Stochastic joint-inversion and uncertainty quantification of seismic and CSEM data

Pankaj K Mishra[1], Adrien Arnulf[2], Mrinal K Sen[3], Zeyu Zhao[4], and Piyoosh Jaysaval[5]

[1]Geological Survey of Finland
[2]Amazon Web Services
[3]Institute for Geophysics, The University of Texas at Austin
[4]School of Earth and Space science, Peking University
[5]Pacific Northwest National Laboratory

September 17, 2024

Abstract

Uncertainty quantification in geophysical inversion is a well-recognized area of research, yet it has not become routine practice. One of the primary challenges is the computational expense of forward solvers, making robust uncertainty quantification methods like Monte Carlo or Markov Chain Monte Carlo (MCMC) impractical, particularly for higher-dimensional problems. This challenge is amplified in the case of joint inversion, where multiple types of forward solvers must be run thousands of times. We propose a stochastic joint inversion framework that integrates the Very Fast Simulated Annealing (VFSA) approach with a generalized fuzzy c-means clustering technique for effective parameter coupling. By incorporating a sparse parameterization strategy and executing multiple VFSA chains with varying initial models, we effectively mitigate VFSA's tendency to converge at the peak of the derived posterior probability density (PPD) function. The approach presented here address the inherent challenge of high computational costs for implementing joint-inversion with nonlinear sampling methods like MCMC by providing a feasible probabilistic joint inversion alternative that can integrate petrophysical information as well as geological constraints.

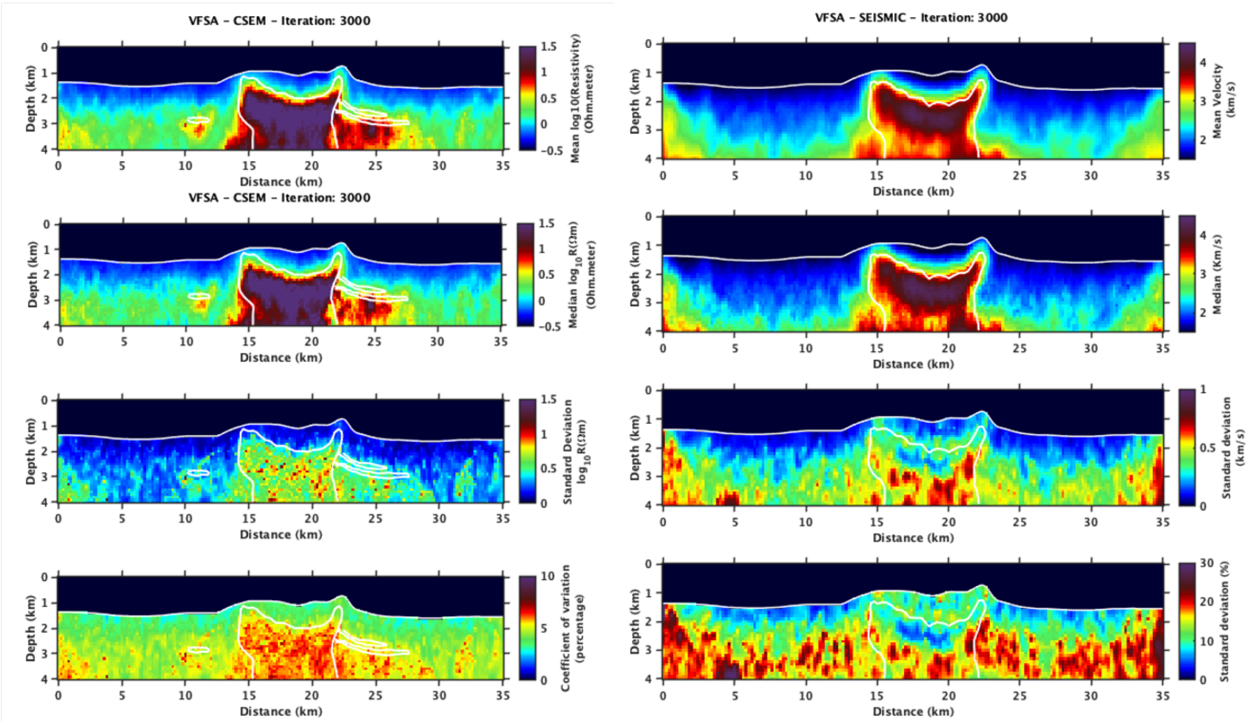

Figure 1: ThMean (a), median (b), standard deviation (c) and percentage coefficient of variance (d) for the estimated resistivity models and the mean (e), median (f), standard deviation (g) and percentage standard deviation (h) for the estimated velocity models from the joint-inversion.

Stochastic joint-inversion and uncertainty quantification of seismic and CSEM data

Pankaj K Mishra[1], Adrien Arnulf[2*], Mrinal K Sen [3], Zeyu Zhao[4†], Piyoosh Jaysaval[5]

[1]Geological Survey of Finland (GTK), Finland
[2]Amazon Web Services, USA
[3]Institute for Geophysics, The University of Texas at Austin, USA
[4]School of Earth and Space science, Peking University, China
[5]Pacific Northwest National Laboratory, USA

*Previously at The University of Texas at Austin
†Previously at The Univeristy of Texas at Austin

Corresponding author: Pankaj K Mishra, pankaj.mishra@gtk.fi

Abstract

Uncertainty quantification in geophysical inversion is a well-recognized area of research, yet it has not become routine practice. One of the primary challenges is the computational expense of forward solvers, making robust uncertainty quantification methods like Monte Carlo or Markov Chain Monte Carlo (MCMC) impractical, particularly for higher-dimensional problems. This challenge is amplified in the case of joint inversion, where multiple types of forward solvers must be run thousands of times. We propose a stochastic joint inversion framework that integrates the Very Fast Simulated Annealing (VFSA) approach with a generalized fuzzy c-means clustering technique for effective parameter coupling. By incorporating a sparse parameterization strategy and executing multiple VFSA chains with varying initial models, we effectively mitigate VFSA's tendency to converge at the peak of the derived posterior probability density (PPD) function. The method presented here addresses the inherent challenge of high computational costs for implementing joint-inversion with nonlinear sampling methods like MCMC by providing a feasible probabilistic joint inversion alternative that can integrate petrophysical information as well as geological constraints.

Plain Language Summary

Understanding uncertainty is important in geophysics, where scientists try to figure out what lies beneath the Earth's surface by interpreting data. However, calculating this uncertainty is very challenging and is not commonly done because it requires a lot of computing power. Traditional methods like Monte Carlo or Markov Chain Monte Carlo (MCMC) are accurate but slow, especially when dealing with complex problems that use multiple types of data at once.

To solve this, we developed a more efficient method that combines a faster technique called Very Fast Simulated Annealing (VFSA) with fuzzy clustering, a machine learning method that helps connect different types of information. We also simplify the process by using fewer points for the inversion through sparse sampling of model parameters, and running multiple simulations with different starting points. This helps avoid issue of getting stuck on one possible solution and provides a more complete picture. Our approach offers a practical way to combine multiple data types and incorporate geological information without the high computational costs of traditional methods.

1 Introduction

Geophysical inverse problems are known to be non-unique, which means, there exist a number of plausible models that would fit the data. The idea behind an integrated inversion is to reduce the number of possible models by using different but complementary geophysical, geological, and petrophysical data in an unified geophysical inversion framework. The term 'joint-inversion' refers to one of the many integrated (coupled) inversion approaches where cost functions of different methods are efficiently combined to construct a joint-objective function, which is minimized while adjusting all the model parameters concurrently. Since all the involved methods contribute to the model update, the inversion artifacts are likely to be reduced in certain subspace of the model, which is sensitive to more than one method (Moorkamp et al., 2016). There are some specific challenges in the development of a efficient joint-inversion algorithm:

1. Although a joint-inversion is likely to narrow-down number of possible solutions, the inversion problem remains non-unique and the estimated model could be sub-optimal.
2. In realistic models, petrophysical relationship(s) among different model parameters could be complicated, which requires an efficient coupling strategy in the joint-inversion algorithm.

