 3(a – c), we see that the simulated AI data partially capture the spatial patterns of the inverted AI data, e.g., in the G segment (the lower right corner). After SHM, data match in the G segment seems further improved, e.g., the dark-blue region with relatively lower AI difference values in Fig. 3(b) becomes light-blue instead in Fig. 3(c), closer to the values in the same region of the inverted AI data map. On the other hand, for the residual maps, it seems that they (partially) help compensate for the differences between the inverted and simulated AI maps. For instance, before SHM, in the G segment of Fig. 3(b), the values of simulated AI differences in the light-yellow region tend to be lower than those of the inverted AI differences in the same region of Fig. 3(a). Accordingly, the corresponding residuals in the same region of Fig. 3(d) have relatively large values. After SHM, since the simulated AI data in the same region of Fig. 3(c) tend to match the inverted AI data better, the values of corresponding residuals in the same region of Fig. 3(e) decrease accordingly.

For further performance validation, Table 4 reports the corresponding data mismatch values with respect to production data. As it is clear therein, for a given η^0, when MEC is introduced to SHM, the ADM of forecast production data becomes lower than the corresponding case without MEC, indicating that the kernel-based residual model indeed tends to help improve the quality of reservoir models estimated through SHM. On the other hand, in terms of STD of data mismatch, those from the case with MEC tends to be larger than those from the case without MEC.

In contrast, when the choice of adopting MEC or not is made, the impact of η^0 on the history matching performance (in terms of ADM of production data) is case dependent. When model errors are not considered, then one obtains lower ADM for forecast production data if a smaller initial step size (at $\eta^0 = 10$) is used for model updates. This

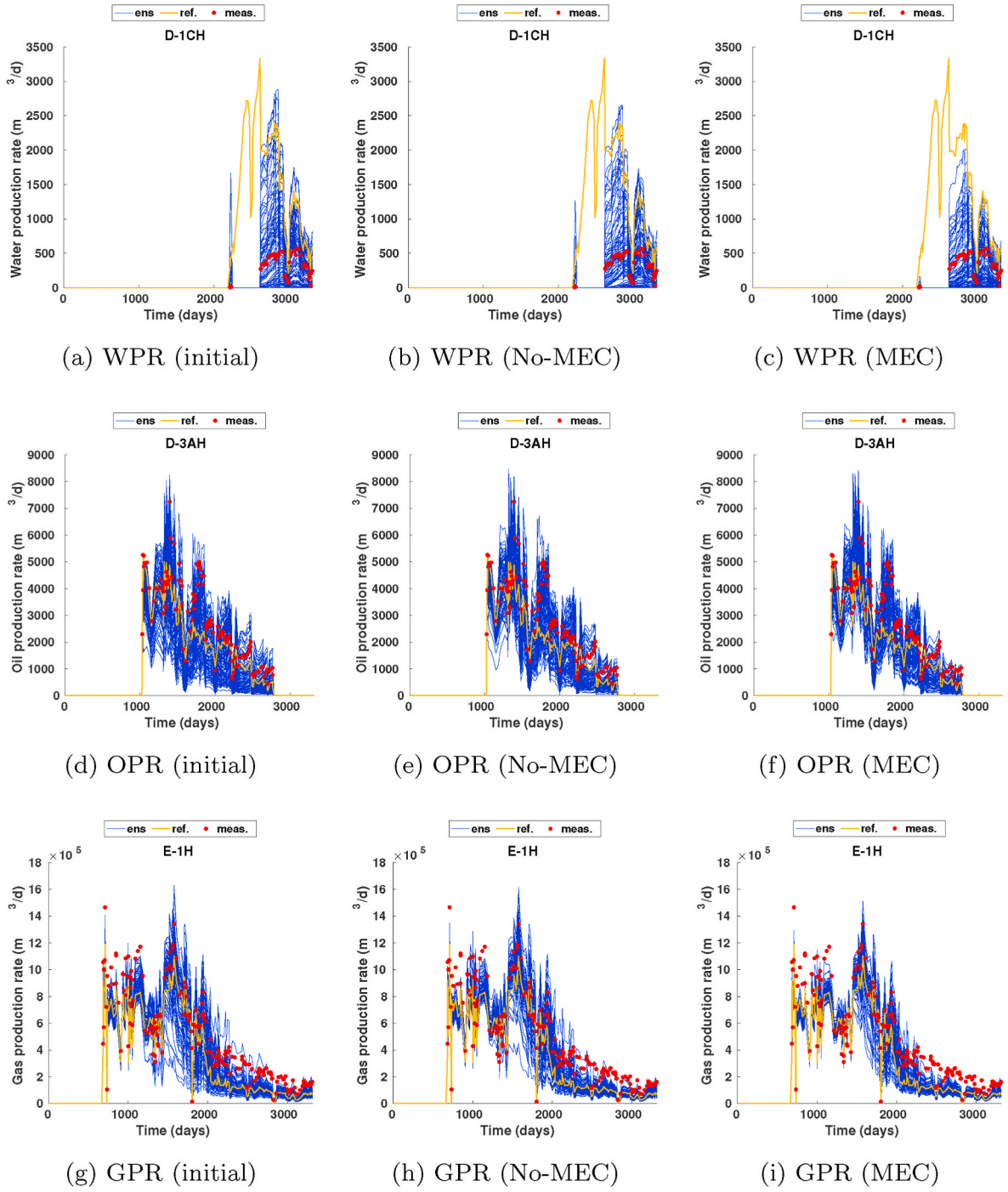


Fig. 6. Production data profiles in some of the wells. In each sub-figure, red dots represent respective observed production data; orange curves stand for the forecast production data from a manually history-matched reservoir model provided in the Norne open dataset; and blue curves are the forecasts from the ensembles of reservoir models, before or after seismic history matching. The top row shows the water production rates (WPR) from well D-1CH; the middle row displays the oil production rates (OPR) from well D-3AH, and the bottom one depicts the gas production rates (GPR) from well E−1H. On the other hand, the first column presents results with respect to the initial ensemble; the second column contains results with respect to the ensemble of reservoir models obtained at the 8th iteration step, which are the best in the case of No-MEC according to Fig. 5; and the last column shows results with respect to the final ensemble of reservoir models in the case of MEC. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

suggests that, in this case, more model changes (due to a smaller η^0) does not necessarily lead to improved quality of reservoir models, which is likely due to the possibility of over-fitting seismic data in order to compensate for the unaccounted model errors, as will be further illustrated later. Meanwhile, when model errors are taken into account, using a smaller η^0 value (hence more model changes) does improve the quality of reservoir models, in terms of ADM of forecast production data.

