In-House Developed Full-Waveform Package and Testing on a Foothill Model

*Vats V[1]., Ghosal D[1]., Roy S[2].
[1] Department of Earth Sciences, Indian Institute of Technology, Kanpur
[2] Department of Computer Science and Engineering, Indian Institute of Technology, Kanpur

*vvats@iitk.ac.in

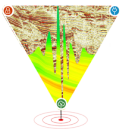
Keywords
Full waveform inversion (FWI), Graphical user interface (GUI), and parallelization.

Abstract
In oil and gas exploration industry, precise subsurface imaging is essential for visualizing underground structures, which ultimately reduce the uncertainties in the selection of drilling sites for exploration wells. However, the task of achieving high-resolution imaging becomes challenging in subsurface regions characterized by complex geological structures and lateral velocity variations. This type of complex structures contributes to a low signal-to-noise ratio (S/N ratio) in seismic data. Complicated geological conditions pose challenges for conventional imaging methods due to the physics on which they are based. Full waveform inversion (FWI) emerges as a powerful technique capable of generating high-resolution images, with a theoretical resolution of half the dominant wavelength, even in the presence of this type of geological feature. Our team has successfully developed a time-domain acoustic full waveform inversion (FWI) code specifically designed to handle the inversion of data for complex subsurface structures. The primary challenge associated with the FWI algorithm is its computational cost apart from non-linearity. In our software package, our central focus has been on optimizing the algorithm through parallelization, achieved by automatically allocating processors and memory to individual shots based on the availability of processors. This dynamic allocation makes the algorithm adaptable to a wide range of computing machines, enhancing its versatility.

Moreover, we've intentionally designed the code in a modular fashion, empowering users to customize the code to their specific needs without disrupting other components. This design flexibility allows users to make the algorithm work optimally for their unique requirements. We have also incorporated a graphical user interface (GUI) into our package, making it easily accessible for users to utilize. To demonstrate the effectiveness of our development, we tested the code on a foothill model developed by SEG advanced modelling (SEAM) phase II corporation that closely resembles the most complicated real-world scenarios. The successful application of our code on this model demonstrates its robustness and capability to handle complex geological scenarios encountered in real-world oil and gas exploration.

Introduction

Seismic imaging is most famous among geophysicists due to its capability to provide highly detailed images of the subsurface. Seismic refraction tomography (Nolet, G. et al., 1987) and reverse time migration (RTM) (Huang, R., et al., 2021) are recognized as two established and widely used conventional methods in the field of seismic imaging. These methods are designed to utilize specific information within the seismogram. For example, in refraction tomography, the focus is primarily on analyzing refracted phases, allowing for the visualization of velocity distribution in the subsurface. RTM is primarily relying on primary reflections to generate a reflectivity image of the subsurface which highlights the reflectivity properties of the subsurface. These methods encounter challenges when applied to complex subsurface conditions, such as areas with gas zones and salt structures. These conditions often produce seismic phases with low signal-to-noise ratios, making it difficult to accurately identify and interpret the desired seismic signals, further complicate the imaging process, as the subsurface lacks sufficient seismic energy to generate clear and reliable images.
Full waveform inversion (FWI) is a deterministic approach that utilizes optimization methods to minimize the discrepancy between observed and predicted seismic data (Tarantola, A. et al., 1984, Virieux, J., et al., 2009). By continuously refining the model parameters, FWI aims to provide a seismic image of the subsurface, this concept was proposed by Tarantola in 1984, and its ability to exhaustively incorporate all possible seismic phases, including reflected, refracted, surface waves, multiples, and



diving waves, within the seismic section set it apart from other conventional imaging methods. This approach encompasses a wide range of seismic information and leads to a more detailed understanding of the subsurface structures. This technique has gained significant attention in recent years and its applications have expanded beyond oil and gas exploration (Huo, Y., et al., 2009), including medical imaging (Guasch, L. et al., 2020), tunnel detection (Llopis, J. L., et al., 2005), archaeology (Köhn, D. et al., 2023), and shallow surface imaging (Borisov, D., et al., 2019).

Various versions of commercial and open-source packages, including Inversion of Full Observed Seismograms (IFOS), Seismic Waveform Inversion Toolbox (SWIT-1.0), PyFWI, PySIT, Devito, and SeisFlows, are available for seismic waveform inversion. These packages predominantly employ MPI for parallelization, which is well-suited for high-performance clusters. However, it's worth noting that this choice introduces certain considerations. MPI-based parallelization may necessitate adjustments when transitioning between different computing architectures, and its explicit communication approach can introduce complexities into the code.