The first challenge would be dominant if we use a deterministic method for the joint-inversion which produces a single 'best' model. A probabilistic method, however, provides many possible models which average to a final model estimation and provide uncertainties in it.

Some previous works in the context of probabilistic joint-inversion have used MC method (Bosch & McGaughey, 2001; Chen et al., 2004; Bosch et al., 2006; Jardani & Revil, 2009; Shen et al., 2013), Co-kriging method (Shamsipour et al., 2012), Markov-Chain Monte-Carlo (MCMC) method(Rosas-Carbajal et al., 2014; Wéber, 2018), trans-dimensional MCMC (Blatter et al., 2019), and Very Fast Simulated Annealing (VFSA) method (Kaikkonen & Sharma, 1998; YANG et al., 2002; Hertrich & Yaramanci, 2002; Santos et al., 2006). Out of these, MC method is the most rigorous probabilistic approach as it samples proposal models randomly in the parameter space. The proposal models are accepted or rejected based on a computed probability using the Metropolis-Hasting criterion (Metropolis et al., 1953). Due to randomness in the selection of the proposal model, MC method provides the most accurate posterior probability distribution (PPD) of the model at an extremely expensive cost (Sen & Stoffa, 2013).

Unlike the MC method which is able to draw independent samples from the distribution, MCMC method draws proposals where the next model is dependent on the existing model by using the Markov-Chain. This allows the algorithms to narrow in on the quantity that is being approximated from the distribution making it a less expensive alternative to the MC method. Sen and Stoffa (1996) discuss several alternatives of MC method and demonstrate that a tweaked VFSA is an affordable alternative (Roy, Sen, Blankenship, et al., 2005; Roy, Sen, McIntosh, et al., 2005), which can be used to derive PPD without using a rigorous sampling method like MC or MCMC. VFSA is an optimization algorithm based on the Metropolis-Hastings criterion.

There are two main differences between MCMC and VFSA as the latter uses a temperature dependent Cauchy-distribution to draw the proposal model, which tends to narrow down the proposal to the previous state as the temperature decreases. Moreover, the probability of accepting a 'bad' model also decreases over number of iterations and becomes sufficiently low near the global minimum. The PPD derived from a single chain of VFSA is inherently biased towards towards the global minimum, therefore, multiple chains of VFSA are needed to get many plausible models for uncertainty quantification. Although, the PPD estimated through rigorous sampling methods are more accurate, the same obtained through the VFSA does provide a sweet-spot between affordability and accuracy.

The second challenge, that is, effective coupling of model parameters has mostly been discussed in the context of deterministic joint-inversion, which can be categorized as (1) structure-based coupling (Haber & Oldenburg, 1997; Gallardo & Meju, 2004) and (2) petrophysical coupling (Koketsu & Nakagawa, 2002; Jegen et al., 2009). A detailed review about different approaches for parameter coupling can be found in (Colombo & Rovetta, 2018). In this paper, we use a guided-fuzzy c-means clustering developed by (Sun & Li, 2012), which is a generalized version of the the method proposed by Lelièvre et al. (2012) and has been effectively used in deterministic joint-inversion in geoscience (Sun & Li, 2016a, 2016b).

In this paper, we introduce a probabilistic joint inversion approach for multi-physics data integration, utilizing the Very Fast Simulated Annealing (VFSA) algorithm combined with a generalized Fuzzy C-Means (FCM) clustering method. Given that a large number of inversion parameters would typically necessitate an impractically high number of VFSA iterations for convergence, we employ a sparse parameterization strategy to randomly distribute inversion points across the model space. We provide a detailed discussion of the VFSA and FCM algorithms, explaining our rationale for their selection in the joint inversion framework. To validate the proposed algorithm, we present numerical experiments on the joint inversion of first-arrival seismic traveltime and controlled-source electromagnetic (CSEM) data for a 2D slice of the SEAM Phase I model, focusing on computing mean models and associated uncertainties. Finally, we show that, the mean P-wave velocity model obtained from the joint inversion can be further refined using Full Waveform Inversion (FWI) (Tarantola, 1984).

2 Forward Modeling and Cost Computation

In recent times, there has been a growing interest in the modeling and inversion of CSEM data for mapping the resistivity of the subsurface (Constable, 2010; Key, 2016; Constable et al., 2019; Lu & Farquharson, 2020). Under the low-frequency assumption (neglecting the displacement current), and quasi-static limit, the Maxwell's equations can be written as the following

$$\nabla \times \nabla \times \mathbf{E} - i\omega\mu\sigma\mathbf{E} - i\omega\mu\mathbf{J} = 0. \quad (1)$$

where $\mathbf{E}$ is the electric field, ω is the angular frequency, μ is the magnetic permeability, σ, is electrical conductivity, and $\mathbf{J}$ is the source-term. The second-order PDE given by equation (1) is the governing equation for CSEM modeling. We discretize it by using the staggered-grid finite-different method (Yee, 1966; Newman & Alumbaugh, 1995). Dirichlet boundary conditions are applied at all boundaries of the computational domain by forcing the electric field values to zero. In order to reduce the spurious data due to the limited finite-domain, the computational domain was stretched by adding a thick ($70km$) highly-resistive ($10^7 Ohm.m$) layer on top of the water layer. This additional domain was discretized using stretched cells with increasing thickness along the z-direction. In principle, as the angular frequency reaches the static limit, the highly-resistive layer (air) causes non-uniqueness in the solution, which further requires a static divergence correction. However, for practical CSEM modeling the angular frequency is usually greater than $0.1Hz$, therefore, CSEM modeling does not require such corrections . A detailed description about the finite-difference discritization of the equation(1) can be found in (Streich, 2009; Jaysaval et al., 2014). The forward modeling kernel is parallelized by shot, frequencies wavenumber in the y-dircetion.

If $\mathbf{m}$ is a conductivity model in a set of $\mathcal{M}$ models, the optimization problem for CSEM can be written as

$$\mathbf{m}^{est} = arg \min_{\mathbf{m}\in\mathcal{M}} \Phi^{CSEM}(\sigma). \quad (2)$$

The cost function $\Phi^{CSEM}(\sigma)$ is given by

$$\Phi^{CSEM}(\sigma) = \sum_{\mathbf{r}_s,\mathbf{r}_r,F,i,f} 0.5 W_i^F (\mathbf{r}_r|\mathbf{r}_s;\mathbf{J},f)\, |\Delta F_i(\mathbf{r}_s,\mathbf{r}_r,F,i,f,\sigma)|^2 \quad (3)$$

where F_i is the component of electric or magnetic field in the direction of x or y ($i = (x,y)$). ΔF is the difference between observed and computed fields at receiver location $\mathbf{r}_r$ due the the source $\mathbf{J}$ at location $\mathbf{r}_s$.

$$W_i^F = \frac{1}{|F_i^{obs}|^2 + \eta^2} \quad (4)$$

The datum weight W_i^F balances the contributions from fields at different offsets to the cost function.

For raytracing of seismic first-arrival paths, we use the shortest path method, which is an efficient and flexible approach to compute the raypaths and traveltimes of first arrivals to all points in the earth simultaneously (Moser, 1991; Arnulf et al., 2011, 2014, 2018). The forward calculation of synthetic first arrivals is fully parallelized for each receiver taking advantage of the source receiver reciprocity. The cost function for seismic

ϕ^{SE} is given by

$$\Phi^{SE} = \sum_{i=1}^{N} \frac{\left(T_i^{pick} - T_i^{cal}\right)^2}{\sigma_i^{err}}, \tag{5}$$

where T^{pick}, and T^{cal} are picked and calculated traveltimes respectively, σ_i^{err} is uncertainty in picking, and N is total number of picks.