By borrowing the concepts of "under-fitting" and "over-fitting" with respect to the training data in supervised learning problems, we provide a possible explanation of the results in Tables 3 and 4 In our problem, seismic data can be considered as an analogy to the "training dataset", whereas production data an analogy to the "test dataset". For the case without MEC, a relatively small η^0 value (e.g., $\eta^0 = 1$) leads to more substantial changes of reservoir models, and this results in an overfitting

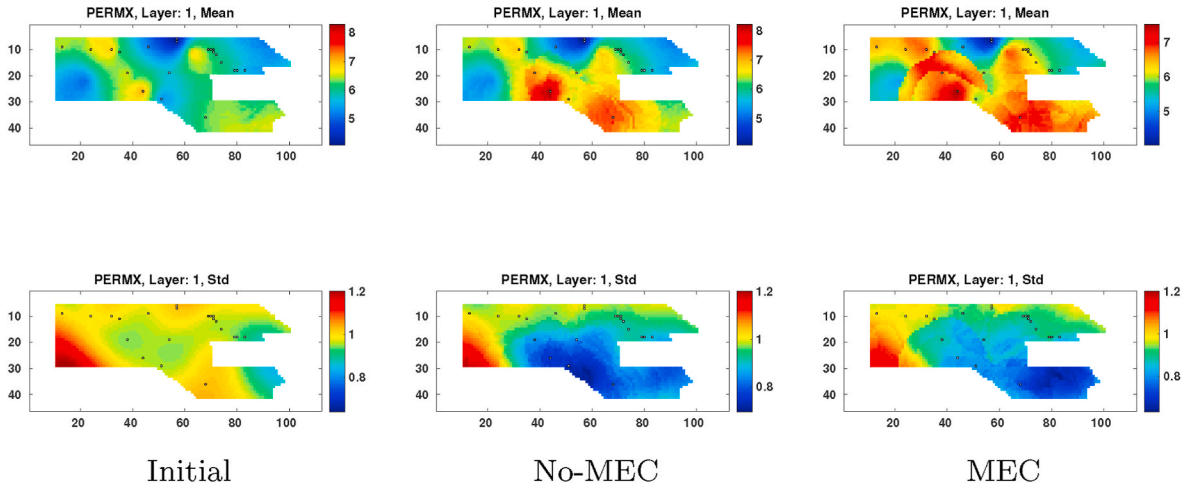


Fig. 7. Mean (top) and STD (bottom) maps of x-directional permeability (PERMX) on Layer 1 of the reservoir model, with respect to the initial ensemble (first column), the ensemble at the 8th iteration step in the case of No-MEC (second column), and the final ensemble in the case of MEC (third column). The small circles on the maps indicate the positions of field wells.

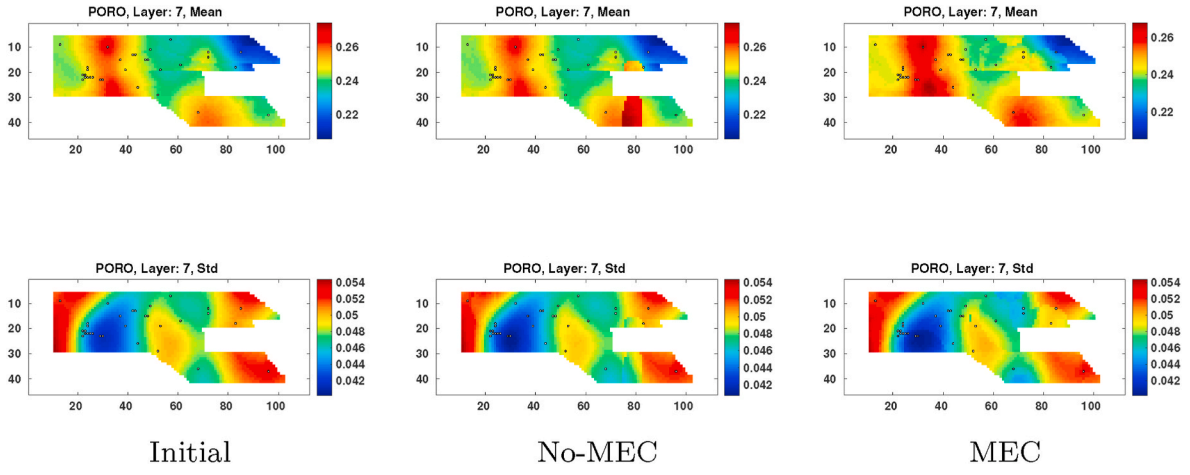


Fig. 8. As in Fig. 7, but for the mean (top) and STD (bottom) maps of porosity (PORO) on Layer 7 of the reservoir model.

