

Towards real-time monitoring: data assimilated time-lapse full waveform inversion for seismic velocity and uncertainty estimation

Chao Huang¹ and Tieyuan Zhu^{1,2}

¹*Department of Geosciences, Pennsylvania State University, University Park, PA 16802, USA*

²*EMS Energy Institute, Pennsylvania State University, University Park, PA 16802, USA. E-mail: tyzhu@psu.edu*

Accepted 2020 July 9. Received 2020 July 2; in original form 2019 November 15

SUMMARY

Rapid development of time-lapse seismic monitoring instrumentations has made it possible to collect dense time-lapse data for tomographically retrieving time-lapse (even continuous) images of subsurface changes. While traditional time-lapse full waveform inversion (TLFWI) algorithms are designed for sparse time-lapse surveys, they lack of effective temporal constraint on time-lapse data, and, more importantly, lack of the uncertainty estimation of the TLFWI results that is critical for further interpretation. Here, we propose a new data assimilation TLFWI method, using hierarchical matrix powered extended Kalman filter (HiEKF) to quantify the image uncertainty. Compared to existing Kalman filter algorithms, HiEKF allows to store and update a data-sparse representation of the cross-covariance matrices and propagate model errors without expensive operations involving covariance matrices. Hence, HiEKF is computationally efficient and applicable to 3-D TLFWI problems. Then, we reformulate TLFWI in the framework of HiEKF (termed hereafter as TLFWI-HiEKF) to predict time-lapse images of subsurface spatiotemporal velocity changes and simultaneously quantify the uncertainty of the inverted velocity changes over time. We demonstrate the validity and applicability of TLFWI-HiEKF with two realistic CO₂ monitoring models derived from Frio-II and Cranfield CO₂ injection sites, respectively. In both 2-D and 3-D examples, the inverted high-resolution time-lapse velocity results clearly reveal a continuous velocity reduction due to the injection of CO₂. Moreover, the accuracy of the model is increasing over time by assimilating more time-lapse data while the standard deviation is decreasing over lapsed time. We expect TLFWI-HiEKF to be equipped with real-time seismic monitoring systems for continuously imaging the distribution of subsurface gas and fluids in the future large-scale CO₂ sequestration experiments and reservoir management.

Key words: Probability distributions; Statistical methods; Time-series analysis; Waveform inversion.

1 INTRODUCTION

Monitoring the dynamic evolution of injected fluids or gas in the subsurface is crucial for CO₂ sequestration, geothermal exploration, or enhanced oil recovery (Lazaratos & Marion 1997; Daley *et al.* 2007; Zhu *et al.* 2019). Time-lapse seismic monitoring is a cost-effective and promising way to provide the spatiotemporal image of seismic changes due to fluids or gas migration (Lumley 2001). However, the poor repeatability of traditional sparse time-lapse seismic surveys will likely decrease the spatial resolution of the time-lapse image because the errors caused by non-repeatable surveys may easily mask subtle seismic changes (e.g. leakages, Arogunmati & Harris 2012). Additionally, the temporal resolution of seismic changes is often limited due to the sparsity of time-lapse surveys (typically year interval). To improve the image spatiotemporal resolution of

physical parameters, repeatable dense time-lapse or even continuous active-source seismic monitoring (CASSM) surveys have been proposed and tested in projects of monitoring CO₂ plume (Daley *et al.* 2007; Zhu *et al.* 2017) and hydraulic fracturing (Ajo-Franklin *et al.* 2011). Sequent time-lapse traveltimes tomographic analysis of CASSM data has been reported to image the dynamic evolution of a reservoir or injected CO₂ plume in the few minutes temporal resolution (Daley *et al.* 2008; Ajo-Franklin *et al.* 2011).

Compared to traveltimes tomography, time-lapse full waveform inversion (TLFWI) can reconstruct high-resolution time-lapse seismic changes by exploiting all information embedded in seismic data. The straightforward strategy to perform TLFWI is utilizing independent FWI to deal with baseline and monitor data, then the velocity changes are derived from the subtraction between the monitored and the baseline velocity models. This subtraction between

two models required two inversions converging to a similar level. Otherwise, the velocity changes could be contaminated by unphysical differences due to different local minima between inversions. On the other hand, the sequential strategy was proposed to use the inverted result from FWI at the previous time lapse as the initial model for FWI at the current time lapse (Zheng *et al.* 2011; Routh *et al.* 2012; Raknes *et al.* 2013; Asnaashari *et al.* 2015; Raknes & Arntsen 2015). Another effective strategy is double-difference TLFWI that minimizes the data differences between two time-lapse data sets to directly obtain time-lapse velocity changes (Watanabe *et al.* 2004; Denli & Huang 2009). The above-described methods have been demonstrated to sparse time-lapse seismic data (baseline and one or two monitoring data sets, e.g. Zhang & Huang 2013; Yang *et al.* 2016), but it is not clear whether they are suitable to deal with dense even continuous seismic monitoring data.

Due to the nature of seismic inversion (e.g. incomplete data, noise, instrument errors, etc.), the uncertainty of the TLFWI results would be crucial for subsequent interpretation of time-lapse changes. Nonetheless, to our best knowledge, very few studies on evaluating uncertainty of time-lapse images have been reported. Following recent studies on the quantification of the FWI uncertainty, we could assess the uncertainty of TLFWI results by calculating the posterior covariance matrix in each time lapse. The posterior covariance is equal to the inverse Hessian when all errors are Gaussian (Tarantola 2005; Fichtner & Trampert 2011; Bui-Thanh *et al.* 2013; Zhu *et al.* 2016). However, the calculation of the inverse Hessian is still computationally challenging in single FWI processing (Virieux & Operto 2009; Zhu *et al.* 2016).

An excellent approach to the aforementioned problems is incorporating the Kalman filter (KF) into geophysical inverse problems. This not only constrains the time-lapse variation (evolution) but also provides the uncertainty estimation. KF is a powerful data assimilation tool based on Bayesian inference for evaluating time-series data of random variables that incorporates knowledge acquired from all previous time-steps into the estimation of parameters at the current time-step (Kalman 1960). Unlike deterministic inversion that only gives a single best estimation in the least-squares sensing, the solutions given by KF represent a range of possible estimates with uncertainty quantification, characterized by the maximum *a posteriori* best estimation and a statistical covariance. The original KF is only suitable to deal with the linear problem. The extended Kalman filter (EKF, Anderson & Moore 1979; Dennis Jr & Schnabel 1996) approximates the nonlinearity between the measurements and state parameters using a first-order Taylor expansion around the current state. The accuracy of the EKF is heavily dependent on the degree of nonlinearity. We can increase the accuracy of EKF by either using high-order Taylor expansions (Gelb 1974) or applying iterative approaches when dealing with nonlinearity problem (Dennis Jr & Schnabel 1996). Recently, Eikrem *et al.* (2019) showed the ability of iterative EKF to assess the uncertainty of 2-D TLFWI. However, the iterative EKF has to explicitly store and calculate the covariance error matrices at every iteration within each time lapse, and the immense storage and computational cost make it only suitable for small-scale seismic problems. Instead of iterative EKF, Huang & Zhu (2019) proposed to use TLFWI to predict a good *a priori* velocity model to reduce the nonlinearity in their TLFWI-EKF method, which is not only far more efficient than iterative or high-order EKF, but also impose additional spatial constraints on predicting the *a priori* velocity model. Nevertheless, because the amount of storage and computational cost of TLFWI-EKF is proportional to the squared number of discretized model parameters, it is still very challenging

for the TLFWI-EKF method to solve large-scale seismic inverse problems, especially 3-D cases.

On the other hand, researchers proposed to incorporate the ensemble KF (EnKF) into FWI for uncertainty estimation (Jin *et al.* 2008; Gineste & Eidsvik 2017; Thurin *et al.* 2019). The EnKF method proposes to replace the full-rank covariance matrix by its low-rank approximation. Based on Monte Carlo sampling theory, the estimation error is proportional to $1/\sqrt{N}$, which means that the ensemble size N increases, the estimation error decreases (Evensen 2003). Besides, the sampling bias in covariance matrices becomes significant when the ensemble size N is smaller than the number of observations (Kepert 2004). Therefore, maintaining small statistical errors requires a large ensemble size for EnKF, although the ensemble size may possibly be reduced by additional optimal methods (Evensen 2003; Aanonsen *et al.* 2009). So far, no attempt has been made to use EnKF for 3-D FWI.

Li *et al.* (2014) introduced a hierarchical matrix (\mathcal{H}^2 -matrices) powered Kalman filter (HiKF) that produces results as accurate as original KF but with a dramatically reduced computational and storage cost. This HiKF method approaches the dense covariance matrices with \mathcal{H}^2 algebra from a data-sparse (low rank) representation, then applies fast multipole methods (Fong & Darve 2009) to accelerate the multiplication referred to those dense covariance matrices in KF. The accuracy and efficiency of the \mathcal{H}^2 -matrices approach was demonstrated in linear large-scale geostatistical inverse problems (Ambikasaran *et al.* 2013; Li *et al.* 2014).

To solve large-scale nonlinear seismic TLFWI problems, here we extend the HiKF method into \mathcal{H}^2 -matrices powered extended Kalman filter (HiEKF). We reformulate the TLFWI in the framework of HiEKF (hereafter referred to as TLFWI-HiEKF) to process dense time-lapse monitoring data. By reformulating, TLFWI instead of the random walk forecast model (Quan & Harris 2008; Li *et al.* 2014) is used to provide a good *a priori* velocity model by imposing spatial constraints on velocity changes while data assimilation HiEKF technique is designed for processing data evolution to maintain temporal constraints on high rate of time-lapse velocity variations. Moreover, HiEKF uses the \mathcal{H}^2 -matrices approach to explore the low rank representation of the covariance matrices at an affordable storage and computational cost in large-scale problems. Another important aspect of this reformulation is to seek the best estimation of time-lapse velocity changes and meanwhile provide the corresponding uncertainty maps of those estimations.

In the following we first illustrate the framework of TLFWI-HiEKF. Then, we present 2-D Frio-II and 3-D Cranfield CO₂ sequestration synthetic examples to prove the validity and applicability of the proposed method on dealing with time-lapse seismic monitoring data. Discussions and conclusions associated to the proposed scheme are made finally.