3 VERY-FAST SIMULATED ANNEALING

Simulated annealing (SA) is an optimization algorithm, which is inspired from the annealing process in Thermodynamics (i.e. (Kirkpatrick et al., 1983); see (Sen & Stoffa, 2013) for geophysical applications). The annealing is a process where a metal is heated until it melts and then cooled in a controlled way to achieve the lowest energy state. During annealing, thermal equilibrium is reached at every temperature where the set of all possible molecular configuration is given by the probability of a particular configuration of particles being in a state i

$$p_i = \frac{e^{\frac{-E_i}{kT}}}{\sum_{j \in \mathcal{S}} e^{\frac{-E_j}{kT}}}. \tag{6}$$

Equation (6) is named the Boltzmann distribution where E_i is the energy state of the configuration i, $\mathcal{S}$ is the set of all possible configuration, k is the Boltzmann constant and T is the temperature. The SA method starts with an initial model and an initial (highest) temperature. In the next step, the temperature is decreased by a predefined cooling schedule, and the algorithm draws a new (proposed) model from a flat-distribution. The objective function is defined as the difference between energy states of the initial and the new model. If the energy state of the proposed model is less than that of the previous model, the model is accepted. On the other hand, if the energy state of the proposed model is greater than that of the previous model, the model is considered as a bad proposal. Since SA is a global optimization algorithm with a purpose of jumping out of a local minimum, it does not discard a bad proposal model but accepts it with a probability called the Metropolis-Hasting criterion, which is given as

$$P_n = e^{\left(-\frac{E(\mathbf{m}_n) - E(\mathbf{m}_{n-1})}{t_n}\right)} \tag{7}$$

where t_n is the temperature at the current step. The probability (P_n) of accepting a bad proposal model is temperature dependent. As the SA algorithm proceeds to higher iterations, the current temperature decreases and so does the probability. This means that at initial iterations, SA would accept most of the bad proposal models but becomes parsimonious when the solution approaches the global minimum. The cooling schedule needs to be carefully defined as it controls the trade off between the computational cost of SA algorithm and its ability of finding the global minimum. A faster cooling would accelerate the algorithm but a slower cooling makes sure the globally optimal solution. Therefore, for a complicated optimization problem the cooling schedule needs to be significantly slower, which can make SA — unpractical.

Ingber (1989) proposed a variant of SA, called Very Fast Simulated Annealing (VFSA), which offers a significant speed-up with minimum sacrifice in the ability to find the globally optimal solution. The main difference between SA and VFSA is that instead of drawing a proposal model from a flat-distribution, the latter draws the proposal model using a temperature-dependent Cauchy-like distribution over the previously accepted model as given by

$$m_i^n = m_i^{n-1} + y_i(m_i^{ub} - m_i^{lb}), \tag{8}$$

where m^{ub} and m^{lb} are maximum and minimum allowed values of the model parameter and

$$y_i = sgn(u_i - 0.5)t_n \left[(1 + 1/t_n)^{|2u_i - 1|} - 1\right], \tag{9}$$

where u is a random number between 0 and 1. The VFSA algorithm offers some unique flexibility over the SA. As such, in the initial stage of the optimization process (i.e. the earlier iterations) the algorithm is set to explore a broad region of the model space. Then, as the optimization process starts to converge towards the global minimum, the algorithm is set to focus on narrower regions of the model space located near the previously accepted model. It is worth to be noted that in the case of multi-parameter optimization, VFSA allows different parameters to have their own cooling schedule and model search space.

4 Parameter Coupling

Clustering is a machine Learning technique that finds groups in a specific dataset, where the data belonging to a group are more similar to each other than the data belonging to another group. These groups in the data are called clusters. A clustering algorithm segments the data into high-density clusters such that inter-cluster members have low similarity and intra-cluster data have high similarity. Clustering is an efficient tool for exploring the relationship(s) between data, which, traditionally, has been a post-inversion process where inversion results from single data inversions are analyzed to interpret the final petrophysics and geology of the subsurface. A cross-plot between inverted parameters explains the statistical relationship between model parameters and often a direct relationship is derived and used as a coupling scheme in the joint-inversion. Statistical coupling gets tricky when the crossplot has complex and more than one relationship between parameters. Recently, (Sun & Li, 2016b, 2016a) proposed an efficient way of parameter coupling by using a generalized fuzzy c-means (FCM) clustering to constrain the statistical relationship of inverted parameters for several deterministic joint-inversion studies. In this paper, we show that FCM can also be incorporated into our probabilistic joint-inversion framework.

FCM (Dunn, 1973; Bezdek, 1981) is one such clustering approach, where instead of belonging to one cluster, the data points can have a certain degree of memberships to different clusters. Clustering methods, essentially, solve local optimization problems where an optimal data segmentation is achieved by minimizing a certain cost function. The cost function for a typical FCM method is given as,

$$\Phi^{FCM} = \sum_{k=1}^{N} \sum_{i=1}^{c} (\mu_{i,k})^m (\mathbf{x}_k - \mathbf{v}_i)^T A (\mathbf{x}_k - \mathbf{v}_i), \tag{10}$$

where $\mu_{i,k}$ are the elements of the 'partition matrix' $\mathbf{U} \in \mathbf{R}^{N \times c}$ and are termed as membership values, which is a measure of of degree of membership of k^{th} data in the i^{th} cluster. The 'fuzzification parameter' $(1 < m < \infty)$ determines the degree of 'fuzziness' in the clustered data, and the vector $\mathbf{v}_i$ is the center of i^{th} cluster. The distance-norm A determines the shape of the clusters. The standard FCM algorithm uses the Euclidean norm $(A = I)$; therefore, it detects only circular clusters in the data. However, if we modify A to include variances adaptive to given data, the FCM can be generalized to detect different shapes of clusters in one cross-plot. A generalized formulation of FCM uses an adaptive (mahalanobis) distance-norm (Gustafson & Kessel, 1979), where each cluster has its own norm-inducing A matrix given by

$$A_i = \left[\rho_i det(F_i)\right]^{1/n} F_i^{-1}, \tag{11}$$

where ρ_i is the determinant of the matrix A_i, and F_i is the fuzzy covariance matrix for i^{th} data as given by

$$F_i = \frac{\sum_{k=1}^{N} (\mu_{i,k})^m (\mathbf{x}_k - \mathbf{v}_i)^T (\mathbf{x}_k - \mathbf{v}_i)}{\sum_{k=1}^{N} (\mu_{i,k})^m}. \tag{12}$$

5 Joint-inversion Workflow

We write the combined objective function for the joint-inversion as follows

$$\Phi = \Phi^{EM} + \Phi^{SE} + \Phi^{FCM} + \lambda \sum_{i=1}^{c} \parallel \mathbf{v}_i - \mathbf{g}_i \parallel_2 . \tag{13}$$

Where Φ^{EM}, Φ^{SE}, and Φ^{FCM} are cost functions of CSEM, seismic, and FCM respectively. To adjust the contributions from each methods, we normalize the individual cost functions for their target errors derived from separate VFSA inversions. This will make sure that all the methods contribute equally near the global minimum and avoids free parameters (weights) in the joint-inversion. The forth-term in equation (13) incorporates prior information about the known geology by minimizing the distance between inverted cluster centers $\mathbf{v}_i$, and known cluster centers $\mathbf{g}_i$. The parameter λ reflects our confidence in the priors. The value of $\lambda = 0$ would imply that there is no prior information about cluster centers and final inverted centers will be determined purely through the inversion. The iterative update scheme (Sun & Li, 2016b) for FCM centers (including the priors) is derived as

$$\mathbf{v}_i = \frac{\sum_{k=1}^{N} (\mu_{i,k})^m \mathbf{x}_k + \lambda \mathbf{g}_i}{\sum_{k=1}^{N} (\mu_{i,k})^m + \lambda}. \tag{14}$$

In the following algorithm, we explain the probabilistic joint-inversion workflow for CSEM and seismic data for one inner loop inside the main VFSA loop. The model parameters $\mathbf{m}_{res}$ and $\mathbf{m}_{vel}$ represent the vertical resistivity (R) and p-wave velocity (V_p), respectively.