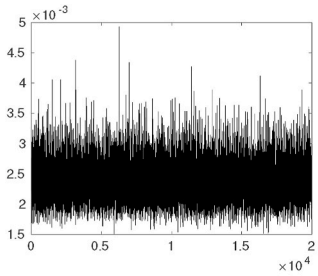
of the training data. As a consequence, the test performance is worsened when the history matching performance is evaluated using the ADM of production data. On the other hand, a larger η^0 value (e.g., $\eta^0 = 10$) produces less changes of reservoir models, which then causes an underfitting of the seismic data, and does not achieve descent performance neither at the test stage. In contrast, for the case with MEC, more proper training is achieved in terms of matching the seismic data for different choices of η^0 values, and this leads to better performance at the test stage.

In the sequel, we present more numerical results. For succinctness, we only compare the results of the cases (with and without MEC, respectively) that achieve the best performance in terms of ADM of forecast production data, i.e., the results from the case without MEC at $\eta^0 = 10$, and those from the case with MEC at $\eta^0 = 1$. Without causing confusion, we simply call them the No-MEC case and the MEC case, respectively.

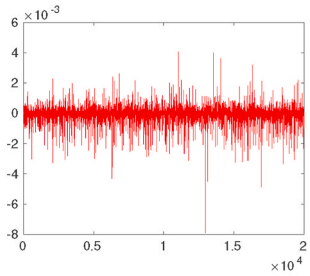
Fig. 4 shows box plots of seismic (top) and production (bottom) data mismatch (in the scale of $\log_{10}$) with respect to the ensembles of reservoir models at different iteration steps, in the cases of MEC and No-MEC. In terms of seismic data mismatch obtained in history matching, the introduction of the residual model appears to result in higher seismic data mismatch than that in the No-MEC case. On the other hand, in terms of production data mismatch for performance test, it seems beneficial to adopt the residual model, as it tends to lead to both lower ADM and spread in the MEC case (also see the results in Table 4).

For better visualization, Fig. 5 reports the changes of ADM of production data, with respect to ADM of production data resulting from the initial ensemble (before SHM), in both cases of MEC (in blue) and No-MEC (in red). As one can see there, in the case of No-MEC, the ADM tends to decrease until it researches the 8th iteration step. After that, the ADM of production data increases instead, indicating a potential over-fitting of seismic data during SHM. In contrast, in the case of MEC, in the presence of the residual model, the ADM of production data monotonically decreases, meaning that the potential issue of over-fitting seismic data is mitigated.

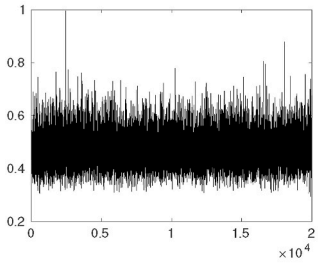
In addition, Fig. 6 shows some of the production data profiles, namely, water production rates (WPR) from well D-1CH, oil production rates (OPR) from well D-3AH and gas production rates (GPR) from well E−1H, with respect to the ensembles of reservoir models before or after SHM. In each sub-figure, red dots stand for observed production data; yellow curves for forecast production data from a manually history-matched reservoir model provided in the Norne open dataset; and blue ones for forecasts generated by respective ensembles of reservoir models. The first column of Fig. 6 corresponds to the results of the initial ensemble of reservoir models; the second column to the forecasts from the ensemble of reservoir models at the 8th iteration step in the case of



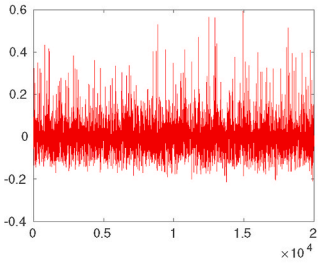
(a) Mean weight parameters (initial)



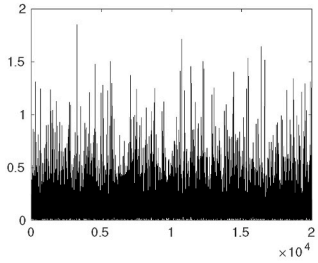
(b) Mean weight parameters (diff)



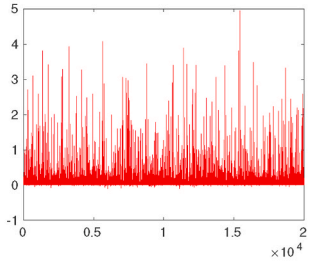
(c) Mean scale parameters (PORO, initial)



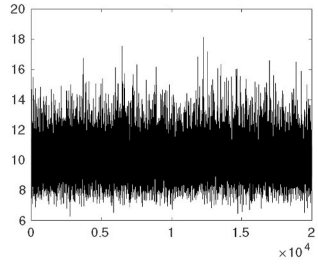
(d) Mean scale parameters (PORO, diff)



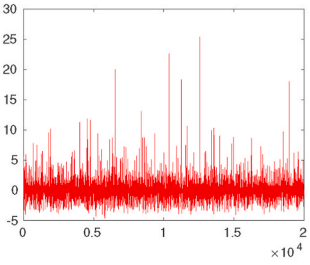
(e) Mean scale parameters (Swat, initial)



(f) Mean scale parameters (Swat, diff)

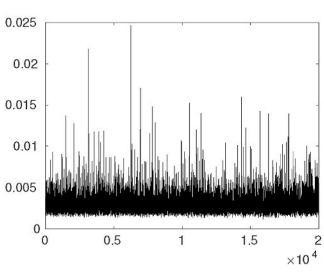


(g) Mean scale parameters (Seis, initial)

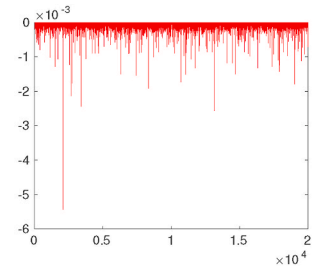


(h) Mean scale parameters (Seis, diff)

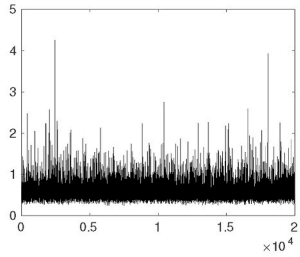
Fig. 9. Left column: Mean kernel parameters with respect to the initial ensemble in the case of MEC. These include the weights c (first row) and scale parameters β associated with porosity (PORO, second row), water saturation (Swat, third row) and seismic data (Seis, fourth row) of individual center points in the residual model, cf. Eq. (17) – (21) of the current work. Right column: for ease of visualization, we show the differences (diff) between the mean kernel parameters of the initial and final ensembles, in the form of mean parameters of the final ensemble minus those of the initial ensemble. Note that for the scale parameters, only their magnitudes matter, so we actually calculate the means of their absolute values, before computing the differences.