Our team has developed a FWI package that enables the application of this technique on seismic data, resulting in highly detailed images of the subsurface. The core code of our package is developed using the 'C' language. To address the high computational costs associated with FWI, we have implemented a shot parallelization strategy by distributing the computational workload across multiple processors using OpenMP API. This capability allows us to effectively utilize multicore machines for this application and effectively manage substantial volumes of seismic data. Our code is built upon the steepest descent gradient optimization technique, one of the key advantages of our code is its flexibility, as it allows users to choose from multiple objective functions (Gómez, L., et al., 2017). These objective functions can include calculations such as the L2 norm and normalized cross-correlation. For accurate forward modelling (Antunes, A. et al., 2014), our approach utilizes a finite difference operator. To ensure numerical stability, we employ a 6th order discretization in space and a 2nd order discretization in time. To mitigate artificial reflections, we have implemented absorbing boundary conditions (Gao, Y., et al., 2017) at the three edges of the model. These boundary conditions effectively absorb outgoing waves, preventing them from reflecting into the computational domain and causing unwanted artifacts. we have also developed a user-friendly graphical user interface (GUI) for easy implementation of our algorithm on different datasets. The GUI incorporates various features like live plotting of results, real-time progress updates, interactive visualization of the inverted model's quality, and the ability to save outputs.

In this article, we have also conducted an experiment using a synthetic foothill model (Michael Oristaglio, et al., 2013) to showcase the effectiveness of our developed package. This model is known for its complexity, making it a challenging test case. This model encompasses geological features, including substantial lateral velocity variations, subsurface undulations, and multiple fold/faults. Our package excels at capturing features in high detail, even when the initial model deviates significantly from the true model. This sets us apart from conventional imaging methods that struggle to provide such detailed images for complex features.

In this article, section 2 presents an explanation of the methodology, followed by Section 3, which delves into the development of the graphical interface. In Section 4, we present the results obtained from our experiments and engage in discussions. Finally, in Section 5, we draw conclusions based on our findings and summarize this study. Through this structured presentation, we aim to provide an understanding of our research and its implications.

Methodology

The FWI is an iterative process that involves three main steps, as illustrated in Figure (1). The first step is forward modelling, where synthetic seismic data is generated based on the initial model. In the second step, gradient computation, the misfit between the observed and predicted data is quantified to determine the direction of model updates. Lastly, in the model updating step, the current model is adjusted based on the computed gradients, aiming to minimize the misfit between the observed and predicted data. These steps form the core framework of FWI.

Forward modelling: FWI starts with comparing the synthetic and observed seismic data. This synthetic data is prepared by simulating the wave through an initial velocity model. The mathematical expression of the acoustic wave is:

In-House Developed Full-Waveform Package and Testing on a Foothill Model

$$\frac{\partial^2 \boldsymbol{U}}{\partial t^2} = \boldsymbol{V}^2 \left(\frac{\partial^2 \boldsymbol{U}}{\partial x^2} + \frac{\partial^2 \boldsymbol{U}}{\partial z^2} \right) + \boldsymbol{S} \qquad (1)$$

Where, the field $\boldsymbol{U} = u(x, z, t)$, velocity $\boldsymbol{V} = v(x, z, t)$, and source $\boldsymbol{S} = s(x, z, t)$ are at a given location (x, z) and time t.

It is impossible to simulate the wave through the real complicated subsurface structure with the help of analytic methods (Alford, R. M., et al., 1974), but this task can be accomplished with numerical modelling such as finite difference, finite element, pseudo spectral methods. In our code we have took the help of the finite difference method where the model is discretized in small grids and field at every grid is contributed by neighbour grids. We used the 6th and 2nd finite difference order for spatial and temporal discretization of Equation (1). The dimension of the model is restricted by the limited computational resources which is responsible for the artificial reflections from the edge of model, so it is necessary to implement the absorbing boundaries at the edge of model which motivated us to use the improved hybrid absorbing boundary conditions (IHABC) (Liu, Y., et al., 2018).

