**Regularized full-waveform inversion with automated salt-flooding**

<table>
<thead>
<tr>
<th>Journal</th>
<th>Geophysics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>GEO-2018-0146.R3</td>
</tr>
<tr>
<td>Manuscript Type</td>
<td>Technical Paper</td>
</tr>
<tr>
<td>Keywords</td>
<td>full-waveform inversion, Salt, salt dome, subsalt</td>
</tr>
<tr>
<td>Area of Expertise</td>
<td>Seismic Inversion</td>
</tr>
</tbody>
</table>
Regularized full-waveform inversion with automated salt-flooding

Mahesh Kalita¹, Vladimir Kazei¹, Yunseok Choi², and Tariq Alkhalifah¹

¹Department of Physical Science and Engineering, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia.

E-mail: mahesh.kalita.2@kaust.edu.sa; vladimir.kazei@kaust.edu.sa; tariq.alkhalifah@kaust.edu.sa

²Formerly, Department of Physical Science and Engineering, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia; Presently, Korea Institute of Geoscience and Mineral Resources, Petroleum and Marine Research Division, Daejeon, South Korea.

Email:ychoi@kigam.re.kr

(March 20, 2019)

Running head: FWI for salt-bodies

ABSTRACT

Full-waveform inversion (FWI) attempts to resolve an ill-posed non-linear optimization problem in order to retrieve the unknown subsurface model parameters from seismic data. Generally, FWI fails to obtain an adequate representation of models with large high-velocity structures over a wide region, like salt bodies and the sediments beneath them, in the absence of low frequencies in the recorded seismic signal, due to non-linearity and non-uniqueness. We alleviate the ill-posedness of FWI associated with datasets affected by salt bodies using model regularization. We split the optimization problem into two parts: first, we minimize the data misfit and the total variation in the...
model, seeking to achieve an inverted model with sharp interfaces; and second, we minimize sharp
velocity drops with depth in the model. Unlike conventional industrial salt flooding, our proposed
technique requires minimal human intervention and no information about the top of the salt. Those
features are demonstrated on datasets of the BP 2004 and Sigsbee2A models, synthesized from a
Ricker wavelet of dominant frequency 5.5 Hz and minimum frequency 3 Hz. We initiate the inver-
sion process with a simple model in which the velocity increases linearly with depth. The model is
well retrieved when the same constant density acoustic code is used to simulate the observed data,
which is still one of the most common FWI tests. Moreover, our technique allows us to reconstruct
a reasonable depiction of the salt structure from the data synthesized independently with the BP
2004 model with variable density. In the Sigsbee2A model, we manage even to capture some of the
fine layering beneath the salt. In addition, we show the versatility of our method on a field dataset
from the Gulf of Mexico.
INTRODUCTION

Velocity models are very challenging to build from seismic data, especially for complex salt provinces. Various factors, such as data enrichment by strong multiples, poor sampling due to limitations in the acquisition process, and incomplete subsurface illumination caused by velocity complexities in the overburden, contribute to the failure of conventional computational practices. As of now, there remain many and varied pitfalls at every stage of the building process for velocity models of salt-affected datasets. Attempts to identify and then mitigate these challenges often require both geophysical skills and geological interpretation (Farmer et al., 1996; Jones, 2012; McCann et al., 2014; Jones and Davison, 2014). The conventional industrial practice ‘migrate-pick-flood’ (Dellinger et al., 2017) is subject to the skill level of human professionals, is reiterative by nature, and often exacerbates the already difficult process of velocity model building, thus affecting the quality of the subsurface images. In addition, it impedes the ultimate goal of seismic tomography: replacing human interpretation with computational methods.

Full-waveform inversion (FWI) aims to recover the unknown model parameters from seismic data by minimizing the misfit between observed and modeled data (Lailly, 1983; Tarantola, 1984; Pratt, 1990). Despite continued improvements in computational power over recent years, FWI still requires that model perturbations be evaluated based on a single scattering assumption, known as the Born approximation. As a result, FWI fails to render the non-linear quantitative parameters from the seismic data, given that the initial model is very far from the true subsurface structure. This limitation is aggravated in the presence of a salt body about whose position and shape the initial model most likely has no a priori knowledge, and whose physical properties (e.g. high velocity and low density) are completely distinct from those of its neighborhood. Therefore, the Born assumption breaks down in this scenario, especially in the absence of ultra low frequencies, resulting in a failure
of the salt and subsalt delineation through FWI. Thus, FWI cannot invert the complicated datasets
acquired in salt and subsalt provinces without ultra low frequencies (Shen et al., 2018; Ovcharenko
et al., 2018b) or a good starting model.

There are few studies in the literature on FWI for salt-bodies, signifying the hidden difficulties
associated with this problem. Recently, Lewis et al. (2012) attempted to invert for salt geometries
using the level-set approach. The authors split the domain of inversion into salt and sediment-
ary regions and focused only on the salt geometries. Unfortunately, this distinction requires the
sedimentary velocity to be well known and, therefore, to remain stationary throughout the update
process. The only non-zero velocity difference between the true and initial models is around the
salt boundary.\textsuperscript{1} However, the limitation due to the parametrization of the level-set method in con-
junction with FWI was mitigated by Kadu et al. (2017), who attempted a joint inversion in order
to retrieve the (simple) background\textsuperscript{2} and salt geometry. The authors acknowledged that the joint
estimations might not be feasible in practice due to the difficulties associated with more realistic
and complicated salt models. Ovcharenko et al. (2018a) attempted to retrieve the salt-body model
using a variance-based method. However, the method was comprised of several adaption stages and
was heavily dependent on the selection of parameters. Esser et al. (2016a) proposed using hinge-
loss-based asymmetric constraints in the FWI realm via primal-dual optimization. In this study, we
push the envelope by proposing a new nonlinear procedure that mimics industrial flooding, but in a
way that is more mathematically robust. Our proposed regularization term, which is primarily gov-
erned by a parameter, $p$, automates the flooding procedure and favors the vicinity of the salt-dome.

Thus, we reduce the impact of human intervention, and do not require any assumptions about the
\textsuperscript{1}The differences between sedimentary velocities of the true and initial models in Lewis and Vigh (2016) is not zero.
The authors prepare the initial model by smoothing and subsequently expanding the true sediment velocity. Although
misplaced, the initial model nevertheless has a priori knowledge about the salt body.
\textsuperscript{2}The examples considered in Kadu et al. (2017) have a very simple sedimentary velocity, i.e., a linearly increasing
which is parametrized by its slope.
sedimentary background model or a priori information about the location of the salt dome. With regard to the flooding part, the inequality constraint term of Esser et al. (2016a) represents a special case of our proposed penalty term.

