Uncertainty Quantification for 3D FWI at an Operational Level: an Ensemble Transform Kalman Filter approach

Alexandre Hoffmann, Romain Brossier, Ludovic Métivier, Alizia Tarayoun

Uncertainty quantification for 3D FWI at an operational level: an Ensemble Transform Kalman Filter approach

A. Hoffmann[1,2], L. Métivier[3,1], R. Brossier[1], A. Tarayoun[1]

[1]Univ. Grenoble Alpes, ISTerre, F-38058 Grenoble, France
[2]now at CEA, Grenoble, France
[3]CNRS, Univ. Grenoble Alpes, LJK, F-38058 Grenoble, France

January 15, 2024

Main objectives

Introducing an uncertainty quantification method for full waveform inversion applicable to 3D field data at an operational level.

New aspects covered

1. Concise formalism for the coupling between Ensemble Kalman Transform Strategy (ETKF) and full waveform inversion (FWI)

2. Implementation of this formalism in the frame of a decomposition of the dataset following a shot subsampling strategy

3. Application of the proposed method to a 3D field data set from the North Sea: interpretation of the variance and covariance estimations

4. Analysis of the covariance estimation sensitivity to the size of the ensemble with 3 experiments ran with ensemble sizes of 10, 50, and 200.

Summary

Gaining knowledge on the variability of the FWI solutions is crucial. It shall help control their quality and guide their geological interpretation. The main obstacle for designing efficient uncertainty quantification strategy remains the computational cost. We propose in this study an adaptation to FWI of a low-rank approach developed in the data assimilation community named Ensemble Transform Kalman Filter. The method works in a predictor-corrector manner for an ensemble of models instead of considering a single model. In our strategy, the prediction step implies solving the FWI problem for each model of the ensemble. The correction step involves only low-rank linear algebra operations and is thus computationally feasible. We present the application of this strategy to a 3D OBC dataset from the North Sea. We show how we gain insight on the variability of the solution through the estimated variance, and on the local resolution through the estimated covariance. We also test the sensitivity of the method with respect to the ensemble size and find that a number of 50 models makes it possible to capture most of the information.

Uncertainty quantification for 3D FWI at an operational level: an Ensemble Transform Kalman Filter approach

Uncertainty quantification (UQ) remains crucial for seismic imaging using full waveform inversion (FWI). The latter is an ill-posed inverse problem with non-unique solutions. UQ should provide necessary tools to control the quality of FWI solutions and give confidence indices to enhance the interpretation of FWI outputs. The main obstacle so far remains the design of a computationally efficient and reliable UQ strategy. Conventional methods either rely on an estimation of the posterior covariance through an approximation of the inverse Hessian operator in the final estimated model (see Mulder and Kuvshinov, 2023, for a recent example), or on global sampling Monte-Carlo based strategies, with limitations on problem sizes (Gebraad et al., 2020). More recently, the Stein Variational gradient (SVG) approach, developed in statistics community, has emerged for solving geophysics inverse problems (Izzatullah et al., 2023; Lomas et al., 2023), showing interesting properties and an ability to target larger scale application than other aforementioned methods.

What we propose in this study is an adaptation to FWI of another UQ method set up in the frame of data assimilation, known as the Ensemble Transform Kalman Filter (ETKF). ETKF relies on the evolution of an ensemble of models, instead of a single one as is usually considered for FWI applications. From this ensemble of models, first-order (mean) and second-order (covariance) statistics can be inferred with a low-rank approximation. This ensemble description of the solution space is shared with the SVG methods. What differs is how the models within the ensemble are updated. In ETKF it is based on a prediction-correction scheme, namely forecast and analysis steps. In the implementation we propose, the forecast amounts to a FWI run for each member of the ensemble. Our methods makes it possible to compute local uncertainty in the basin of attraction associated with a given initial model. It means that it can be ran as soon as a FWI workflow is set up to converge towards a meaningful estimate of the subsurface mechanical parameters. We summarize first the methodology, then we illustrate its application on a 3D OBC dataset from the North Sea.