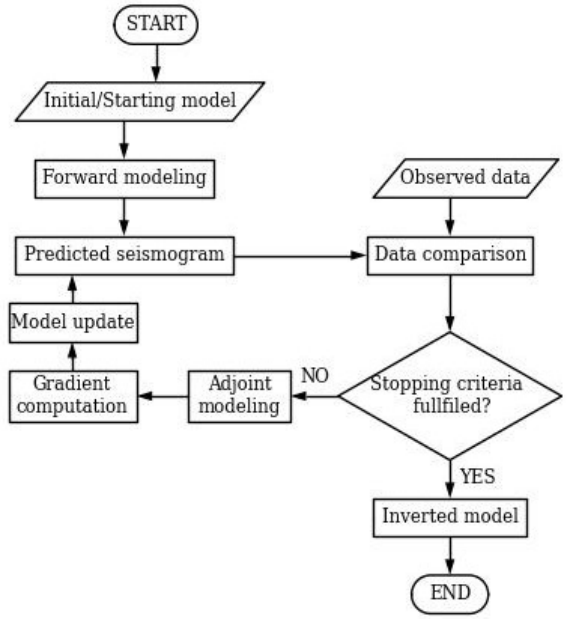


Figure 1: Flow chart of Full waveform inversion (FWI).

Gradient computation: Initial model is updated with the gradient such that the difference between the observed and synthetic data can be minimized. This gradient is calculated by cross-corelated the forward and adjoint field (Plessix, R. E. et al., 2006) according to the Equation 3. The adjoint field is calculated according to Equation 2.

$$\frac{\partial^2 \boldsymbol{A}}{\partial t^2} = \boldsymbol{V}^2 \left(\frac{\partial^2 \boldsymbol{A}}{\partial x^2} + \frac{\partial^2 \boldsymbol{A}}{\partial z^2} \right) + \boldsymbol{R}_{rec,src} \qquad (2)$$

Where, the adjoint field $\boldsymbol{A} = a(x, z, t)\ at\ x\ and\ z\ spatial\ location$, and the residual source $\boldsymbol{R}_{rec,src} = r(t)$ at time t having the same location as receiver for that shot.

$$\boldsymbol{g_n} = \sum_{src} \sum_{rec} \int_0^T dt(\boldsymbol{U}.\ \boldsymbol{\varphi}), \qquad (3)$$

Where, $\boldsymbol{g_n}$ is the gradient correspond to $\boldsymbol{n^{th}}$ shot, $\boldsymbol{U}$ forward propagating wavefield, and $\boldsymbol{\varphi}$ is the field arising from the data residuals propagating backward in time Equation (2).

In our code, there is a flexibility to use either the L2 norm or cross-correlation objective function both have advantage and disadvantage. In Cross-correlation objective function only phases are matched due to which it is sensitive to cycle skipping, amplitude of the signals is not matched in this it makes it useful in real case because sometimes it is hard to get the actual amplitude of the source for real data, at the same time L2 norm match the whole signal including amplitude.

Model updating: The gradient calculated in the last step is scaled with the step length before updating the model such that the convergence gets faster (Schuster, G. T.). This scaled gradient is added to the model parameters in last iteration or initial model. The model is updated according to the Equation 3.

$$\boldsymbol{m}_{n+1} = \boldsymbol{m}_n - \alpha_n \boldsymbol{g_n} \qquad (4)$$

Where, $\boldsymbol{m}_n$ is the model parameter vector at n^{th} iteration, and α is the step length.

This updated model is closer to the actual model. This model updating process is repeated until a stopping criterion is reached.

Interface

In-House Developed Full-Waveform Package and Testing on a Foothill Model

A GUI is designed to enhance the implementation of the FWI algorithm by providing a user-friendly experience as shown in Figure (2). It is divided into four components to streamline the user's interaction and facilitate the execution of FWI:

1. Input console: This panel of the interface serves as the input point for all essential components needed to initiate the inversion process. These include the initial model, observed data, source signature, acquisition geometry, and parameter file.
2. Data Visualization Panel: This allows for real-time monitoring of both the initial and inverted models. Here, users can plot and observe the models as they evolve throughout the inversion process.
3. Progress report and control Panel: The panel serves as a control centre, featuring essential buttons such as start, quit, and restart for managing the inversion process. Additionally, a progress bar is provided to allow users to monitor the live progress of the inversion. Within this window, users can also access important outputs such as receiver information, source information, and updates regarding the misfit. These outputs are displayed in a dedicated text window, ensuring convenient access to relevant information throughout the inversion process.
4. Result Analysis Panel: This enables the user to select a 1D depth-velocity profile. In this window, the velocity profiles for both the initial and inverted results are plotted, allowing for a visual comparison at a specific location within the model. This feature offers a convenient way to assess and analyse the differences between the initial and inverted velocities at a particular position of model, aiding in the evaluation of the inversion process.

