Efficient big data assimilation through sparse representation: A 3D benchmark case study in petroleum engineering

Xiaodong Luo¹*, Tuhin Bhakta¹, Morten Jakobsen¹,², Geir Nævdal¹

1 International Research Institute of Stavanger (IRIS), Bergen, Norway, 2 University of Bergen, Bergen, Norway

* xiaodong.luo@iris.no

Abstract

Data assimilation is an important discipline in geosciences that aims to combine the information contents from both prior geophysical models and observational data (observations) to obtain improved model estimates. Ensemble-based methods are among the state-of-the-art assimilation algorithms in the data assimilation community. When applying ensemble-based methods to assimilate big geophysical data, substantial computational resources are needed in order to compute and/or store certain quantities (e.g., the Kalman-gain-type matrix), given both big model and data sizes. In addition, uncertainty quantification of observational data, e.g., in terms of estimating the observation error covariance matrix, also becomes computationally challenging, if not infeasible. To tackle the aforementioned challenges in the presence of big data, in a previous study, the authors proposed a wavelet-based sparse representation procedure for 2D seismic data assimilation problems (also known as history matching problems in petroleum engineering). In the current study, we extend the sparse representation procedure to 3D problems, as this is an important step towards real field case studies. To demonstrate the efficiency of the extended sparse representation procedure, we apply an ensemble-based seismic history matching framework with the extended sparse representation procedure to a 3D benchmark case, the Brugge field. In this benchmark case study, the total number of seismic data is in the order of \(O(10^6)\). We show that the wavelet-based sparse representation procedure is extremely efficient in reducing the size of seismic data, while preserving the salient features of seismic data. Moreover, even with a substantial data-size reduction through sparse representation, the ensemble-based seismic history matching framework can still achieve good estimation accuracy.

Introduction

Data assimilation is an important discipline in geosciences that aims to combine the information contents from both prior geophysical models and observational data (observations) to obtain improved model estimates [1]. The advance of modern technologies has led to a
Funding: XL acknowledges partial financial supports from the CIPR/IRIS cooperative research project “4D Seismic History Matching” (Research Council of Norway project number 243860), which is funded by industry partners, Eni (www.eni.com), Petrobras (www.petrobras.com), and Total (www.total.com), as well as the Research Council of Norway (PETROMAKS) (www.forskningradet.no). All authors acknowledge financial supports from the 4D seismic project (Research Council of Norway project number 230303), which is supported by the Research Council of Norway and the 4D seismic project (Research Council of Norway project number 243680), which supports from the CIPR/IRIS cooperative research collaboration. The authors received funding for this study from the following commercial companies: Eni; Petrobras; Total; ConocoPhillips Skandinavia AS (www.conocophillips.no); Aker BP ASA, Eni Norge AS (www.eni.com.norway.no); Maersk Oil Norway AS (www.maersk.com); Ørsted AS (https://orsted.com/en); Equinor Petroleum AS (www.equinor.com); Lundin Norway AS (lundin-norway.no); Halliburton AS (www.halliburton.no); DEA Norge AS (www.deanorge.com); Neptune Energy Norge AS (www.neptuneenergy.com); Schlumberger Norge AS (www.slb.com); Wintershall Norge AS (www.wintershall.no)-of The National IOR Centre of Norway.. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors received funding for this study from the following commercial companies: Eni; Petrobras; Total; ConocoPhillips Skandinavia; Aker BP ASA; Eni Norge AS; Maersk Oil Norway AS; Ørsted AS; Equinor AS; Neptune Energy Norge AS; Lundin Norway AS; Halliburton AS; Schlumberger Norge AS; and Wintershall Norge AS. There are no patents, products in development or marketed products to declare. This does not alter our adherence to all PLOS ONE policies on sharing data and materials.

massive growth of high-resolution observational data in geosciences [2, 3]. For instance, in the petroleum industry, Permanent Reservoir Monitoring (PRM) system is the cutting-edge technology used to collect 4-dimensional (4D) seismic data. The frequent multiple vintages of 4D seismic result in huge datasets, with the size of total data often in the order of hundreds of millions, or even higher. Therefore, there is a high demand from the petroleum industry for efficient methods of big data analytics to extract, analyze and utilize the information from big 4D seismic data. Similar problems are also faced in many other fields that involve abundant data obtained through, for example, satellite remote sensing [4], medical imaging [5], geophysical surveys [6, 7], and so on. As a result, big (geophysical) data assimilation has become an important topic in practice.

In this study, we focus on seismic data assimilation problems (also known as seismic history matching problems in the petroleum industry). In the petroleum industry, seismic is one of the most important tools used for reservoir exploration, monitoring, characterization and management. Compared to conventional production data used in history matching, seismic data is less frequent in time, but much denser in space. Therefore, complementary to production data, seismic data provide valuable additional information for reservoir characterization.

The main contribution of the current study is to extend an ensemble-based seismic history matching (SHM) framework in [8] to handle 3D reservoir characterization problems, and such an extension will be an important step towards real field applications. Compared to similar existing frameworks in reservoir engineering community, our proposed SHM framework consists of some relatively new ingredients, in terms of the seismic data in use, wavelet multiresolution analysis for the chosen seismic data and related data noise estimation, and the use of recently developed iterative ensemble history matching algorithms.

The core of our proposed SHM framework resides in two aspects. One is to reduce the size of big data, while retaining sufficient information for data assimilation; and the other is to quantify data uncertainties in the course of data-size reduction. Data reduction is a natural (and often necessary) step in big data analytics, as can be found in many other fields like geosciences [9], business management [10], and social or geo-spatial network [11] described by complex network theory [12–16]. Meanwhile, quantifying the uncertainties of observational data is a key step towards successful data assimilation. Consequently, devising a workflow that achieves simultaneous data reduction and uncertainty quantification will render an efficient big data assimilation framework. In the sequel, we elaborate how such a framework can be established in the context of ensemble-based data assimilation.

In the data assimilation community, ensemble-based methods [17] are among the state-of-the-art assimilation algorithms. The first ensemble-based method, the ensemble Kalman filter (EnKF), was proposed in [18]. The EnKF can be considered as a Monte Carlo implementation of the classic Kalman filter (KF) [19]. For large-scale data assimilation problems, the EnKF is computationally much more efficient than the KF in terms of propagating forward in time model error covariance matrices (which are needed for model updates using observational data). In addition, the EnKF is very simple to implement, since it is derivative-free (e.g., no need to linearize the dynamical model and/or the observation system as in the extended Kalman filter [20]), and works even when the dynamical model and the observation system are black-box systems. Apart from the EnKF, the ensemble implementations of fixed-lag and fixed-interval smoothers are also developed in [21, 22]. These methods thus form a family of ensemble-based assimilation algorithms that can be deployed for different types of estimation (filtering or smoothing) problems. Due to the practical conveniences and good performances in real-world applications, in the past decades, ensemble-based methods have widely spread into various disciplines of geosciences, e.g., meteorology [23], oceanography [24], hydrology [25], reservoir engineering [26, 27], to name but a few.
A number of ensemble-based seismic history matching frameworks are already proposed in the literature to assimilate seismic data into reservoir models. For instance, [28–33] adopt the ensemble Kalman filter (EnKF) or a combination of the EnKF and ensemble Kalman smoother (EnKS), whereas [8, 34–36] employ the ensemble smoother with multiple data assimilation (ES-MDA), and regularized Levenburg-Marquardt (RLM) based iterative ensemble smoother (RLM-MAC, see [37]), respectively.

There are different types of seismic data that one can use in SHM. Fig 1 provides an outline of the relation of some types of seismic data to reservoir petro-physical parameters (e.g., permeability and porosity) in forward simulations. As indicated there, using petro-physical parameters as the inputs to reservoir simulators, one generates fluid saturation and pressure fields. Through a petro-elastic model (PEM), one obtains acoustic and/or shear impedance (or equivalently, compressional and/or shear velocities (velocity)) and formation density based on fluid saturation and/or pressure, and porosity. Finally, amplitude versus angle (AVA) data are computed by plugging impedance (or velocities and density) into an AVA equation (e.g., Zoeppritz equation, see, for example, [38]).

To reduce the computational cost in forward simulations, many SHM studies use inverted seismic attributes that are obtained through seismic inversions. Such inverted properties can be, for instance, acoustic impedance (see, for example, [29, 32, 34, 35]) or fluid saturation fronts (see, for example, [28, 31, 33]). One issue in using inverted seismic attributes as the observational data is that, they may not provide uncertainty quantification for the observation errors, since inverted seismic attributes are often obtained using certain deterministic inversion algorithms. Quantifying the uncertainties in inverted seismic attributes is thus not a trivial issue [39, 40].

Since the volume of seismic data is typically huge, SHM often constitutes a “big data assimilation” problem. For ensemble-based history matching algorithms, a big data size may lead to
certain numerical problems, e.g., ensemble collapse and high costs in computing and storing Kalman gain matrices \[27, 34\]. In addition, many history matching algorithms are developed for under-determined inverse problems (in which there are infinitely many solutions that could match the observational data exactly), whereas a big data size could make the inverse problem become over-determined instead (such that there is no solution that could match the observational data exactly). The presence of big data thus changes the nature of the inverse problem, and may affect the performance of the history matching algorithm in use, as demonstrated in \[8\].