ETKF-FWI method

A complete description of the ETKF-FWI method is proposed in Thurin et al. (2019) and Hoffmann et al. (2024). We propose here a brief summary. The first step is to decompose the data into K subsets. This decomposition can be done in many ways: according to the frequency band (sequence of increasing frequency data), time-offset windowing, or shot subsampling for instance. ETKF-FWI considers an ensemble of N_e models $m^1, m^2, \dots m^{N_e}$. The statistics, *i.e.* mean model $\overline{m}$ and covariance P are extracted from this ensemble of models following:

$$\overline{m} = \frac{1}{N_e}\sum_{i=1}^{N_e} m^i, \;\; P = MM^T, \;\; \text{where} \;\; M = \left[m^1 - \overline{m}, \; \dots, \; m^{N_e} - \overline{m}\right], \tag{1}$$

under Gaussian assumptions. Consider one model m^i as a discrete vector in $\mathbb{R}^N$. The covariance P is a square matrix of size N. The matrix M is a square root of P, with N rows and N_e columns. The kth iteration of the ETKF-FWI strategy consists in the interpretation of the kth subset of data. This interpretation is done in two steps: a forecasting step, and an analysis step. We define the forecasting operator as one FWI on the considered subset of observed data $d_{\text{obs},k}$, starting from an initial guess m_{init}:

$$\min_m \frac{1}{2}\|d_{\text{cal},k}[m] - d_{\text{obs},k}\|^2, \;\; \text{starting from a given initial model} \;\; m_{\text{init}}, \tag{2}$$

where $d_{\text{cal},k}[m]$ corresponds to data calculated through the solution of partial differential equations representing the wave propagation in a subsurface model m. The forecasting step at iteration k consists in solving approximately the minimization problem (2) for each model m^i of the ensemble, considering m^i is the initial model m_{init}.

The analysis step takes into account the information from each member and performs a balance between the fit to the data and the deviation to the mean model $\overline{m}$. From a mathematical standpoint, each member of the population is updated following

$$m^i = m^i + \sqrt{N_e - 1}MA^{1/2}, \;\; i = 1, \dots, N_e, \tag{3}$$

where $A^{1/2}$ is the transformation matrix such that

$$A^{-1} = \left((N_e - 1) I_{N_e} + Y^T R^{-1} Y\right), \tag{4}$$

with

$$Y = \left[d_{\text{cal}}[m^1] - \overline{d_{\text{cal}}}, \ldots, d_{\text{cal}}[m^{N_e}] - \overline{d_{\text{cal}}}\right], \quad \overline{d_{\text{cal}}} = \frac{1}{N_e} \sum_{n=1}^{N_e} d_{\text{cal}}[m^i]. \tag{5}$$

In equation (5), $\overline{d_{\text{cal}}}$ is the mean calculated data, and Y is the zero-mean calculated data matrix. Considering $d_{\text{cal}}[m^i]$ as a discrete vector in $\mathbb{R}^{N_{\text{obs}}}$, the matrix Y has N_{obs} rows and N_e columns. In equation (4), I_{N_e} is the identity matrix of size N_e and R is a square matrix of size N_{obs} called measurement noise matrix. In this work we assume it is diagonal and estimate it from the signal-over-noise ratio of the observed data. The matrix A is a square matrix of size N_e. In practice, we first compute A^{-1}, then perform a SVD of A^{-1}, and build the square root $A^{1/2}$. The essence of ETKF-FWI is to have $N_e \ll N$, typically N_e to the order of few tens, while N can reach hundreds of millions for 3D FWI applications to field data. This ensures computationally feasible linear algebra operations to perform the analysis step.

Application to a 3D OBC data from the North Sea

The 3D OBC data we consider in this study has been acquired in a shallow water environment in the North Sea. We select 2048 receivers deployed along cables and 50 000 shot positions. We use reciprocity to decrease the computation cost. We also focus on a narrow frequency band between 3 - 5 Hz to keep a coarse 70 m grid for modeling and inversion. The ETKF-FWI workflow starts by the definition of an initial ensemble. To generate it, we start with a smooth reflection tomography velocity model to which we add zero mean random perturbations, such that the initial mean $\overline{m}$ is equal to the reflection tomography model. The perturbations are spatially correlated to make sure the wavenumber content of the perturbed model is compatible with the considered frequency band. The amplitude of the perturbations is controlled so as to generate no cycle skipping. Doing so, we ensure each model of the ensemble converge towards the same main basin of attraction. The aim of ETKF-FWI is to sample this attraction basin. The initial mean model and the corresponding initial variance (the diagonal of the estimated covariance matrix P) are presented in Figure 1. We consider here and in the following an ensemble size $N_e = 50$.