The interface also includes a menu bar, offering additional functionalities through its "Tool" and "Help" sections. Under the "Tool" section, users can access tools to generate the acquisition geometry,

which aids in setting up the required input data. Additionally, the "Parameter" section provides options for selecting parameters crucial to the inversion process. These parameters include information about boundary conditions, the type of objective function to be used, and details regarding the processors involved

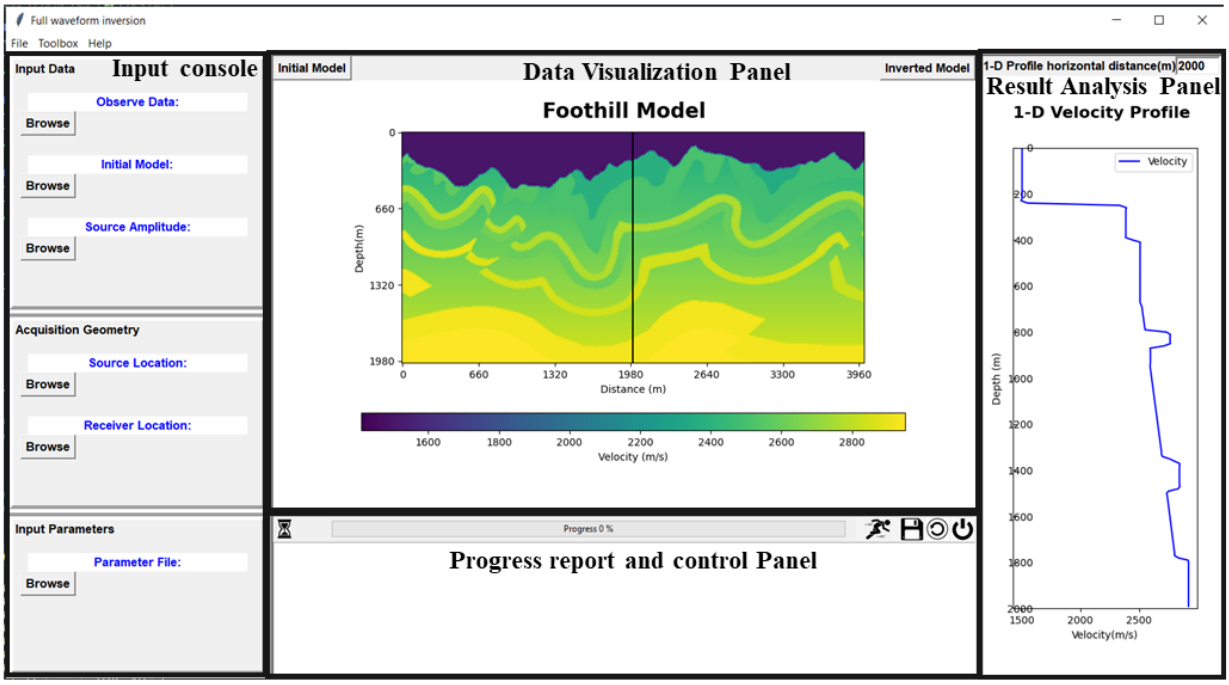


Figure 2: Interface of the Developed Package The figure displays the graphical user interface (GUI) of the developed package, where each section is clearly marked for easy navigation and interaction. Additionally, a visualization of the real foothill model is presented, providing a visual representation of the geological features within the interface.

in the process. These features within the menu bar enhance the flexibility and customization of the inversion process, allowing users to tailor it according to their specific requirements. The "Help" section is available to aid and guidance to users throughout their interaction with the interface.

Results and discussion

To validate the developed package, we utilized a foothill model featured in the SEAM Phase II consortium. This model stands out as one of the most complex models, characterized by significant variations in seabed depths, reaching up to *350 m* having multiple fault/folds and thin beds. In this experiment, our focus is on a specific subset of the model that encompasses significant geological structures. The dimensions of this segment measure *4.0 X 2.0 km*, as depicted in Figure (3a).