Initialize maximum temperature t_0, number of VFSA iterations, number of loops inside VFSA iteration, random initial models: $\mathbf{m}^0_{res}$, $\mathbf{m}^0_{vel}$, initial random centers $\mathbf{v}^0_i$
repeat for $n = 0, 1, 2, ...$, number of VFSA iteration
| Compute the current temperature t_n using a cooling schedule
| Draw proposal models $\mathbf{m}^{n+1}_{res}$, and $\mathbf{m}^{n+1}_{vel}$ using equation (8)
| Compute memberships $\mu^{n+1}_{i,k}$ using $FCM(\mathbf{m}^n_{res}, \mathbf{m}^n_{vel}, \mathbf{v}^n_i)$
| Update cluster centers $\mathbf{v}^{n+1}_i$ using equation (14)
| Compute $\delta E = \Phi^{n+1} - \Phi^n$
| **If** $\delta E \leq 0$
| Accept $\mu^{n+1}_{i,k}$, $\mathbf{v}^{n+1}_i$, $\mathbf{m}^{n+1}_{res}$, $\mathbf{m}^{n+1}_{vel}$
| $\mathbf{m}^n_{res} = \mathbf{m}^{n+1}_{res}$; $\mathbf{m}^n_{vel} = \mathbf{m}^{n+1}_{vel}$
| **else**
| $P = \exp\left(-\frac{\delta E}{t_n}\right)$
| Draw a random number $\{u : u \in (0, 1)\}$
| **If** $P > u$
| Accept $\mu^{n+1}_{i,k}$, $\mathbf{v}^{n+1}_i$, $\mathbf{m}^{n+1}_{res}$, $\mathbf{m}^{n+1}_{vel}$
| $\mathbf{m}^n_{res} = \mathbf{m}^{n+1}_{res}$; $\mathbf{m}^n_{vel} = \mathbf{m}^{n+1}_{vel}$
| **end**
| **end**
| n = n+1
end

We have normalized individual cost functions for CSEM and seismic with their respective target misfit and do not use relative weights.

6 Test case

We apply the proposed joint-inversion workflow for CSEM and seismic travel-time data generated on a subset of SEAM Phase I model (Pangman, 2007). The subsurface model built on the SEAM Phase I data set mimics a realistic geology of a salt-containing region in the Gulf of Mexico (Fehler & Keliher, 2011). There is a massive salt body with steep flanks embedded into a layered sediment environment. The velocity model has complex geometrical structures and strong velocity variations that makes seismic imaging below salt: challenging. The top boundary of the salt is rugose and has a thin layer of muddy salt having velocity slightly lower than the main salt body. Since we use using only first-arrival travel-time data for this numerical test, we restrict our area of interest to 4 km depth. The SEAM model for this test is a subset of a 2D slice of the original 3D model (at north=23900 m) having the dimension of 35 $km \times 4$ km. The model has a seawater layer of 0.3125 Ωm vertical resistivity and 1490 m/s p-wave velocity, and the thickness of the seabed varies from 0.7269 km to 1.606 km. The true synthetic models for this experiment are shown in figure (1). The sediments on either side of the salt body have some interesting formations, which are not visible in the velocity model but are prominent in the vertical resistivity model. A preliminary cluster analysis of this crossplot between true model parameters shows that the geology of the model can reasonably be described with five clusters. We'll assume these cluster centers as a prior geological information about the facies in the model. The goal of this numerical test is to perform the joint-inversion of seismic and CSEM data over this SEAM model using the given petrophysical and geological constraints and quantify the uncertainty in the estimated models.

For the seismic data, we assumed a typical ocean bottom seismometers (OBSs) profile with 34 receivers uniformly distributed every 1 km, with the seismic wavefield downward extrapolated to the seafloor (Arnulf et al., 2011, 2014). For the seismic modeling we took advantage of the source receiver reciprocity. As such we are modeling 34 shots, uniformly distributed at the ocean bottom between $x = 1$ km to $x = 34$ km and receivers at 50 m interval. For CSEM modeling, we have used 17 sources between $x = 3$ km and 31 km and receivers at every 500 m. The CSEM source is an $x-$ oriented horizontal electric dipole, which is towed 30 m above the seabed, and receivers are at seabed depth. We use two frequencies 0.1 Hz and 0.25 Hz and set their corresponding maximum offset to 10 km and 8 km, respectively. Forward modeling for both methods have been done on regular grids (200 $m \times 100$ m); however, for the inversion, we use a sparse parameterization approach. That is, we interpolate models on 400 randomly generated points for VFSA inversion. Once a model is accepted, we transform the model back to orthogonal grid for forward computations. The interpolation of the model on the sparse grid uses a linear radial basis interpolation. Choice of number of points for the sparse parameters is a trade-off between how well it can capture the features of the model and how long does it take for VFSA algorithm to converge.

Since the water layer is known as a prior, we perturb models only below that. For sparse parameterization, we fix 400 inversion points (same for both models) for one chain and do scattered data interpolation to transform the perturbation to regular modeling grids.

For each VFSA chain, the sparse parametrization was randomly generated (see supplementary information). As such, each starting model sampled a different spatial location of the model space. For this experiment, we have run 15 different chains (the initial models and inversion points are shown in the supplementary material). For FCM parameters, we assume 4 clusters (not including the water layer) in the model and provide prior centers $\mathbf{g}_i$ (with prior weight $\lambda = 1000$) as deduced from the true models. For a real dataset, these centers would be inferred using the prior knowledge about the subsurface. Figure (2) shows the resistivity and velocity model recovered in one-chain of the joint-inversion. The probabilistic nature of the joint-inversion workflow allows us to generate a number of models, which can be used to compute uncertainty in the model via statistical analysis. We compute mean, median, and uncertainty in the joint-inversion for fifteen independent chains of VFSA for 3000 iterations. Figure (3a) and (3b) show

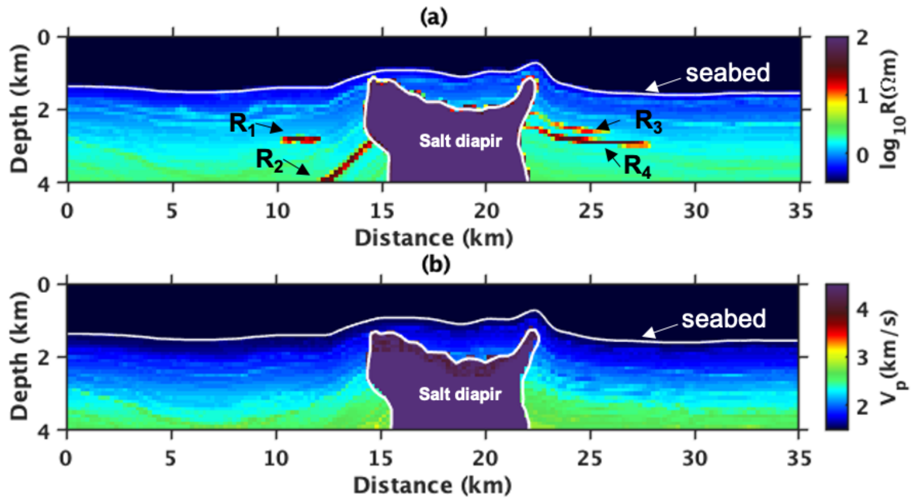

Figure 1. A 2D slice of SEAM Phase I model (a) vertical resistivity (R) on log scale and (b) p-wave velocity (V_p). The model consist of a salt diapir in a sedimentary basin below the sea-floor (white line). There are thin reservoirs on the both side of the salt (R_1, R_2, R_3, and R_4) diapir, which are visible only in the vertical resistivity model.