To avoid the extra uncertainties arising from a seismic inversion process, 2D AVA attributes are adopted as the seismic data in \[8\]. A wavelet-based sparse representation procedure is then adopted to transform the 2D datasets into the wavelet domain for data-size reduction and uncertainty quantification. Fig 2 explains the idea behind the wavelet-based sparse representation procedure. Given a set of 2D seismic data, one first applies a multilevel 2D discrete wavelet transform (DWT) to the data. 2D DWT is adopted there for the following two purposes: one is to reduce the size of seismic data by exploiting the sparse representation nature of wavelet basis functions, and the other is to exploit its capacity of noise estimation in the wavelet domain \[41\]. Based on the estimated standard deviation (STD) of the noise, one can construct the corresponding observation error covariance matrix that is needed in ensemble-based history matching algorithms.

For a chosen family of wavelet basis functions, seismic data are represented by the corresponding wavelet coefficients. When dealing with 2D data, 2D DWT is similar to singular
value decomposition (SVD) applied to a matrix. In the latter case, the matrix is represented by
the corresponding singular values in the space spanned by the products of associated left
and right singular vectors. Likewise, in the 2D case, one can also draw similarities between
wavelet-based sparse representation and truncated singular value decomposition (TSVD).
Indeed, in 2D DWT, small wavelet coefficients are typically dominated by noise, whereas large
coefficients mainly carry signal information [42]. Therefore, as demonstrated in [8], it is possi-
ble for one to use only a small subset of leading wavelet coefficients to capture the main fea-
tures of the 2D signal, while significantly reducing the number of wavelet coefficients. We
remark that TSVD-based sparse representation is not a suitable choice in the context of history
matching, since the associated basis functions (i.e., products of left and right singular vectors)
are data-dependent, meaning that in general it is not meaningful to compare and match the
singular values of observed and simulated data.

Wavelet-based sparse representation involves suppressing noise components in the wavelet
domain. To this end, one first estimates the STD of the noise in wavelet coefficients, and then
computes a threshold value that depends on both the noise STD and data size. One can sub-
stantially reduce the data size by only keeping leading wavelet coefficients above the threshold,
while setting those below the threshold value to zero. The leading wavelet coefficients are then
taken as the (transformed) seismic data, and are history-matched using a certain algorithm. In
this way, one can achieve data-size reduction and uncertainty quantification simultaneously.

In real field case studies, seismic data are typically 3D datasets. Therefore, in this work, we
focus on extending the 2D sparse representation procedure used in [8] to 3D problems. In this
regard, one may first decompose a 3D dataset into a collection of 2D maps, and then apply the
2D sparse representation procedure to each 2D map, similar to the approach used in [43]. By
doing so, however, the spatial correlations along a certain dimension may not be efficiently
coded into the wavelet domain, and this may reduce the efficiency of data compaction. As a
result, we follow the ideas in [44, 45] and directly apply 3D DWT to 3D seismic datasets. To
accommodate this 3D extension, the method of noise STD estimation in the wavelet domain is
also extended accordingly, following the rationale behind wavelet denoising [42].

This work is organized as follows. First, we introduce three key components of the pro-
posed SHM framework, which includes: (1) forward AVA simulation, (2) 3D DWT based
sparse representation procedure, and (3) regularized Levenburg-Marquardt based iterative
ensemble smoother. Then, we apply the proposed framework to the 3D Brugge benchmark
case, and investigate the performance of the proposed framework in various situations. Finally,
we draw conclusions based on the results of our investigation and discuss possible future
work.

The proposed framework

The proposed framework consists of three key components (see Fig 3), namely, forward AVA
simulation, sparse representation (in terms of leading wavelet coefficients) of both observed
and simulated AVA data, and the history matching algorithm. It is expected that the proposed
framework can also be extended to other types of seismic data, and more generally, geophysical
data with spatial correlations.

Forward AVA simulation

AVA data describes variations of seismic amplitudes with the change of distances between
sources and receivers. Typically, the variations in AVA data can be attributed to the changes in
lithology and/or fluid contents inside hydrocarbon reservoirs. As a result, the variations in
AVA data are often used by geophysicists to estimate static and dynamical properties (e.g.,
porosity, pressure and fluid saturations) and/or their changes in hydrocarbon reservoirs [46–49]. In the context of SHM, the forward AVA simulation involves several steps as shown in Fig 1. First, pore pressure and fluid saturations are generated through reservoir flow simulation that takes petro-physical (e.g., permeability and porosity) and other parameters as the inputs. The generated pressure and saturation values are then used to calculate seismic parameters, such as P- and S-wave velocities and densities of reservoir and overburden formations, through a petro-elastic model (PEM). Finally, a certain AVA equation is adopted to compute the AVA attributes at different angles or offsets.

In practice, constructing a proper PEM is crucial to the success of SHM. To interpret the changes in seismic response over time, an in-depth understanding of rock and fluid properties is required [50]. A proper PEM is often initially built upon our basic understanding of the properties of reservoir formation, and then adjusted and calibrated to match (typically a limited number of) observational data from well logs and/or other sources [51]. In this study, since our focus is on validating the performance of the proposed SHM framework in a 3D problem, we assume that the PEM is perfect. Here, we use the soft sand model as the PEM [51]. The model assumes that, the cement is deposited away from the grain contacts. It further considers that, the initial framework of the uncemented sand rock is densely random pack of

![Diagram of the proposed 4D seismic history matching framework.](https://doi.org/10.1371/journal.pone.0198586.g003)
spherical grains with the porosity (denote by $\phi$ hereafter) around 36%, which is the maximum porosity value that the rock could have before suspension. For convenience of discussion later, we denote this value as the critical porosity ($\phi_c$) [52, 53]. The dry bulk modulus ($K_{HM}$) and shear modulus ($\mu_{HM}$) at critical porosity can then be computed using the Hertz-Mindlin model [54] below

$$K_{HM} = \frac{C_p^2(1 - \phi_c)^2\mu_s^2}{18\pi^2(1 - \nu_s)^2P_{eff}}^\frac{1}{n},$$

and

$$\mu_{HM} = \frac{5 - 4\nu_s}{5(2 - \nu_s)}\frac{3C_p^2(1 - \phi_c)^2\mu_s^2}{2\pi^2(1 - \nu_s)^2P_{eff}}^\frac{1}{n},$$

where $\mu_s$, $\nu_s$, $P_{eff}$ are grain shear modulus, Poisson’s ratio, and effective stress, respectively. The coordination number $C_p$ denotes the average number of contacts per sphere, and $n$ is the degree of root. Here, $C_p$ and $n$ are set to 9 and 3, respectively.

To find the effective dry moduli for a porosity value less than $\phi_c$, the modified Lower Hashin-Shtrikman (MLHS) bound can be used [51]. The MLHS connects two end points in the elastic modulus-porosity plane. One end point, ($K_{HM}$, $\mu_{HM}$), corresponds to critical porosity $\phi_c$. The other end point corresponds to zero porosity, taking the moduli of the solid phase, i.e. quartz mineral ($K_s$, $\mu_s$). For a porosity value $\phi$ between zero and $\phi_c$, the lower bound for dry rock effective bulk ($K_d$) and shear ($G_d$) moduli can be expressed as

$$K_d = \left[\frac{\phi}{\phi_c} + \frac{1 - \phi}{\phi_c}\frac{4}{3}\frac{K_s}{\mu_{HM}}\right]^{-1} - \frac{4}{3}\mu_{HM},$$

and

$$G_d = \left[\frac{\phi}{\mu_{HM} + \frac{\mu_{HM}}{6}Z} + \frac{1 - \phi}{\mu_s - \frac{\mu_{HM}}{6}Z}\right]^{-1} - \frac{\mu_{HM}}{6}Z,$$

respectively, where $K_s$ is solid/mineral bulk modulus and $Z = (9K_{HM} + 8\mu_{HM})/(K_{HM} + 2\mu_{HM})$.

Further, the saturation effect is incorporated using the Gassmann model [55]. The saturated bulk modulus ($K_{sat}$) and shear modulus ($\mu_{sat}$) can be expressed as

$$K_{sat} = K_d + \frac{(1 - K_d)}{\phi(1 - \phi)\frac{K_s}{K_d} - \frac{K_s}{K_d}},$$

and

$$\mu_{sat} = \mu_d,$$

respectively, where $K_f$ is the effective fluid bulk modulus, and is estimated using the Reuss
average [56]. For an oil-water mixture (as is the case in the Brugge field), $K_f$ is given by

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

(7)

where $K_w$, $K_o$, $S_w$ and $S_o$ are bulk modulus of water/brine, bulk modulus of oil, saturation of water/brine and saturation of oil, respectively.

Further, the saturated density [51] can be written as (for the water-oil mixture)

$$\rho_{sat} = (1 - \phi)\rho_m + \phi S_w \rho_w + \phi S_o \rho_o,$$

(8)

where $\rho_{sat}$, $\rho_m$, $\rho_w$ and $\rho_o$ are saturated density of rock, mineral density, water/brine density and oil density, respectively.

Using the above equations, we can obtain P- and S-wave velocities given by [51]

$$V_p = \sqrt{\frac{K_{sat} + \frac{4}{3}\mu_{sat}}{\rho_{sat}}},$$

(9)

and

$$V_s = \sqrt{\frac{\mu_{sat}}{\rho_{sat}}},$$

(10)

where $V_p$ and $V_s$ represent P- and S-wave velocities, respectively.

After seismic parameters are generated by plugging reservoir parameters into the PEM, we can then simulate seismogram based on these seismic parameters. First, the Zoeppritz equation is used to calculate the reflection coefficient at an interface between two layers. For multilayer cases, we need to calculate reflectivity series as a function of two-way travel time (see, for example, [47, 51]). Here, travel time is computed from the P-wave velocity and vertical thickness of each grid block. We then convolve the reflectivity series with a Ricker wavelet of 45 Hz dominant frequency to obtain the desired seismic AVA data. In the experiments later, we generate AVA data at angles of 10˚ and 20˚, and call them near- and far-offset attributes, respectively. We choose not to further convert these AVA data into other attributes (e.g., intercept and gradient) to avoid introducing extra nonlinearities in the course of attribute conversion.