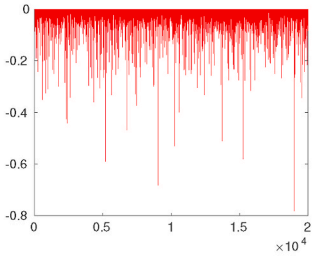
(a) STD of weight parameters (initial)



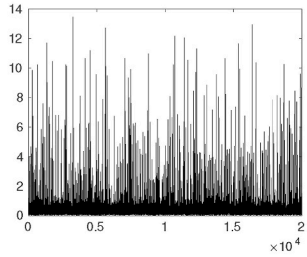
(b) STD of weight parameters (diff)



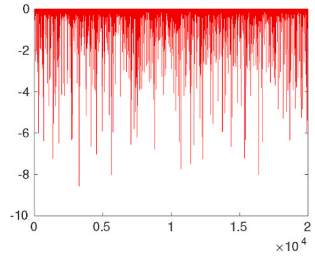
(c) STD of scale parameters (PORO, initial)



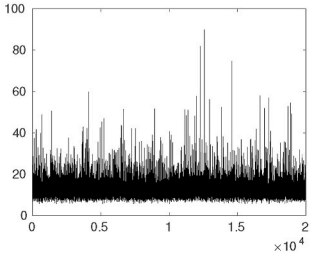
(d) STD of scale parameters (PORO, diff)



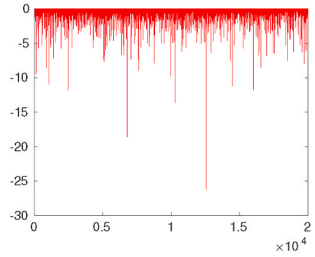
(e) STD of scale parameters (Swat, initial)



(f) STD of scale parameters (Swat, diff)



(g) STD of scale parameters (Seis, initial)



(h) STD of scale parameters (Seis, diff)

Fig. 10. As in Fig. 9, but for STDs of kernel parameters.

No-MEC, which are the best possible results according to Fig. 5; and finally, the third column to the results with respect to the final ensemble of reservoir models in the case of MEC.

Comparing the results in the first two columns of Fig. 6, we see that after SHM, the production data profiles in the case of No-MEC tend to match the observed data better than those from the initial ensemble. However, since production data are only reserved for performance test, rather than used in history matching, such improvements do not appear to be significant, although noticeable. In contrast, in the case of MEC, the

Table 5
Computational costs in four sets of history-matching studies.

	With MEC	No MEC
$\eta^0 = 1$	$\sim$ 5 days	$\sim$ 3 days
$\eta^0 = 10$	$\sim$ 5 days	$\sim$ 3 days

improvements appear to be more substantial. For instance, if one looks at the forecasts of WPR from well D-1CH (the top row), around day 2200, the observed data indicates that there should be no water breakthrough. The forecasts from the initial ensemble and the ensemble in the case of No-MEC both state that there are relatively high WPR values (cf the blue spikes in Fig. 6(a and b)). On the other hand, the ensemble in the case of MEC predicts much lower WPR values, which are much closer to the observed data. In comparison to the case of No-MEC, similar data-match improvements in the case of MEC are also spotted in other sub-figures, although tending to be less substantial in general.

Fig. 7 reports the mean (top row) and STD (bottom row) maps of x-directional permeability (PERMX) values distributed on the active gridblocks of Layer 1 of the reservoir model, with regard to the initial ensemble (first column), the ensemble at the 8th iteration step in the case of No-MEC (second column), and the final ensemble in the case of MEC (third column), respectively. Comparing to the maps of the initial ensemble, there are significant differences in those with respect to the cases of MEC and No-MEC. The maps in the cases of MEC and No-MEC exhibit similarities in certain regions, but also with substantial differences in others. For instance, for the mean maps, we see some clear distinctions in the G segment (lower right corner of the figures). In some of the maps, e.g., the mean PERMX map with respect to the case of MEC (upper right plot of Fig. 7), certain artificial effects (discontinuity) are present, which is largely due to the hard thresholding strategy adopted in correlation-based adaptive localization.

Similar to Figs. 7 and 8 shows mean and STD maps of porosity (PORO) values distributed on the active gridblocks of Layer 7 of the reservoir model. Different from the situation in Fig. 7, the maps in Fig. 8 appear more similar to each other. One possible implication of this observation is that, in this particular case study, the residual model has more impact on certain reservoir model variables (e.g., permeability) than the others (e.g., porosity).

In addition, Fig. 9 reports some of the mean kernel parameters with respect to the initial and final ensembles in the case of MEC. The left column shows the means (over members of the initial ensemble) of the weight parameters c (first row), the scale parameters $\boldsymbol{\beta}$ with respect to porosity (second row), water saturation (third row) and seismic data (fourth row), which are associated with 20,000 center points in the residual model, as described in the previous section (cf. Eq. (17) – (21) therein). The right column provides respective information corresponding to the final ensemble. For ease of visualization, however, we choose to present the differences between the mean kernel parameters of the initial and final ensembles, in the form of mean parameters of the final ensemble minus those of the initial one. We note that, for the scale parameters $\boldsymbol{\beta}$, it is their magnitudes that really matter in the implementation of the residual model (cf. Eq. (17) – (19)). Therefore, we actually calculate the means of their absolute values before computing the differences.