2 THEORY

2.1 Time-lapse full waveform inversion

FWI of seismic data can be expressed as the minimization of a cost function of the difference between the modelled data u and the observed data d as

$$E = \frac{1}{2} \|d - u(\mathbf{m})\|^2, \quad (1)$$

where \mathbf{m} is the model parameter (e.g. P - or S -wave velocity, density) to be recovered. In this paper, we focus on P -wave velocity, hence we replace \mathbf{m} with \mathbf{v} hereafter. $u(\mathbf{v})$ is the synthetic data

generated by numerically solving the acoustic wave equation with an initial velocity model \mathbf{v} . It can be expressed in the frequency domain:

$$u(\mathbf{v}) = G(\mathbf{v}) S(\omega) \delta(\mathbf{x} - \mathbf{x}_r), \quad (2)$$

where G is the Green's function, S is the source function, ω is the angular frequency, and \mathbf{x}_r is the position of receivers. Since FWI is a strongly nonlinear optimization problem, the initial velocity model (*a priori* velocity model) needs to be in the vicinity of the solution for the global minimum. Assuming the initial velocity model is adequately defined, iterative gradient-based searching methods can be used to solve this optimization problem, (Mora 1989; Vireux & Operto 2009; Huang *et al.* 2015), mathematically written as

$$\mathbf{v}_{i+1} = \mathbf{v}_i + \alpha_i \mathbf{g}_i, \quad (3)$$

where α_i is the step length. \mathbf{g}_i is the gradient of the cost function E with respect to \mathbf{v} at the i th iteration and can be derived from the multiplication of forward wavefield and backpropagated data residual wavefield in the frequency domain using the adjoint-state method (Plessix 2006). We then iteratively update the velocity model \mathbf{v} using eq. (3) until the cost function reaches its minimum. The final velocity model \mathbf{v} is the output of FWI.

The purpose of TLFWI is to reveal changes of \mathbf{v} over time. TLFWI often involves two or more FWIs on baseline and monitoring data. The time-lapse images of changes are obtained by subtracting the background velocity model from the inverted velocity models obtained from FWIs of monitoring data. We perform the TLFWI in the frequency domain, in which a frequency group strategy is applied to perform from low- to high-frequency group (Sirgue & Pratt 2004).

2.2 Extended Kalman filter for time-lapse seismics

In data assimilation, EKF uses the state \mathbf{x} and measurements \mathbf{z} to describe the dynamic evolution of a system according to

$$\mathbf{x}_{k+1} = F\mathbf{x}_k + \mathbf{w}_{k+1} \quad (4)$$

$$\mathbf{z}_{k+1} = h(\mathbf{x}_{k+1}) + \mathbf{e}_{k+1}, \quad (5)$$

where F is the state transition matrix, h is a measurement function which related measurements \mathbf{z} to the state \mathbf{x} , $\mathbf{w}_{k+1} \sim N(0, Q_{k+1})$ and $\mathbf{e}_{k+1} \sim N(0, R_{k+1})$ are the state error and measurement noise at time lapse $k+1$ with zero mean and known covariances, respectively. In the context of seismic forward modelling, eq. (5) can be replaced by eq. (2), and the commonly used random walk model (Vauhkonen *et al.* 1998) is applied to describe the velocity model evolution. Then, the time-lapse seismics can be expressed as a dynamic evolution problem using the following equations

$$\mathbf{v}_{k+1} = \mathbf{v}_k + \mathbf{w}_{k+1} \quad (6)$$

$$\mathbf{d}_{k+1} = G(\mathbf{v}_{k+1}) S(\omega) \delta(\mathbf{x} - \mathbf{x}_r) + \mathbf{e}_{k+1}, \quad (7)$$

where \mathbf{v} is the subsurface velocity model and \mathbf{d} is the observed seismic data in the frequency domain, where $\mathbf{w}_{k+1} \sim N(0, Q_{k+1})$ and $\mathbf{e}_{k+1} \sim N(0, R_{k+1})$ are the velocity model error and observed data noise at time lapse $k+1$, respectively.

To solve eqs (6) and (7), the classical EKF schemes are formulated in two steps: predict and update. In the prediction step, EKF

predicts a velocity model at the current time lapse using information from the previous time lapse. Therefore, the estimation from the prediction step is an *a priori* estimation. Then in the update step, EKF combines the *a priori* velocity model and the observed data at the current time lapse together to update the velocity model. This update can be treated as a weighted average between the *a priori* and the measured models. Hence, the estimation from the update step is an *a posteriori* estimation. The formulation of the two steps are described as follow:

predict

$$\hat{\mathbf{v}}_{k+1}^- = \hat{\mathbf{v}}_k \quad (8)$$

$$\mathbf{P}_{k+1}^- = \mathbf{P}_k + \mathbf{Q}_{k+1} \quad (9)$$

update

$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1}^- \mathbf{H}_{k+1}^T (\mathbf{H}_{k+1} \mathbf{P}_{k+1}^- \mathbf{H}_{k+1}^T + \mathbf{R}_{k+1})^{-1} \quad (10)$$

$$\hat{\mathbf{v}}_{k+1} = \hat{\mathbf{v}}_{k+1}^- + \mathbf{K}_{k+1} (\mathbf{d}_{k+1} - G(\hat{\mathbf{v}}_{k+1}^-) S(\omega) \delta(\mathbf{x} - \mathbf{x}_r)) \quad (11)$$

$$\mathbf{P}_{k+1} = (\mathbf{I} - \mathbf{K}_{k+1} \mathbf{H}_{k+1}) \mathbf{P}_{k+1}^- \quad (12)$$

where $\hat{\mathbf{v}}_{k+1}^-$ and $\hat{\mathbf{v}}_{k+1}$ are the *a priori* and *a posteriori* estimation of velocity models. \mathbf{P}_{k+1}^- and \mathbf{P}_{k+1} are the *a priori* and *a posteriori* covariance matrices at time lapse $k+1$ with the dimension of $m \times m$, where m is the total number of discretized subsurface model parameters. \mathbf{K}_{k+1} is the Kalman gain which acts as the weights between the *a priori* and measured model, and \mathbf{H}_{k+1} is the derivative of wavefield at the locations of observations with respect to velocity \mathbf{v} , which is the Fréchet kernel matrix. Based on the Born approximation, we can derive the explicit expression of the Fréchet kernel function. Then, we calculate the Fréchet kernel matrix \mathbf{H} using the scattering-integral approach method (Liu *et al.* 2015). \mathbf{Q}_{k+1} is the model error covariance with the dimension of $m \times m$, and \mathbf{R}_{k+1} is the observation error covariance with the dimension of $n \times n$, where n is the number of receivers (measurements). \mathbf{R} is always set to be a diagonal matrix with a constant value equal to the noise variance in the observed seismic data.

From eqs (8)–(12), it is clear that the covariance matrices \mathbf{P} and \mathbf{Q} need be stored and calculated at each step. The amount of storage for the covariance matrices \mathbf{P} and \mathbf{Q} is proportional to the squared number (m^2) of discretized model parameters. Taking a 3-D velocity model with discretized grid size of $100 \times 100 \times 100$ for example, the number of model parameters m equals to 10^6 . Hence, the size of matrices \mathbf{P} or \mathbf{Q} is $10^6 \times 10^6$ and to store \mathbf{P} or \mathbf{Q} in memory takes up to about 4 TB. Such a high memory cost is still unaffordable for most of current computer clusters. Therefore, the EKF algorithm is only suitable to small-scale 2-D problems (Eikrem *et al.* 2019; Huang & Zhu 2019).

2.3 \mathcal{H}^2 -matrices powered extended Kalman filter

To avoid explicitly storing and calculating the covariance matrices \mathbf{P} and \mathbf{Q} , following the \mathcal{H}^2 -matrices approach proposed by Li *et al.* (2014), we define the *a priori* and *a posteriori* cross-covariance matrices \mathbf{C}^- and \mathbf{C} as

$$\mathbf{C}_{(k+1)}^- = \mathbf{P}_{(k+1)}^- \mathbf{H}_{(k+1)}^T \text{ and } \mathbf{C}_{(k+1)} = \mathbf{P}_{(k+1)} \mathbf{H}_{(k+1)}^T, \quad (13)$$

and the model error cross-covariance matrix \mathbf{A}

$$\mathbf{A}_{k+1} = \mathbf{Q}_{k+1} \mathbf{H}_{k+1}^T. \quad (14)$$

Then, we multiply \mathbf{H}_{k+1}^T to eq. (9) to obtain

$$\mathbf{C}_{k+1}^- = \mathbf{P}_k \mathbf{H}_{k+1}^T + \mathbf{A}_{k+1}. \quad (15)$$

Assume $\mathbf{H}_{k+1}^T \cong \mathbf{H}_k^T$ given that the seismic velocity would not change rapidly from one time lapse to the next time lapse. Based on our numerical tests, this replacement $\mathbf{H}_{k+1}^T \cong \mathbf{H}_k^T$ barely influences the results. Hence, eq. (15) is rewritten as

$$\mathbf{C}_{k+1}^- \cong \mathbf{P}_k \mathbf{H}_k^T + \mathbf{A}_{k+1} = \mathbf{C}_k + \mathbf{A}_{k+1}. \quad (16)$$

Substituting eqs (13) and (14) into eqs (10)–(12), combining with eqs (8) and (16), yields the \mathcal{H}^2 -matrices powered extended Kalman filter (we call it HiEKF hereafter) equations as follows:

predict

$$\hat{\mathbf{v}}_{k+1}^- = \hat{\mathbf{v}}_k \quad (17)$$

$$\mathbf{C}_{k+1}^- = \mathbf{C}_k + \mathbf{A}_{k+1} \quad (18)$$

update

$$\mathbf{K}_{k+1} = \mathbf{C}_{k+1}^- (\mathbf{H}_{k+1} \mathbf{C}_{k+1}^- + \mathbf{R}_{k+1})^{-1} \quad (19)$$

$$\hat{\mathbf{v}}_{k+1} = \hat{\mathbf{v}}_{k+1}^- + \mathbf{K}_{k+1} (\mathbf{d}_{k+1} - G(\hat{\mathbf{v}}_{k+1}^-) S(\omega) \delta(\mathbf{x} - \mathbf{x}_r)) \quad (20)$$

$$\mathbf{C}_{k+1} = (\mathbf{I} - \mathbf{K}_{k+1} \mathbf{H}_{k+1}) \mathbf{C}_{k+1}^-. \quad (21)$$

According to the rules of multiplication between two matrices, the posterior variance δ^2 that is the diagonal of covariance matrix \mathbf{P} in eq. (12) can be calculated using the following equation:

$$\delta_{k+1}^2 = \delta_{k+1}^{-2} - \sum_{j=1}^n (K_{k+1})_{ij} (\mathbf{C}_{k+1}^-)_{ij}, \quad (22)$$

where δ_{k+1}^{-2} is the diagonal of covariance matrix \mathbf{P}_{k+1}^- . In the new formulation, rather than storing and updating the covariance matrices \mathbf{P} and \mathbf{Q} with the size of $m \times m$ in the original EKF (eqs 8–12), we instead store and update the cross-covariance matrices \mathbf{C} and \mathbf{A} with the size of $m \times n$ (where $n \ll m$).