The model was discretized into small square grids, with each grid measuring *10 meters* in each direction. To ensure numerical stability and minimize numerical dispersion, the sampling interval for the forward field calculation was set at *1 millisecond* for a duration of *2.5 second*. The choice of sampling interval and grid size was carefully made to maintain the stability conditions and avoid numerical dispersion respectively. The seismic data generation process involved a total of *60* sources and *320* receivers positioned along the surface. In Figure (3a), the sources are denoted by red stars, while the green triangles represent the locations of the receivers. This configuration ensured coverage of all significant features of the subsurface, the inversion process is started with an initial model which was prepared by applying a Gaussian filter to smooth the true model. Figure 3(b) visually represents this initial model. To mitigate the issue of cycle skipping, we implemented a multiscale strategy during the inversion process. This approach involved gradually increasing the frequency range of the data being inverted. Initially, we started with a peak frequency of *5 Hz* and gradually increased it by *2 Hz*. The inversion process involved running *20* iterations for each frequency range, which took approximately *3.5 hours* to complete for a single frequency. When employing parallel processing with *6* processors on an Intel Xeon platform, the entire process required a cumulative runtime of *32.17 hours*. The result of the inversion, conducted at a peak frequency of *21 Hz*, is depicted in Figure 3(c).

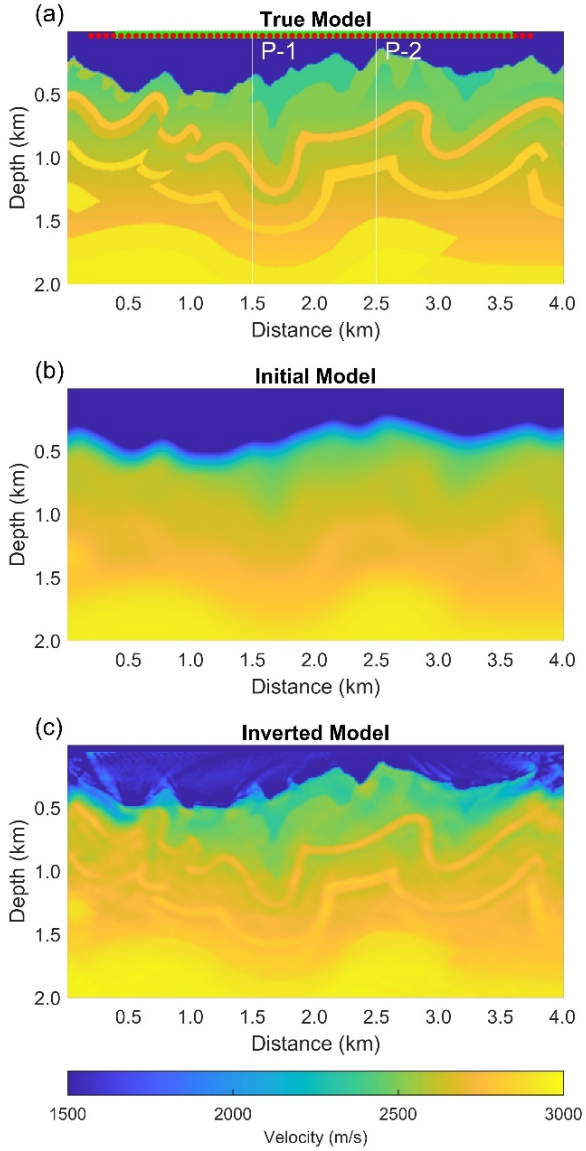


Figure 3: (a) Foothill Model: The figure shows the true foothill model with fold and faults, featuring a maximum seabed undulation of 350 meters. The red star represents the source location, while the green triangle represents the receivers positioned along the surface. Two specific locations, denoted as P-1 and P-2, are selected on the model for comparison of 1D depth-velocity slices from the true, initial, and inverted models. (b) Initial Model: This plot displays the initial model, which is obtained by smoothening the true model using a Gaussian filter. (c) Inverted Model: The figure illustrates the inverted model, demonstrating significant recovery of the geological features.

In-House Developed Full-Waveform Package and Testing on a Foothill Model

To analyze the inverted result in more detail we have selected two 1D velocity-depth slices at depths of *1500 meters* and *2500 meters*. These slices encompassed the initial, true, and inverted models, and were visualized in Figure (4a, b). This figure provided a clear and comparative view of the results, facilitating an assessment of the agreement between the inverted model and the true model. Upon examination, it became evident that our result closely resembled the true model, particularly for shallow depth features. This observation indicates the successful matching of the inverted model with the true model, highlighting the effectiveness of our inversion process in capturing and reproducing the subsurface structures with satisfactory accuracy, especially in shallow depth regions. As part of our analysis, we also computed the Normalized Root Mean Square Error (NRMS) to quantify the improvement achieved in the inversion process. The NRMS was calculated using Equation (5), resulting in values of *13.26%* and *9.3%* for the initial model and inverted model, respectively. These percentages indicate a significant improvement in the inverted model compared to the initial model. The lower NRMS value for the inverted model demonstrates a reduction in the overall misfit, indicating a closer match between the inverted model and the true model. This quantification further supports the effectiveness of our inversion process in achieving a notable enhancement in the accuracy of the results.