In this article, we attempt to mitigate the perennial question of seismic industry: can the velocity models of salt and subsalt provinces be retrieved from seismic datasets without human intervention? We propose regularizing FWI by penalizing the velocity drops with depth. To do so, we split the minimization problem into two cascading stages: first, we match the data, preserving the sharp edges of the model; and second, we flood from the top of the salt (without picking), which is a post-processing optimization procedure. This split approach is extremely convenient, because it is easy to implement in existing FWI schemes. Moreover, multiple regularization weights can be efficiently tested in the flooding. We begin the article with a theory section on FWI with model regularization, which retrieves the sharp edges of the salt body and subsequently, floods it. Next, we discuss the gradient direction behavior associated with the flooding process. We demonstrate the versatility of our proposed method on the synthetic datasets of the BP 2004 and Sigsbee2A models in which the lowest available frequency is 3 Hz. We begin the inversion process with a very poor model whose velocity increases linearly with the depth and that contains no information of the salt body. The model is well reconstructed when the same constant density acoustic code is used to simulate the observed data. Moreover, our technique still manages to delineate most of the BP 2004 velocity model from the data synthesized independently with the variable density model. Also, we demonstrate its application on a complicated field dataset (for FWI) from Gulf of Mexico.
FWI has summoned enormous attention in the past decade, owing to its potential to recover unknown model parameters from seismic data (Pratt et al., 1998; Sigure and Pratt, 2004; Virieux and Operto, 2009). Despite intrinsic difficulties such as cycle-skipping issues, computational overburden, and so on, the capability of FWI to recover parameters from field datasets has been confirmed by numerous studies, e.g., Vigh et al. (2009); Plessix et al. (2010); Prieux et al. (2013); Warner et al. (2013); Ramos-Martinez et al. (2013); Choi and Alkhalifah (2015); Kalita and Alkhalifah (2017); Routh et al. (2017); Oh et al. (2017). Despite its success, FWI, like many geophysical-estimation problems, is an ill-posed mathematical problem that is operated on seismic field data corrupted by noise and inaccurate experiments. Several studies have used model regularization schemes to alleviate this ill-posedness by reinforcing some geologically plausible characteristics in the inverted model. Total-variation (TV) is one of the most popular regularization techniques in seismic tomography (Anagaw and Sacchi, 2012; Loris and Verhoeven, 2012; Esser et al., 2016b). TV is a non-quadratic regularization method that imposes sharp edges without over-smoothing the reconstructed model (Rudin et al., 1992). Alternatively, Kazei et al. (2017) applied a regularization based on minimum gradient support (MGS) (Portniaguine and Zhdanov, 1999) and the Sobolev space norm \(W^2_p\) in order to maintain a balance between the sharp and smooth updates in FWI. Here, we implement two regularization terms in the context of a salt-body inversion: one that preserves the edge and another that penalizes for the velocity drop, as discussed in the following section.

**Problem statement**

Typically, salt bodies are steeply-dipping complex shaped structures with higher velocity and lower density compared to its sedimentary neighborhood. Therefore, an edge-preserving technique, such
as TV, MGS, or $W^1_2$, is required to capture the sharp boundaries. The accuracy of their position depends on the overburden velocities. For this reason, the bottom of the salt-body and the subsalt are more challenging to recover than the top salt due to notable erroneous velocity estimations in the vicinity of the salt body. Unfortunately, in the absence of sufficient low enough frequencies and a good starting model, the gradient evaluation of the FWI misfit function, often continues to update the velocity in the wrong directions because of its sinusoidal nature of update. This is especially pronounced beneath the top of the salt, as described in Esser et al. (2016a). Anticipating this undesired behavior, we solve two cascading minimization problems to determine the velocity field from the salt-affected dataset, where the output of first problem is the input for the second problem, and vice versa. Mathematically, they are formulated as follows:

**Problem 1 : FWI+TV**

$$J_{FWI+TV}(m) = \min_m \{ \| d_0 - d_m \|^2_2 + \lambda_{TV} \| m \|_{TV} \} \quad (1)$$

$d_0$ : observed data

d_m : synthetic data that follows the wave equation

$$\| m \|_{TV} = |\nabla m| = \sqrt{m_x^2 + m_z^2} \text{ with}$$

$$\nabla m = (m_x, m_z) = \left( \frac{\partial m}{\partial x}, \frac{\partial m}{\partial z} \right) \quad (2)$$

$\lambda_{TV}$ : regularization coefficient of TV regularization;

**Problem 2 : Flooding**

$$J_F(q) = \min_q \{ \| q - u \|^2_2 + \lambda_F \| q \|_F \} \quad (3)$$
\[
\mathbf{u} : \arg \min_{\mathbf{u}} J_{FWI+TV}
\]

\[
\| \mathbf{q} \|_F : \text{flooding regularization defined as }
\]

\[
\| \mathbf{q} \|_F = \frac{1}{2p} \| (|q_z| - q_z)^p \|_2^2 \quad \text{with} \quad p > 1/2
\]  

\[
\lambda_{TV} : \text{regularization coefficient of } \| \mathbf{q} \|_F.
\]

The first minimization problem, ‘FWI+TV’, recovers the velocity field by promoting limited variation regions in the model. Here, \( \lambda_{TV} \) is a positive regularization parameter that controls the trade-off between the data misfit term and the amount of interface-preservation in the inversion process. The second optimization problem, flooding, starts with the final output of ‘FWI+TV’ and penalizes the velocity drop in the model. We call it flooding because the velocity field is smeared along the vertical direction after many iterations. In addition to the advantage of minimal human intervention, flooding is applied anywhere and everywhere since it is governed by the model under consideration instead of the top of the salt. Thus, the high velocity is smeared across any region where there is a drop in the velocity field, rather than only the salt body region. This might have a detrimental effect in the presence of low-velocity zones, especially over the top of the salt. In that case, a minimal human intervention might be required to flood from the top of the salt. However as mentioned later, the \( p \) value in equation (4) plays the most significant role in the automated selection of a flooding region in the model.