To calculate the cross-covariance \mathbf{A} in eq. (14), we adopt a \mathcal{H}^2 -matrices approach (Li *et al.* 2014) to avoid an explicit expression of the model error covariance \mathbf{Q} . The \mathcal{H}^2 -matrices approach is a data-sparse representation method which relies on the \mathcal{H}^2 -matrices structure [We refer readers to Ambikasaran *et al.* (2013) and Fong & Darve (2009) for details of \mathcal{H}^2 -matrix approach]. By observation, we first rewrite the cross-covariance \mathbf{A} using matrix decomposition as

$$\mathbf{A} = \mathbf{Q} \mathbf{H}^T = [\mathbf{Q} h^{(1)} \mathbf{Q} h^{(2)} \dots \mathbf{Q} h^{(n)}], \quad (23)$$

where $h^{(k)}$, $k \in \{1, 2, \dots, n\}$ is the column vector of matrix \mathbf{H}^T , and n is the number of observations. Since \mathbf{Q} represents the interaction between well-separated clusters of gridpoints, the matrix-vector products can be written as

$$\mathbf{Q} h_i^{(k)} = \sum_{j=1}^m K(x_i, x_j) h_j^{(k)} \quad (24)$$

where $i \in \{1, 2, \dots, m\}$, m is the number of model parameters, and $K(x_i, x_j)$ is the interaction between gridpoints x_i and x_j . Solving

eq. (24), we apply the fast multipole method (Fong & Darve 2009) to compute and store the matrix-vector products at a reduced storage and computational cost by using Chebyshev interpolation to obtain the low-rank factorizations of $K(x_i, x_j)$. And the rank is determined by the desired accuracy.

Therefore, HiEKF avoids the explicit usage of covariance matrices \mathbf{P} and \mathbf{Q} by replacing them with cross-covariance \mathbf{C} and \mathbf{A} . The maximum matrices storage size reduces from $m \times m$ to $m \times n$ and the total computational cost decreases from $\mathcal{O}(nm^2)$ to $\mathcal{O}(mn^2)$. Since $m \gg n$ is often true in large-scale geophysical problems, the storage and computational cost of HiEKF are dramatically reduced compared to original EKF.

C++ software packages related to the fast multipole method used in this paper, including BBFMM2D (available at <https://github.com/sivaramambikasaran/BBFMM2D>) and PBBFMM3D (available at <https://github.com/ruoxi-wang/PBBFMM3D>), can be operated as a black box. To handle the calculation of matrices, the Eigen package (available at http://eigen.tuxfamily.org/index.php?title=Main_Page) is used in this paper.

2.4 TLFWI plus \mathcal{H}^2 -matrices powered extended Kalman filter

Due to the first-order Taylor expansion approach in EKF, HiEKF is only suitable to weak nonlinearity problems. To extend the applicability of HiEKF to nonlinear problems, following our previous study (Huang & Zhu 2019), we modify eq. (6) by adding a velocity perturbation term as follows:

$$\mathbf{v}_{k+1} = \mathbf{v}_k + \delta \mathbf{v}_{k+1} + \mathbf{w}_{k+1}, \quad (25)$$

where $\delta \mathbf{v}_{k+1}$ is the velocity change produced from TLFWI at time lapse $k+1$. By this modification, rather than the random walk model in eq. (6), the high spatial resolution velocity change $\delta \mathbf{v}_{k+1}$ generated from TLFWI can help HiEKF to predict a good *a priori* velocity model, which reduces the nonlinearity of the targeted problems and is likely to guarantee the weak nonlinearity requirement of the HiEKF. To solve the new dynamic evolution problem defined by eqs (25) and (7), we only need to revise the eq. (17) in the HiEKF formulation (eqs 13 and 14, and 17–22) as

$$\hat{\mathbf{v}}_{k+1}^- = \hat{\mathbf{v}}_k + \delta \mathbf{v}_{k+1}. \quad (26)$$

Together with eqs (13), (14) and (18)–(22), we name them as the TLFWI–HiEKF algorithm.

Fig. 1 shows the workflow of the TLFWI–HiEKF method, which includes three main parts: TLFWI, prediction step and update step. In the first step, TLFWI is conducted to calculate the velocity changes $\delta \mathbf{v}_{k+1}$ at time lapse $k+1$. Then, the calculated velocity changes $\delta \mathbf{v}_{k+1}$ are used to predict *a priori* velocity model $\hat{\mathbf{v}}_{k+1}^-$ (eq. 26) in the prediction step. The update step of TLFWI–HiEKF also includes two parts: one part is treated as an additional inversion (eq. 20) which uses the *a priori* velocity model $\hat{\mathbf{v}}_{k+1}^-$ (eq. 26) and the Kalman gain (eq. 19) to weight the velocity model on the basis of seismic data residuals. The other is the quantification of the variance (uncertainty) of the inverted results (eqs 21 and 22). Given that TLFWI instead of the random walk model can invert the velocity changes $\delta \mathbf{v}_{k+1}$ at time lapse $k+1$, leading to a better prediction of velocity $\hat{\mathbf{v}}_{k+1}^-$, the TLFWI–HiEKF not only reduces the nonlinearity in the subsequent update step but also speed up the convergence of the inversion. Therefore, the proposed TLFWI–HiEKF method integrates the advantages of TLFWI and HiEKF together, which not only provide

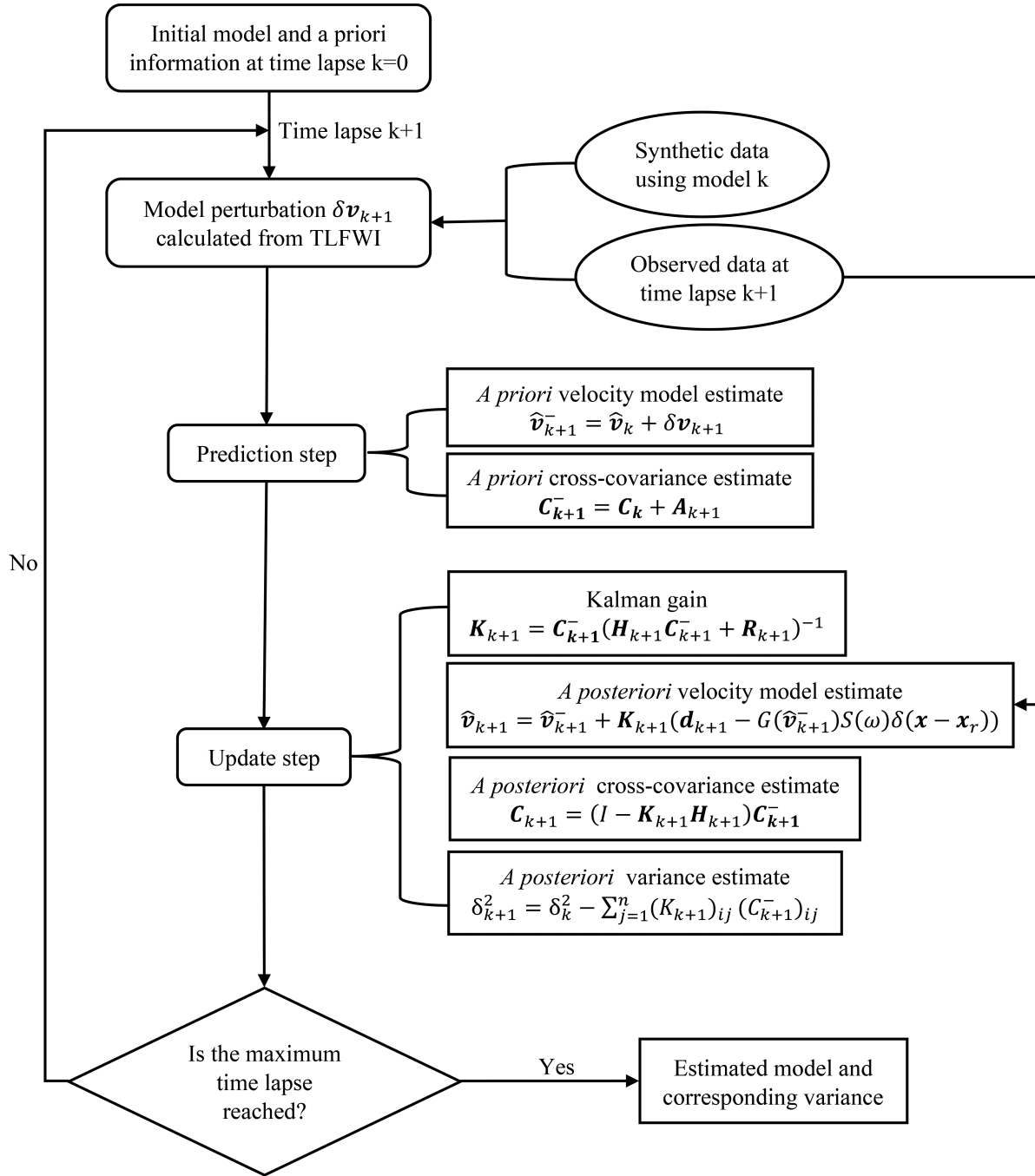


Figure 1. Workflow for the TLFWI-HiEKF method.

high-resolution subsurface velocity models at each time lapse, but also quantifies the variance (uncertainty) of the inverted velocity results.

3 SYNTHETIC EXAMPLES

This section will present synthetic tests of the scheme of TLFWI-HiEKF in two different time-lapse seismic monitoring data sets collected in two CO₂ sequestration sites: Frio-II and Cranfield. In Frio-II, dense 2-D crosswell time-lapse seismic surveys were carried out, while relatively sparse 3-D surface time-lapse seismic surveys were conducted in Cranfield.

3.1 2-D Frio-II model

3.1.1 Site background

The Frio-II CO₂ pilot was a small-scale injection of supercritical CO₂ into a high permeable reservoir at the depth of 1650 m to test geological storage in saline aquifers (Daley *et al.* 2007). The baseline *P*-wave velocity model (i.e. a velocity map before CO₂ injection) of the Frio-II site is derived from 2-D extrapolation of logs acquired in the Frio injection well with a local dip determined from gamma ties (Fig. 2a). The velocity values range from 2650 to 2765 m s⁻¹ with a background velocity of 2700 m s⁻¹ outside the Blue Sand. Fig. 2(b) shows the simulated

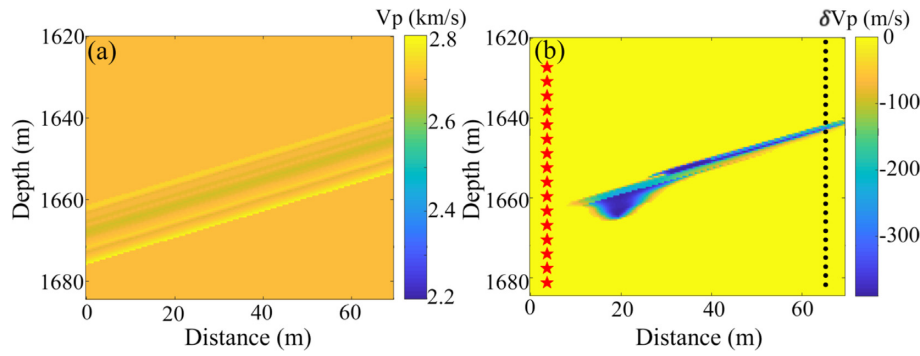


Figure 2. (a) Baseline velocity model for Frio-II tests; and (b) time-lapse velocity changes after 120 hr CO₂ injection (red stars indicate sources and black dots stand for receivers).