From the left column of Fig. 9, we see that the weight parameters c have relatively low magnitudes. The scale parameters for seismic data tend to have the largest magnitudes, and the magnitudes of scale parameters for porosity and water saturation are in the middle. In the residual model (cf. Eq. (17) – (21)), however, given the same amount of change of an input variable, a larger associated scale parameter implies less influence on the output. As a result, one can tell that initially the residual model is more sensitive to porosity and water saturation than to seismic data. From the second column, we see that, after SHM, the magnitudes of changes of the weight parameters are smaller than those of the other three presented scale parameters. Overall, the scale parameters for seismic data still tend to have the largest magnitudes. On the other hand, in the final ensemble, the scale parameters for water saturation become larger than those in the initial ensemble in most of the time, meaning that the residual model becomes less sensitive to water saturation after SHM. In contrast, for the scale parameters associated with porosity, the changes contain both positive and negative values, implying that the changes of their influence on the residual model are mixed.

Similar to Figs. 9 and 10 indicates the STDs of some of the kernel parameters in the initial ensemble (left column), and the changes of the STD values after SHM. As expected, through history matching, the STDs with respect to the final ensemble are reduced in comparison to those of the initial ensemble. On the other hand, the STDs of the kernel parameters do not appear to concentrate around the zero value, meaning that ensemble collapse do not appear to be a problem during history matching.

Finally, Table 5 reports the computational costs (in number of days) of the history-matching studies. The way for us to carry out the numerical computations is as follows: The forward reservoir (through ECLIPSE©) and seismic (through MATLAB©) simulations are conducted in a parallel way, using a computer server with 24 AMD Opteron 6180 CPUs, and 144 GB memory in total, but model updates through the IES have to wait until the forward simulations with respect to all ensemble members are finished. Under this setting, the computational costs of the history-matching studies do not appear sensitive to the hyper-parameter η^0 of the IES algorithm, but are more affected by the choice of whether MEC is conducted or not. When no MEC is introduced, the total computational time is around 3 days for both the cases $\eta^0 = 1$ and $\eta^0 = 10$. In contrast, with MEC introduced to the IES, the total computational time rises to around 5 days.

This 67% increment of the computational time is largely due to the bottleneck of the total computer memory, which is shared by all the sessions of parallel forward simulations. Indeed, when MEC is adopted, at each seismic survey time, the residual model in each forward seismic simulation needs to generates 44,927 AI data points distributed on active gridblocks, based on a set of 140,000 input kernel parameters. As such, generating all the AI data points simultaneously would require us to store a matrix in the dimension of 44927×140000 for each seismic survey time in each forward seismic simulation, which is not possible for our computational platform since we run multiple (24) forward simulations simultaneously. To address this problem, instead of simultaneously generating all the AI data points, we choose to do it sequentially in multiple (40) times, each time with a decreased amount of computer memory, but at the cost of increased computational time overall. Bearing this setting in mind, we expect that the computational overhead of adopting MEC would be reduced by conducting parallel numerical computations in a more powerful platform. As such, the integrated SHM workflow (with kernel-based MEC) will be applicable to (possibly much) larger reservoir models, which remains to be tested in future work.

6. Discussion and conclusion

In this work, we present an integrated seismic history matching (SHM) framework that consists of a few functional modules, namely, forward seismic simulations, seismic data processing and ensemble-based history matching. As the focus of the current work, we consider how to handle model errors arising in forward seismic simulations. To this end, we propose to introduce a parametrized residual model to the rock physics model, from a perspective of machine learning. During seismic history matching, the parameters associated with the residual model are then jointly estimated together with the reservoir model variables under estimation. In this way, the addition of the residual model to the SHM framework would only require some very minor

changes of the history matching algorithm, making it convenient for practical implementations.

For demonstration, we apply the integrated SHM framework to the full Norne field case study. We document the details in implementing the integrated SHM framework, and conduct comparison studies to examine the impact of the introduced residual model on history matching performance. To this end, we use the production data to cross-validate the qualities of the estimated reservoir models. The numerical results indicate that, in this particular case study, introducing the residual model to SHM leads to better history matching performance, in terms of the average data mismatch of forecast production data, compared to the case that the residual model is not adopted. On the other hand, a consequence of introducing the residual model is the increased computational overhead, which, in the current study, appears moderate and acceptable, and can be further mitigated by carrying out the computations in more powerful facilities.

As previously mentioned, using a machine-learning-based residual model represents a data-driven approach to compensating for the effect of model errors during history matching. In general, it would not be able to help identity the sources of model errors for model improvements. For this reason, we recommend that its use should be after domain knowledge, expertise and experience are properly and sufficiently exercised.

Related to the recent work of Lu and Chen (2020), we expect that it is possible to extend the idea in the current work, and build a residual model to account for model errors in the forward reservoir simulator which generates forecast production data based on the input reservoir models. Investigations in this regard will be considered in our future work.

CRediT authorship contribution statement

Xiaodong Luo: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing - original draft, Writing - review & editing. **Rolf J. Lorentzen:** Data curation, Investigation, Software, Validation. **Tuhin Bhakta:** Data curation, Methodology, Software, Supervision, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to thank anonymous reviewers for their valuable suggestions, and acknowledge financial support from the Research Council of Norway through the Petromaks-2 project DIGIRES (RCN no. 280473) and the industrial partners AkerBP, Wintershall DEA, Vår Energi, Petrobras, Equinor, Lundin and Neptune Energy. The authors would also like to thank Schlumberger for providing academic licenses to ECLIPSE©.