**Sparse representation and noise estimation in the wavelet domain**

Let $\mathbf{m}^r \in \mathbb{R}^m$ denote the reference reservoir model. In the current study, we consider 3D AVA attributes (near- and far-offset traces) in the form of $p_1 \times p_2 \times p_3$ arrays (tensors), where $p_1$, $p_2$ and $p_3$ represent the numbers of inline, cross-line and time slices in seismic surveys, respectively. Accordingly, let $g : \mathbb{R}^m \rightarrow \mathbb{R}^{p_1 \times p_2 \times p_3}$ be the forward simulator of AVA attributes. The observed AVA attributes $\mathbf{d}^o$ are supposed to be the forward simulation $g(\mathbf{m}^r)$ with respect to the reference model, plus certain additive observation errors $\epsilon$, that is,

$$\mathbf{d}^o = g(\mathbf{m}^r) + \epsilon.$$   

(11)

For ease of discussion below, suppose at the moment that all the tensors in Eq (11), i.e., $\mathbf{d}^o$, $g(\mathbf{m}^r)$ and $\epsilon$, are reshaped into vectors with $p_1 \times p_2 \times p_3$ elements. Throughout this study, we assume that, for a given AVA attribute, the elements of $\epsilon$ are independently and identically distributed (i.i.d) Gaussian white noise, with zero mean but unknown variance $\sigma^2$, where $\sigma$ will be estimated through wavelet multiresolution analysis below. More generally, one may also assume that $\epsilon$ follows a Gaussian distribution with zero mean and covariance $\mathbf{R}$, where $\mathbf{R}$ is...
not known a priori in practice. In this case, one may still apply awavelet-denoising-based method to estimate the noise STD in the wavelet domain, see, for example, [40].

As shown in Fig 2, wavelet-based sparse representation involves the following steps: (I) Apply DWT to seismic data; (II) Estimate noise STD of wavelet coefficients; and (III) Compute a threshold value that depends on both noise STD and data size, and do thresholding accordingly. Sparse representation (either wavelet-based or not) is also used for other purposes in petroleum engineering. For instance, it is adopted in [57–64] for re-parametrization of reservoir models, rather than size reduction of observational data. [57] also used wavelet-based sparse representation to represent time-lapse saturation maps. However, [57] does not consider the estimation of noise STD of wavelet coefficients, which is an issue to be addressed below.

To handle 3D seismic datasets, we need to extend the 2D sparse representation procedure used in [8] to 3D problems. To this end, instead of applying 2D DWT to a collection of 2D seismic maps as suggested in [43], we directly apply 3D DWT to the 3D datasets, in light of the works [44, 45].

For brevity, an introduction to wavelet theory is omitted in this work. Interested readers are referred to, for example, [65]. Instead, we focus more on explaining how wavelet transform is used for sparse representation and noise estimation in the wavelet domain. Before we proceed, let us first discuss the differences between 2D and 3D DWTs. Fig 4 illustrates single level decompositions in 2D (left) and 3D (right) DWTs, respectively. For convenience of discussion,
we label the dataset dimensions by $x$ and $y$ in 2D DWT, and $x$, $y$ and $z$ in 3D DWT. Without loss of generality, we assume that the data is $d^x_{j_{ll}l}$ in a 3D DWT ($d^x_{j_l}$ in a 2D DWT), which can be either the original observation $d^x$ (when $j = 0$) or the partial data recovered using the wavelet coefficients of the sub-band $LLL$ in 3D DWT ($LL$ in 2D DWT) when $j > 0$. Conventionally, a multiple dimensional DWT is implemented by sequentially applying 1D DWTs along different directions [42]. In each 1D DWT, there are both low-pass (L) and high-pass (H) filter banks, and the transform results in one “L” and one “H” sub-band of wavelet coefficients (intuitively, the “H” sub-band corresponds to high frequency components in the wavelet domain, while the “L” sub-band to low frequency ones). Therefore, given the dataset dimension $dim$, a single level wavelet decomposition would result in $2^{dim}$ sub-bands at the next level, e.g., four ($dim = 2$) for 2D DWT and eight ($dim = 3$) for 3D DWT. For convenience, one can tag these sub-bands by a combination of “H” and “L” labels. Here, let us assume $j = 0$, such that Fig 4 corresponds to the first level 3D (or 2D) wavelet transform. For 3D DWT, after a single level DWT, there are 8 sub-bands in the wavelet domain, which are labeled as $LLL_1$, $LLH_1$, $LHL_1$, $LHH_1$, $HLL_1$, $HLH_1$, $HHL_1$ and $HHH_1$, respectively. The sub-band $LLL_1$ ($HHH_1$) results only from low-pass (high-pass) filters, while the others from mixtures of low- and high-pass ones.

One can continue the 3D DWT to the next level by applying the transform to the data $d^x_{j_{ll}l}$ that corresponds to the sub-band $LLL_1$. This leads to a set of new sub-bands of wavelet coefficients (labelled as $LLL_2$, $LLH_2$, $LHL_2$, $LHH_2$, $HLL_2$, $HLH_2$, $HHL_2$ and $HHH_2$, respectively, and so on. The situation in 2D DWT is similar, but with a less number of sub-bands (labeled by $LL_p$, $LH_p$, $HL_p$ and $HH_p$ in general) at each level of wavelet decomposition.

For our purpose in the current study, we focus on 3D DWT in the sequel. In this case, since $HHH_1$ corresponds to the high frequency (typically noise) components of the original data $d^x$, it can be used to infer noise STD in the wavelet domain, similar to the idea used in [8]. Specifically, let $W$ and $T$ denote orthogonal wavelet transform and thresholding operators, respectively; $d^x = W(d^x)$ stands for the whole set of wavelet coefficients corresponding to the original data $d^x$, and $d^x_{HHH_1}$ for the wavelet coefficients in the sub-band $HHH_1$. After DWT and thresholding, the effective observation system becomes

$$ T \circ W(d^x) = T \circ (W \circ g(m^{ref}) + W(\epsilon)) . $$(12)

As will be discussed below, for leading wavelet coefficients (those above the threshold value), $T$ is an identity operator such that it does not modify the values of leading wavelet coefficients. The reason for us to require orthogonal $W$ is as follows: If $W$ is orthogonal, then the wavelet transform preserves the energy of Gaussian white noise (e.g., the Euclidean norm of the noise term $\epsilon$). In addition, like the power spectral distribution of white noise in frequency domain, the noise energy in the wavelet domain is uniformly distributed among all wavelet coefficients [42]. This implies that, if one can estimate the noise STD $\sigma$ of small wavelet coefficients (e.g., those in $HHH_1$), then this estimation can also be used to infer the noise STD of leading wavelet coefficients used in history matching. Similar to our previous study [8], the noise STD $\sigma$ is estimated using the median absolute deviation (MAD) estimator [41]:

$$ \sigma = \frac{\text{median}(\text{abs}(d^x_{HHH_1})))}{0.6745}, $$

(13)

where abs($\bullet$) is an element-wise operator, and takes the absolute value of an input quantity.
After estimating $\sigma$ in an $n$-level wavelet decomposition, we apply hard thresholding and select leading wavelet coefficients on the element-by-element basis, in a way such that

$$T(\tilde{d}_i^n) = \begin{cases} 0 & \text{if } |\tilde{d}_i^n| < \lambda, \\ \tilde{d}_i^n & \text{otherwise}, \end{cases}$$

(14)

where, without loss of generality, the scalar $\tilde{d}_i^n \in \tilde{d}^n$ represents an individual wavelet coefficient, and $\lambda$ is a certain threshold value to be computed later. Eq (14) means that, for leading wavelet coefficients above (or equal to) the threshold $\lambda$, their values are not changed, whereas for those below $\lambda$, they are set to zero. Note that in [8], hard thresholding is not applied to the coarsest sub-band (i.e., the $LL_n/LLL_n$ sub-band for an $n$-level 2D/3D DWT) in light of the fact that the wavelet coefficients in this sub-band correspond to low-frequency components, which are typically dominated by the signal. As a result, applying thresholding to this sub-band may lead to certain loss of signal information in history matching. However, for an AVA attribute in this study, we have observed that its corresponding $LLL_n$ sub-band may contain a large amount of wavelet coefficients (e.g., in the order of $10^4$). To have the flexibility of efficiently reducing the data size, we lift the restriction such that thresholding can also be applied to the sub-band $LLL_n$. By doing so, however, some extra (relatively small) low-frequency wavelet coefficients are discarded, such that one may end up with more loss of data information content.

In [8], the threshold value $\lambda$ is computed using the universal rule [66]

$$\lambda = \sqrt{2 \ln(\#\tilde{d}^n)} \sigma,$$

(15)

with $\#\tilde{d}^n$ being the number of elements in $\tilde{d}^n$. The rationale behind the universal rule is as follows: suppose that all the elements of a set $\tilde{d}^n$ (e.g., wavelet coefficients) are noise samples drawn from the Gaussian distribution $N(0, \sigma^2)$, then the probability $pr(\max(abs(\tilde{d}^n)) < \lambda) \rightarrow 1$ as the number of elements of $\tilde{d}^n$ tends to $\infty$ [65]. Therefore, if an element of $\tilde{d}^n$ is larger than the threshold $\lambda$, then it is more likely to be a signal component, than to be a noise component. This viewpoint thus underpins the thresholding strategy in Eq (14).