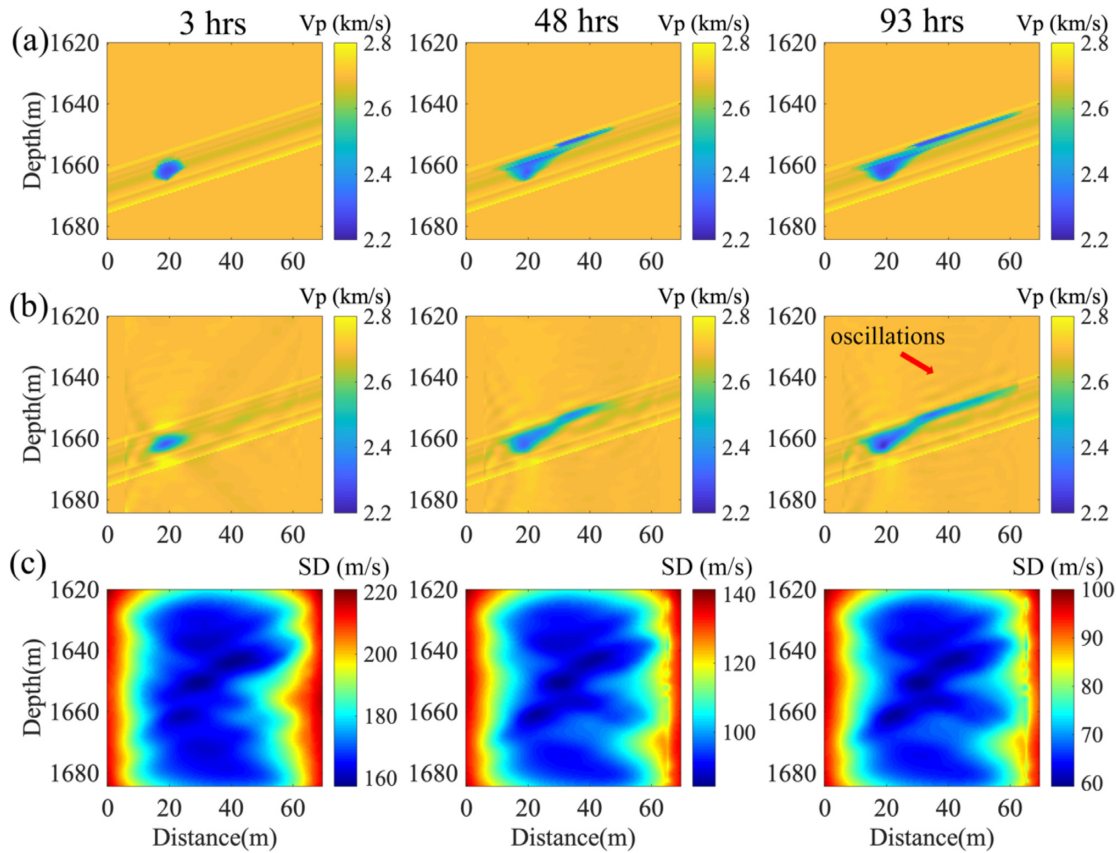


Figure 3. (a) True time-lapse velocity models, (b) inverted results and (c) the corresponding SD (in m s⁻¹) distribution at the injection time of 3, 48 and 93 hr, respectively.

CO₂-induced velocity reduction after five-days of injection in Frio-II site. The seismic velocity changes due to CO₂ plumes are predicted by rock physics modelling of velocity and saturation from a 3-D multiphase flow model (Daley *et al.* 2011). We have totally 41 time-lapse seismic models with 3-hr intervals. Fig. 3(a) shows three time-lapse *P*-wave velocity (v_p) models at 3, 48 and 93 hr.

3.1.2 Time-lapse seismic data acquisition and inversion parameter selection

To mimic the field survey in our test, we simulate synthetic time-lapse seismic crosswell data every 3 hr. We use 32 time-lapse models

for forward simulations. In each time-lapse experiment, 32 active sources are deployed at the injection well and 140 receivers are deployed in the observation well (Fig. 2b). The observed seismic data is simulated by a finite-difference scheme with eighth-order accuracy in space and second order in time. Gaussian white noise is added in the observed seismic data, and the signal-to-noise ratio is defined as

$$\text{SNR} = 10 \log_{10} \frac{\|d_2^2\|}{\|\sigma_2^2\|}, \quad (27)$$

where d is the observed data, and σ^2 is the noise variance. The signal-to-noise ratio of observed data is 25 after adding noise. The source function is a Ricker wavelet with the dominant frequency

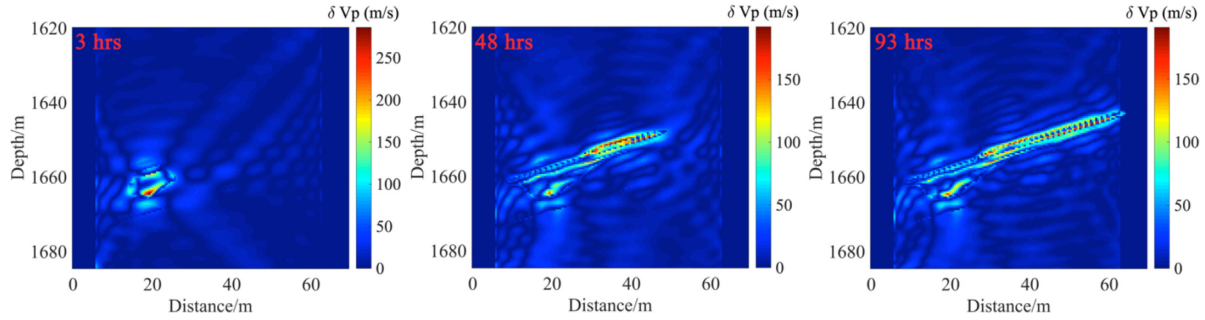


Figure 4. The velocity model error distribution at the injection time of 3, 48 and 93 hr, respectively.

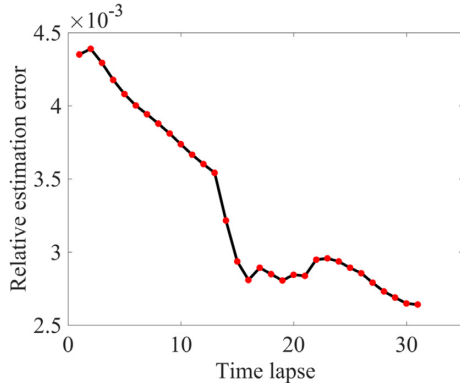


Figure 5. Estimated model error relative to the true solution verse time lapses.

of 250 Hz. The grid size is 0.45 m × 0.45 m and the model size is about 69.3 m × 64.35 m.

During the test, we use the baseline velocity model as the initial model for the proposed TLFWI–HiEKF method at the first time lapse. The frequency selection for TLFWI is from 13.5 to 313.5 Hz with an interval of 12.5 Hz. These frequencies are divided into three groups: the first group is ranging from 13.5 to 151 Hz, the second is [151 Hz, 238.5 Hz] and the third is [238.5 Hz, 313.5 Hz]. The initial value of δ^2 is set as 90 000 (m · s⁻¹)² assuming we have no prior information about the model. The noise matrix \mathbf{R} in the seismic data is set diagonal with a constant value equal to the noise variance. The model error covariance matrix of \mathbf{Q} is constructed from an exponential function (Li *et al.* 2014)

$$\mathbf{Q}(r) = \exp\left(-\sqrt{\frac{r}{L}}\right), \quad (28)$$

where r is the distance between two points, and L is the correlation length.

3.1.3 TLFWI–HiEKF inversion results

After processing simulated time-lapse seismic data in 32 time frames using TLFWI–HiEKF, we reconstruct the corresponding maps of velocity changes induced by CO₂ injection. Fig. 3(b) shows the inverted results at three time lapses (3, 48 and 93 hr). Comparing the inverted results to true velocity models, the primary velocity changes (blue zones in Figs 3a and b) induced by CO₂ injection are captured with a comparable resolution. The perturbation results indicate the spatiotemporal changes of the injected CO₂ gas. We note some oscillations (pointed by arrow in Fig. 3b) in the inverted results. The oscillations are likely caused by the limited frequency

band we used for inversion. It is possible that we can broaden the inversion frequency band to improve the quality of inverted results, but large number of frequencies would cause extra computational cost.

Fig. 3(c) shows the estimated standard deviation (SD) δ (the square root of diagonal of the posterior covariance matrix) distributions in which the SD value represents the degree of uncertainty, that is small SD indicates low uncertainty, vice versa. We found that the uncertainty distribution is somehow consistent to the illumination distribution of the crosswell geometry of seismic surveys. Overall, the estimated velocity between the two wells have less uncertainty compared to the remainder of the model. We also observed that the SD is decreasing over lapsed time. It indicates that with more time-lapse seismic data, the accuracy of the recovered model is increasing, which proves that our method can impose temporal constraints on velocity changes and update the model well over time. For comparison, the model error distributions are also plotted in Fig. 4. The error maps show the absolute value differences between true and inverted models while the uncertainty (SD) maps show how much we can trust on our inverted results. We can see that the values of model error and the SD are at the same level.

To measure the accuracy of the inverted result, we calculate relative model error e_t by

$$e_t = \frac{\|\mathbf{v}_{\text{est}} - \mathbf{v}_{\text{true}}\|_2}{\|\mathbf{v}_{\text{true}}\|_2}. \quad (29)$$

The \mathbf{v}_{est} and \mathbf{v}_{true} are the inverted and true velocity models at time lapse t . Fig. 5 plots the relative estimation error as a function of time. As shown, the predictions become increasingly accurate over time as more data have been assimilated, which is consistent with the decreasing trend of the SD values in Fig. 3(c).

To measure the accuracy of the results in terms of data matching, we compare seismic shot gather using inverted and true velocity models in Fig. 6. We first generate two synthetic seismograms using the final velocity model (blue wiggles in Fig. 6b) and its corresponding initial model (black wiggles in Fig. 6a) at the time lapse of 3 hr, then compare two synthetic seismograms with observed seismic waveforms (red wiggles in Fig. 6), respectively. We can see that the synthetic data with the final velocity model matches the observed data much better than when using the initial model. We repeat comparisons at other two other time lapses (48 and 93 hr) in Figs 6(c)–(f). By comparing the data matching over time (from left to right in Fig. 6), it seems that the residual between the synthetic and observed data is decreasing over time lapses, which confirms that the TLFWI–HiEKF method is able to increase the accuracy of results over time by assimilating more observed time-lapse data with effective temporal constraints.