the mean and median resitivity models. The top of the main salt diaper and the flanks are recovered. The reservoir R_1 is also clearly visible however R_2 and the left flank of the salt are not clearly resolved. Similarly, the reservoirs on the right-side of the salt R_3 and R_4 are recovered together and not clearly distinguished. The background sediments are well-recovered. Due to the lack of EM signal in the bottom corners as well as inside the salt body, we see higher uncertainties in those areas as shown in figure (3c) and (3d). We notice that the reservoirs in the inverted models are slighly deeper than their location in the true resistivity model. As far as the velocity model is concerned, the top boundary of the salt is well resolved. The salt boundary is clearly visible as shown in the mean and median models in figure (3e) and figure (3f), respectively. The background sediments are well recovered except the bottom corners and lower part of the salt, which is due to lack of rays passing through these areas. Given that we started with random initial models, the estimated models from the joint-inversion show excellent agreement with the true synthetic models.

The joint-inversion framework allows us to manually decide the weight on the prior cluster centers by adjusting the value of the parameter λ. A smaller value of λ shows less prior constrains and final cluster centers are mostly recovered through the inversion. A higher value of λ, on the other hand, does not let the centers in the proposal model to move too far away from the prior center by forcing a high prior constrains. For example, figure (4a) shows the cross-plot between velocity and resistivity of the true synthetic model clustered by using FCM with five centers. Using these centers as priors, figure (4b) and figure (4c) show the recovered petrophysics from the joint-inversion with $\lambda = 10$, and $\lambda = 1000$.

Figure (5) shows the posterior probability density of five vertical profiles in both estimated resistivity (top row) and estimated p-wave velocity model (bottom row). Assuming the estimated values at each location (not the estimated models themselves) in all the chains have the Gaussian distribution, the PPD has been computed using histograms. In the resistivity models, the profile at $x = 12$ km passes through the reservoir R_1 between 2.7$-$3.0 km depth. The uncertainty at the top of R_1 is less than that of the bottom of R_1, which means that the upper part of R_1 is better resolved than the lower part. The vertical profile at $x = 14$ km passes through a part of the reservoir R_2 between 3.0−3.2 km depth. Since R_2 is close to the salt diapir, it is not as well resolved as R_1. The vertical profile at $x = 18$ km passes through the salt diapir between $2.0 - 4.0$ km depth. This profiles shows that the uncertainties are lower near the boundary of the salt and higher as the observation point goes towards the center of the salt. The uncertainty between R_3 ($2.4 - 2.6$ km depth) and R_4 ($2.7 - 2.9$ km depth) in the vertical profile at $x = 24$ km have lower uncertainty bounds, however uncertainties inside the reservoir are relatively higher.

Figure(6) show the convergence of individual (CSEM and seismc) as well as total (joint) cost function for 3000 iterations of 15 different chains of VFSA. The individual costs of CSEM and seismic are normalized by their target errors i.e. 1 and 0.01 respectively. The convergence plots show that the joint-inversion converges in 3000 iteration of VFSA and uses approximately equal weight of individual cost functions. This shows that VFSA is a more affordable alternative to MC or MCMC methods, which require thousands of iterations to reach convergence for posterior analysis.

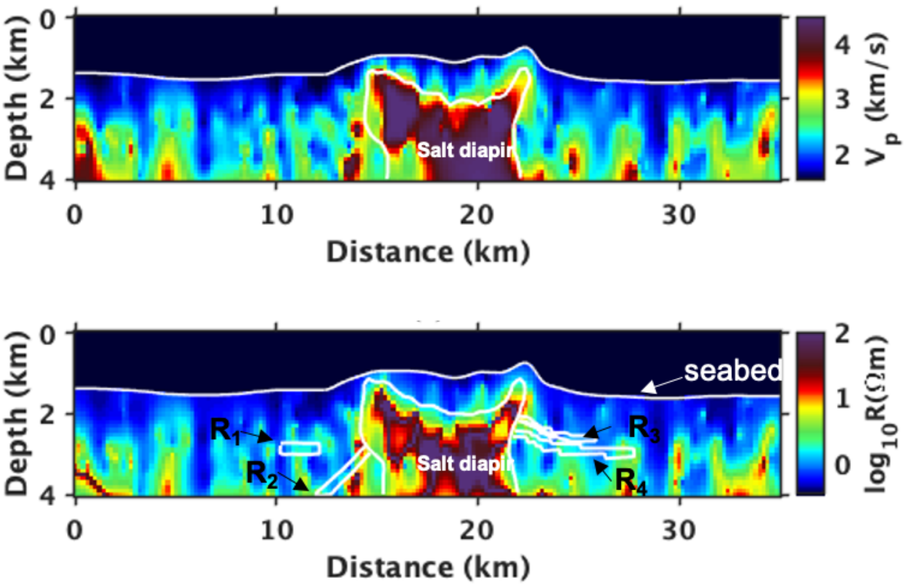

Figure 2. Inversion results for one chain: p-wave velocity (V_p) and vertical resistivity (R) models.

On of the advantage of the joint-inversion of CSEM and travel-time seismic data is that the mean p-wave velocity model estimated from the joint-inversion is smooth and fairly close to the true model, therefore, it can be a good starting model for FWI. We modeled 16 sources placed every 1167 m over the length of the model, and receivers were laid every 50 m over the length of the model. We generated synthetic data using a 4 Hz Ricker wavelet. We used JetPackWaveFD.jl[1] for forward modeling and Optim.jl[2] for FWI. The initial model (mean model from the joint-inversion) is shown in figure (7a) (sampled at 10 m). We run a low frequency FWI at the same initial model sampled at 50 m.

[1] https://github.com/ChevronETC/JetPackWaveFD.jl
[2] https://github.com/JuliaNLSolvers/Optim.jl

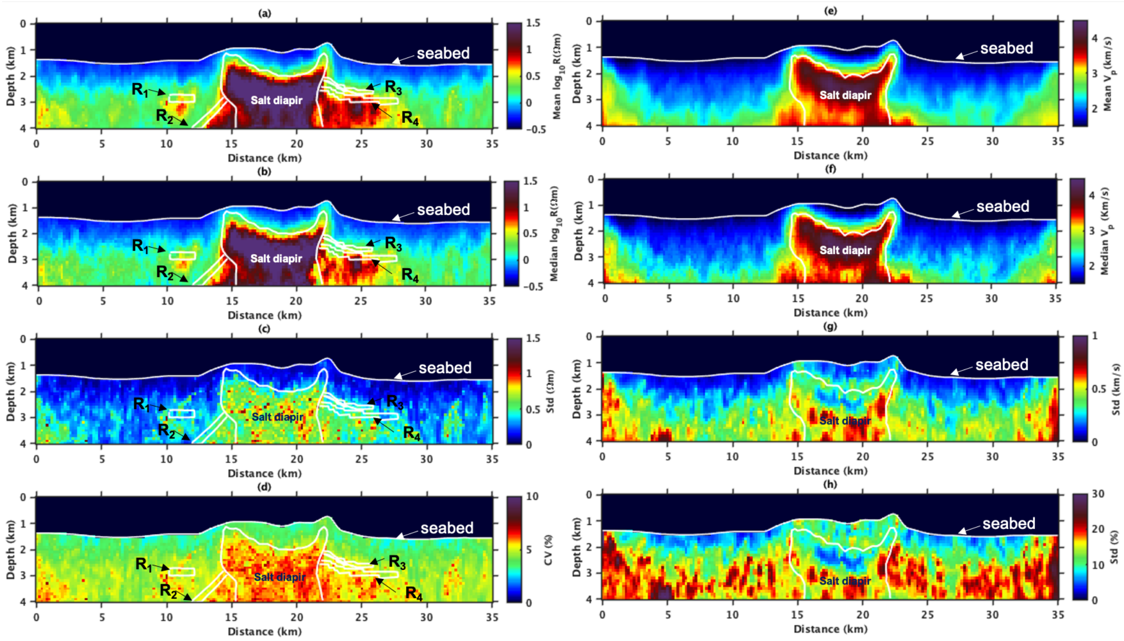

Figure 3. Mean (a), median (b), standard deviation (c) and percentage coefficient of variance (d) for the estimated resistivity models and the mean (e), median (f), standard deviation (g) and percentage standard deviation (h) for the estimated velocity models from the joint-inversion. We notice that the reservoirs in the inverted models are slighly deeper than their location in the true resistivity model.