The second step is to decompose the available dataset into subsets which we interpret sequentially in the ETKF-FWI scheme. It is usual to rely on shot subsampling strategy when addressing 3D field data applications to decrease computational cost. This method can be readily employed for ETKF-FWI.

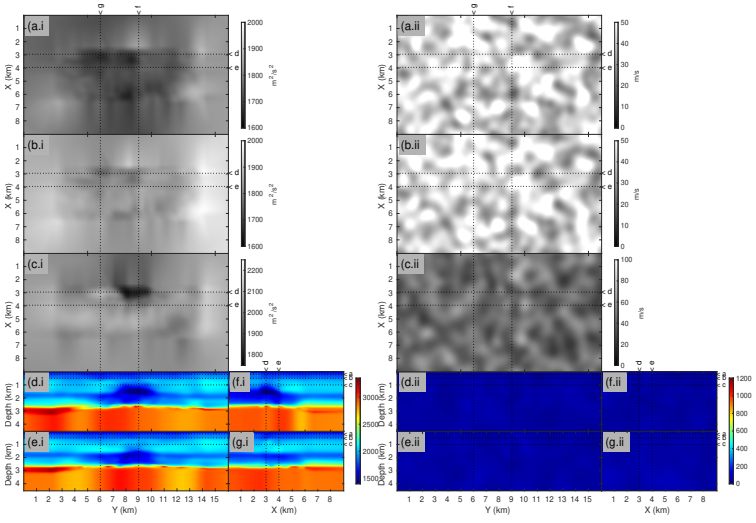

Figure 1 Left: sections of the average initial model obtained by reflection tomography. Right, sections of the variance of the initial normally distributed random models. (a-c) Horizontal sections at (a) 200 m depth, (b) 500 m depth, (c) 1 km depth. (d and e) Inline vertical sections for (d) x= 2.95 and (e) x= 3.95 km. (f and g) Crossline vertical sections for (f) y=9 km and (g) y= 6 km.

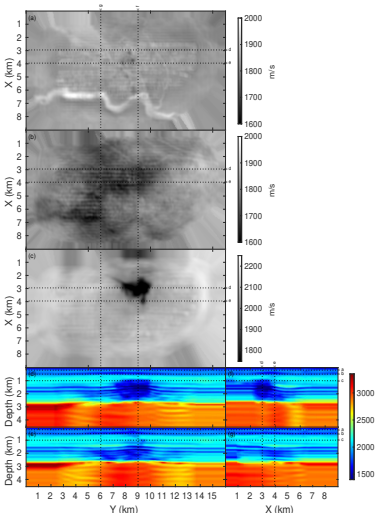

Figure 2 Final mean model after ETKF-FWI in the same display as in Figure 1.

The initial pool of 2048 shots (after reciprocity) is divided into $K = 16$ batches of 128 randomly selected shots. With this decomposition, the forecast step at stage k of ETKF-FWI consists in performing 3 FWI iterations for each of the ensemble members to invert for the kth batch of shots.

The final mean model obtained using the ETKF-FWI scheme is presented in Figure 2. It exhibits expected features in this frequency band considering previous work on the same dataset (see Pladys et al., 2022, for instance), in particular a low-velocity-anomaly at 1000 m depth appearing in black in the black and white constant depth section. This low velocity anomaly is also visible in the inline and crossline vertical sections as the blue layered structure. This result show that the ETKF-FWI scheme is able to produce a mean model similar to what would be obtained performing a conventional FWI.

The estimated variance in the inline vertical section at x=2.95 km (crossing the low velocity anomaly) is presented in Figure 3. We present together the mean, variance, and the distribution of models for three different pixels in this section. The variance helps identifying zones where the reconstruction is more uncertain. Variance peaks corresponds to zones where the spread of the model is higher as in examples b(iii) and c(iii). Note that the model distribution approximately follows a Gaussian distribution, indicating we satisfy the Gaussian assumption on which relies the ETKF-FWI scheme.