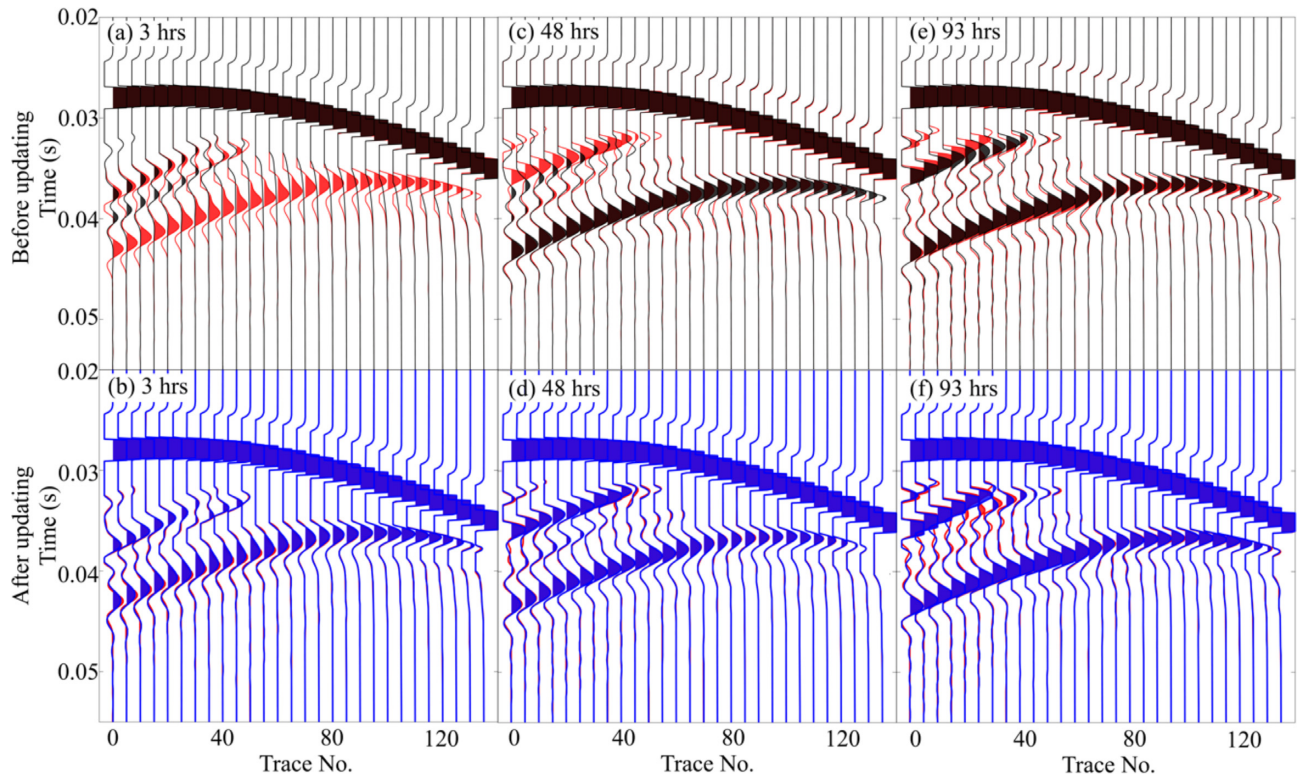


Figure 6. Waveform comparisons before inversion (a), (c) and (e) at the injection time of 3, 48 and 93 hr. Waveform comparisons after inversion (b), (d) and (f) at the injection time of 3, 48 and 93 hr (red stands for the observed data, black stands for the predicted data before inversion and blue stands for the predicted data after inversion at different time lapses).

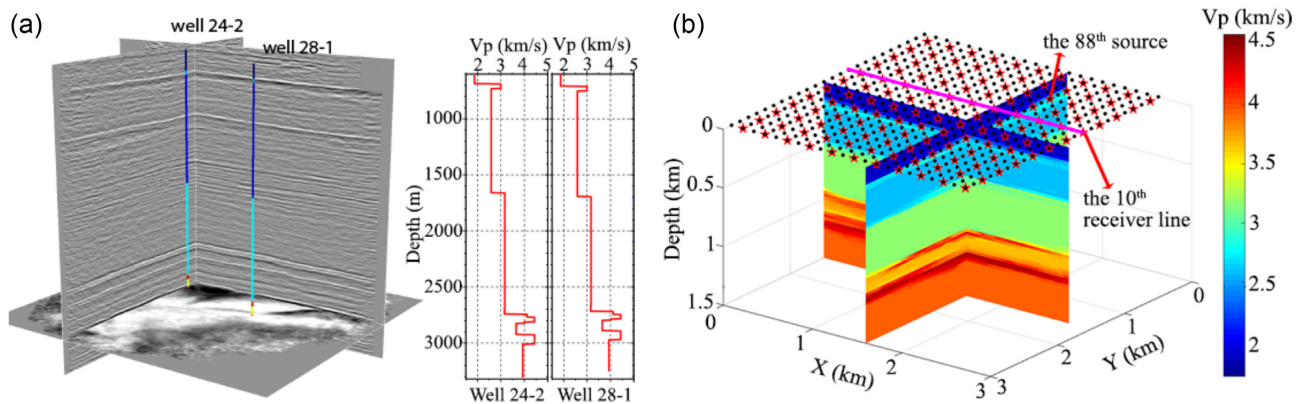


Figure 7. (a) 3-D seismic migration data and two P -wave velocity logs for Cranfield; and (b) 3-D baseline velocity model with arrays of seismic sources and receivers (black dots represent receivers and red stars stand for sources).

3.2 3-D Cranfield model

In this section, we demonstrate the applicability of the proposed TLFWI-HiEKF method in the 3-D Cranfield CO_2 injection time-lapse models.

3.2.1 Site background

The Cranfield site is located at Denbury Resources Cranfield Field in southwest Mississippi. It was discovered in 1946 and abandoned in 1965 at the end of primary oil and gas production. The reservoir was redeveloped by Denbury Onshore LLC and brought on as a CO_2 -EOR field in 2008 using CO_2 transported via pipeline

from a natural accumulation near Jackson, MS, USA. The main CO_2 injection test was conducted in the water leg of the structure beyond producible hydrocarbon accumulations (Ajo-Franklin *et al.* 2013). The baseline seismic P -wave velocity model is built from the 3-D baseline seismic migration section and two well logs (Fig. 7a) based on seismic structural and stratigraphic features (Wu 2017). To meet the memory limits in our high-performance computing cluster, we modify the size of the original velocity model ($2.75 \text{ km} \times 3.0 \text{ km} \times 1.5 \text{ km}$) as shown in Fig. 7(b). Six time-lapse velocity models are manually constructed (Wu 2017) and shown in Fig. 9, where the layered structures are in the colour scale and black blocks drawn by isosurface indicate the velocity changes induced by CO_2 plumes.

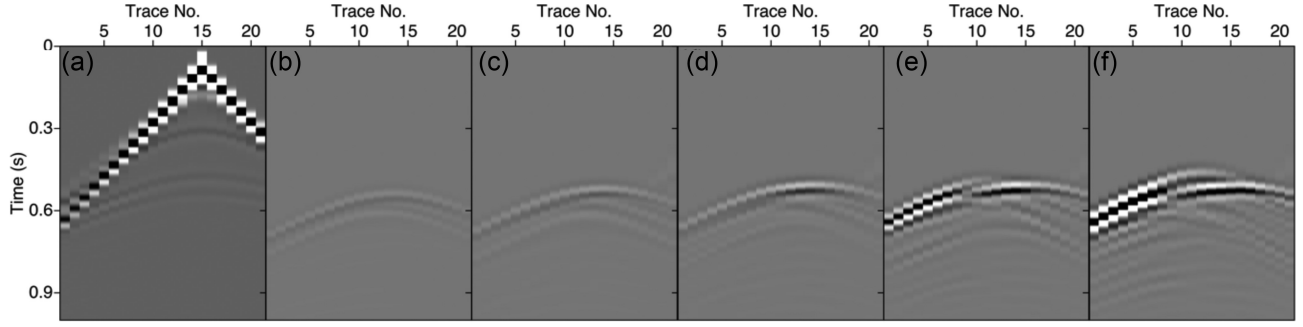


Figure 8. A time-lapse shot gather profile from a selected receiver line (the 10th receiver line in Fig. 6b). (a) Baseline data. (b)–(f) Data residual by subtracting baseline data (a) from five time-lapse data. (b) First time lapse; (c) second time lapse; (d) third time lapse; (e) fourth time lapse and (f) fifth time lapse. Seismic reflections are caused by injected CO₂.

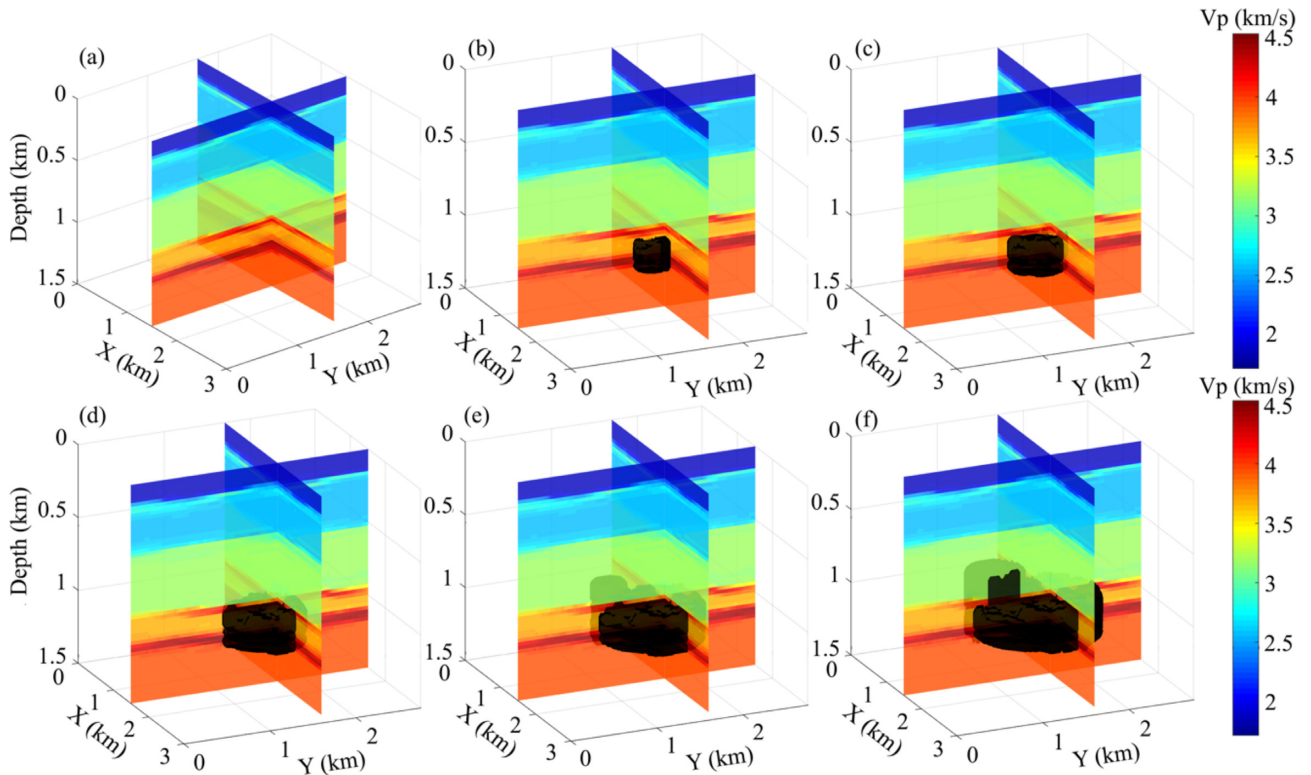


Figure 9. True velocity models at different time lapses with an isosurface plot of time-lapse velocity changes (black blocks) induced by CO₂ plumes. (a) Baseline model; (b) first time lapse; (c) second time lapse; (d) third time lapse; (e) fourth time lapse and (f) fifth time lapse.