**Gradient evaluation**

In practice, the gradient direction of the data misfit term in FWI is evaluated using the adjoint-state method (Lailly, 1983; Plessix, 2006). The gradient of the TV functional (equation (2)) cannot be
computed at the origin \( \nabla m = 0 \). Therefore, we are required to introduce a smoothing parameter, \( \epsilon_{TV} \), to the functional as follows:

\[
J_{TV}(m) = \sqrt{m_x^2 + m_z^2 + \epsilon_{TV}}. \tag{5}
\]

Accordingly, its gradient direction is given by

\[
\frac{\delta J_{TV}}{\delta m} = -\nabla \cdot \frac{\nabla m}{J_{TV}}. \tag{6}
\]

The role of \( \epsilon_{TV} \) in the functional is very well known; more details can be found in Kazei et al. (2017). Next, the gradient computation of the flooding regularization, equation (4), is given by

\[
\frac{\delta ||q||_F}{\delta q} = \begin{cases} 
-\frac{(2p - 1) 4p q_z^{2(p-1)} \frac{\partial q_z}{\partial z}}{q_z^{2}} & \text{for } q_z < 0 \\
0 & \text{otherwise.}
\end{cases} \tag{7}
\]

Equation (7) suggests that the gradient direction of the flooding regularization is primarily governed by the quantity \(-q_z^{2(p-1)} \frac{\partial q_z}{\partial z}\). Careful observation of this quantity reveals that the parameter \( p > 1 \) favors the gradient direction towards a region of steeper velocity drops. A value of \( 1/2 < p \leq 1 \) sets \( q_z^{2(1-p)} \) as a denominator in equation (7), thus, producing an impartial gradient, albeit an undesired one in the context of salt flooding. Please refer to the appendix for a detailed derivation of the gradient computation using calculus of variations.
Gradient preconditioning of flooding regularization

We consider a dummy 1D velocity profile, given by the sine function of depth, to analyze the behavior of the gradient direction during flooding regularization. The velocity profile (red) and the update direction (negative gradient) (green) are displayed in Figure 1. As expected, the update is zero while the velocity increases with depth. However, when the velocity drops, the non-zero update is not always positive, which reverses the direction. Anticipating this, we update the velocity field only when the update direction is positive, as shown by the black profile in Figure 1.

Role of $p$

We study the functional behavior of the flooding regularization with respect to $p$ and $q_z$ in Figure 2. We see that the functional given by equation (4) is biased towards a steeper velocity drop for large $p$, and is almost insensitive to smaller decreases in velocity. We verify this observation by evaluating the gradient of the velocity profile shown in Figure 3. The velocity profile contains three noticeable low-velocity zones, marked as (1), (2), and (3) according to their ascending order of the steepness.

For $p = 1$, the update direction responds to even minute changes in the velocity profile. On the other hand, there are two non-zero regions for $p = 2$ and only one for $p = 4$ in the update direction, confirming that the functional with large $p$ is immune to small-scale changes in the velocity field. Therefore, in the context of our study, we maneuver the update direction towards a region of steeper velocity drops by setting $p$ high enough. A steeper drop, fortunately, occurs just below the top of the salt-body for the numerical problems under our consideration.
Implementation strategy

As mentioned above, we decouple the full problem into the two parts: ‘FWI+TV’ and ‘flooding’. Both of them can be solved using any method of the non-linear optimization framework, such as steepest-descent or L-BFGS, etc. We opt for the non-linear conjugate gradient method and provide the pseudo-code in Algorithm 1 and 2.

Algorithm 1 PROBLEM 1 : FWI+TV

Input: $d_{0}$, source wavelet, acquisition geometries
Input: $m_{0}$ : initial velocity model.
Input: Initialize $\lambda_{TV}$ & $\epsilon_{TV}$
Output: $m^{k}$ will be fed back as the initial velocity model for the Algorithm 2

1: $k = 0$
2: while halting criterion false do
3: \[ g_{FW}^{k} \leftarrow \arg\min_{m} ||d_{m} - d_{0}||^{2} \]
4: Normalize $g_{FW}^{k}$
5: \[ g_{TV}^{k} \leftarrow \arg\min_{m} ||m||_{TV} \]
6: Normalize $g_{TV}^{k}$
7: \[ g_{FW+TV}^{k} = g_{FW}^{k} + \lambda_{TV} g_{TV} \]
8: Normalize $g_{FW+TV}^{k}$
9: \[ cg^{k} \leftarrow \text{conjugate gradient of } g_{FW+TV}^{k} \text{ and } g_{FW+TV}^{k-1} \]
10: \[ m^{k+1} = m^{k} + sl cg^{k} \]
11: $k = k + 1$
12: end while

In any non-linear regularization problem, the smoothing parameter $\epsilon$ and the regularization coefficient $\lambda$ control the success and convergence rate of the problem. A large $\epsilon$ leads to a smoother solution, which is undesirable since we want to preserve the sharp edges in a model with high resolution. The regularization coefficient $\lambda$ implies the confidence level between the terms involved in the functional. Therefore, an appropriate selection of $\epsilon$ and $\lambda$ is very important for the effective application of model regularization. Here, we set $\epsilon_{TV}$ to be very small, depending on the model parameters. For example, we set $\epsilon_{TV}$ to 0.1% of $\frac{\nabla m, \nabla m}{3}$.

\[ \frac{a, b}{3} \] signifies the dot product of $a$ and $b$ divided by the total number of elements in the vector $a$. 
Algorithm 2 PROBLEM 2 : FLOODING

Input: $u = m^k$
Input: Initialize $\lambda_F$ & $\epsilon_F$
Output: $q^i$ will be fed back as the initial velocity model for the Algorithm 1

1: $i = 0$
2: $q^0 = u$
3: while halting criterion false do
4:   $g_{first}^i \leftarrow \text{argmin}_{q} ||q^i - u||_2^2$
5:   Normalize $g_{first}^i$
6:   $g_F^i \leftarrow \text{argmin}_{q} ||q^i||_F$
7:   Normalize $g_F^i$
8:   $g_{total}^i = g_{first}^i + \lambda_F g_F^i$
9:   Normalize $g_{total}^i$
10:  $c g^i \leftarrow$ conjugate gradient of $g_{total}^i$ and $g_{total}^{i-1}$
11:  $q^{i+1} = q^i + sl c g^i$ ; $sl$ : step length
12:  $i = i + 1$
13: end while

Although several robust selection procedure methods for $\lambda_{TV}$ exist (Solo, 1999; Lin et al., 2010; Guitton, 2012; Lin and Huang, 2014), we apply the very simple selection procedure of assigning a stationary number for each bandwidth of the data. For example, $\lambda_{TV} = 0.5$ for the first bandwidth, $\lambda_{TV} = 0.3$ for the second bandwidth of data, and so on. Therefore, the data misfit term in ‘FWI+TV’ gradually dominates the inversion, which produces results that are less biased by the initial values of the regularization parameters. In addition, the final stage of the inversion using the full-bandwidth of the data is purely data driven without any model regularization. In other words, we solve only ‘FWI+TV’ with $\lambda_{TV} = 0$ for the full bandwidth of the data. For the second optimization problem, ‘flooding’, we assign $p = 2$ and $\lambda_F = 10 \times \lambda_{TV}$.