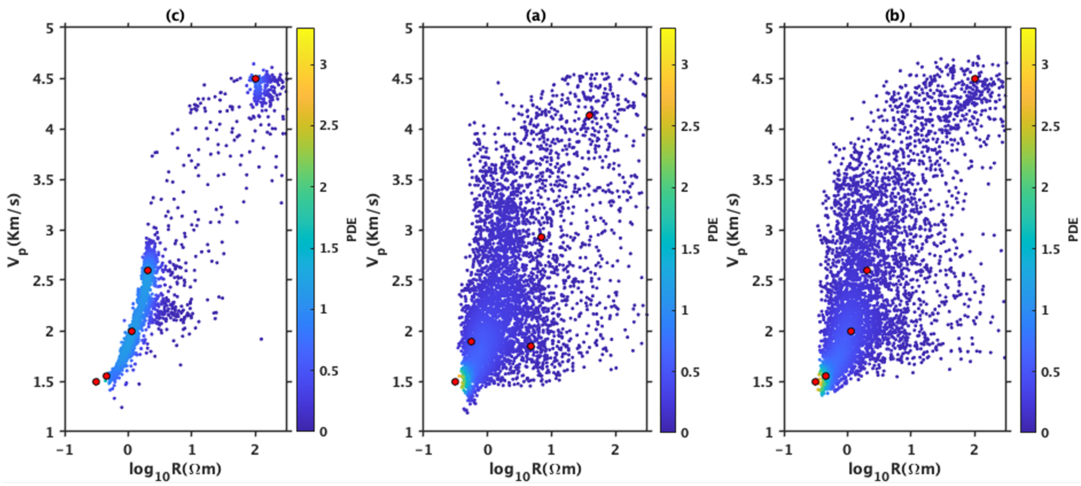

Figure 4. True (a) and recovered petrophysics with $\lambda = 10$ (b), and $\lambda = 1000$ (c). Each point is colored by the probability density estimate (PDE) for each point using kernel smoothing over its nearby points. Red dots represent the prior cluster center included in the joint-inversion as constraints. For a smaller value of λ, the centers (red dots) can move and the distance between the centers obtained from joint-inversion and prior centers are minimized over the iterations. For a larger value of λ, however, the inverted centered are forced near prior centers even in the early stages of the joint-inversion. This figure represents joint-inversion evolution for one-chain.

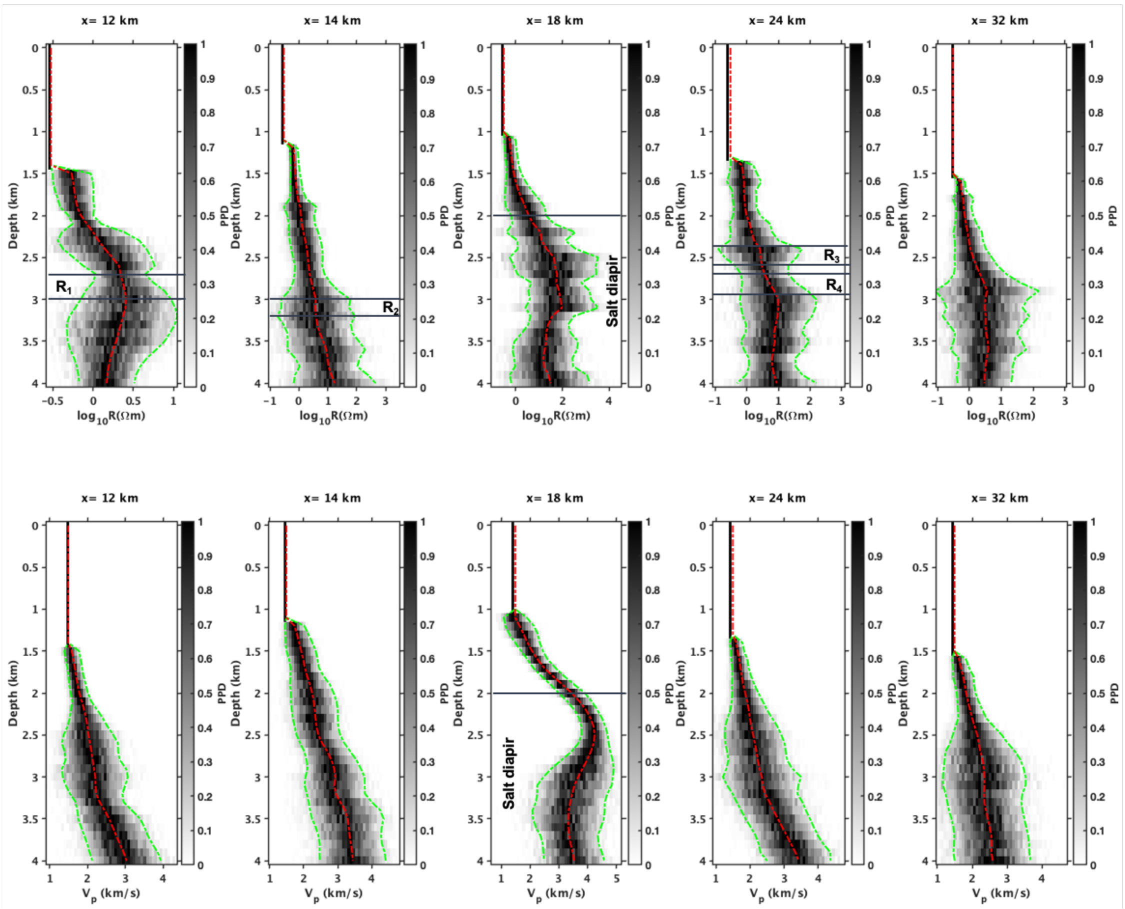

Figure 5. Posterior probability density (PPD) along vertical profiles at locations $x = (12, 14, 18, 24, 32)$ km for resistivity (top row) and the velocity models (bottom row). The red line shows the mean of of the models obtained in 15 different chains and green lines show the upper and lower bound of the PPD (± 2 standard deviation). The uncertainty at the top of R_1 is less than that of the bottom of R_1, which means that the upper part of R_1 is better resolved than the lower part. Since R_2 is close to the salt diapir, it is not as well resolved as R_1. The uncertainties are lower near the boundary of the salt and higher as the observation point goes towards the center of the salt. The uncertainty between R_3 and R_4 have lower uncertainty bounds however uncertainties inside the reservoir are relatively higher.

Figure (7b) shows the recovered model after 25 iterations of FWI and is in very good agreement with the true synthetic model shown in figure (7c). The modeled data and their comparison with true data have been shown in figure (7e). We flip the direction of the residual and modeled data in order to help display the match with the true data. Note that the data modeled in the initial model lacks a lot of reflectivity that is evident in the data modeled in the true model. These missing reflectivities are recovered with FWI. The FWI results from this experiemnt show that the mean p-wave velocity model is a good starting model for FWI.