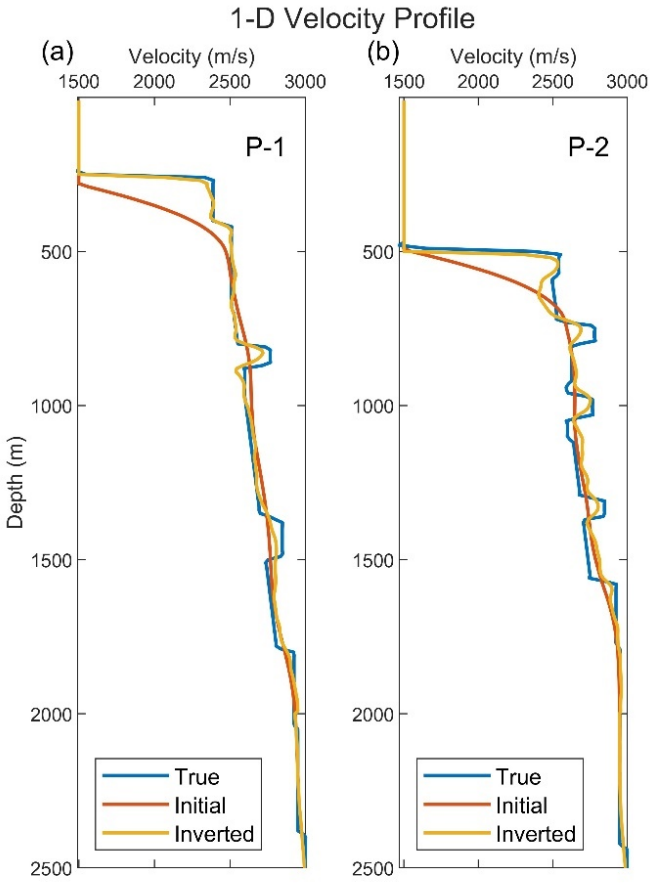


Figure 4: 1D Velocity-Depth Profiles at (a) 1500 m and (b) 2500 m. The true, initial, and inverted sections are represented by the blue, red, and yellow lines, respectively. From this figure, it is evident that the inversion process begins with a smooth initial model, where the geological structures are not clearly visible. However, after the completion of the inversion process, all the features become discernible, indicating the successful recovery of the subsurface structures through the inversion process.

$$NRMS = \frac{1}{\max(\boldsymbol{m}) - \min(\boldsymbol{m})} \sqrt{\frac{1}{N}\sum_{i=1}^{M}\left(\mathbf{m}_{\mathrm{i}} - \boldsymbol{m}_i^{est}\right)^2}, \quad (5)$$

Where, $\mathbf{m}$ is the model parameters, N is the total number of parameters, $\boldsymbol{m}_i^{est}$ is the estimated model, and $\mathbf{m}_{\mathrm{i}}$ is the real model.

Figure 5 presents the seismogram used to assess the quality of the recovered model. To enhance the visibility of low-amplitude phases, specifically the direct arrivals within the first 0.35 seconds for nearby receivers, they have been muted. In Figure 5(a), we observe the seismogram corresponding to the true model, where all seismic phases justify the structures in the foothill model. Notably, Figure 5(b) illustrates the initial data, which lacks significant phases because of low resolution image severs as an initial model for inversion process. However, when employing FWI for model inversion, Figure 5(c) demonstrates the successful restoration of major structures, clearly evidenced by the presence of corresponding phases. To validate the authenticity of these regenerated structures, Figure 5(d) illustrates the differences between the inverted and true data. This comparison shows that the retrieved structures support the observed phases in the true data.