In the current work, when using Eq (15) to select the threshold value, it is found that the resulting number of leading wavelet coefficients may still be very large. As a result, in the experiments later, we select the threshold value according to

$$\lambda = c \sqrt{2 \ln(\#\tilde{d}^n)} \sigma,$$

(16)

where $c > 0$ is a positive scalar, and the larger the value of $c$, the less the number of leading wavelet coefficients. Therefore, the scalar $c$ can be used to control the total number of leading wavelet coefficients.

Combining Eqs (12)–(16), the effective observation system in history matching becomes

$$\tilde{d}^n = \mathcal{W} \circ g(m_{\mathcal{W}}) + \mathcal{W}(\epsilon), \text{ for } \tilde{d}^n \geq \lambda,$$

(17)

where now $\tilde{d}^n$ is a vector containing all selected leading wavelet coefficients, and $\mathcal{W}(\epsilon)$ the corresponding noise component in the wavelet domain, with zero mean and covariance $\mathcal{C}_\epsilon = \sigma^2 I$ (here $I$ is the identity matrix with a suitable dimension).

We use an example to illustrate the performance of sparse representation and noise estimation in 3D DWT. In this example, we first generate a reference AVA far-offset trace using the forward AVA simulator. The dimension of this trace is $139 \times 48 \times 251$, therefore the data size is 1, 674, 672. Fig 5A plots slices of the AVA trace at $X = 40, 80, 120$, and $Z = 50, 100, 150$ and 200, respectively. We then add Gaussian white noise to obtain the noisy AVA trace, with the
noise level being 30% (we also tested the performance of Eq (13) in estimating noise STD with different noise levels, e.g., 0% and 10%, and obtained results similar to what are presented below). Here, noise level is defined as:

\[
\text{Noise level} = \frac{\text{variance of noise}}{\text{variance of pure signal}}.
\]  

(18)

Fig 5B shows slides of the noisy AVA trace at the same locations as in Fig 5A.
We apply a three-level 3D DWT to the noisy data using Daubechies wavelets with two vanishing moments [65], and use hard thresholding combined with the universal rule (Eqs (13)–(15)) to select leading wavelet coefficients. After thresholding, the number of leading wavelet coefficients reduces to 33, 123, only 2% of the original data size. On the other hand, by applying Eq (13), the estimated noise STD is 0.0105, and it is very close to the true noise STD 0.0104. By applying an inverse 3D DWT to leading wavelet coefficients, we obtain the reconstructed AVA trace (Fig 5C plots slices of this trace at the same places as Fig 5A and 5B). Comparing Fig 5A–5C, one can see that, using leading wavelet coefficients that amounts to only 2% of the original data size, the slices of reconstructed AVA trace well capture the main features in the corresponding slices of the reference AVA data.

Fig 5D–5F show wavelet coefficients in the sub-bands \( HHL_1 \) of the reference, noisy and reconstructed AVA traces, respectively. From these figures, one can see that, after applying thresholding to wavelet coefficients of noisy data (Fig 5E), the modified coefficients (Fig 5F) preserve those with large amplitudes in the reference case (Fig 5D). In general, the modified coefficients appear similar to those of the reference case, whereas certain small coefficients of the reference case are suppressed due to thresholding.

Finally, Fig 5G–5I depict slices of reference and estimated noise, and their difference, respectively, at the same places as in Fig 5A. Here, reference noise is defined as noisy AVA data (Fig 5B) minus reference AVA data (Fig 5A), estimated noise as noisy AVA data minus reconstructed AVA data (Fig 5C), and noise difference as estimated noise minus reference noise. The estimated noise appears very similar to the reference noise, although there are also certain differences according to Fig 5I. This might be largely due to the fact that some small wavelet coefficients of the reference data are smeared out after thresholding, as aforementioned.

The ensemble history matching algorithm

We adopt the RLM-MAC algorithm [37] in history matching. Without loss of generality, let \( d^o \) denote \( p \)-dimensional observations in history matching, which stands for values in the ordinary data space (e.g., 3D AVA attributes by reshaping 3D arrays into vectors), or their sparse representations in the transform domain (e.g., leading wavelet coefficients in the wavelet domain). The observations \( d^o \) are contaminated by Gaussian noise with zero mean and covariance \( C_d \) (denoted by \( d^o \sim N(0, C_d) \)). Also denote by \( g \) the forward simulator that generates simulated observations \( d^g(m) \) given an \( m \)-dimensional reservoir model \( m \).

In the RLM-MAC algorithm, let \( M^i \equiv \{ m^i_j \}_{j=1}^{N_e} \) be an ensemble of \( N_e \) reservoir models obtained at the \( i \)-th iteration step, based on which we can construct two square root matrices used in the RLM-MAC algorithm. One of the matrices is in the form of

\[
S^i_m = \frac{1}{\sqrt{N_e-1}} \left[ m^i_1 - \bar{m}^i, \ldots, m^i_{N_e} - \bar{m}^i \right]; \quad \bar{m}^i = \frac{1}{N_e} \sum_{j=1}^{N_e} m^i_j, \quad (19)
\]

and is called model square root matrix, in the sense that \( C_m^i = S^i_m (S^i_m)^T \) equals the sample covariance matrix of the ensemble \( M^i \). The other, defined as

\[
S^i_d = \frac{1}{\sqrt{N_e-1}} \left[ g(m^i_1) - g(\bar{m}^i), \ldots, g(m^i_{N_e}) - g(\bar{m}^i) \right], \quad (20)
\]

is called data square root matrix for a similar reason.
The RLM-MAC algorithm updates $M_i$ to a new ensemble $M_{i+1} = \{m_j^{i+1}\}_{j=1}^{N_e}$ by solving the following minimum-average-cost problem

$$\arg\min_{\{m_j^{i+1}\}_{j=1}^{N_e}} \frac{1}{N_e} \sum_{j=1}^{N_e} \left[ (d_j - g(m_j^{i+1}))^T C_d^{-1} (d_j - g(m_j^{i+1})) + \gamma (m_j^{i+1} - m_j^i)^T (C_m)^{-1} (m_j^{i+1} - m_j^i) \right],$$

(21)

where $d_j^i$ ($j = 1, 2, \cdots, N_e$) are perturbed observations generated by drawing $N_e$ samples from the Gaussian distribution $N(d_j^i, C_d)$, and $\gamma$ a positive scalar that can be used to control the step size of an iteration step, and is automatically chosen using a procedure similar to back-tracking line search [37]. Through linearization, the MAC problem is approximately solved through the following iteration:

$$m_j^{i+1} = m_j^i + S_m(S_d)^T (S_d(S_d)^T + \gamma C_d)^{-1} (d_j - g(m_j^i)), \text{ for } j = 1, \cdots, N_e.$$  

(22)

The stopping criteria have substantial impact on the performance of an iterative inversion algorithm [67]. [37] mainly used the following two stopping conditions for the purpose of run-time control:

(C1) RLM-MAC stops if it reaches a maximum number of iteration steps;
(C2) RLM-MAC stops if the relative change of average data mismatch over two consecutive iteration steps is less than a certain value.

For all the experiments later, we set the maximum number of iterations to 20, and the limit of the relative change to 0.01%.

Let

$$\Xi^i \equiv \frac{1}{N_e} \sum_{j=1}^{N_e} \left[ (d_j^i - g(m_j^i))^T C_d^{-1} (d_j^i - g(m_j^i)) \right]$$

(23)

be the average (normalized) data mismatch with respect to the ensemble $M_i$. Following Proposition 6.3 of [67], a third stopping condition is introduced and implemented in [8]. Concretely, we also stop the iteration in Eq (22) when

$$\Xi^i < 4 \rho$$

(24)

for the first time, where the factor 4 is a critical value below which the iteration process starts to transit from convergence to divergence [67]. Numerical results in [8] indicate that, in certain circumstances, equipping RLM-MAC with the extra stopping condition (24) may substantially improve its performance in history matching. Readers are referred to [8] for more details. In the current study, however, the impact of the stopping criterion (24) is not as substantial as that in [8]. Nevertheless, we prefer to keep this stopping criterion as an extra safeguard procedure.

**Numerical results in the Brugge benchmark case**

We demonstrate the performance of the proposed workflow through a 3D Brugge benchmark case study. Table 1 summarizes the key information of the experimental settings. Readers are referred to [68] for more information of the benchmark case study.

The Brugge field model has 9 layers, and each layer consists of $139 \times 48$ gridblocks. The total number of gridblocks is 60048, whereas among them 44550 are active. The data set of the original benchmark case study does not contain AVA attributes, therefore we generate
synthetic seismic data in the following way: The benchmark case contains an initial ensemble of 104 members. We randomly pick one of them as the reference model (which turned out to be the member “FN-SS-KP-1-92”), and use the rest 103 members as the initial ensemble in this study. The model variables to be estimated include porosity (PORO) and permeability (PERMX, PERMY, PERMZ) at all active gridblocks. Consequently, the total number of parameters in estimation is 178200.

There are 20 producers and 10 water injectors in the reference model, and they are controlled by the liquid rate (LRAT) target. The production period is 10 years, and in history matching we use production data at 20 report times. The production data consist of oil production rates (WOPR) and water cuts (WWCT) at 20 producers, and bottom hole pressures (WBHP) at all 30 wells. Therefore the total number of production data is 1400. Gaussian white noise is added to production data of the reference model. For WOPR and WWCT data, their noise STD are taken as the maximum values between 10% of their magnitudes and $10^{-6}$ (the latter is used to prevent the numerical issue of division by zero), whereas for WBHP data, the noise STD is 1 bar.