From the low-rank posterior covariance matrix, we build the correlation matrix as

$$C = \mathrm{diag}\,(P)^{-\frac{1}{2}}\, P \mathrm{diag}\,(P)^{-\frac{1}{2}}\,. \tag{6}$$

The rows of the correlation matrix encodes the information about the links between the pixel associated with the row and all other pixels. A correlation coefficient close to 1 indicates that the pixels are positively correlated. When it is close to -1 it indicates an anticorrelation. When it is close 0 it indicates no correlation. We see how we can use this information as a proxy for a local resolution analysis. In Figure 4 we present three rows of the correlation matrix, for threes points belonging to a given layer. For the orange and green points, the strong correlation with the pixels in the same layer indicates that these layers are well resolved. For the red point, the strong correlation with the large zone above the layer indicates a local lower resolution.

Finally, we investigate the effect of the parameter N_e on the mean, variance, and covariance estimation. This parameter is crucial as it controls the whole computational cost of the ETKF-FWI scheme. Basically, the computational cost of ETKF-FWI compared to a conventional FWI is N_e times larger, as N_e FWI have to be ran instead of 1. This is the reason why it is important that N_e remains in an acceptable limit. We thus perform an experiment where N_e is equal to $10, 50$ and 200. In Figure 5, we can observe that the mean model does not change and is already quite well estimated with $N_e = 10$. However, the variance for $N_e = 10$ is clearly underestimated, an expected feature for ETKF strategies. The corresponding distribution of models for the selected pixel also exhibit a very poor sampling of a Gaussian distribution. What is interesting to note is that when N_e increases to 50 and 200, the distribution converges to a Gaussian distribution, indicating that the Gaussian assumption behind the ETKF-FWI scheme holds.

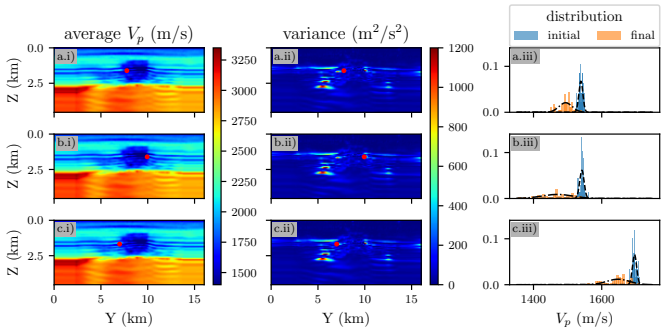

Figure 3 (a.i-f-i) final mean V_p model, (a.ii-f.ii) final variance, (a.iii-f.iii) model distribution at a given point marked in red.

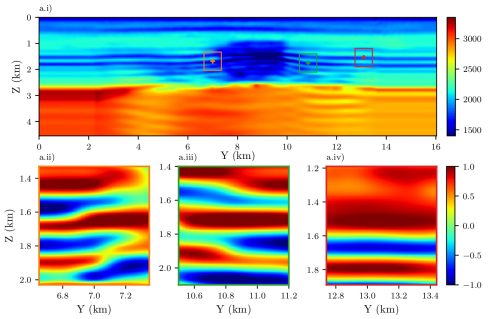

Figure 4 (a.i) Inline vertical section of the mean model, (a.ii-a.iv) local correlation maps from the correlation matrix C in orange, green and red boxes.

On top of that, the variation of the variance from $N_e = 50$ to $N_e = 200$ is mild, indicating that $N_e = 50$ provides already a good approximation of the variance and covariance. For this experiment, the number of discrete points in the model space is $N \simeq 2 \times 10^6$. It is worth noting that $N_e = 50$ provides an already acceptable low-rank representation of the covariance, meaning that there exists a low dimensional subspace encoding most of the model features in the considered frequency band.