Features of our flooding procedure

In this section, we shed some light on the contrasting methodologies in this article and of Esser et al. (2016a). In their study, Esser et al. dealt with the constrained optimization problem as a set of...
‘hard-constraints’, which mathematically can be formulated as:

\[
J(m) = \min_m \{ \| d_o - d_m \|^2 \}, \text{which is}
\]

subject to \( \|m\|_{TV} \leq \tau_1 \) \hspace{1cm} (9)

and \( \| \min(0, m_z) \|_1 \leq \tau_2 \), \hspace{1cm} (10)

where, \( \{ \tau_1, \tau_2 \} \) are the sizes of the TV and hinge-loss (asymmetric TV) norm balls, respectively.

However, both methodologies promote regions of limited variation in the model by implementing TV minimization: ours as a ‘penalty’, and theirs as a ‘hard-constraint’. Here, we choose to apply TV with the simple implementation of the ‘penalty’ term because it does not require any projection methods, unlike the ‘hard constraint’ implementation. Moreover, applying TV as a ‘hard constraint’ reduces the model space and, in the case of a salt-body, it can prevent the inversion from predicting the fine scale features of the model.\(^4\) Both methods start placing progressively more confidence in the data misfit term than the model regularization terms. However, the following distinctions between the two approaches are worth mentioning:

\(^4\) Since the final inverted results should be governed by the data fidelity term more than the model regularization, we believe the both implementations of the model regularization will render the same results, especially from a field dataset.
In contrast to their formulation, we split the optimization problem into two separate segments: ‘FWI+TV’ and ‘flooding’. We activate the flooding regularization on a more mature and refined velocity model in which the top of the salt region has already been reconstructed, rather than at every iteration of FWI as Esser et al. (2016a) do. This concept of a preconstructed top of the salt before flooding is essentially in line with the industrial strategy, but the lack of assumption about the homogeneous velocity inside the salt dome and the picking make it mathematically more robust.

Thanks to our split approach, we can efficiently facilitate the trial and error strategy that is necessary for the computation of the regularization weights in the flooding part without any incurring significant computational costs.

EXAMPLES

**BP 2004 velocity model**

We demonstrate an application of the proposed method to the west part of the BP 2004 velocity model (Billette and Brandsberg-Dahl, 2005). As is often done with this model in inversion tests, we adjust the size to half of its original size, but we keep the aspect ratio unaltered for our study. It contains a simple background with a complex, rugose salt body. There are several low-velocity subsalt anomalies and complex overhangs, which attempt to imitate the deep-water prospect areas in the Gulf of Mexico. Figure 4 displays the true synthetic model of our consideration.

The model of consideration has 801 and 216 grid points along the $x$ and $z$ directions, respectively, at equal intervals of 0.020 km. There are 248 sources placed between $x = 0.5$ km and $x = 15.44$ km at intervals of 0.06 km. All the sources are placed at a depth of $z = 0.02$ km. The data are recorded by a streamer of 16.0 km fixed length, which is placed at the same depth as the
source line for 10 seconds. A Ricker wavelet of 5.5 Hz dominant frequency and 3 Hz minimum frequency is used to synthesize the observed data with the free surface boundary condition. Figures 5(a) and 5(b) show the source wavelet and its frequency content, respectively. Also, we display two shot gathers from sources placed at \( x = 0.5 \) km and \( x = 8.0 \) km in Figures 6(a) and 6(b), respectively. We display the frequency content of the data from Figure 16(b) in Figure 16(c).

We kick off the inversion process with the velocity model displayed in Figure 7. The velocity increases linearly with the depth from the water bottom. More importantly, the model contains no prior knowledge of the salt’s position or velocity. We follow the multiscale approach described in Bunks et al. (1995), starting the inversion at the lower frequency bandwidth and progressively including the higher frequencies. We consider frequencies less than 3.5 Hz, followed by 4 Hz, 6 Hz and full bandwidth of the data in the inversion scheme. Figure 8 shows the final inverted model of the first optimization problem, i.e., ‘FWI+TV’. It manages to retrieve the top of the salt, but fails to retrieve the vicinity of the salt as pointed out by the encircled region in Figure 8. We subsequently subject the final inverted model to flooding. A couple snapshots of the flooding process are displayed in Figure 9. The flooding manages to wipe out the sudden drop in velocity below the top of the salt. However, it pays the price of the flooding at other regions as well, e.g., the areas designated by blue circles in Figure 9(b). We feed the final output of the flooding process back into the ‘FWI+TV’ scheme. The data misfit term rescues the region from over-flooding, and yields a model close to the true one, as displayed in Figure 10. Next, this velocity model is again subjected to the ‘flooding’ optimization problem. As shown in Figure 11, this procedure removes velocity drops, especially beneath the salt; thus, it damages the already reconstructed subsalt zone. However, a subsequent ‘FWI+TV’ retrieves back those velocity drops as shown in Figure 12. As expected, the inversion starts by building the top of the salt correctly, then slowly propagates further throughout its vicinity. Once the salt body has been retrieved, it reattempts to reconstruct the subsalt region. Since the
velocity models shown in Figure 10 and the bottom panel of Figure 12 look similar, we do not stimulate further ‘flooding’. Instead, we move on to ‘FWI+TV’ with $\lambda TV = 0$, i.e., a conventional FWI, to obtain the final inverted model (Figure 13). As indicated by the blue circle in Figure 13, our proposed method is able to retrieve even small intrusions inside the salt dome, which are often referred to dirty salt.\(^5\) A comparison between Figures 13 and 14 confirms that the method proposed in this article reconstructs the salt and subsalt regions much better than FWI alone.

Next, we perform the optimization problem without the inverse crime. We incorporate a variable density (Figure 15) in the simulation for the observed data using a program named `sfawefd2d` from the open-source software package named Madagascar (Fomel et al., 2013). We retain the same source wavelet and the acquisition geometry in order to prepare the dataset. Figures 16(a) and 16(b) display two shot gathers from sources placed at $x = 0.5$ km and $x = 8$ km, respectively. We display the average frequency spectrum of the data from Figure 16(b) in Figure 16(c). However, we replace the $L^2$ norm of data misfit with the normalized cross-correlation objective function (Routh et al., 2011; Choi and Alkhalifah, 2012). By performing our two-problem strategy with constant density acoustic wave-equation as above, we obtain the inverted model (Figure 17). As expected, inversion cannot reconstruct the model as good as in the constant density data case (Figure 13). Nevertheless, we managed to delineate most of the salt velocity model. Thus, it demonstrates the algorithm’s potential adaptability to a realistic setting where the wave propagation is much more complex.

\(5\)In general, industrial manual flooding assumes that the velocity is homogeneous inside the salt dome; hence, it fails to retrieve the dirty salts.