In the experiments, there are three seismic surveys taking place on day 1 (base), day 991 (1st monitor), and day 2999 (2nd monitor), respectively. At each survey, we apply forward AVA simulation described in the previous section to the static (porosity) and dynamic (pressure and saturation) variables of the reference model, and generate AVA attributes at two different angles: 10° (near-offset) and 20° (far-offset). Each AVA attribute is a 3D ($139 \times 48 \times 9$) cube, and consists of around $1.17 \times 10^6$ elements. Therefore the total number of 4D seismic data is around $3 \times 2 \times 1.17 \times 10^6 = 7.02 \times 10^6$. For convenience of discussion later, we label the dimensions of the 3D cubes by $X$, $Y$ and $Z$, respectively, such that $X = 1, 2, \cdots, 139$, $Y = 1, 2, \cdots, 48$ and $Z = 1, 2, \cdots, 176$. In history matching, we add Gaussian white noise to each reference AVA attribute, with the noise level being 30%. Here we do not assume the noise STD is known. Instead, we first apply three-level 3D DWT to each AVA attribute using Daubechies wavelets with two vanishing moments, and then use Eq (13) to estimate noise STD in the wavelet domain.

Table 1. Summary of experimental settings in the Brugge benchmark case study.

| Parameter | Value |
|-----------|-------|
| Model dimension | $139 \times 48 \times 9$ (60048 gridblocks), with 44550 out of 60048 being active cells |
| Parameters to estimate | PORO, PERMX, PERMY, PERMZ. Total number is $4 \times 44550 = 178200$ |
| Gridblock size | Irregular. Average $\Delta X \approx 93 \text{m}$, $\Delta Y \approx 91 \text{m}$, and average $\Delta Z \approx 5 \text{m}$ |
| Reservoir simulator | ECLIPSE 100 (control mode LRAT) |
| Number of wells | 10 injectors and 20 producers |
| Production period | 3647.5 days (with 20 report times) |
| Production data | Production wells: BHP, OPR and WCT; Injection wells: BHP. Total number: $20 \times 70 = 1400$ |
| Seismic survey time | Base: day 1; Monitor (1st): day 991; Monitor (2nd): day 2999 |
| 4D seismic data | AVA data from near- and far- offsets at three survey times. Total number: $\sim 7.02 \times 10^6$ |
| DWT (seismic) | Three-level decomposition using 3D Daubechies wavelets with two vanishing moments |
| Thresholding | Hard thresholding based on Eqs (14) and (16) |
| History matching method | iES (RLM-MAC) with an ensemble of 103 reservoir models |

https://doi.org/10.1371/journal.pone.0198586.t001
In what follows, we consider three history matching scenarios that involve: (S1) production data only; (S2) 4D seismic data only; and (S3) both production and 4D seismic data. Because of the huge volumes of AVA attributes, in scenarios (S2) and (S3), we cannot directly use the 4D seismic data in the original data space (although this was done in the 2D problem of [8]). Indeed, in one of our experiments that attempted to use the original seismic dataset, we encountered the problem of running out of computer memory. Therefore, in the current work, to examine the impact of data size on the performance of SHM, in each scenario (S2 or S3), we consider two cases that have different numbers of leading wavelet coefficients. This is achieved by letting the scalar $c$ of Eq (16) be 1 and 5, respectively. In contrast, for production data in (S1) and (S3), no wavelet transform is applied as the data size is already small.

Results of scenario S1 (using production data only)

Fig 6 shows the boxplots of data mismatch as a function of iteration step. The average data mismatch of the initial ensemble (iteration 0) is around $5.65 \times 10^9$. After 20 iteration steps, the average data mismatch is reduced to 5431.97, lower than the threshold value $4 \times 1400 = 5600$ in (24) for the first time. In this particular case, the stopping step selected according to the criterion (24) coincides with the maximum number of iteration steps. Therefore, we take the ensemble at the 20th iteration step as the final estimation.

In this synthetic study, the reference reservoir model is known. As a result, we use root mean squared error (RMSE) in the sequel to measure the $\ell_2$-distance (up to a factor) between an estimated model and the reference one. More specifically, let $v^\prime$ be the $\ell$-dimensional reference property, and $\hat{v}$ an estimation, then the RMSE $e_v$ of $\hat{v}$ with respect to the reference $v^\prime$ is

![Fig 6. Boxplots of production data mismatch as a function of iteration step (scenario S1). The horizontal dashed line indicates the threshold value ($4 \times 1400 = 5600$) for the stopping criterion (24). For visualization, the vertical axis is in the logarithmic scale. In each box plot, the horizontal line (in red) inside the box denotes the median; the top and bottom of the box represent the 75th and 25th percentiles, respectively; the whiskers indicate the ranges beyond which the data are considered outliers, and the whiskers positions are determined using the default setting of MATLAB R2019b, while the outliers themselves are plotted individually as plus signs (in red).](https://doi.org/10.1371/journal.pone.0198586.g006)
defined by
\[ e_v = \frac{\| \mathbf{v} - \mathbf{v}' \|_2}{\sqrt{\ell}}, \]  
where \( \| \cdot \|_2 \) denotes the \( \ell_2 \) norm.

For brevity, Fig 7 reports the boxplots of RMSEs in estimating PERMX (in the natural log scale) and PORO, at different iteration steps, whereas the results for PERMY and PERMZ are similar to that for PERMX. As can be seen in Fig 7, the average RMSEs of both log PERMX and PORO tend to reduce as the number of iteration steps increases.

Fig 8 shows the profiles of WBHP, WOPR and WWCT of the initial (1st row) and final (2nd row) ensemble at the producer BR-P-5. It is evident that, through history matching, the final ensemble matches the production data better than the initial one, and this is consistent with the results in Fig 6.

For illustration, Fig 9 presents the reference log PERMX and PORO at layer 2 (1st column), the mean of log PERMX and PORO at layer 2 from the initial ensemble (2nd column), and the mean of log PERMX and PORO at layer 2 from the final ensemble (3rd column). A comparison between the initial and final estimates of log PERMX and PORO indicates that the final estimates appear more similar to the reference fields, in consistence with the results in Fig 7.

Results of scenario S2 (using seismic data only)

To examine the impact of data size on the performance of SHM, we consider two cases with different threshold values chosen through Eq (16). In the first case, we let \( c = 1 \), such that Eq (16) reduces to the universal rule in choosing the threshold value [66]. Under this choice, the number of leading wavelet coefficients is 178332, around 2.5% of the original AVA data size (7.04 \times 10^6). In the second case, we increase the value of \( c \) to 5, such that the number of leading wavelet coefficients further reduces to 3293, which is more than 4000 times reduction in comparison to the original data size. Let \( \mathbf{d}^0 \) be the data reconstructed through an inverse DWT using the leading wavelet coefficients of the noisy data \( \mathbf{d}^o \) (while setting the rest of the wavelet coefficients of \( \mathbf{d}^o \) to zero). Then to
evaluate information loss in the sparse representation procedure, we adopt the following measure

\[ \mathcal{L} = \frac{\|d^{\text{ref}} - d^{\text{rec}}\|_2}{\|d^{\text{ref}}\|_2} \times 100\%. \]  

(26)

In our experiments, the information loss is 38.7% when \( c = 1 \), whereas it creases to 91.4% at \( c = 5 \). The gap in information loss with different \( c \) values will be further illustrated later.

Fig 10 indicates the boxplots of seismic data mismatch as functions of iteration step. In either case, seismic data mismatch reduces fast at the first few iteration steps, and then changes slowly afterwards. The stopping criterion (C2), monitoring the relative change of average data mismatch, becomes effective in both cases, such that the iteration stops at the 11th step when \( c = 1 \), and at the 19th when \( c = 5 \). Accordingly, the ensembles at iteration step 11 and 19, respectively, are taken as the final estimates in these two cases. In addition, it appears that
ensemble collapse takes place in both cases, although this phenomenon is somewhat mitigated in the case \( c = 5 \), in comparison to the case \( c = 1 \). The mitigation of ensemble collapse is even more evident when we further increase \( c \) to 8, and accordingly, reduce the data size to 534. By doing so, however, the history matching performance is deteriorated (results not included here), largely due to the fact that such a significant reduction of data size leads to substantial loss of information content in the seismic data, a point to be elaborated soon.

Fig 11 shows boxplots of RMSEs of log PERMX (1st column) and PORO (2nd column) as functions of iteration step. It is clear that the RMSEs of the final ensembles are lower than those of the initial ones, even at \( c = 5 \), the case in which data size is reduced more than 2000 times. On the other hand, when \( c = 1 \) (top row), the RMSEs of both log PERMX and PORO in

---

**Fig 9.** Log PERMX (top row) and PORO (bottom row) of the reference reservoir model (1st column) and the means of the initial (2nd column) and final (3rd column) ensembles at Layer 2 (scenario S1). The black dots in the figures represent the locations of injection and production wells (top view).

https://doi.org/10.1371/journal.pone.0198586.g009

---

**Fig 10.** Boxplots of seismic data mismatch as functions of iteration step (scenario S2). Case (a) corresponds to the results with \( c = 1 \), for which choice the number of leading wavelet coefficients is 178332, roughly 2.5% of the original data size; Case (b) to the results with \( c = 5 \), for which choice the number of leading wavelet coefficients is 3293, more than 2000 times reduction in data size.

https://doi.org/10.1371/journal.pone.0198586.g010
the final ensemble are lower than those at \( c = 5 \) (bottom row). This indicates that, better history matching performance is achieved at \( c = 1 \), with more information content captured in the leading wavelet coefficients.