3.2.2 3-D time-lapse seismic data acquisition and inversion parameter selection

Similar to field surveys, we adopt a surface seismic acquisition survey to collect 3-D time-lapse data. We deploy 24 receiver lines along the Y -direction with an interval of 125 m and 22 receivers are evenly distributed in each receiver line from 0 to 2.75 km along the X -direction (black dots in Fig. 7b). At each time lapse, we design 12 shot lines along the Y -direction with an interval of 250 m with 11 shots evenly distributed along each shot line from 0 to 2.75 km in X -direction on the surface. In total, 132 sources are fired in each time frame (red stars in Fig. 7b). A 3-D finite-difference scheme with eighth-order accuracy in space and second-order in time is used to generate the observed seismic data using the true velocity models in each time lapse. Again, we contaminate the synthetic observed seismic data with Gaussian noise (SNR = 25). The source

function is a Ricker wavelet with the dominant frequency of 15 Hz. The grid size is 25 m \times 25 m \times 25 m. Baseline shot gather data and time-lapse data at different time lapses are shown in Fig. 8 and time-lapse waveform differences are caused by the CO₂ injection (Figs 8b–f). TLFWI–HiEKF has been applied from 1 to 25 Hz with an interval of 1 Hz. Totally two frequency groups are used for TLFWI in this test: one is from 1 to 12 Hz and the other is from 12 to 25 Hz. The initial settings of δ^2 , the noise matrix \mathbf{R} and the cross-covariance \mathbf{A} are identical to those in the Frio-II test.

3.2.3 TLFWI–HiEKF inversion results

Comparing the ground truth models in Fig. 9, with the results shown in Fig. 10, where black blocks represent inverted time-lapse velocity

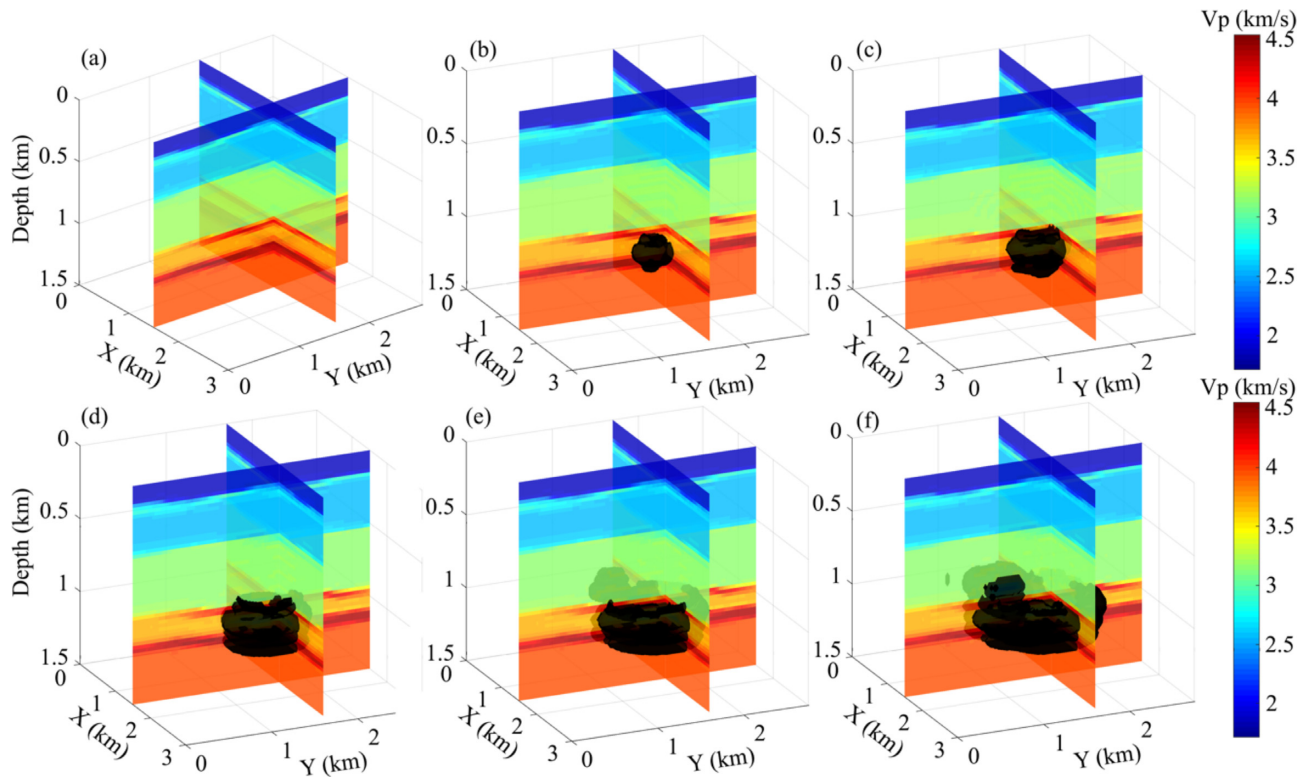


Figure 10. Inverted velocity models at different time lapses with an isosurface plot of time-lapse velocity changes (black blocks) induced by CO₂ plumes. (a) Baseline model; (b) first time lapse; (c) second time lapse; (d) third time lapse; (e) fourth time lapse and (f) fifth time lapse.

changes at different time lapses, time-lapse velocity changes (the CO₂ plume in the subsurface) are well captured by 3-D TLFWI–HiEKF. Fig. 11 compares the true velocity models (left-hand column) with the inverted results (middle column) from the first time lapse to the fifth time lapse. Similar observations are that the inverted results successfully capture the velocity changes induced by the CO₂ plume. In addition, the corresponding SD map (the square root of diagonal of the posterior covariance matrix) at each time lapse is shown in Fig. 11 (right-hand column). It is clear that the SD maps (time lapse 1–5) are all dominated by surface seismic wave illumination, and the uncertainty increases toward the depth where poor illumination is expected. From the model error distributions in Fig. 12, we can see the trend of decreasing values over time. To compare the SD over time, we sum the SD matrix and plot these values in Fig. 13. We found that the SD values decrease over time lapse, similar to 2-D Frio tests, which indicate that TLFWI–HiEKF is able to continuously update the results with more incoming time-lapse seismic data.

Fig. 14 presents the comparison between the true and inverted CO₂-induced velocity perturbations at the fifth time lapse. The 3-D velocity changes induced by CO₂ injection are well captured, showing the successful applicability of the TLFWI–HiEKF method in a 3-D time-lapse seismic experiment.

3.3 Memory usage and computational cost comparison in 2-D and 3-D examples

In this section, we compare the memory usage and computational cost of the TLFWI–HiEKF method and other KF methods including, original EKF- and EnKF-based TLFWI methods. Since the procedure of TLFWI is nearly the same in all of those methods, we

first compare the memory usage of EKF, EnKF and HiEKF in the aforementioned 2-D and 3-D examples. For EnKF, the accuracy of results is proportional to the ensemble size N , and N should be larger than the size of observations n to reduce the sampling bias in covariance matrices (Kepert 2004; Li *et al.* 2014; Thurin *et al.* 2019). To compare the memory cost of three types of KF methods (EKF, EnKF and HiEKF) with maintaining the same accuracy, we take large N as an example for computing the memory cost of EnKF, for example 600 in the Frio 2-D example and 6000 the Cranfield 3-D model. Table 1 lists the approximately memory requirements. In 2-D case, the number of model parameters m is 155×144 and the number of observations n is 140; in 3-D case, the m is $111 \times 121 \times 61$ and n is 132. We can see that the HiEKF method requires the minimum memory storage. For 3-D, the original EKF requires as much as 5 TB storage cost which is unaffordable in most high-performance computing architectures. HiEKF only requires approximately 3.5 GB storage cost which is over 1000 times less than original EKF and over 10 times less than EnKF. The total computational cost for original EKF is $\mathcal{O}(m^2)$ because it explicitly stores and calculates the covariance matrix of size $m \times m$. Since HiEKF operates on the cross-covariance matrix of size $m \times n$ where $n \ll m$, the overall computational cost is $\mathcal{O}(m)$. The storage and computational costs of EnKF also scale linearly with the number of parameters m . However, for each ensemble in EnKF, FWI should be implemented independently (Thurin *et al.* 2019), which means N ensemble size requires N times of FWI while HiEKF only needs one at each time lapse. That is why EnKF is computationally more expensive than HiEKF even if the ensemble size is equal to the number of receivers. In summary, due to its high accuracy and low memory usage and computational cost, HiEKF is superior to other KF methods for 3-D time-lapse seismic problems.

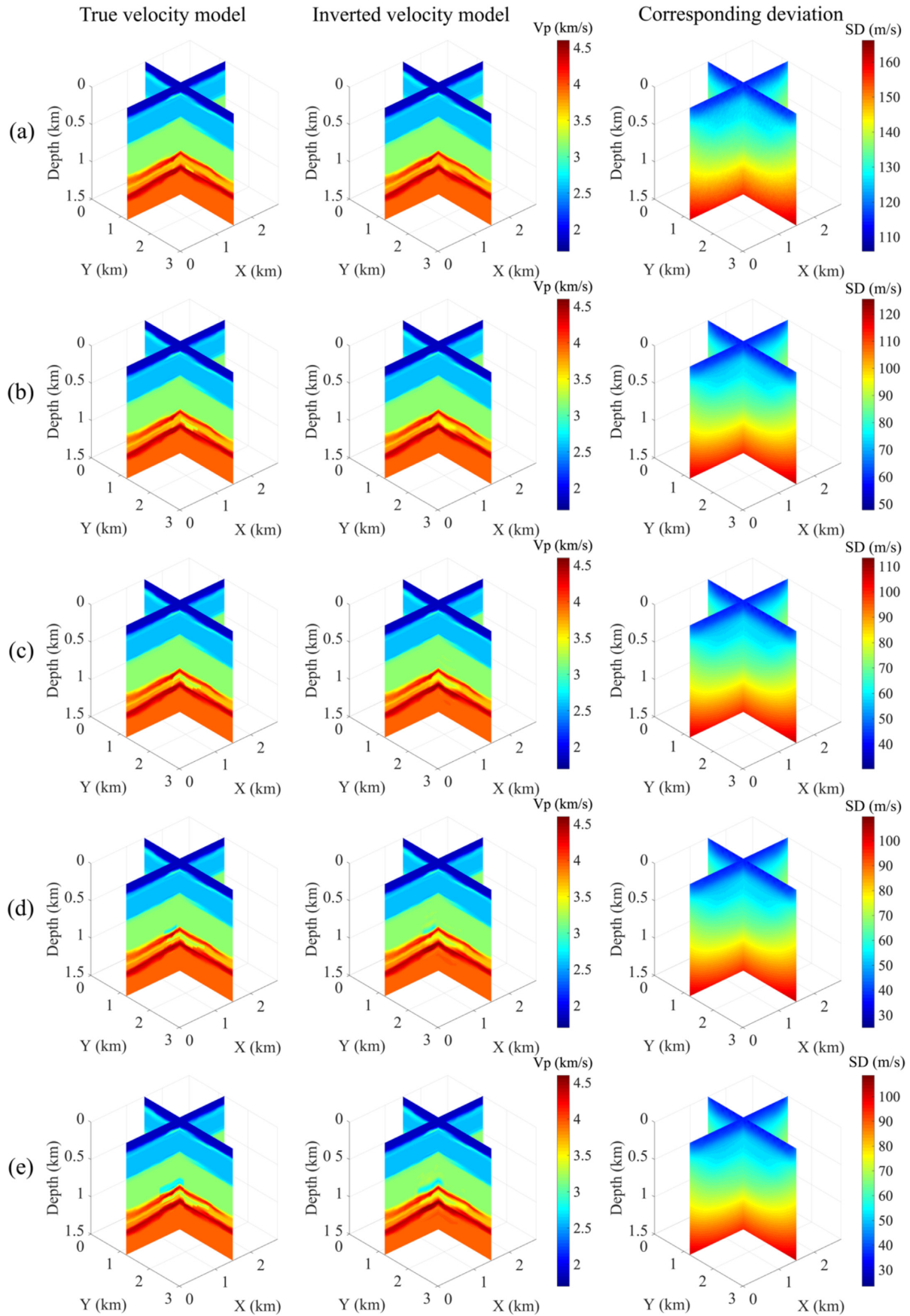


Figure 11. Slices of true velocity models (left-hand column), inverted results (middle column) and corresponding SD (right-hand column) at (a) first, (b) second, (c) third, (d) fourth and (e) fifth time lapses.