Sigsbee2A

We consider the Sigsbee2A model (Paffenholz et al., 2002) to demonstrate the versatility of our proposed method. However, we reduce the thickness of the water column of the original model by...
1.5 km. Figure 18 shows the true model of our consideration. Besides the complex salt structure with its variable size and shape, this model contains a sedimentary sequence broken up by a number of normal and thrust faults. The source wavelet shown in Figure 5(a) is used to prepare the observed data, which is recorded in a 16 km long streamer, starting from 0.16 km at the near offset, placed at a depth of 0.016 km. Unlike the previous example, we set the streamer length to vary with the position of the sources in order to emulate the typical marine acquisition geometry. Figure 19 confirms the poor acquisition coverage towards the left part of the model. We distribute 360 sources between $x = 0.480$ km to $x = 23.52$ km at the surface, $z = 0.016$ km. Figures 20(a) and 20(b) show two shot gathers from sources placed at $x = 3.04$ km and $x = 23.52$ km, respectively. Figure 20(c) displays the average frequency content of the shot-gather shown in Figure 20(b). We prepare the initial model (Figure 21(a)) by increasing the velocity linearly from the water bottom.

The final inverted model, displayed in Figure 21(b), shows that we retrieved the salt-body very well. The inversion tries to render the smooth version of the subsalt, which looks very promising for the imaging purpose. In order to retrieve the reflectors in high resolution, we run basic FWI on the model displayed in Figure 21(b) by preconditioning the gradient to zero whenever the velocity exceeds 4.2 km/s. Figure 21(c) displays the high resolution result. An enlarged area from Figure 21(c) is shown in Figure 22 for comparison. By taking into account the innate limitations of FWI, such as poor subsurface illumination, and limited acquisition aperture, we are able to retrieve most of the model.

**Field data application**

This example contains a vintage 2D marine seismic dataset acquired in a region of Gulf of Mexico (GOM) during the decade of 1990. The data were recorded using a short streamer of length 4.8 km,
containing 180 receivers with the near offset of 0.1 km. The lowest usable frequency in the dataset is around 7.0 Hz. We consider the model to be 21.6 km long and 4.5 km deep. Figure 23 displays one shot gather from the source placed at \( x = 11.0 \) km and its frequency content.

We start the inversion process with a model shown in Figure 24. We keep the water layer velocity unchanged throughout the inversion scheme. As expected, ‘FWI+TV’ scheme retrieves the position of the top of the salt (Figure 25(a)). Next, this model is subjected to the ‘flooding’ optimization problem. It removes the velocity drops beneath the recovered top of the salt as shown in Figure 25(b). A couple of additional snapshots of inversion procedure are shown in Figure 25 to portray the evolution of the model. Finally, we retrieve the inverted model as shown in Figure 26. We manage to recover the top and bottom boundaries of the salt dome. Unlike the previous synthetic examples of consideration, in particular of BP 2004 model (Figure 12), this seismic dataset did not manage to wipe out the undue flooding as shown by the blue arrows in Figure 26 because of its innate limitations such as very limited offset (only 4.8 km) and the absence of low frequencies (7.0 Hz). Low wavenumber components in the update are needed to correct the flooding influence automatically.

We validate the inversion result by the following QC measures:

- **Data comparison:** We interleave observed and predicted seismic data from a source located at \( x = 11.0 \) km. Figure 27(a) shows the matching of data from the initial model and Figure 27(b) from the final model. The inverted model produces a data set which resembles the observed recording better than the initial model does.

- **Migration image comparison:** As shown in Figure 28, we observe improvements in the reverse time migration (RTM) image using the final inverted model. The initial velocity

\[^6\text{In general, this type of acquisition geometry and the available frequency content of the dataset are not best suited for successful FWI given a poor starting model.}\]
model fails to delineate the bottom of the salt region as shown in Figure 28(a), whereas the final velocity model clearly accentuates the shape of the salt-dome as shown in Figure 28(b).

- **Angle domain common image gather (ADCIG) comparison:** The erroneous velocities produce non-flat events in ADCIGs. We have computed ADCIGs at an interval of 0.7 km, starting from 6.0 km to 17.0 km in Figure 29. It shows some of the events, as pointed out by the arrows, are flatter with the inverted model compared to those using the initial model.

With this challenging field data example for FWI, we managed to highlight some of the limitations of this flooding based FWI, which is synonymous with the limitations we expect from FWI. A closer look at the inverted model suggests that the bottom of the salt might have been captured. However, the ability of the inversion to correct the velocity below the salt hinged on the availability of long model wavenumber updates capable of reversing the flooding. In this dataset, the lack of low frequencies and the limited available offsets contributed to this limitation. In fact, those limitations can be explained with the help of synthetic examples presented in this article:

- In the case of BP-2004 model, we have 16.0 km long receiver line which in hindsight proves to be long enough for recovering the subsalt region as shown in Figure 13. On the other hand, the 16 km long streamer is not good enough for the Sigsbee 2A model. For instance, we could not remove the over-flooding from the subsalt region as indicated by the ellipses in Figure 21. Under such a thick salt, the data, in spite of the long streamer, could not illuminate the subsalt with enough low model wavenumber energy to undue the flooding. Thus, the correction has had only high wavenumber content and was focused to capture the bottom of the Salt. Unfortunately, the field dataset of consideration has only 4.8 km long spread.

- The observed data for the synthetic examples compared to the GOM field data are richer
in low-frequency bandwidth. Therefore, they enjoy deeper penetration in the model, and, hence offer more in the retrieval process of subsalt region compared to those of the field data example.

- We have shown the versatility of our proposed method on a seismic dataset governed by different physics than the underlying constant density acoustic wave equation of our inversion scheme (Figure 17). To make full use of the limited GOM field dataset, however, we cannot overrule the application of better suitable physics such as including elastic/anisotropic behavior of wave propagation to the inversion machinery. In addition, the utilization of multi-scattered events, which are prevalent in any salt-flank affected dataset, might expedite the recovery process of residual features in the inverted model (Alkhalifah and Wu, 2016). A detailed investigation is required towards that goal.

DISCUSSIONS

Instead of picking the top of the salt body for flooding, in the proposed method we cast the problem into a better, automated mathematical framework that involves a minimization problem with model regularization. We note that our TV and flooding regularizations also require some user intervention to follow the very simple workflow used to select the parameters, which depend on the model properties, though this intervention is negligible compared to the amount involved in industrial flooding. We still manage to retrieve most, if not all, of the model. Alternate efficient algorithms like the split-Bregman method could reduce the dependency of the model on the smoothing parameter and accelerate its convergence to the true solution (Huang et al., 2008; Goldstein and Osher, 2009; Lin and Huang, 2014). However, we stick with the basic implementations of model regularization since the data misfit term, sooner or later, does take over the inversion scheme, which makes the
results less biased by the initial values of the regularization parameters. It is worth mentioning that although TV regularization helps FWI reconstruct missing wavenumbers in the model, it has its own limitations (Alkhalifah et al., 2018; Kazei and Alkhalifah, 2018).