In-House Developed Full-Waveform Package and Testing on a Foothill Model

Conclusion and future work

This study focused on the results obtained using our developed package. We modify the algorithm to operate effectively on OpenMP, incorporating meticulous parallelization strategies to ensure smooth performance even on modest hardware configurations. Furthermore, we embrace a modular framework, partitioning the algorithm into discrete modules, each assigned to distinct segments of the FWI algorithm. This design empowers users to adjust these modules to align with their needs and specifications. The findings with this package indicate that FWI can produce good results even for complex models. The graphical user interface (GUI) of our developed package significantly facilitates its implementation, making it user-friendly and accessible to researchers and

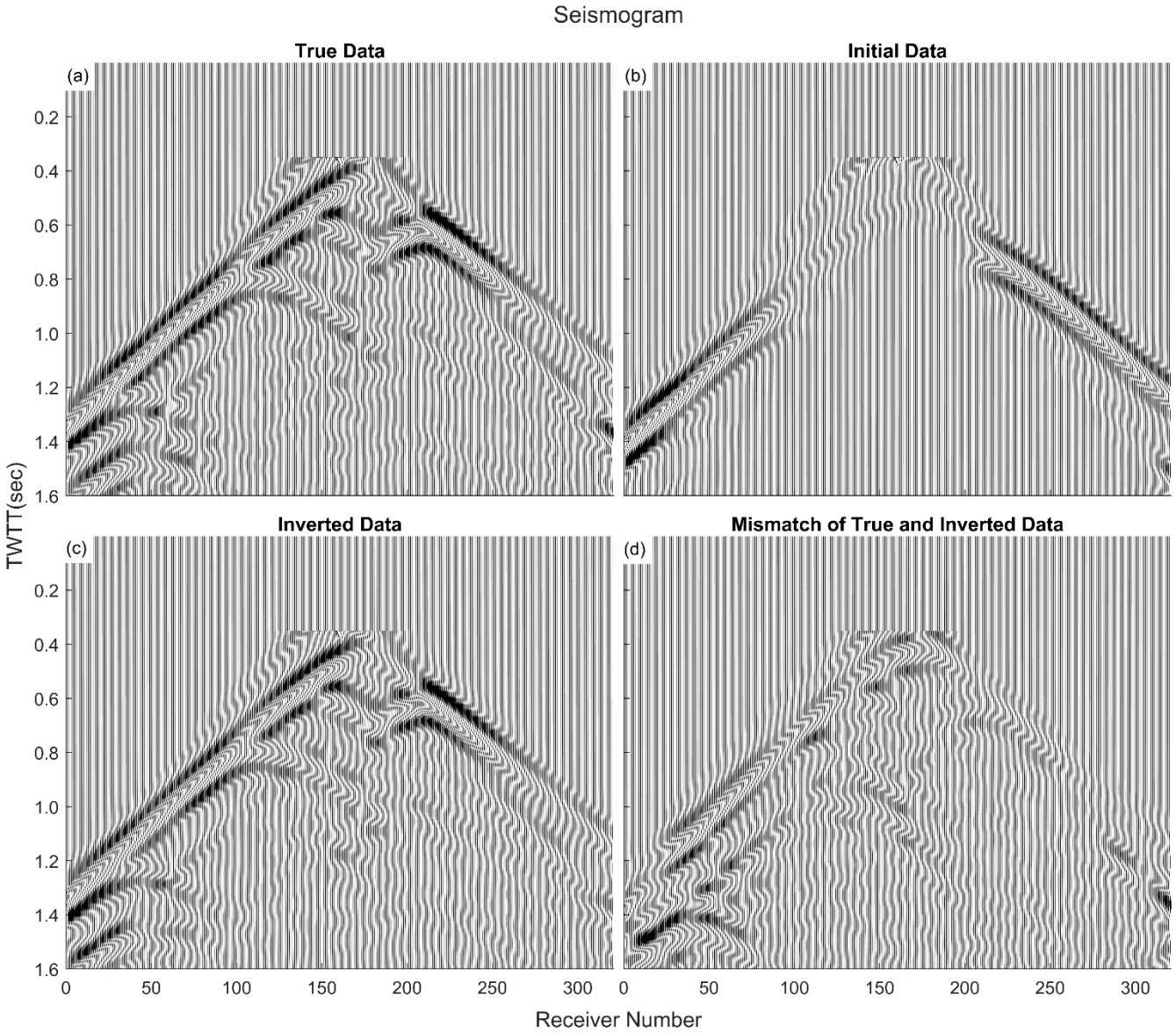


Figure 5: The seismogram of models for a shot located at center of model. Panel (a) displays the seismogram obtained from the true model, while panel (b) represents the seismogram from the initial model. In panel (c), the seismogram corresponds to the inverted model. The mismatch between the seismogram data from the true and recovered model is depicted in panel (d).

practitioners. This work highlights the effectiveness of FWI and the usability of our package in obtaining accurate subsurface model representations, providing valuable insights for various applications in the field of geophysics.