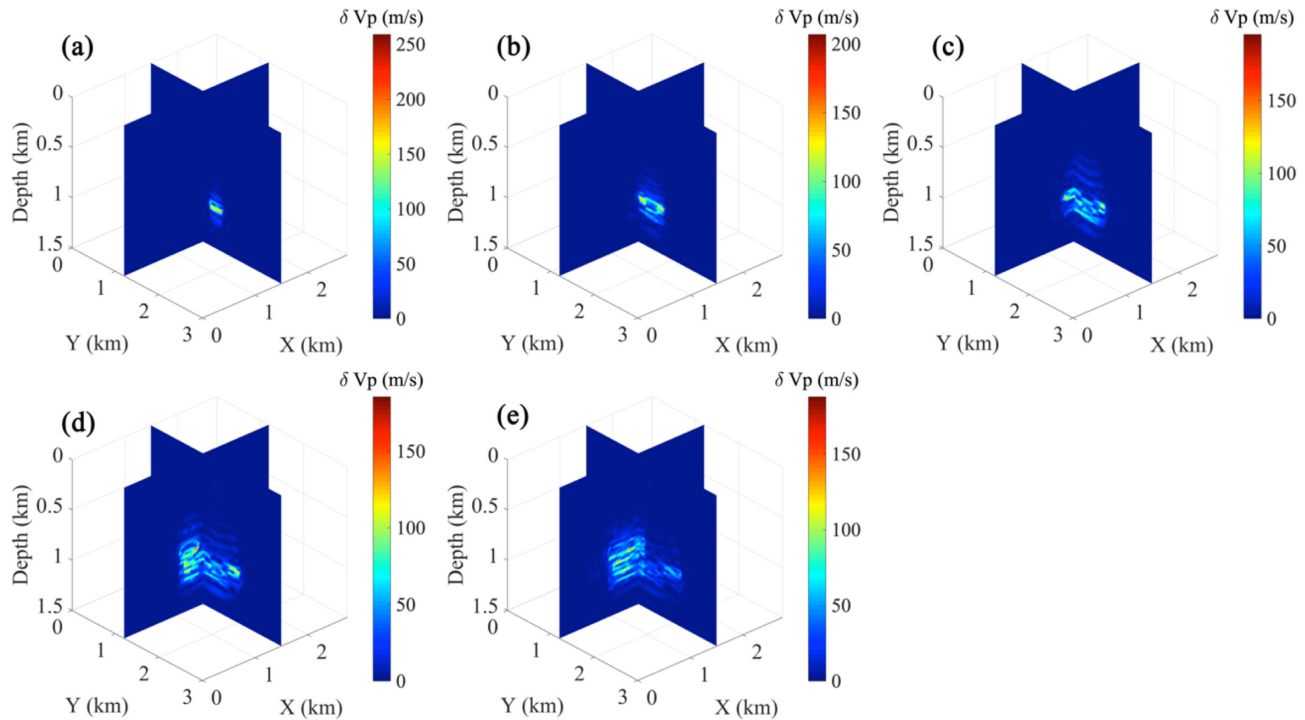


Figure 12. The velocity model error distribution at (a) first, (b) second, (c) third, (d) fourth and (e) fifth time lapses.

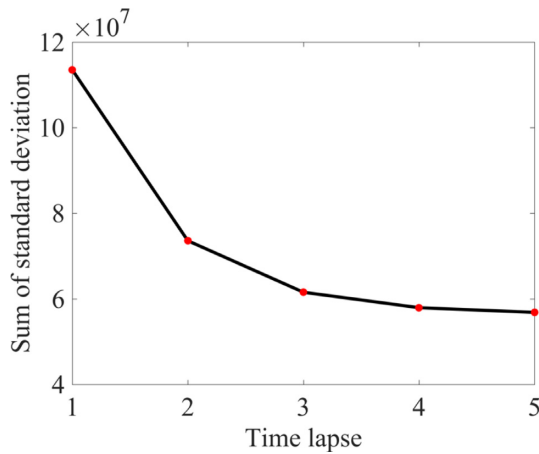


Figure 13. Summed values of the SD matrix over five time lapses (red dots stand for the values of SD at each time lapse).

4 DISCUSSION AND CONCLUSION

We have presented the TLFWI-HiEKF method to process dense time-lapse seismic data. We showed that the velocity estimate at current time lapse by TLFWI rather than the random walk model can serve as a better guess of the *a priori* velocity model in the update step. This step reduces the nonlinearity in the subsequent update step but also speed up the convergence of the inversion. Moreover, HiEKF allows to store and update a data-sparse representation of the cross-covariance matrices using \mathcal{H}^2 -matrices algebra and propagate model errors without expensive operations involving covariance matrices. As a result, the proposed TLFWI-HiEKF algorithm can handle large-scale 3-D problems with affordable memory and computational requirements. The proposed

HiEKF can be easily modified to 3-D FWI by applying multi-scale FWI from low to high frequency and treating seismic data as a dynamic evolution over the frequency axis (Thurin *et al.* 2019).

Taking advantage of both TLFWI and HiEKF, the proposed TLFWI-HiEKF method can impose both spatial and temporal constraints for inverting spatiotemporal subsurface time-lapse velocity changes while simultaneously allowing to quantify the uncertainty of the inverted velocity changes over time. The SD maps provide us a straightforward way of evaluating the results. It is worth to note that the proposed method can continuously update the subsurface velocity model and the accuracy of the model is increasing over time by assimilating more input time-lapse data. This is advantageous over other time-lapse strategies.

In the HiEKF algorithm, we have to calculate and store the Fréchet kernel matrix during inversion which is the most memory consuming part in the TLFWI-HiEKF implementation. The Fréchet kernel matrix \mathbf{H} in our algorithm is calculated by the scattering-integral approach method (Liu *et al.* 2015). Based on the fact that the number of observations is often much smaller (at least one or two orders) than that of model parameters, storing the Fréchet kernel matrix is possibly affordable in most 3-D seismic cases. Rather than calculating the Fréchet kernel matrix itself, however, the adjoint state method is to calculate the product of Fréchet derivative with a vector (data residuals), which is the gradient in FWI.

We have demonstrated that the TLFWI-HiEKF is feasible and efficient in two different seismic time-lapse surveys: 3-D surface seismic sparse surveys in Cranfield and 2-D crosswell dense seismic surveys in Frio-II. In terms of the CASSM that has extremely high temporal sampling rates and can reveal the corresponding rate of changes in the monitored physical parameters (Daley *et al.* 2007), our perspective is to equip the proposed full waveform processing (TLFWI-HiEKF) with the CASSM system for long-term

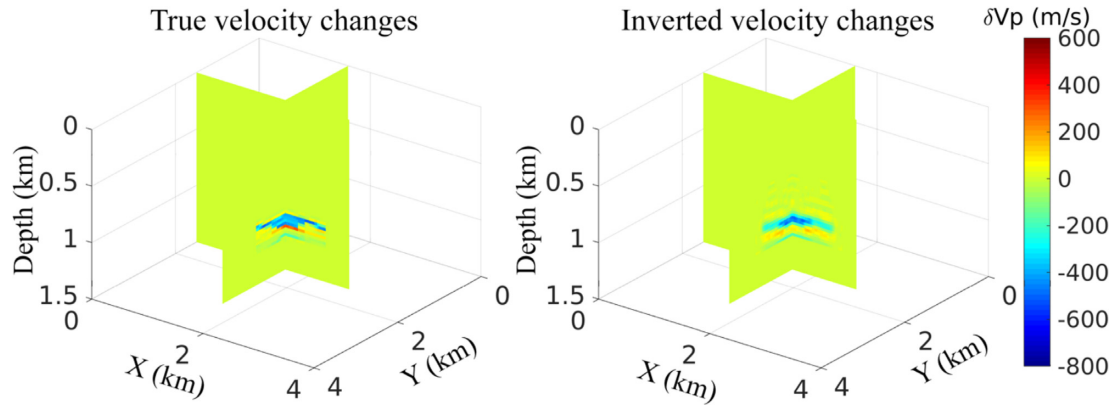


Figure 14. Comparison between true and the inverted perturbations at the fifth time lapse.

Table 1. Comparisons of memory usage and computational complexity of three KF methods in 2-D and 3-D examples (m equals to the grid numbers).

	Grid numbers	Storage cost (GB)		
		EKF	EnKF	HiEKF
2-D Frio test	155 × 144	3.99	0.11 (600 samples)	0.02
3-D Cranfield test	111 × 121 × 61	5369.9	39.32 (6000 samples)	3.46
Computational complexity		$\mathcal{O}(m^2)$	$\mathcal{O}(m)$	$\mathcal{O}(m)$

and large-scale CO₂ monitoring. Although the early demonstration of CASSM is restricted to the crosswell acquisition due to limited seismic source power (Daley *et al.* 2007), newly developed seismic sources and sensors including the Accurately Controlled Routinely Operated Signal System (Ikeda *et al.* 2017), distributed acoustic sensing sensors (Daley *et al.* 2013) and distributed sensor networks (Song *et al.* 2019) offer a path for surface CASSM surveys. Moreover, with the capability of seismically inverting elasticity and attenuation (Xing & Zhu 2019), multiparameters TLFWI including elasticity and attenuation will be incorporated in the future development. We anticipate that the proposed method will become a companion processing tool for processing CASSM data in future large-scale subsurface monitoring applications (e.g. CO₂ sequestration, geothermal monitoring and reservoir monitoring).