The large computational cost of the first problem, ‘FWI+TV’, which involves more than 1000 non-linear conjugate gradient iterations, appear to be the bottleneck of this method, especially for large 3D problems. Unlike the ‘FWI+TV’ problem, the flooding process requires negligible computational time because it does not require wavefield extrapolation. Other methods, such as adaptive waveform inversion (Warner and Guasch, 2016), FWI with approximate Gauss-Newton methods (Hou and Symes, 2016), multiscale FWI using a flux-corrected transport technique (Kalita and Alkhalifah, 2018, 2019), or modification in the objective function like the optimal transport distance (Métilvier et al., 2016; Yang and Engquist, 2017; Sun and Alkhalifah, 2018) or scattering-angle based preconditioning of FWI gradients (Alkhalifah, 2014; Kazei et al., 2016) have the potential to expedite convergence in the context of inverting for salt-body and remain a topic for future study.

CONCLUSIONS

We have attempted to address the perennial question on feasibility of ‘automated’ velocity model building process, especially for a salt-body-affected seismic dataset. In that regard, we have developed FWI with model regularization. We split the minimization problem into two parts, namely ‘FWI+TV’ and ‘flooding’. The ‘FWI+TV’ promotes the condensed variations in the model obtained from the data misfit function. The ‘flooding’ smears the high velocity vertically in a controlled manner wherever there is a drop. Any undesired flooding will be eradicated by the data misfit term in the subsequent ‘FWI+TV’ stage of our cascading strategy. It is mathematically more robust and less cumbersome than the standard ‘migrate-pick-flood’ method. We have shown the application
of the proposed method to parts of the BP 2004 and Sigsbee2A models, in which the lowest avail-
able frequency of the observed synthetic datasets is 3 Hz. Moreover, we have started the inversion
process with a very simple initial model. Furthermore, we have validated the susceptibility of our
method in an inverse crime free framework. Moreover, we have demonstrated an application of our
methodology to a challenging field dataset.

ACKNOWLEDGMENTS

We would like to thank King Abdullah University of Science and Technology (KAUST) for its
support and all members of the Seismic Wave Analysis Group (SWAG) for fruitful discussions. For
computer time, this research used the resources of the Supercomputing Laboratory and IT Research
Computing at KAUST. We also thank Antoine Guitton for his valuable suggestions.

APPENDIX

This section shows how to derive the gradient of model regularization using variational calculus.

We define the functional as follows:

\[ J(f) = \int dx \, L(x, f(x), f'(x)). \]  

(11)

The functional derivative of \( J(f) \), denoted by \( \frac{\delta J}{\delta f} \), is given by

\[ \frac{\delta J}{\delta f} = \frac{\partial L}{\partial f} - \frac{d}{dx} \frac{\partial L}{\partial f'}. \]  

(12)

If \( L \) doesn’t depend explicitly on \( f \), then first term in equation (12) vanishes and the variational
The gradient becomes

$$\frac{\delta J}{\delta f} = -\frac{d}{dx} \frac{\partial L}{\partial f}.$$  \hfill (13)

**TV**

We consider the TV functional, defined as

$$J_{TV} = \sqrt{m_x^2 + m_z^2}; \quad m_x = \frac{\partial m}{\partial x}; \quad m_z = \frac{\partial m}{\partial z}.$$  \hfill (14)

Therefore, the gradient of (14) with respect to \(m\) (using equation (13)) is given by

$$\frac{\delta J_{TV}}{\delta m} = -\frac{\partial}{\partial x} \frac{\partial \sqrt{m_x^2 + m_z^2}}{\partial m_x} - \frac{\partial}{\partial z} \frac{\partial \sqrt{m_x^2 + m_z^2}}{\partial m_z}.$$  \hfill (15)

Since \(J_{TV}\) is not differentiable at the origin \(\nabla m = 0\), we need to add one parameter, \(\epsilon_{TV}\),

$$J'_{TV} = \sqrt{m_x^2 + m_z^2} + \epsilon_{TV}.$$  \hfill (16)

Therefore,

$$\frac{\delta J'_{TV}}{\delta m} = -\frac{\partial}{\partial x} \frac{m_x}{J'_{TV}} - \frac{\partial}{\partial z} \frac{m_z}{J'_{TV}}$$  \hfill (17)

$$= -\nabla \cdot \frac{\nabla m}{J'_{TV}}.$$  \hfill (18)
Flooding regularization

We consider the following functional that penalizes the velocity drop with increasing \( z \):

\[
|\mathbf{m}|_F = \frac{1}{2p} \| (|m_z| - m_z)^p \|_2^2 ; \quad p > 1/2
\]

\[
= \frac{1}{2p} \int \text{d}x \ (|m_z| - m_z)^{2p}.
\] (19)

Suppose, \( f = |m_z| - m_z \); \( \Rightarrow \) \( \frac{\partial f}{\partial m_z} = \text{sgn}(m_z) - 1 \).

Now, the gradient becomes:

\[
\delta|\mathbf{m}|_F \delta m = -\frac{1}{2p} \frac{\partial}{\partial z} \frac{\partial f^{2p}}{\partial m_z} = -\frac{1}{2p} \frac{\partial}{\partial z} \left[ 2p f^{2p-1} \left( \frac{\partial|m_z|}{\partial m_z} - 1 \right) \right]
\]

\[
= -\frac{1}{2p} \frac{\partial}{\partial z} \left[ f^{2p-1} \left( \frac{\partial|m_z|}{\partial m_z} - 1 \right) \right]
\]

\[
= -\left( \frac{\partial|m_z|}{\partial m_z} - 1 \right) \frac{\partial f^{2p-1}}{\partial z} - f^{2p-1} \frac{\partial}{\partial z} \left( \frac{\partial|m_z|}{\partial m_z} - 1 \right)
\]

\[
= -\left( \text{sgn}(m_z) - 1 \right) (2p - 1) f^{2p-2} \left( \text{sgn}(m_z) - 1 \right) \frac{\partial m_z}{\partial z} - f^{2p-1} \frac{\partial}{\partial z} \text{sgn}(m_z)
\]

\[
= -(2p - 1) f^{2p-2} \left( \text{sgn}(m_z) - 1 \right)^2 \frac{\partial m_z}{\partial z} - f^{2p-1} \frac{\partial}{\partial z} \text{sgn}(m_z)
\]

\[
\Rightarrow \delta|\mathbf{m}|_F \delta m = \begin{cases} 
-(2p - 1) 4^p m_z^{2(p-1)} \frac{\partial m_z}{\partial z} & \text{for } m_z < 0 \\
0 & \text{otherwise}.
\end{cases}
\] (20)

Equation (20) suggests that the magnitude of the gradient for \( p > 1 \) is very large in a region of steeper velocity drops compared to the rest of the model.
REFERENCES


———, 2015, Unwrapped phase inversion with an exponential damping: Geophysics, 80, no. 5, R251–R264.