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.petrol.2020.107961.

Appendix A. Details of the rock physics model

Various RPM have been developed over the years to address various geological formations and their depositional environments. They vary from simple empirical relations derived from laboratory measurements to more complex physical models based on granular media theory, inclusion theory or other theories (Mavko et al., 2009).

In this work, we use the average of the Hashin-Shtrikman lower (LHS) and upper (UHS) bounds (Hashin and Shtrikman, 1963) for mineral constituents:

$$[K_s, G_s] \leftarrow \mathrm{HS}\left(K_{\mathrm{quartz}}, G_{\mathrm{quartz}}, K_{\mathrm{clay}}, G_{\mathrm{clay}}, V_{\mathrm{clay}}\right), \tag{A.1}$$

where K_s and G_s are the effective bulk and shear moduli of mineral mixing, respectively. Similarly, we define K_q and G_q, $q \in \{\mathrm{quartz}, \mathrm{clay}\}$, as the effective bulk and shear moduli of quartz and clay minerals, respectively. Furthermore, $V_{\mathrm{clay}} = 1 - \mathrm{NTG}$ is the volume fraction of clay, where NTG stands for the net-to-gross ratio of the formation layers. Here, we assume that the volume fraction of quartz (V_{quartz}) is equal to NTG. On the other hand, the dry bulk ($K^{\mathrm{ini}}_{\mathrm{dry}}$) and shear ($G^{\mathrm{ini}}_{\mathrm{dry}}$) moduli are computed as a function of porosity under initial reservoir conditions:

$$\left[K^{\mathrm{ini}}_{\mathrm{dry}}, G^{\mathrm{ini}}_{\mathrm{dry}}\right] \leftarrow f^{Norne}(\Phi), \tag{A.2}$$

where f^{Norne} stands for certain empirical relations specific for the Norne field, which are provided by Equinor but cannot be revealed here due to the confidentiality agreement between Equinor and NORCE. In addition, the pore-pressure effect is incorporated based on laboratory measurements derived from various core plugs of the Norne Field:

$$\left[K_{\mathrm{dry}}, G_{\mathrm{dry}}\right] \leftarrow g^{Norne}\left(P_{\mathrm{ini}}, P, K^{\mathrm{ini}}_{\mathrm{dry}}, G^{\mathrm{ini}}_{\mathrm{dry}}\right), \tag{A.3}$$

where P_{ini} and P stand for pore-pressures under initial reservoir conditions and those at a later time, respectively. The pressure model is denoted as g^{Norne}, and is calibrated using seismic observations in such a way that no pressure effect is observed below 310 bar (Huang et al., 2013). The fluid saturation effect on the dry moduli is then introduced by using the Gassmann model (Gassmann, 1951):

$$[K_{\mathrm{sat}}, G_{\mathrm{sat}}] \leftarrow \mathrm{Gassmann}\left(K_s, K_{\mathrm{dry}}, G_{\mathrm{dry}}, K_{fl}, \Phi\right), \tag{A.4}$$

here, K_{sat} and G_{sat} are saturated bulk and shear moduli, respectively. K_{fl} is the effective fluid bulk modulus, which is estimated by the Reuss average (Reuss, 1929): For the three-phase fluid mixture, K_{fl} is calculated as

$$K_{fl} = \left(\frac{S_w}{K_w} + \frac{S_o}{K_o} + \frac{S_g}{K_g}\right)^{-1}, \tag{A.5}$$

where K_w, K_o, K_g, S_w, S_o and S_g are bulk modulus of water/brine, bulk modulus of oil, bulk modulus of gas, saturation of water/brine, saturation of oil and saturation of gas, respectively. Further, the saturated density (Mavko et al., 2009) is calculated as (for the three-phase fluid)

$$\rho_{\text{sat}} = (1-\Phi)\rho_m + \Phi\left(S_w\rho_w + S_o\rho_o + S_g\rho_g\right), \tag{A.6}$$

where ρ_{sat}, ρ_m, ρ_w, ρ_o and ρ_g are saturated density of rock, mineral density, water/brine density, oil density and gas density, respectively. The mineral density is expressed as the weighted average between clay and quartz minerals, i.e., $\rho_m = V_{\text{clay}}\rho_{\text{clay}} + V_{\text{quartz}}\rho_{\text{quartz}}$.

The above equations is then used to calculate pressure (P)- and shear (S)-wave velocities (Mavko et al., 2009) as

$$v_p = \sqrt{\frac{K_{\text{sat}} + \frac{4}{3}G_{\text{sat}}}{\rho_{\text{sat}}}}, \tag{A.7}$$

and

$$v_s = \sqrt{\frac{G_{\text{sat}}}{\rho_{\text{sat}}}}, \tag{A.8}$$

where v_p and v_s represent P- and S-wave velocities, respectively. Moreover, we compute acoustic and shear impedances (z_p and z_s, respectively) using the following equations: $z_p = v_p \cdot \rho_{\text{sat}}$ and $z_s = v_s \cdot \rho_{\text{sat}}$. For more details about the RPM, readers are referred to Bhakta et al. (2016); Mavko et al. (2009) and the references therein.

References

Abadpour, A., Bergey, P., Piasecki, R., 2013. 4D seismic history matching with ensemble Kalman filter-assimilation on Hausdorff distance to saturation front. SPE Reservoir Simulation Symposium. Society of Petroleum Engineers. SPE-163635-MS.

Alfonzo, M., Oliver, D.S., 2020. Seismic data assimilation with an imperfect model. Comput. Geosci. 24, 889–905. https://doi.org/10.1007/s10596-019-09849-0. doi:10.1007/s10596-019-09849-0. URL:

Bhakta, T., Avseth, P., Landrø, M., 2016. Sensitivity analysis of effective fluid and rock bulk modulus due to changes in pore pressure, temperature and saturation. J. Appl. Geophys. 135, 77–89.