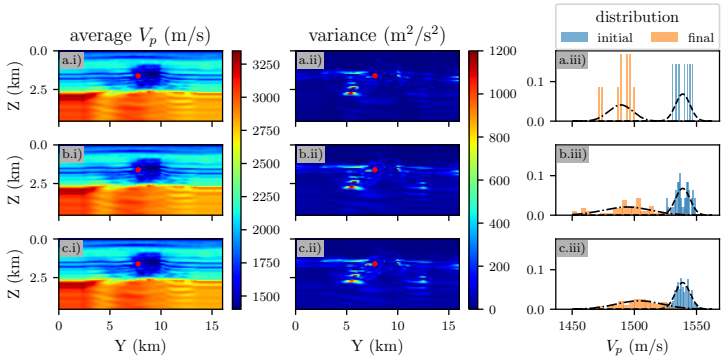

Figure 5(a.i-c.i) Inline vertical section of the final mean V_p model, (a.ii-c.ii) final variance, (a.iii-c.iii) distribution of the models at a given point marked in red. The results are obtained with three ensemble sizes (a) $N_e = 10$, (b) $N_e = 50$, (c) $N_e = 200$.

Conclusions

We present an uncertainty quantification scheme for FWI based on an Ensemble Transform Kalman filter we name ETKF-FWI. Instead of running a single FWI from an initial model, an ensemble of models perturbed around the initial models are inverted jointly following a data assimilation strategy. From the final ensemble, we can infer first-order and second-order statistics: mean and covariance, in the Gaussian approximation. We show on a 3D OBC data from the North Sea how this method can help better qualifying the final results in terms of local variability and resolution. For a problem involving approximately 2×10^6 discrete points in the model space, we find that a small ensemble of $N_e = 50$ members already captures most of the required features to provide a reliable covariance estimation. The computational cost of the ETKF-FWI is N_e times larger than for a single FWI however the N_e FWI required at each forecast step can be ran in parallel: in the forecast they are independent process. This makes the method a good candidate for the incoming generation of exascale HPC structures.

Acknowledgments

This study was partially funded by the SEISCOPE consortium (`http://seiscope2.osug.fr`), sponsored by AKER BP, CGG, DUG, EXXONMOBIL, GEOLINKS, JGI, PETROBRAS, SHELL, SINOPEC and TOTALENERGIES. This study was granted access to the HPC resources provided by the GRICAD infrastructure (https://gricad.univ-grenoble-alpes.fr), Cray Marketing Partner Network (`https://partners.cray.com`) and IDRIS/TGCC/CINES under the allocation 046091 made by GENCI.

References

Gebraad, L., Boehm, C. and Fichtner, A. [2020] Bayesian Elastic Full-Waveform Inversion Using Hamiltonian Monte Carlo. *Journal of Geophysical Research: Solid Earth*, **125**(3), e2019JB018428. E2019JB018428 10.1029/2019JB018428.

Hoffmann, A., Brossier, R., Métivier, L. and Tarayoun, A. [2024] Uncertainty quantification for 3D time-domain full waveform inversion with ensemble Kalman filters: application to the Valhall OBC dataset. *Geophysical Journal International*, **in revision**.

Izzatullah, M., Alkhalifah, T., Romero, J., Corrales, M., Luiken, N. and Ravasi, M. [2023] Plug-and-Play Stein variational gradient descent for Bayesian post-stack seismic inversion. In: *84th Annual EAGE Meeting (Vienna)*, 2023. European Association of Geoscientists & Engineers.

Lomas, A., Luo, S., Irakarama, M., Johnston, R., Vyas, M. and Shen, X. [2023] 3D Probabilistic Full Waveform Inversion: Application to Gulf of Mexico Field Data. In: *84th Annual EAGE Meeting (Vienna)*, 2023. European Association of Geoscientists & Engineers.

Mulder, W. and Kuvshinov, B. [2023] Estimating large-scale uncertainty in the context of full-waveform inversion. In: *84th Annual EAGE Meeting (Vienna)*, 2023. European Association of Geoscientists & Engineers.

Pladys, A., Brossier, R., Kamath, N. and Métivier, L. [2022] Robust FWI with graph space optimal transport: application to 3D OBC Valhall data. *Geophysics*, **87**(3), 1–76.

Thurin, J., Brossier, R. and Métivier, L. [2019] Ensemble-based uncertainty estimation in Full Waveform Inversion. *Geophysical Journal International*, **219**(3), 1613–1635.