Review, 8, no. 1, 50.


Paffenholz, J., J. Stefani, B. McLain, and K. Bishop, 2002, Sigsbee 2a synthetic subsalt dataset-image quality as function of migration algorithm and velocity model error: 64th annual international meeting, european association of geoscience engineers: Presented at the Extended Abstracts B.


LIST OF FIGURES

1 A dummy velocity profile (red); its update direction (green) is not free from undesired directions; the black line shows the controlled update.
2 Behavior of the functional $\|q\|_F$ with $p$. At large $p$, the functional is biased towards steeper velocity drops compared to smaller ones.
3 A dummy velocity profile (black) and its update directions with $p = \{1, 2, 4\}$. The large $p$ makes the update direction immune to small-scale changes in velocity.
4 The west part of the BP 2004 velocity model.
5 (a) The Ricker wavelet and (b) its frequency content in order to simulate the observed dataset. The lowest available frequency in the wavelet is 3.0 Hz.
6 Shot records from the sources placed at (a) $x = 0.5$ km and (b) $x = 8.0$ km, respectively; (c) average spectrum of (b).
7 Initial model to invert for the BP 2004 velocity.
8 Results of the first optimization problem (‘FWI + TV’). The blue circle indicates the region where FWI starts failing to update the velocity in the right direction.
9 Snapshots from the second optimization problem (‘flooding’). The flooding procedure succeeded in wiping out the wrong velocity of the reconstructed model of ‘FWI+TV’ (Figure 8). However, we observe undue flooding in some regions, which is indicated by the ellipses in the bottom panel.
10 Output of the ‘FWI+TV’ scheme with the input velocity from the bottom panel of Figure 9. ‘FWI+TV’ removed the residuals of undue flooding, and recovered most, including the subsalt, regions of the model.
11 Result of the ‘flooding’ procedure shown in Figure 10 velocity.
12 (From top to bottom) Evolution of the velocity model using ‘FWI+TV’. We decrease the
regularization coefficient of TV with higher frequency bandwidths.

13 Final reconstructed model of ‘FWI+TV’ with $\lambda_{TV} = 0$ for the full bandwidth of the dataset. We retrieved not only the salt and subsalt regions, but also the dirty salt as indicated by the blue circle.

14 Reconstructed model with no regularization, i.e., conventional FWI with multi-scale approach (Bunks et al., 1995). Compare with Figure 13.

15 Variable density of the BP 2004 model. The density of the salt body is lower than that of the sedimentary neighborhood.

16 Simulated observed data using acoustic wave equation with variable density. The shot gathers for the sources placed at (a) $x = 0.5$ km and (b) $x = 8.0$ km, respectively; (c) average frequency spectrum of data shown in (b). Compare with Figure 6.

17 Final velocity model reconstructed from the observed dataset, displayed in Figure 16.

18 Sigsbee2A velocity model (Paffenholz et al., 2002).

19 Available streamer length for each shot (maximum=16.0 km, minimum= 0.8 km). The acquisition coverage towards the left part of the model is very poor.

20 Shot gathers from the sources placed at (a) $x = 2.9$ km and (b) $x = 23.0$ km; (c) average frequency spectrum of the data shown in (b).

21 (a) Initial model, (b) inverted model, and (c) final inverted model using Figure 21(b) as the initial model, with the gradient preconditioned to zero whenever the velocity is larger than 4.2 km/s. The inversion process could not remove the unwanted over-flooding from the subsalt region as indicated by the ellipses.

22 Enlarged sections of the (a) true, (b) initial, (c) intermediate (Figure 21(b)), and (d) final inverted (Figure 21(c)) models.

23 (a) A shot gather from the source placed at $x = 11.0$ km and (b) its frequency content.
There is no useful signal below 7 Hz.

Initial velocity model to invert available GOM dataset.

Evolution of the velocity model through our proposed method: the left panel shows the output of ‘FWI+TV’ scheme, which are subjected to ‘flooding’ procedure to obtain the right panel.

For ‘flooding’, we assign $p = 4$.

Final inverted model using our proposed methodology. The blue arrows indicate the unexpected over flooding residual.

Shot gather displaying interleaved predicted and observed data using the (a) initial velocity and (b) final velocity model. We interleave 20 traces starting form the predicted followed by observed data. The blue ellipses confirm that the final inverted velocity model predicts a dataset that match the observed ones better than the initial model does.

RTM images using (a) initial velocity and (b) final inverted model. We notice clear boundaries of the salt-dome using the final inverted model, whereas erroneous initial model cannot image those.

ADCIGs ($0^\circ - 30^\circ$) at an interval of 0.7 km, starting from $x = 6.0$ km to $x = 17.0$ km using (a) initial and (b) final velocity models. We notice events are flatter using the final model as indicated by the arrows.
Figure 1: A dummy velocity profile (red); its update direction (green) is not free from undesired directions; the black line shows the controlled update.
Figure 2: Behavior of the functional $\| q \|_F$ with $p$. At large $p$, the functional is biased towards steeper velocity drops compared to smaller ones.
Figure 3: A dummy velocity profile (black) and its update directions with $p = \{1, 2, 4\}$. The large $p$ makes the update direction immune to small-scale changes in velocity.
Figure 4: The west part of the BP 2004 velocity model.

185x75mm (300 x 300 DPI)
Figure 5: (a) The Ricker wavelet and

272x120mm (300 x 300 DPI)
5(b) its frequency content in order to simulate the observed dataset. The lowest available frequency in the wavelet is 3.0 Hz.
Figure 6: Shot records from the sources placed at (a) $x = 0.5$ km and (b) $x = 8.0$ km, respectively; (c) average spectrum of (b).

195x259mm (300 x 300 DPI)
Figure 7: Initial model to invert for the BP 2004 velocity.

185x75mm (300 x 300 DPI)
Figure 8: Results of the first optimization problem ('FWI + TV'). The blue circle indicates the region where FWI starts failing to update the velocity in the right direction.
Figure 9: Snapshots from the second optimization problem ('flooding'). The flooding procedure succeeded in wiping out the wrong velocity of the reconstructed model of ‘FWI+TV’ (Figure 8). However, we observe undue flooding in some regions, which is indicated by the ellipses in the bottom panel.