As aforementioned, each 3D seismic dataset is in the dimension of 139 \( \times \) 48 \( \times \) 176. For illustration, the top row of Fig 12 indicates the slices of far-offset AVA attributes at \( X = 80 \), with respect to the base survey (1st column), the 1st monitor survey (2nd column) and the 2nd monitor survey (3rd column), respectively, whereas the middle and bottom rows show the corresponding slices reconstructed using the leading wavelet coefficients (while setting other coefficients to zero) at \( c = 1 \) and \( c = 5 \), respectively. Compared to figures in the top row, it is clear that the reconstructed ones at \( c = 1 \) capture the main features of the observed slices, while removing the noise component. Therefore in this case, although the universal rule (corresponding to \( c = 1 \)) still leads to a relatively large data size, it achieves a good trade-off between data size reduction and feature preservation. In contrast, at \( c = 5 \), the seismic data size is

Fig 11. Boxplots of RMSEs of log PERMX (1st column) and PORO (2nd column) as functions of iteration step, with \( c \) being 1 (top) and 5 (bottom), respectively (scenario S2).

https://doi.org/10.1371/journal.pone.0198586.g011
significantly reduced. However, the reconstructed slices in the bottom row only retain a small portion of the strips in the observed slices of the top row, meaning that the data size is reduced at the cost of losing substantial information content of the seismic data. Nevertheless, even with such an information loss, using the leading wavelet coefficients at $c = 5$ still leads to significantly improved model estimation in comparison to the initial ensemble, and this will become more evident when both production and seismic data are used in scenario S3.
For brevity, in what follows we only present the results with respect to the case $c = 1$. In the top row of Fig 13, we show the slices (at $X = 80$) of differences between two groups of reconstructed far-offset AVA attributes. One group corresponds to the reconstructed far-offset AVA attributes at three survey times, using the leading wavelet coefficients ($c = 1$) of the observed far-offset AVA attributes. The other group contains the reconstructed far-offset AVA attributes at three survey times, using the corresponding leading wavelet coefficients ($c = 1$) of the mean simulated seismic data of the initial ensemble. Therefore the slices of differences in the top row can be considered as a reflection of the initial seismic data mismatch in Fig 10A. Here, we use the slices of differences for ease of visualization, as the reconstructed slices of the observed and the mean simulated AVA attributes look very similar. Similarly, in the bottom row, we show the slices of differences between the reconstructed far-offset AVA attributes of the observed seismic data, and the reconstructed far-offset AVA attributes of the mean simulated seismic data of the final ensemble. In this case, the slices of differences can be considered as a reflection of the final seismic data mismatch in Fig 10A. Comparing the top and bottom rows at a given survey time, one can observe certain distinctions, which, however, are not very significant in general. This is in line with the results in Fig 10A, where the initial and final
seismic data mismatch remain in the same order, in contrast to the substantial reduction of production data mismatch in scenario S1 (Fig 6).

Similar to Fig 9, Fig 14 depicts the reference log PERMX and PORO at layer 2 (1st column), the mean of initial log PERMX and PORO at layer 2 (2nd column), and the mean of final log PERMX and PORO at layer 2 (3rd column). Compared to the initial mean estimates, the final mean log PERMX and PORO show clear improvements, in terms of the similarities to the reference fields. In addition, an inspection on the 3rd columns of Figs 9 and 14 reveals that the final mean estimates in S2 capture the geological structures of the reference fields better, especially in areas where there is neither injection nor production well (well locations are represented by black dots in Figs 9 and 14).

Results of scenario S3 (using both production and seismic data)

In scenario S3, production and seismic (in terms of leading wavelet coefficients) data are assimilated simultaneously. Fig 15 reports the boxplots of production (top) and seismic (bottom) data mismatch as functions of iteration step. Because of the simultaneous assimilation of production and seismic data, the way to use the seismic data (in terms of the value of $c$ in Eq (16)) will affect the history matching results. This becomes evident if one compares the first and second columns of Fig 15. Indeed, when $c = 1$, because the relatively large data size, it is clear that ensemble collapse takes place in Fig 15A and 15C. Also, the iteration stops at step 14, due to the stopping criterion (C2). By increasing $c$ to 5, the size of seismic data is reduced from 178332 to 3293, and ensemble collapse seems mitigated to some extent, especially for production data, while the final iteration step is 19, due to the stopping criterion (C2). On the other hand, by comparing Figs 6, 10 and 15, it is clear that, in S3, the presence of both production and seismic data makes the reduction of data mismatch different from the case of using either production or seismic data only. For instance, in the presence of seismic data, the production data mismatch (see Fig 15A and 15B) tend to be higher than that in Fig 6. On the other hand, with the influence of production data, the occurrence of ensemble collapse seems to be delayed in Fig 15C and 15D, in comparison to those in Fig 10.

Fig 16 shows boxplots of RMSEs of log PERMX (1st column) and PORO (2nd column) as functions of iteration step. Again, the RMSEs of the final ensembles are lower than those of the initial ones, for either $c = 1$ or $c = 5$. On the other hand, a comparison of Figs 7, 11 and 16...
indicates that the RMSEs of log PERMX and PORO (and similarly, the RMSEs of log PERMY and log PERMZ) are the lowest when using both production and seismic data in history matching. Using $c = 5$ in scenario S3 (Fig 16C and 16D) leads to higher RMSEs than using $c = 1$. Nevertheless, they are still better than the RMSEs in scenario S1, and close to (for PORO) or better than (for log PERMX) those in scenario S2 with $c = 1$ (see Fig 11A and 11B). This suggests that, in this particular case, reasonably good history matching performance can still be achieved, even though the data size is reduced more than 2000 times (at $c = 5$) through the wavelet-based sparse representation procedure.

Again, for brevity, in what follows we only report the results with respect to the case $c = 1$. Fig 17 shows the production data profiles at the producer BR-P-5 in scenario S3 with $c = 1$. Clearly, compared to the initial ensemble, the final one matches the observed production data (red dots) better. Nevertheless, a comparison of the bottom rows of Figs 8 and 17 indicates that, the ensemble spreads of simulated production data (blue curves) tend to be under-
estimated, such that the reference production data (yellow curves) are outside the profiles of simulated production data at certain time instances.

Similar to Fig 13, in Fig 18 we show the slices (at X = 80) of differences between the reconstructed far-offset AVA attributes of observed and mean simulated seismic data, at three survey times. Again, compared to the slices with respect to the initial ensemble (top), there are some visible distinctions in the slices with respect to the final ensemble (bottom). However, if one compares the bottom rows of Figs 13 and 18, it seems that these slices look very similar to each other. This is consistent with the results in Figs 10A and 15C, where the final seismic data mismatch of S2 and S3 remains close.

Finally, Fig 19 compares the reference, initial and final mean log PERMX and PORO fields at layer 2. Again, the final mean estimates improve over the initial mean fields, in terms of the closeness to the references. In addition, a comparison of the final estimated fields (the 3rd columns) of Figs 9, 14 and 19 shows that the final mean estimates in S3 best capture the geological

---

**Fig 16.** Boxplots of RMSEs of log PERMX (1st column) and PORO (2nd column) as functions of iteration step, with c being 1 (top) and 5 (bottom), respectively (scenario S3).

https://doi.org/10.1371/journal.pone.0198586.g016
structures of the reference fields (the same observation is also obtained at \( c = 5 \)). This indicates the benefits of using both production and seismic data in history matching.

**Discussion and conclusion**

In this work, we extend a 2D wavelet-based sparse representation procedure used in [8] to handle 3D seismic datasets, and integrate it into an ensemble-based seismic history matching framework. To demonstrate the efficiency of the integrated workflow, we apply it to a 3D benchmark case, the Brugge field case. The seismic data used in this study are near- and far-offset amplitude versus angle (AVA) attributes, with the data size being more than 7 million.

Through numerical experiments, we show that the 3D sparse representation procedure is very efficient for data-size reduction, and works well for noise STD estimation in the wavelet domain. Via sparse representation, the size of seismic data (in the form of leading wavelet coefficients) can be conveniently controlled through a threshold value. A relatively large threshold value means more reduction in data size, which is desirable for the history matching algorithm.
but at the cost of extra information loss. In contrast, a relatively small threshold value results in a larger number of leading wavelet coefficients, hence better preserves the information content of observed data. In this case, however, the history matching algorithm may become more vulnerable to certain practical issues like ensemble collapse. As a result, the best practice would

Fig 18. As in Fig 13, but for the slices (at X = 80) of differences in scenario S3 with c = 1.

https://doi.org/10.1371/journal.pone.0198586.g018

Fig 19. As in Fig 9, but for scenario S3 with c = 1.

https://doi.org/10.1371/journal.pone.0198586.g019
need to achieve a trade-off between reduction of data size and preservation of data information. Another observation from the experiment results is that, in this particular case, a combined use of production and seismic data in history matching leads to better estimation results than the cases of using either production or seismic data only.

Ensemble collapse is clearly visible when seismic data is used in history matching. This phenomenon can be mitigated to some extent by increasing the threshold value (hence reducing the seismic data size), but it cannot be completely avoided. A possible remedy to this problem is to also introduce localization (see, for example, [69, 70]) to the iterative ensemble smoother. In the presence of the sparse presentation procedure, however, seismic data are transformed into the wavelet domain, and the concept of “physical distance” are not valid any more. As a result, localization will need to be adapted to this change. We will investigate this issue in our future study.