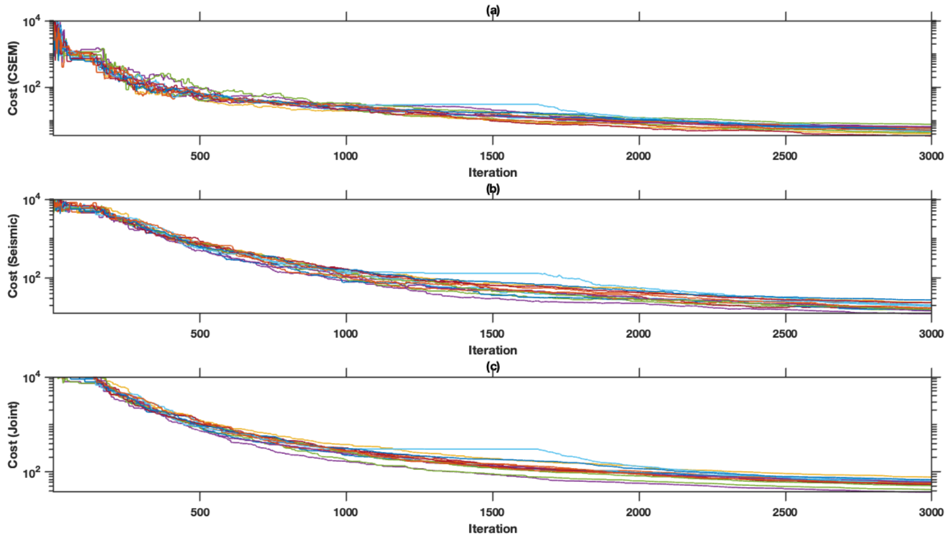

Figure 6. Convergence of (a) CSEM cost function (b) seismic cost function and (c) total cost function in the joint-inversion for 15 VFSA chains. The individual costs of CSEM and seismic are normalized by their target errors i.e. 1 and 0.01 respectively. The convergence plots show that the joint-inversion converges in 3000 iterations of VFSA and uses approximately equal weight of individual cost functions.

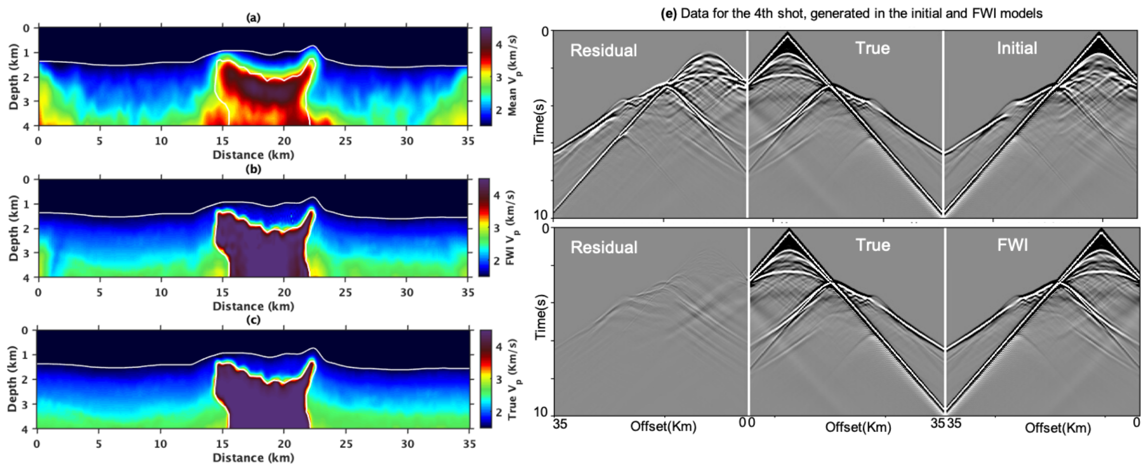

Figure 7. (a) Mean p-wave velocity model from the joint-inversion of CSEM and first-arrival traveltime data, sampled at 10 m grid. The recovered model after 25 iterations of FWI (b) and is in very good agreement with the true synthetic model (c). Comparison (e) of modeled data with true data for the initial model as well as for the FWI model shows that FWI recovered significant missing reflectivities. Note the flip in the direction of the residual and modeled data in order to help display the match with the true data. (e) shows data misfit for one shot.

7 Conclusions

We have proposed a probabilistic workflow for joint inversion and uncertainty estimation, incorporating petrophysical and geological constraints. We applied this workflow to the joint inversion of CSEM and seismic synthetic data from the SEAM Phase I model. The workflow efficiently integrates petrophysical constraints and prior geological knowledge of the model. With better priors, such as facies interpreted from existing well logs, one can assign a significantly higher prior weight, causing the joint inversion to more rigorously honor the geological information.

We have demonstrated that VFSA with sparse parameterization converges faster and enables the affordable computation of multiple chains, which in turn provide uncertainty estimates in the model. The generalized FCM approach can accommodate different distance measures, which are necessary for efficient clustering based on the statistical relationships between model parameters. We believe that VFSA offers a good trade-off between the speed of deterministic methods and the robustness of computationally expensive sampling techniques.

However, we acknowledge that the required number of iterations, although significantly fewer than those needed for MCMC methods, are still substantially higher than those typically used in deterministic approaches. For 3D inversion, a more practical application of this stochastic joint inversion approach would be to estimate starting models for deterministic inversion methods given the high computational cost of the forward solvers.

Open Research Section

The resistivity and velocity models used in the test case can be openly accessed from Fehler and Keliher (2011). A MATLAB function for fuzzy c-means clustering used in this paper is freely available at (Balasko et al., 2005). An open-source code of the full-waveform inversion used can be freely accessed at: https://github.com/ChevronETC/Examples.
A julia implementation for VFSA joint inversion will be available at
https://github.com/JuliaGeophysics/MultiphysicsInversion.jl

Acknowledgments

The research was funded by TOTAL E and P, Houston, USA. The first author was partially supported by Research Council of Finland

References

Arnulf, A., Harding, A., Kent, G., Singh, S., & Crawford, W. (2014). Constraints on the shallow velocity structure of the lucky strike volcano, mid-atlantic ridge, from downward continued multichannel streamer data. *Journal of Geophysical Research: Solid Earth*, *119*(2), 1119–1144.

Arnulf, A., Harding, A., Kent, G., & Wilcock, W. (2018). Structure, seismicity, and accretionary processes at the hot spot-influenced axial seamount on the juan de fuca ridge. *Journal of Geophysical Research: Solid Earth*, *123*(6), 4618–4646.

Arnulf, A., Singh, S., Harding, A., Kent, G., & Crawford, W. (2011). Strong seismic heterogeneity in layer 2a near hydrothermal vents at the mid-atlantic ridge. *Geophysical Research Letters*, *38*(13).

Balasko, B., Abonyi, J., & Feil, B. (2005). Fuzzy clustering and data analysis toolbox. *Department of Process Engineering, University of Veszprem, Veszprem*.

Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms.* USA: Kluwer Academic Publishers.

Blatter, D., Key, K., Ray, A., Gustafson, C., & Evans, R. (2019). Bayesian joint inversion of controlled source electromagnetic and magnetotelluric data to image freshwater aquifer offshore new jersey. *Geophysical Journal International*, *218*(3), 1822–1837.

Bosch, M., & McGaughey, J. (2001). Joint inversion of gravity and magnetic data under lithologic constraints. *The leading edge*, *20*(8), 877–881.

Bosch, M., Meza, R., Jiménez, R., & Hönig, A. (2006). Joint gravity and magnetic inversion in 3d using monte carlo methods. *Geophysics*, *71*(4), G153–G156.