185x152mm (300 x 300 DPI)
Figure 10: Output of the 'FWI+TV' scheme with the input velocity from the bottom panel of Figure 9. 'FWI+TV' removed the residuals of undue flooding, and recovered most, including the subsalt, regions of the model.

185x75mm (300 x 300 DPI)
Figure 11: Output of the ‘FWI+TV’ scheme with the input velocity from the bottom panel of Figure 9. ‘FWI+TV’ removed the residuals of undue flooding, and recovered most, including the subsalt, regions of the model.

185x75mm (300 x 300 DPI)
Figure 12: (From top to bottom) Evolution of the velocity model using ‘FWI+TV’. We decrease the regularization coefficient of TV with higher frequency bandwidths.

185x228mm (300 x 300 DPI)
Figure 13: Final reconstructed model of 'FWI+TV' with $\lambda_{TV} = 0$ for the full bandwidth of the dataset. We retrieved not only the salt and subsalt regions, but also the dirty salt as indicated by the blue circle.

185x75mm (300 x 300 DPI)
Figure 14: Reconstructed model with no regularization, i.e., conventional FWI with multi-scale approach (Bunks et al., 1995). Compare with Figure 13.

185x75mm (300 x 300 DPI)
Figure 15: Variable density of the BP 2004 model. The density of the salt body is lower than that of the sedimentary neighborhood.

190x75mm (300 x 300 DPI)
Figure 16: Simulated observed data using acoustic wave equation with variable density. The shot gathers for the sources placed at (a) $x = 0.5$ km and (b) $x = 8.0$ km, respectively; (c) average frequency spectrum of data shown in (b). Compare with Figure 6.
Figure 17: Final velocity model reconstructed from the observed dataset, displayed in Figure 16.
Figure 18: Sigsbee2A velocity model (Paffenholz et al., 2002).

184x81mm (300 x 300 DPI)
Figure 19: Available streamer length for each shot (maximum=16.0 km, minimum= 0.8 km). The acquisition coverage towards the left part of the model is very poor.

203x128mm (300 x 300 DPI)
Figure 20: Shot gathers from the sources placed at (a) $x = 2.9$ km and (b) $x = 23.0$ km; (c) average frequency spectrum of the data shown in (b).

195x259mm (300 x 300 DPI)
Figure 21: (a) Initial model, (b) inverted model, and (c) final inverted model using Figure 21(b) as the initial model, with the gradient preconditioned to zero whenever the velocity is larger than 4.2 km/s. The inversion process could not remove the unwanted over-flooding from the subsalt region as indicated by the ellipses.

185x232mm (300 x 300 DPI)
Figure 22: Enlarged sections of the (a) true, (b) initial, (c) intermediate (Figure 21(b)), and (d) final inverted (Figure 21(c)) models.
Figure 23a: (a) A shot gather from the source placed at $x = 11.0$ km and (b) its frequency content. There is no useful signal below 7 Hz.
Figure 23: (a) A shot gather from the source placed at $x = 11.0$ km and (b) its frequency content. There is no useful signal below 7 Hz.
Figure 24: Initial velocity model to invert available GOM dataset.

176x58mm (300 x 300 DPI)
Figure 25P1: Evolution of the velocity model through our proposed method: the left panel shows the output of ‘FWI+TV’ scheme, which are subjected to ‘flooding’ procedure to obtain the right panel. For ‘flooding’, we assign $p = 4$.  

176x58mm (300 x 300 DPI)
Figure 25P2: Evolution of the velocity model through our proposed method: the left panel shows the output of ‘FWI+TV’ scheme, which are subjected to ‘flooding’ procedure to obtain the right panel. For ‘flooding’, we assign $p = 4$. 

176x58mm (300 x 300 DPI)
Figure 25P3: Evolution of the velocity model through our proposed method: the left panel shows the output of ‘FWI+TV’ scheme, which are subjected to ‘flooding’ procedure to obtain the right panel. For ‘flooding’, we assign $p = 4$.

176x58mm (300 x 300 DPI)
Figure 25P4: Evolution of the velocity model through our proposed method: the left panel shows the output of ‘FWI+TV’ scheme, which are subjected to ‘flooding’ procedure to obtain the right panel. For ‘flooding’, we assign $p = 4$. 

176x58mm (300 x 300 DPI)
Figure 25P5: Evolution of the velocity model through our proposed method: the left panel shows the output of ‘FWI+TV’ scheme, which are subjected to ‘flooding’ procedure to obtain the right panel. For ‘flooding’, we assign $p = 4$. 

176x58mm (300 x 300 DPI)
Figure 25P6: Evolution of the velocity model through our proposed method: the left panel shows the output of ‘FWI+TV’ scheme, which are subjected to ‘flooding’ procedure to obtain the right panel. For ‘flooding’, we assign $p = 4$. 

176x58mm (300 x 300 DPI)
Figure 26: Final inverted model using our proposed methodology. The blue arrows indicate the unexpected over flooding residual.

178x58mm (300 x 300 DPI)
Figure 27P1: Shot gather displaying interleaved predicted and observed data using the (a) initial velocity and (b) final velocity model. We interleave 20 traces starting from the predicted followed by observed data. The blue ellipses confirm that the final inverted velocity model predicts a dataset that match the observed ones better than the initial model does.

239x211mm (300 x 300 DPI)
Figure 27P2: Shot gather displaying interleaved predicted and observed data using the (a) initial velocity and (b) final velocity model. We interleave 20 traces starting from the predicted followed by observed data. The blue ellipses confirm that the final inverted velocity model predicts a dataset that match the observed ones better than the initial model does.
Figure 28P1: RTM images using (a) initial velocity and (b) final inverted model. We notice clear boundaries of the salt-dome using the final inverted model, whereas erroneous initial model cannot image those.

364x114mm (300 x 300 DPI)
Figure 28P2: RTM images using (a) initial velocity and (b) final inverted model. We notice clear boundaries of the salt-dome using the final inverted model, whereas erroneous initial model cannot image those.
Figure 29P1: ADCIGs (0°−30°) at an interval of 0.7 km, starting from x = 6.0 km to x = 17.0 km using (a) initial and (b) final velocity models. We notice events are flatter using the final model as indicated by the arrows.
Figure 29: ADCIGs (0°–30°) at an interval of 0.7 km, starting from x = 6.0 km to x = 17.0 km using (a) initial and (b) final velocity models. We notice events are flatter using the final model as indicated by the arrows.
DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.