ACKNOWLEDGEMENTS

The authors thank Jonathan Ajo-Franklin (Rice) for providing us the seismic Frio models, and Xinming Wu (USTC) for generating the Cranfield seismic time-lapse models. The authors also thank Natalie Accardo (NewGen) for reading the draft. Funding for this project is provided by the U.S. Department of Energy's (DOE) National Energy Technology Laboratory (NETL) under award no. DE-FE0031544.

REFERENCE

- Aanonsen, S.I., Nævdal, G., Oliver, D.S., Reynolds, A.C. & Vallès, B., 2009. The ensemble Kalman filter in reservoir engineering—a review, *SPE J.*, **14**(03), 393–412.
- Ambikasaran, S., Li, J.Y., Kitanidis, P.K. & Darve, E., 2013. Large-scale stochastic linear inversion using hierarchical matrices, *Comput. Geosci.*, **17**(6), 913–927.
- Anderson, B. & Moore, J., 1979. *Optimal Filtering*, vol. **11**, Prentice Hall, Englewood Cliffs, NJ.
- Ajo-Franklin, J., Daley, T., Butler-Veytia, B., Peterson, J., Wu, Y., Kelly, B. & Hubbard, S., 2011. Multi-level continuous active source seismic monitoring (ML-CASSM): mapping shallow hydrofracture evolution at a TCE contaminated site[M], in *SEG Technical Program Expanded Abstracts, 2011*, (pp. 3727–3731), Society of Exploration Geophysicists.
- Ajo-Franklin, J.B., Peterson, J., Doetsch, J. & Daley, T.M., 2013. High-resolution characterization of a CO₂ plume using crosswell seismic tomography: Cranfield, MS, USA, *Int. J. Greenhouse Gas Contr.*, **18**, 497–509.
- Arogunmati, A. & Harris, J.M., 2012. An effective crosswell seismic traveltime-estimation approach for quasi-continuous reservoir monitoring, *Geophysics*, **77**, M17–M26.
- Asnaashari, A., Brossier, R., Garambois, S., Audebert, F., Thore, P. & Virieux, J., 2015. Time-lapse seismic imaging using regularized full-waveform inversion with a prior model: which strategy?, *Geophys. Prospect.*, **63**, 78–98.
- Bui-Thanh, T., Ghattas, O., Martin, J. & Stadler, G., 2013. A computational framework for infinite-dimensional Bayesian inverse problems Part I: the linearized case, with application to global seismic inversion, *SIAM J. Sci. Comput.*, **35**(6), A2494–A2523.
- Daley, T.M., Ajo-Franklin, J.B. & Doughty, C., 2011. Constraining the reservoir model of an injected CO₂ plume with crosswell CASSM at the Frio-II brine pilot, *Int. J. Greenhouse Gas Contr.*, **5**, 1022–1030.
- Daley, T.M., Freifeld, B.M., Ajo-Franklin, J., Dou, S., Pevzner, R., Shulakova, V. & Lueth, S., 2013. Field testing of fiber-optic distributed acoustic sensing (DAS) for subsurface seismic monitoring, *Leading Edge*, **32**(6), 699–706.
- Daley, T.M., Myer, L.R., Peterson, J.E., Majer, E.L. & Hoversten, G.M., 2008. Time-lapse crosswell seismic and VSP monitoring of injected CO₂ in a brine aquifer, *Environ. Geol.*, **54**(8), 1657–1665.
- Daley, T.M., Solbau, R.D., Ajo-Franklin, J.B. & Benson, S.M., 2007. Continuous active-source seismic monitoring of CO₂ injection in a brine aquifer, *Geophysics*, **72**, A57–A61.
- Dennis, J.E., Jr & Schnabel, R.B., 1996. Numerical methods for unconstrained optimization and nonlinear equations, *SIAM*, **16**, doi: 10.1137/1.9781611971200.
- Denli, H. & Huang, L., 2009. Double-difference elastic waveform tomography in the time domain, in *SEG Technical Program Expanded Abstracts, 2009*, (pp. 2302–2306), Society of Exploration Geophysicists.

- Eikrem, K.S., Nævdal, G. & Jakobsen, M., 2019. Iterated extended kalman filter method for time-lapse seismic full-waveform inversion, *Geophys. Prospect.*, **67**(2), 379–394.
- Evensen, G., 2003. The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean Dyn.*, **53**(4), 343–367.
- Fichtner, A. & Trampert, J., 2011. Resolution analysis in full waveform inversion, *Geophys. J. Int.*, **187**(3), 1604–1624.
- Fong, W. & Darve, E., 2009. The black-box fast multipole method, *J. Comput. Phys.*, **228**(23), 8712–8725.
- Gelb, A.(Ed.), 1974. *Applied Optimal Estimation*. MIT Press.
- Gineste, M. & Eidsvik, J., 2017. Seismic waveform inversion using the ensemble Kalman smoother, in 79th EAGE Conference and Exhibition, Volume 2017, p.1–5, European Association of Geoscientists & Engineers.
- Huang, C., Dong, L.G. & Chi, B.X., 2015. Elastic envelope inversion using multicomponent seismic data with filtered-out low frequencies, *Appl. Geophys.*, **12**(3), 362–377.
- Huang, C. & Zhu, T., 2019. Time-lapse full waveform inversion plus extended Kalman filter for high-resolution seismic models and uncertainty estimation, in *SEG Technical Program Expanded Abstracts 2019*(pp. 5239–5244). c
- Ikeda, T. et al. 2017. Temporal variation of the shallow subsurface in the Aquistore CO₂ storage site associated with environmental influences using a continuous and controlled seismic source, *J. geophys. Res.*, **122**, 2859–2872.
- Jin, L., Sen, M.K. & Stoffa, P.L., 2008. One-dimensional prestack seismic waveform inversion using Ensemble Kalman Filter, in *SEG Technical Program Expanded Abstracts 2008*(pp. 1920–1924). Society of Exploration Geophysicists.
- Kalman, R.E., 1960. A new approach to linear filtering and prediction problems[J], *J. Basic Eng.*, **82**(1), 35–45.
- Keptert, J.D., 2004. On ensemble representation of the observation-error covariance in the Ensemble Kalman Filter, *Ocean Dyn.*, **54**(6), 561–569.
- Li, J.Y., Ambikasaran, S., Darve, E.F. & Kitanidis, P.K., 2014. A kalman filter powered by h2-matrices for quasi-continuous data assimilation problems, *Water Resour. Res.*, **50**, 3734–3749.
- Lazaratos, S.K. & Marion, B.P., 1997. Crosswell seismic imaging of reservoir changes caused by CO₂ injection, *Leading Edge*, **16**(9), 1300–1308.
- Liu, Y., Yang, J., Chi, B. & Dong, L., 2015. An improved scattering-integral approach for frequency-domain full waveform inversion, *Geophys. J. Int.*, **202**(3), 1827–1842.
- Lumley, D.E., 2001. Time-lapse seismic reservoir monitoring, *Geophysics*, **66**, 50–53.
- Mora, P., 1989. Inversion = migration+ tomography, *Geophysics*, **54**(12), 1575–1586.
- Plessix, R.E., 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, *Geophys. J. Int.*, **167**(2), 495–503.
- Quan, Y. & Harris, J., 2008. Stochastic seismic inversion using both waveform and traveltimes data and its application to time-lapse monitoring, in *SEG Technical Program Expanded Abstracts, 2008*, (pp. 1915–1919), Society of Exploration Geophysicists.
- Raknes, E.B. & Arntsen, B., 2015. A numerical study of 3D elastic time-lapse full-waveform inversion using multicomponent seismic data, *Geophysics* **80**, R303–R315.
- Raknes, E.B., Weibull, W. & Arntsen, B., 2013. Time-lapse full waveform inversion: Synthetic and real data examples, in *SEG Technical Program Expanded Abstracts, 2013*, (pp. 944–948), Society of Exploration Geophysicists.
- Routh, P., Palacharla, G., Chikichev, I. & Lazaratos, S., 2012. Full wavefield inversion of time-lapse data for improved imaging and reservoir characterization, in *SEG Technical Program Expanded Abstracts, 2012*, (pp. 1–6), Society of Exploration Geophysicists.
- Sirgue, L. & Pratt, R.G., 2004. Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies, *Geophysics*, **69**, 231–248.
- Song, W., Li, F., Valero, M. & Zhao, L., 2019. Toward creating subsurface camera, *Sensors*, **19**(2), 301, doi: 10.3390/s19020301.
- Tarantola, A., 2005. Inverse problem theory and methods for model parameter estimation, *SIAM*, **89**, doi:10.1137/1.9780898717921.
- Thurin, J., Brossier, R. & Métivier, L., 2019. Ensemble-based uncertainty estimation in full waveform inversion, *Geophys. J. Int.*, **219**(3), 1613–1635.
- Vauhkonen, M., Karjalainen, P.A. & Kaipio, J.P., 1998. A kalman filter approach to track fast impedance changes in electrical impedance tomography, *IEEE Trans. Biomed. Eng.*, **45**, 486–493.
- Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, *Geophysics*, **74**, WCC1–WCC26.
- Watanabe, T., Shimizu, S., Asakawa, E. & Matsuoka, T., 2004. Differential waveform tomography for time-lapse crosswell seismic data with application to gas hydrate production monitoring, in *SEG Technical Program Expanded Abstracts, 2004*, (pp. 2323–2326), Society of Exploration Geophysicists.
- Wu, X., 2017. Building 3D subsurface models conforming to seismic structural and stratigraphic features, *Geophysics*, **82**(3), IM21–IM30.
- Xing, G. & Zhu, T., 2019. Fréchet kernels based on a fractional viscoacoustic wave equation, in *SEG Technical Program Expanded Abstracts, 2019*, (pp. 1455–1459), Society of Exploration Geophysicists.
- Yang, D., Liu, F., Morton, S., Malcolm, A. & Fehler, M., 2016. Time-lapse full-waveform inversion with ocean-bottom-cable data: application on Valhall field, *Geophysics*, **81**(4), R225–R235.
- Zhang, Z. & Huang, L., 2013. Double-difference elastic-waveform inversion with prior information for time-lapse monitoring, *Geophysics*, **78**(6), R259–R273.
- Zheng, Y., Barton, P. & Singh, S., 2011. Strategies for elastic full waveform inversion of time-lapse ocean bottom cable (OBC) seismic data, in *SEG Technical Program Expanded Abstracts, 2011* (pp. 4195–4200), Society of Exploration Geophysicists.
- Zhu, H., Li, S., Fomel, S., Stadler, G. & Ghattas, O., 2016. A Bayesian approach to estimate uncertainty for full-waveform inversion using a priori information from depth migration, *Geophysics*, **81**(5), R307–R323.
- Zhu, T., Ajo-Franklin, J.B. & Daley, T.M., 2017. Spatiotemporal changes of seismic attenuation caused by injected CO₂ at the Frio-II pilot site, Dayton, TX, USA, *J. geophys. Res.: Solid Earth*, **122**(9), 7156–7171.
- Zhu, T., Ajo-Franklin, J., Daley, T.M. & Marone, C., 2019. Dynamics of geologic CO₂ storage and plume motion revealed by seismic coda waves, *Proc. Natl. Acad. Sci.*, **116**(7) 2464–2469.