Chen, J., Hoversten, G. M., Vasco, D., Rubin, Y., & Hou, Z. (2004). Joint inversion of seismic avo and em data for gas saturation estimation using a sampling-based stochastic model. In *Seg technical program expanded abstracts 2004* (pp. 236–239). Society of Exploration Geophysicists.

Colombo, D., & Rovetta, D. (2018). Coupling strategies in multiparameter geophysical joint inversion. *Geophysical Journal International*, *215*(2), 1171–1184.

Constable, S. (2010). Ten years of marine csem for hydrocarbon exploration. *Geophysics*, *75*(5), 75A67–75A81.

Constable, S., Orange, A., & Myer, D. (2019). Marine controlled-source electromagnetic of the scarborough gas field—part 3: Multicomponent 2d magnetotelluric/controlled-source electromagnetic inversions. *Geophysics*, *84*(6), B387–B401.

Dunn, J. C. (1973). A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters.

Fehler, M., & Keliher, P. J. (2011). *Seam phase 1: Challenges of subsalt imaging in tertiary basins, with emphasis on deepwater gulf of mexico.* Society of Exploration Geophysicists.

Gallardo, L. A., & Meju, M. A. (2004). Joint two-dimensional dc resistivity and seismic travel time inversion with cross-gradients constraints. *Journal of Geophysical Research: Solid Earth*, *109*(B3).

Gustafson, D. E., & Kessel, W. C. (1979). Fuzzy clustering with a fuzzy covariance matrix. In *1978 ieee conference on decision and control including the 17th symposium on adaptive processes* (pp. 761–766).

Haber, E., & Oldenburg, D. (1997). Joint inversion: a structural approach. *Inverse problems*, *13*(1), 63.

Hertrich, M., & Yaramanci, U. (2002). Joint inversion of surface nuclear magnetic resonance and vertical electrical sounding. *Journal of Applied Geophysics*, *50*(1-2), 179–191.

Ingber, L. (1989). Very fast simulated re-annealing. *Mathematical and computer modelling*, *12*(8), 967–973.

Jardani, A., & Revil, A. (2009). Stochastic joint inversion of temperature and self-potential data. *Geophysical Journal International*, *179*(1), 640–654.

Jaysaval, P., Shantsev, D., & de la Kethulle de Ryhove, S. (2014). Fast multimodel finite-difference controlled-source electromagnetic simulations based on a schur complement approach. *Geophysics*, *79*(6), E315–E327.

Jegen, M. D., Hobbs, R. W., Tarits, P., & Chave, A. (2009). Joint inversion of marine magnetotelluric and gravity data incorporating seismic constraints: Preliminary results of sub-basalt imaging off the faroe shelf. *Earth and Planetary Science Letters*, *282*(1-4), 47–55.

Kaikkonen, P., & Sharma, S. (1998). 2-d nonlinear joint inversion of vlf and vlf-r data using simulated annealing. *Journal of Applied Geophysics*, *39*(3), 155–176.

Key, K. (2016). Mare2dem: a 2-d inversion code for controlled-source electromagnetic and magnetotelluric data. *Geophysical Journal International*, *207*(1), 571–588.

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *science*, *220*(4598), 671–680.

Koketsu, K., & Nakagawa, K. (2002). Joint inversion of refraction and gravity data for the three-dimensional topography of a sediment–basement interface. *Geophysical Journal International*, *151*(1), 243–254.

Lelièvre, P. G., Farquharson, C. G., & Hurich, C. A. (2012). Joint inversion of seismic traveltimes and gravity data on unstructured grids with application to mineral exploration. *Geophysics*, *77*(1), K1–K15.

Lu, X., & Farquharson, C. G. (2020). 3d finite-volume time-domain modeling of geophysical electromagnetic data on unstructured grids using potentials. *Geophysics*, *85*(6), E207–E226.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, *21*(6), 1087–1092.

Moorkamp, M., Lelièvre, P. G., Linde, N., & Khan, A. (2016). Integrated imaging of the earth: Theory and applications. , *218*.

Moser, T. (1991). Shortest path calculation of seismic rays. *Geophysics*, *56*(1), 59–67.

Newman, G. A., & Alumbaugh, D. L. (1995). Frequency-domain modelling of airborne electromagnetic responses using staggered finite differences1. *Geophysical Prospecting*, *43*(8), 1021–1042.

Pangman, P. (2007). Seam launched in march. *The Leading Edge*, *26*(6), 718–720.

Rosas-Carbajal, M., Linde, N., Kalscheuer, T., & Vrugt, J. A. (2014). Two-dimensional probabilistic inversion of plane-wave electromagnetic data: methodology, model constraints and joint inversion with electrical resistivity data. *Geophysical Journal International*, *196*(3), 1508–1524.

Roy, L., Sen, M. K., Blankenship, D. D., Stoffa, P. L., & Richter, T. G. (2005). Inversion and uncertainty estimation of gravity data using simulated annealing: An application over lake vostok, east antarctica. *Geophysics*, *70*(1), J1–J12.

Roy, L., Sen, M. K., McIntosh, K., Stoffa, P. L., & Nakamura, Y. (2005). Joint inversion of first arrival seismic travel-time and gravity data. *Journal of Geophysics and Engineering*, *2*(3), 277–289.

Santos, F. M., Sultan, S., Represas, P., & Sorady, A. E. (2006). Joint inversion of gravity and geoelectrical data for groundwater and structural investigation: application to the northwestern part of sinai, egypt. *Geophysical Journal International*, *165*(3), 705–718.

Sen, M. K., & Stoffa, P. L. (1996). Bayesian inference, gibbs' sampler and uncertainty estimation in geophysical inversion 1. *Geophysical Prospecting*, *44*(2), 313–350.

Sen, M. K., & Stoffa, P. L. (2013). *Global optimization methods in geophysical inversion.* Cambridge University Press.
Shamsipour, P., Marcotte, D., & Chouteau, M. (2012). 3d stochastic joint inversion of gravity and magnetic data. *Journal of Applied Geophysics*, *79*, 27–37.
Shen, W., Ritzwoller, M. H., Schulte-Pelkum, V., & Lin, F.-C. (2013). Joint inversion of surface wave dispersion and receiver functions: a bayesian monte-carlo approach. *Geophysical Journal International*, *192*(2), 807–836.
Streich, R. (2009). 3d finite-difference frequency-domain modeling of controlled-source electromagnetic data: Direct solution and optimization for high accuracy. *Geophysics*, *74*(5), F95–F105.
Sun, J., & Li, Y. (2012). Joint inversion of multiple geophysical data: A petrophysical approach using guided fuzzy c-means clustering. In *Seg technical program expanded abstracts 2012* (pp. 1–5). Society of Exploration Geophysicists.
Sun, J., & Li, Y. (2016a). Joint inversion of multiple geophysical and petrophysical data using generalized fuzzy clustering algorithms. *Geophysical Journal International*, *208*(2), 1201–1216.
Sun, J., & Li, Y. (2016b). Joint inversion of multiple geophysical data using guided fuzzy c-means clustering. *Geophysics*, *81*(3), ID37–ID57.
Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, *49*(8), 1259–1266.
Wéber, Z. (2018). Probabilistic joint inversion of waveforms and polarity data for double-couple focal mechanisms of local earthquakes. *Geophysical Journal International*, *213*(3), 1586–1598.
YANG, H., WANG, J.-L., WU, J.-S., YU, P., & WANG, X.-M. (2002). Constrained joint inversion of magneto-telluric and seismic data using simulated annealing algorithm. *Chinese Journal of Geophysics*, *45*(5), 764–776.
Yee, K. (1966). Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media. *IEEE Transactions on antennas and propagation*, *14*(3), 302–307.