In our plans, we aim to implement hybrid parallelization where we exhaust OpenMP and graphical processing unit (GPU) to deal with the computational cost and enhance the functionality of our interface by incorporating additional modules that allow for data editing. This ongoing development will enable us to utilize our package for a wider range of FWI applications and enable to implement FWI within the feasible time which improve their ability to extract valuable information from various types of geophysical data.

Acknowledgement

We are thankful to ONGC, India for supporting the project financially and providing the datasets. We are especially thankful to Prof. M. K. Sen and P. Jaiswal for their suggestions. The resources provided by the Param Sangana under the NSM, GoI are gratefully acknowledged.

References

Alford, R. M., Kelly, K. R., Recent, B., International, A., Meeting, S. E. G., January, E., & Co, A. P. (1974). *Analytical solution*. *39*.

Antunes, A. J. M., Leal-Toledo, R. C. P., Da Silveira Filho, O. T., & Toledo, E. M. Ã. (2014). Finite difference method for solving acoustic wave equation using locally adjustable time-steps. *Procedia Computer Science*, *29*(December), 627–636. https://doi.org/10.1016/j.procs.2014.05.056

Borisov, D., Gao, F., Williamson, P., Simons, F. J., & Tromp, J. (2019). *Robust surface-wave full-waveform inversion*. 5005–5009.

Gao, Y., Song, H., Zhang, J., & Yao, Z. (2017). Comparison of artificial absorbing boundaries for acoustic wave equation modelling. *Exploration Geophysics*, *48*(1), 76–93. https://doi.org/10.1071/EG15068

Gómez, L., & Pestana*, R. C. (2017). *Full-waveform inversion using alternative objective functions in the presence of noise and uncertainties of source signature*. *1*, 296–301. https://doi.org/10.1190/sbgf2017-058

Guasch, L., Calderón Agudo, O., Tang, MX. *et al.* (2020). Full-waveform inversion imaging of the human brain. *npj Digit. Med.* **3**, 28 https://doi.org/10.1038/s41746-020-0240-8

Huang, R., Zhang, Z., Wu, Z., Wei, Z., Mei, J., & Wang, P. (2021). Full-waveform inversion for full-wavefield imaging: Decades in the making. *Leading Edge*, *40*(5), 324–334. https://doi.org/10.1190/tle40050324.1

Huo, Y., & Zhang, M. (2009). Full waveform inversion of gas hydrate reflectors in Northern South China Sea. *Acta Geophysica*, *57*(3), 716–727. https://doi.org/10.2478/s11600-009-0011-z

Köhn, D., Zolchow, M., Mecking, R., Wilken, D., Wunderlich, T., De Nil, D., & Rabbel, W. (2023). Seismic full waveform inversion in archaeological prospecting. *Advances in On- and Offshore Archaeological Prospection*, 31–40. https://doi.org/10.38072/978-3-928794-83-1/p4

Liu, Y., & Sen, M. K. (2018). An improved hybrid absorbing boundary condition for wave equation modeling. *Journal of Geophysics and Engineering*, *15*(6), 2602–2613. https://doi.org/10.1088/1742-2140/aadd31

Llopis, J. L., Dunbar, J. B., Wakeley, L. D., Corcoran, M. K., & Butler, D. K. (2005). Tunnel detection along the southwest U.S. border. *Proceedings of the Symposium on the Application of Geophyics to Engineering and Environmental Problems, SAGEEP*, *1*(January 2005), 624–637. https://doi.org/10.4133/1.2923487

Michael Oristaglio, (2013), "SEAM Update," *The Leading Edge* 32: 1020-1024.

https://doi.org/10.1190/tle32091020.1

Nolet, G. (1987). Seismic wave propagation and seismic tomography. In: Nolet, G. (eds) Seismic Tomography. Seismology and Exploration Geophysics, vol 5. Springer, Dordrecht. https://doi.org/10.1007/978-94-009-3899-1_1

Plessix, R. E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, *167*(2), 495–503. https://doi.org/10.1111/j.1365-246X.2006.02978.x

Schuster, G. T. (n.d.). *Seismic Inversion* (Issue 20).

Tarantola, A. (1984). in the Acoustic Approximation. *Geophysics*, *49*(8), 1259–1266.

Virieux, J., & Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics*, *74*(6). https://doi.org/10.1190/1.3238367