Box, G.E., Luceno, A., del Carmen Paniagua-Quinones, M., 2011. Statistical Control by Monitoring and Adjustment, ume 700. John Wiley & Sons.

Canchumuni, S.W., Emerick, A.A., Pacheco, M.A.C., 2019. History matching geological facies models based on ensemble smoother and deep generative models. J. Petrol. Sci. Eng. 177, 941–958.

Chen, Y., Oliver, D., 2013. Levenberg-Marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification. Comput. Geosci. 17, 689–703.

Chen, Y., Oliver, D.S., 2010. Cross-covariances and localization for EnKF in multiphase flow data assimilation. Comput. Geosci. 14, 579–601.

Chen, Y., Oliver, D.S., 2014. History matching of the Norne full-field model with an iterative ensemble smoother. SPE Reservoir Eval. Eng. 17, 244–256. SPE-164902-PA.

Dadashpour, M., Landrø, M., Kleppe, J., 2008. Nonlinear inversion for estimating reservoir parameters from time-lapse seismic data. J. Geophys. Eng. 5, 54.

Dee, D.P., 1995. On-line estimation of error covariance parameters for atmospheric data assimilation. Mon. Weather Rev. 123 (4), 1128–1145.

Donoho, D.L., Johnstone, J.M., 1994. Ideal spatial adaptation by wavelet shrinkage. Biometrika 81, 425–455.

Emerick, A., Reynolds, A., 2011. Combining sensitivities and prior information for covariance localization in the ensemble Kalman filter for petroleum reservoir applications. Comput. Geosci. 15, 251–269.

Emerick, A.A., 2016. Analysis of the performance of ensemble-based assimilation of production and seismic data. J. Petrol. Sci. Eng. 139, 219–239.

Emerick, A.A., Reynolds, A.C., 2012. Ensemble smoother with multiple data assimilation. Comput. Geosci. 55, 3–15.

Evensen, G., 2019. Accounting for model errors in iterative ensemble smoothers. Comput. Geosci. https://doi.org/10.1007/s10596-019-9819-z. doi:10.1007/s10596-019-9819-z. URL:

Evensen, G., Eikrem, K.S., 2018. Conditioning reservoir models on rate data using ensemble smoothers. Comput. Geosci. 1 https://doi.org/10.1007/s10596-018-9750-8.

Gassmann, F., 1951. Über die Elastizität poröser Medien. Vierteljahresschrift der Nat. Gesellschaft 96, 1–23.

Gosselin, O., Aanonsen, S., Aavatsmark, I., Cominelli, A., Gonard, R., Kolasinski, M., Ferdinandi, F., Kovacic, L., Neylon, K., 2003. History matching using time-lapse seismic (HUTS). SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers. SPE-84464-MS.

Hashin, Z., Shtrikman, S., 1963. A variational approach to the theory of the elastic behaviour of multiphase materials. J. Mech. Phys. Solid. 11 (2), 127–140.

He, J., Reynolds, A.C., Tanaka, S., Wen, X.H., Kamath, J., 2020. Calibrating field-scale uncertainties to local data: is the learning being overgeneralized? SPE J. https://doi.org/10.2118/191480-PA. SPE-191480-PA.

Huang, Y., Alsos, T., Sørensen, H.M., Tian, S., 2013. Proving the value of 4D seismic data in the late-life field–Case study of the Norne main field. First Break 31 (9), 57–67.

Iterative Ensemble Smoother master, 2019. URL: https://github.com/lanhill/iterative-ensemble-smoother https://github.com/lanhill/Iterative-Ensemble-Smoother.

Ketineni, S.P., Kalla, S., Oppert, S., Billiter, T., et al., 2018. Quantitative integration of 4D seismic with reservoir simulation. SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers.

Köpke, C., Irving, J., Elsheikh, A.H., 2018. Accounting for model error in Bayesian solutions to hydrogeophysical inverse problems using a local basis approach. Adv. Water Resour. 116, 195–207.

Lee, K., Lim, J., Yoon, D., Jung, H., 2019. Prediction of shale-gas production at duvernay formation using deep-learning algorithm. SPE J. 24, 2423–2437. SPE-195698-PA.

Lorentzen, R., Bhakta, T., Grana, D., Luo, X., Valestrand, R., Nævdal, G., 2020. Simultaneous assimilation of production and seismic data: application to the Norne field. Comput. Geosci. 24, 907–920. https://doi.org/10.1007/s10596-019-09900-0.

Lorentzen, R., Luo, X., Bhakta, T., Valestrand, R., 2019. History matching the full Norne field model using seismic and production data. SPE J. 24, 1452–1467. SPE-194205-PA.

Lorentzen, R.J., 2018. Norne initial ensemble. https://github.com/rolfjl/Norne-Initial-Ensemble. (Accessed 27 September 2020).

Lu, M., Chen, Y., 2020. Improved estimation and forecasting through residual-based model error quantification. SPE J. https://doi.org/10.2118/199358-PA. SPE-199358-PA.

Luo, X., 2019. Ensemble-based kernel learning for a class of data assimilation problems with imperfect forward simulators. PloS One 14, e0219247.

Luo, X., Bhakta, T., 2017. Estimating observation error covariance matrix of seismic data from a perspective of image denoising. Comput. Geosci. 21, 205–222.

Luo, X., Bhakta, T., 2020. Automatic and adaptive localization for ensemble-based history matching. J. Petrol. Sci. Eng. 184, 106559.