In the present study, the AVA datasets are 3D attributes, and there are no missing data points. In this case, the datasets exhibit regular spatial coverages, and the conventional wavelets can be directly used. When extending the sparse representation procedure to other types of datasets, e.g., acoustic impedance distributed over active reservoir gridblocks and remotely-sensed ocean or atmosphere satellite images, the datasets of interest may have irregular spatial coverages. In this case, due to the irregularities of data distribution fields, the conventional wavelets may not be applicable. To address this issue, there are two possible strategies. One is to add some artificial data (e.g., mean field values) in such a way that, with the extra data points, the expanded datasets have regular spatial coverages. The other (more sophisticated) strategy is to modify the lifting scheme in the wavelets, such that the corresponding DWT is applicable to datasets with irregular spatial distributions [71]. We have tested the first simple idea (adding artificial data points), and observed that it works in some preliminary studies (results not reported here). Implementing the second idea (modifying the wavelet lifting scheme) is more complicated in general, and has not been fully investigated by us yet.

In practice, the observation errors of certain seismic data, such as (inverted) acoustic impedance, may exhibit spatial correlations and thus should be treated as colored noise. In this case, it may not be appropriate to only use the wavelet coefficients in the sub-band resulted from the high-pass filters to estimate the noise level. Instead, it is advocated to estimate the noise level for each wavelet sub-band [40, 41]. Based on this notion, an enhancement of the wavelet-based sparse representation procedure used in the current study will be implemented to handle seismic data with colored noise, and the relevant results will be reported elsewhere.

Acknowledgments

We would like to thank Dr. Takemasa Miyoshi and two anonymous reviewers for their valuable comments on the work, and Schlumberger for providing us academic software licenses to ECLIPSE. XL acknowledges partial financial supports from the CIPR/IRIS cooperative research project “4D Seismic History Matching” (Research Council of Norway project number 243680), which is funded by industry partners, Eni (www.eni.com), Petrobras (www.petrobras.com), and Total (www.total.com), as well as the Research Council of Norway (PETROMAKS) (www.forskningsradet.no). All authors acknowledge financial supports from the 4D seismic project (Research Council of Norway project number 230303), which is supported by the Research Council of Norway and the industry partners-ConocoPhillips Skandinavia AS (www.conocophillips.no), Aker BP ASA, Eni Norge AS (www.eninorge.com), Maersk Oil Norway AS (www.maerskoi.com), Ørsted AS (https://orsted.com/en/), EquiNor Petroleum AS (www.equinor.com), Lundin Norway AS (lundin-norway.no), Halliburton AS (www.halliburton.no), DEA Norge AS (www.dea-norge.com), Neptune Energy...
Norge AS (www.neptuneenergy.com), Schlumberger Norge AS (www.slb.com), Wintershall Norge AS (www.wintershall.no)-of The National IOR Centre of Norway,. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

Conceptualization: Xiaodong Luo.
Data curation: Tuhin Bhakta, Morten Jakobsen.
Formal analysis: Xiaodong Luo.
Funding acquisition: Geir Nævdal.
Investigation: Xiaodong Luo.
Methodology: Xiaodong Luo, Geir Nævdal.
Project administration: Geir Nævdal.
Software: Tuhin Bhakta.
Supervision: Geir Nævdal.
Validation: Xiaodong Luo, Tuhin Bhakta, Morten Jakobsen, Geir Nævdal.
Visualization: Xiaodong Luo, Tuhin Bhakta.
Writing – original draft: Xiaodong Luo, Tuhin Bhakta.
Writing – review & editing: Xiaodong Luo.

References

1. Talagrand O. Assimilation of Observations, an Introduction. Journal of the Meteorological Society of Japan Ser II. 1997; 75:191–209. https://doi.org/10.2151/jmsj1965.75.1B_191
2. Miyoshi T, Kuni M, Ruiz J, Lien GY, Satoh S, Ushio T, et al. “Big data assimilation” revolutionizing severe weather prediction. Bulletin of the American Meteorological Society. 2016; 97:1347–1354. https://doi.org/10.1175/BAMS-D-15-00144.1
3. Miyoshi T, Lien GY, Satoh S, Ushio T, Bessho K, Tomita H, et al. “Big Data Assimilation” toward post-peta-scale severe weather prediction: An overview and progress. Proceedings of the IEEE. 2016;104:2155–2179.
4. Pan M, Wood EF, Wójcik R, McCabe MF. Estimation of regional terrestrial water cycle using multi-sensor remote sensing observations and data assimilation. Remote Sensing of Environment. 2008; 112:1282–1294. https://doi.org/10.1016/j.rse.2007.02.039
5. George AK, Butala MD, Frazin RA, Kamalabadi F, Bresler Y. Time-resolved cardiac CT reconstruction using the ensemble Kalman filter. In: Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on. IEEE; 2008. p. 1489–1492.
6. Katterbauer K, Arango S, Sun S, Hotell I. Multi-data reservoir history matching for enhanced reservoir forecasting and uncertainty quantification. Journal of Petroleum Science and Engineering. 2015; 128:160–176. https://doi.org/10.1016/j.ptole.2015.02.016
7. Katterbauer K, Arango S, Sun S, Hotell I. Integrating gravimetric and interferometric synthetic aperture radar data for enhancing reservoir history matching of carbonate gas and volatile oil reservoirs. Geophysical Prospecting. 2017; 65:337–364. https://doi.org/10.1111/1365-2478.12371
8. Luo X, Bhakta T, Jakobsen M, Nævdal G. An Ensemble 4D-Seismic History-Matching Framework With Sparse Representation Based On Wavelet Multiresolution Analysis. SPE Journal. 2016;SPE-180025-PA.
9. Spanias AS, Jonsson SB, Steams SD. Transform methods for seismic data compression. IEEE Transactions on Geoscience and Remote Sensing. 1991; 29(3):407–416. https://doi.org/10.1109/36.79431
10. ur Rehman MH, Chang V, Batool A, Wah TY. Big data reduction framework for value creation in sustainable enterprises. International Journal of Information Management. 2016; 36(6):917–928. https://doi.org/10.1016/j.ijinfomgt.2016.05.013

11. Stai E, Karyoti S, Papavassiliou S. A hyperbolic space analytics framework for big network data and their applications. IEEE Network. 2016; 30(1):11–17. https://doi.org/10.1109/MNET.2016.7389825

12. Zhang J, Small M. Complex network from pseudoperiodic time series: Topology versus dynamics. Physical review letters. 2006; 96(23):238701. https://doi.org/10.1103/PhysRevLett.96.238701 PMID: 16803415

13. Zhang J, Sun J, Luo X, Zhang K, Nakamura T, Small M. Characterizing pseudoperiodic time series through the complex network approach. Physica D: Nonlinear Phenomena. 2008; 237(22):2856–2865. https://doi.org/10.1016/j.physd.2008.05.008

14. Gao ZK, Cai Q, Yang YX, Dang WD, Zhang SS. Multiscale limited penetrable horizontal visibility graph for analyzing nonlinear time series. Scientific reports. 2016; 6:35622. https://doi.org/10.1038/srep35622 PMID: 27759088

15. Gao ZK, Yang YX, Fang PC, Zou Y, Xia CY, Du M. Multiscale complex network for analyzing experimental multivariate time series. EPL (Europhysics Letters). 2015; 109(3):30005. https://doi.org/10.1209/0295-5075/109/30005

16. Gao ZK, Small M, Kurths J. Complex network analysis of time series. EPL (Europhysics Letters). 2016; 116(5):50001. https://doi.org/10.1209/0295-5075/116/50001

17. Evensen G. Data Assimilation: The Ensemble Kalman Filter. Springer Science & Business Media; 2009.

18. Evensen G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J Geophys Res. 1994; 99:10143–10162. https://doi.org/10.1029/94JC00572

19. Kalman R. A new approach to linear filtering and prediction problems. Trans ASME, Ser D, J Basic Eng. 1960; 82:35–45. https://doi.org/10.1115/1.3662552

20. Jazwinski AH. Stochastic Processes and Filtering Theory. Academic Press; 1970.

21. Evensen G, van Leeuwen PJ. An Ensemble Kalman Smoother for Nonlinear Dynamics. Mon Wea Rev. 2000; 128:1852–1867. https://doi.org/10.1175/1520-0493(2000)128%3C1852:AESFNL%3E2.0.CO;2

22. Van Leeuwen PJ, Evensen G. Data assimilation and inverse methods in terms of a probabilistic formulation. Mon Wea Rev. 1996; 124:2898–2913. https://doi.org/10.1175/1520-0493(1996)124%3C2898:DAIMII%3E2.0.CO;2

23. Houtekamer P, Zhang F. Review of the ensemble Kalman filter for atmospheric data assimilation. Monthly Weather Review. 2016; 144:4489–4532. https://doi.org/10.1175/MWR-D-15-0440.1

24. Evensen G. The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dyn. 2003; 53:343–367. https://doi.org/10.1007/s10236-003-0036-9

25. Reichle RH, McLaughlin DB, Entekhabi D. Hydrologic data assimilation with the ensemble Kalman filter. Monthly Weather Review. 2002; 130:103–114. https://doi.org/10.1175/1520-0493(2002)130%3C0103:HDAAWE%3E2.0.CO;2

26. Naevdal G, Mannseth T, Vefring EH. Near-well reservoir monitoring through ensemble Kalman filter. In: SPE/DOE Improved Oil Recovery Symposium. Tulsa, Oklahoma, USA; 2002. p. 13–17. SPE-75235-MS.