Luo, X., Bhakta, T., Jakobsen, M., Nævdal, G., 2017. An ensemble 4D-seismic history-matching framework with sparse representation based on wavelet multiresolution analysis. SPE J. 22, 985–1010. https://doi.org/10.2118/180025-PA; SPE-180025-PA.

Luo, X., Bhakta, T., Jakobsen, M., Nævdal, G., 2018a. Efficient big data assimilation through sparse representation: a 3D benchmark case study in petroleum engineering. PloS One 13, e0198586.

Luo, X., Bhakta, T., Nævdal, G., 2018b. Correlation-based adaptive localization with applications to ensemble-based 4D seismic history matching. SPE J. 23, 396–427. https://doi.org/10.2118/185936-PA. SPE-185936-PA.

Luo, X., Lorentzen, R.J., Valestrand, R., Evensen, G., 2019. Correlation-based adaptive localization for ensemble-based history matching: applied to the Norne field case study. SPE Reservoir Eval. Eng. 22, 1084–1109. https://doi.org/10.2118/191305-PA. SPE-191305-PA.

Luo, X., Stordal, A., Lorentzen, R., Nævdal, G., 2015. Iterative ensemble smoother as an approximate solution to a regularized minimum-average-cost problem: theory and applications. SPE J. 20, 962–982. https://doi.org/10.2118/176023-PA; SPE-176023-PA.

Ma, Z., Leung, J.Y., 2020. Integration of deep learning and data analytics for SAGD temperature and production analysis. Comput. Geosci. 24, 1239–1255.

Mavko, G., Mukerji, T., Dvorkin, J., 2009. The Rock Physics Handbook: Tools for Seismic Analysis of Porous Media. Cambridge University Press.

McKay, M.D., Beckman, R.J., Conover, W.J., 1979. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21 (2), 239–245.

Mosser, L., Dubrule, O., Blunt, M.J., 2017. Reconstruction of three-dimensional porous media using generative adversarial neural networks. Phys. Rev. 96, 043309 https://doi.org/10.1103/PhysRevE.96.043309. URL: https://link.aps.org/doi/10.1103/PhysRevE.96.043309.

Neto, G.M.S., Davolio, A., Schiozer, D.J., 2020. 3D seismic data assimilation to reduce uncertainties in reservoir simulation considering model errors. J. Petrol. Sci. Eng. 106967.

Obidegwu, D., Chassagne, R., MacBeth, C., 2017. Seismic assisted history matching using binary maps. J. Nat. Gas Sci. Eng. 42, 69–84.

Raanes, P.N., Stordal, A.S., Evensen, G., 2019. Revising the stochastic iterative ensemble smoother. Nonlinear Process Geophys. 26 (3), 325–338. https://doi.org/10.5194/npg-26-325-2019. URL: https://www.nonlin-processes-geophys.net/26/325/2019/.

Rammay, M.H., Elsheikh, A.H., Chen, Y., 2019. Quantification of prediction uncertainty using imperfect subsurface models with model error estimation. J. Hydrol. 576, 764–783. https://doi.org/10.1016/j.jhydrol.2019.02.056.

Raniolo, S., Dovera, L., Cominelli, A., Callegaro, C., Masserano, F., 2013. History match and polymer injection optimization in a mature field using the ensemble Kalman filter. IOR 2013-17th European Symposium on Improved Oil Recovery.

Reuss, A., 1929. Berechnung der fließgrenze von mischkristallen auf grund der plastizitätsbedingung für einkristalle. ZAMM-J. Appl. Math. Mech. Zeitschrift für Angewandte Math. Mech. 9, 49–58.

Shams, R., Masihi, M., Boozarjomehry, R.B., Blunt, M.J., 2020. Coupled generative adversarial and auto-encoder neural networks to reconstruct three-dimensional multi-scale porous media. J. Petrol. Sci. Eng. 186, 106794.

Skjervheim, J.A., Evensen, G., Aanonsen, S.I., Ruud, B.O., Johansen, T.A., 2007. Incorporating 4D seismic data in reservoir simulation models using ensemble Kalman filter. SPE J. 12, 282–292. SPE-95789-PA.

Soares, R., Luo, X., Evensen, G., Bhakta, T., 2020. 4D seismic history matching: Assessing the use of a dictionary learning based sparse representation method. J. Petrol. Sci. Eng. 195, 107763. https://doi.org/10.1016/j.petrol.2020.107763.

Souza, R., Lumley, D., Shragge, J., Davolio, A., Schiozer, D., 2019. 4d seismic bandwidth and resolution analysis for reservoir fluidflow model applications. ASEG Extend. Abstract. 2019 (1), 1–5.

Stephen, K.D., 2007. Scale and process dependent model errors in seismic history matching. Oil & Gas Sci. Technol. Revue de l'IFP 62 (2), 123–135.

Taha, T., Ward, P., Peacock, G., Heritage, J., Bordas, R., Aslam, U., Walsh, S., Hammersley, R., Gringarten, E., 2019. History matching using 4D seismic in an integrated multi-disciplinary automated workflow. SPE Reservoir Characterisation and Simulation Conference and Exhibition. Society of Petroleum Engineers.

Tahmasebi, P., Kamrava, S., Bai, T., Sahimi, M., 2020. Machine learning in geo-and environmental sciences: from small to large scale. Adv. Water Resour. 142, 103619.

Trani, M., Arts, R., Leeuwenburgh, O., 2012. Seismic history matching of fluid fronts using the ensemble Kalman filter. SPE J. 18, 159–171. SPE-163043-PA.

Williamson, D., Blaker, A.T., Hampton, C., Salter, J., 2015. Identifying and removing structural biases in climate models with history matching. Clim. Dynam. 45 (5–6), 1299–1324.

Zhang, Y., Leeuwenburgh, O., 2017. Image-oriented distance parameterization for ensemble-based seismic history matching. Comput. Geosci. 21 (4), 713–731.