27. Aanonsen S, Naevdal G, Oliver D, Reynolds A, Vallès B. The Ensemble Kalman Filter in Reservoir Engineering: a Review. SPE Journal. 2009; 14:393–412. SPE-117274-PA. https://doi.org/10.2118/117274-PA

28. Abadpour A, Bergey P, Pialecki R. 4D Seismic History Matching With Ensemble Kalman Filter-Assimilation on Hausdorff Distance to Saturation Front. In: SPE Reservoir Simulation Symposium. Society of Petroleum Engineers; 2013. SPE-163635-MS.

29. Fahimuddin A, Aanonsen S, Skjervheim JA. Ensemble Based 4D Seismic History Matching—Integration of Different Levels and Types of Seismic Data. In: 72nd EAGE Conference & Exhibition; 2010.

30. Katterbauer K, Hoteit I, Sun S. History matching of electromagnetically heated reservoirs incorporating full-wavefield seismic and electromagnetic imaging. SPE Journal. 2015; 20:923–941. SPE-173896-PA. https://doi.org/10.2118/173896-PA

31. Leeuwenburgh O, Arts R. Distance parameterization for efficient seismic history matching with the ensemble Kalman Filter. Computational Geosciences. 2014; 18:535–548. https://doi.org/10.1007/s10596-014-9434-y
32. Skjervheim JA, Evensen G, Aanonsen SI, Ruud BO, Johansen TA. Incorporating 4D seismic data in reservoir simulation models using ensemble Kalman filter. SPE Journal. 2007; 12:282–292. SPE-95789-PA. https://doi.org/10.2118/95789-PA

33. Trani M, Arts R, Leeuwenburgh O. Seismic history matching of fluid fronts using the ensemble Kalman filter. SPE Journal. 2012; 18:159–171. SPE-163043-PA. https://doi.org/10.2118/163043-PA

34. Emerick AA, Reynolds AC. History matching time-lapse seismic data using the ensemble Kalman filter with multiple data assimilations. Computational Geosciences. 2012; 16:639–659. https://doi.org/10.1007/s10596-012-9275-5

35. Emerick AA, Reynolds AC, et al. History-matching production and seismic data in a real field case using the ensemble smoother with multiple data assimilation. In: SPE Reservoir Simulation Symposium. Society of Petroleum Engineers; 2013. SPE-163675-MS.

36. Emerick AA. Analysis of the performance of ensemble-based assimilation of production and seismic data. Journal of Petroleum Science and Engineering. 2016; 139:219–239. https://doi.org/10.1016/j.petrol.2016.01.029

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

38. Avseth P, Mukerji T, Mavko G. Quantitative seismic interpretation: Applying rock physics tools to reduce interpretation risk. Cambridge University Press; 2010.

39. Aanonsen S, Aavatsmark I, Barkve T, Cominelli A, Gonard R, Gosselin O, et al. Effect of scale dependent data correlations in an integrated history matching loop combining production data and 4D seismic data. In: SPE Reservoir Simulation Symposium; 2003.

40. Luo X, Bhakta T. Estimating observation error covariance matrix of seismic data from a perspective of image denoising. Computational Geosciences. 2016; 21:205–222. https://doi.org/10.1007/s10596-016-9605-0

41. Donoho DL, Johnstone IM. Adapt to unknown smoothness via wavelet shrinkage. Journal of the American Statistical Association. 1995; 90:1200–1224. https://doi.org/10.1080/01621459.1995.10476626

42. Jansen M. Noise reduction by wavelet thresholding. vol. 161. Springer Science & Business Media; 2012.

43. Górszczyk A, Malinowski M, Bellefleur G. Enhancing 3D post-stack seismic data acquired in hardrock environment using 2D curvelet transform. Geophysical Prospecting. 2015; 63:903–918. https://doi.org/10.1111/1365-2478.12234

44. Starck J-L, Fadili JM, Digel S, Zhang B, Chiang J. Source detection using a 3D sparse representation: application to the Fermi gamma-ray space telescope. Astronomy & Astrophysics. 2009; 504:641–652. https://doi.org/10.1051/0004-6361/200811388

45. Shi L, Hu Y, Chen Y, Yin X, Shu H, Luo L, et al. Improving low-dose cardiac CT images based on 3D sparse representation. Scientific reports. 2016; 6.

46. Bhakta T, Landrø M. Estimation of pressure-saturation changes for unconsolidated reservoir rocks with high VP/VS ratio. Geophysics. 2014; 79:M35–M54. https://doi.org/10.1190/geo2013-0434.1

47. Buland A, Omre H. Bayesian linearized AVO inversion. Geophysics. 2003; 68:185–198. https://doi.org/10.1190/1.1543206

48. Dadashpour M, Landrø M, Kleppe J. Nonlinear inversion for estimating reservoir parameters from time-lapse seismic data. Journal of Geophysics and Engineering. 2008; 5:54. https://doi.org/10.1088/1742-2132/5/1/006

49. Landrø M. Discrimination between pressure and fluid saturation changes from time-lapse seismic data. Geophysics. 2001; 66:836–844.

50. Jack I. The coming of age for 4D seismic. First Break. 2001; 19:24–28.

51. Mavko G, Mukerji T, Dvorkin J. The rock physics handbook: Tools for seismic analysis of porous media. Cambridge University Press; 2009.

52. Nur A, Marion D, Yin H. Wave velocities in sediments. In: Shear waves in marine sediments. Springer; 1991. p. 131–140.

53. Nur A, Mavko G, Dvorkin J, Galmudi D. Critical porosity: A key to relating physical properties to porosity in rocks. The Leading Edge. 1998; 17(3):357–362. https://doi.org/10.1190/1.1437977

54. Mindlin RD. Compliance of elastic bodies in contact. Journal of Applied Mechanics. 1949; 16:259–268.

55. Gassmann F. Über die Elastizität poröser Medien. Vierteljahresschrift der Naturforschenden Gesellschaft. 1951; 96:1–23.
56. Reuss A. Berechnung der Fließgrenze von Mischkristallen auf Grund der Plastizitätsbedingung für Einkristalle. ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik. 1929; 9:49–58. https://doi.org/10.1002/zamm.19290090104

57. Awotunde AA. A multiresolution adjoint sensitivity analysis of time-lapse saturation maps. Computational Geosciences. 2014; 18:677–696. https://doi.org/10.1007/s10596-014-9414-2

58. Chen Y, Oliver DS. Multiscale parameterization with adaptive regularization for improved assimilation of nonlocal observation. Water Resources Research. 2012; 48:W04503. https://doi.org/10.1029/2011WR011144

59. Elsheikh AH, Wheeler MF, Hotelt I. Sparse calibration of subsurface flow models using nonlinear orthogonal matching pursuit and an iterative stochastic ensemble method. Advances in Water Resources. 2013; 56:14–26. https://doi.org/10.1016/j.adwatres.2013.02.002

60. Jafarpour B, Goyal VK, McLaughlin DB, Freeman WT. Compressed history matching: exploiting transform-domain sparsity for regularization of nonlinear dynamic data integration problems. Mathematical Geosciences. 2010; 42:1–27. https://doi.org/10.1007/s11004-009-9247-z

61. Gentilhomme T, Oliver DS, Mannseth T, Caumon G, Moyen R, Doyen P. Ensemble-based multi-scale history-matching using second-generation wavelet transform. Computational Geosciences. 2015; 19:999–1025. https://doi.org/10.1007/s10596-015-9517-4

62. Khaninezhad MM, Jafarpour B. Prior model identification during subsurface flow data integration with adaptive sparse representation techniques. Computational Geosciences. 2014; 18:3–16. https://doi.org/10.1007/s10596-013-9378-7

63. Sahni I, Horne RN. Multiresolution wavelet analysis for improved reservoir description. SPE Reservoir Evaluation & Engineering. 2005; 8:53–69. SPE-87820-PA. https://doi.org/10.2118/87820-PA

64. Sana F, Ravanelli F, Al-Naffouri TY, Hotelt I. A Sparse Bayesian Imaging Technique for Efficient Recovery of Reservoir Channels With Time-Lapse Seismic Measurements. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2016; 9:2242–2254. https://doi.org/10.1109/JSTARS.2016.2563163

65. Mallat S. A wavelet tour of signal processing. Academic press; 2008.

66. Donoho DL, Johnstone JM. Ideal spatial adaptation by wavelet shrinkage. Biometrika. 1994; 81:425–455. https://doi.org/10.1093/biomet/81.3.425

67. Engl HW, Hanke M, Neubauer A. Regularization of Inverse Problems. Springer; 2000.

68. Peters L, Arts R, Brouwer A, Geel C, Cullick S, Lorentzen RJ, et al. Results of the Brugge benchmark study for flooding optimization and history matching. SPE Reservoir Evaluation & Engineering. 2010; 13:391–405. https://doi.org/10.2118/119094-PA

69. Emerick A, Reynolds A. Combining sensitivities and prior information for covariance localization in the ensemble Kalman filter for petroleum reservoir applications. Computational Geosciences. 2011; 15:251–269. https://doi.org/10.1007/s10596-010-9198-y

70. Chen Y, Oliver DS. Cross-covariances and localization for EnKF in multiphase flow data assimilation. Computational Geosciences. 2010; 14:579–601. https://doi.org/10.1007/s10596-009-9174-6

71. Sweldens W. The lifting scheme: A construction of second generation wavelets. SIAM journal on mathematical analysis. 1998; 29:511–546. https://doi.org/10.1137/S0036141095289051