geoconvention
Calgary • Canada • June 17-19 2024

Toward Realistic Modelling, Imaging and Inversion Testing

Daniel O. Trad and Ivan Sanchez***
**CREWES*
*** Universidad Nacional de Santander*

Summary

Research in academia often suffers from a limitation in the number of data sets employed for testing, resulting in a lack of feedback diversity that is crucial for comprehensive analysis. Applied research necessitates engagement with a broad spectrum of datasets, which significantly enriches research and development projects. Obtaining real datasets for publication in academia is not only challenging but also involves time-consuming preprocessing, making the pursuit of testing diversity a difficult task. Consequently, numerous tests are conducted on modelled data, often generated using similar algorithms employed in inversion processes, thereby giving rise to the "inverse crime scenario". Predominance of synthetic data testing in academia also comes as a consequence of the substantial difference in computational resources with industrial environments. Software developed in academia often lacks the capability in handling intensive computations with large seismic files with irregular acquisitions, capabilities that are required to work with real data sets used in industry. The consequence is a large gap between toy examples used in academia and realistic examples required for industrial use.

This paper details the implementation advancements made in our seismic libraries, illustrating tests aimed at enhancing the reliability of results in diverse environments, including large models, salt environments, topography settings, and physical models. Furthermore, we discuss the characteristics of inverse crime scenarios. The outcome of this study includes the ability to circumvent the inverse crime problem and conduct tests in a variety of environments. Moreover, we anticipate that this type of research will increase collaboration with industry and deepen our understanding of the practical capabilities of novel techniques.

The inverse crime

The development of migration and inversion algorithms starts invariably by creating synthetic data in a simple small model where we know the exact answer. This is essential to create new code that works correctly and to detect subtle implementation errors. Only by knowing the exact answers we can check that every step is correct. We are obligated to proceed with modelled data for most of the development process. The problem, however, is that in complex algorithms development never ends, in particular in academia. There is a tendency to continue testing with synthetic datasets and we have become so used to it that we become oblivious to the difference with real data.

The most important problem we fall into when testing with synthetic data is the inverse crime (IC) scenario (Schuster, 2017). This refers to the case when the data are modelled with the same algorithm that we used for migration and inversion. Since Reverse Time Migration (RTM) and Full Waveform Inversion (FWI) typically have a forward modeling phase with finite differences (FD) to create the source wavefield, it is common to use the same code to create a modeling program. In other methods that do not have an FD engine, like Kirchhoff migration for example, the IC is less common because Kirchhoff uses ray tracing for modeling instead of FD. It is not impossible although not common, however, to see testing Kirchhoff migration by Kirchhoff modeling which is also an inverse crime scenario.

The inverse crime problem is well known but in general not that well understood. Sometimes researchers add random noise to address it but that random noise is not an issue. The following are what we believe are the most important pitfalls that occur as a consequence of the IC.

1.Artifacts and noise generated during the forward modelling become data. For example, artifacts reflecting from the border of the model are equivalent to a geometry that surrounds the model, converting a reflection into a transmission problem, which is significantly easier to solve. There are a few ways to improve on this, for example changing the cell size from the modeling to the inversion steps and therefore making the artifacts irreversible (Trad, 2021).

2.When the algorithm involves an iterative data fitting, like in least squares migration (LSMIG) or FWI, the residuals in the inverse crime scenario will contain information only about the model errors, producing a perfect gradient estimation and therefore facilitating convergence during the inversion. In real data, the residuals contain, in addition to model parameter errors, also deficiencies in the modeling algorithm and acquisition noise that led to gradient errors too. We can improve on this by using a more complete physics for forward modeling program than for inversion and adding realistic coherent noise to the data.

3.Most synthetic tests typically create data in a regular layout with shots and receivers on a surface at constant intervals. In these geometries, each shot illuminates every receiver, leading to very large offsets (as long as the model). Real data have geometries with a shot patch that changes along the survey, with variable shot and receivers intervals and elevations. To work with real data is necessary to make our geometries have an independent shot/receiver position for each seismic trace.

4.Noise in real data can be very difficult to simulate properly with synthetics. Random noise does not represent the complexities of coherent noise generated on the near- surface. One way to address this is by algorithms to simulate Rayleigh waves and scattering noise on the surface as in Sanchez-Galvis (2023).

5.Data contains internal and surface multiples. These can be included in modeling but if we also generate them during the inversion then they become additional data. Although, much work has been done to use multiples as data, using them in the inverse crime scenario is just equivalent to prove that our algorithm is reversible, not that we can actually use them in real cases

Numerical Experiments

One of the ways we can make more realistic tests, and also understand until which extend, we are being affected by the inverse crime problem (IC), is to duplicate tests by running once on external data set, the other in synthetic surveys created on the given geometry (the same as in the external dataset). For example, Figure 1 shows a RTM tests performed on the Pluto salt model (top left). The original data are synthetic but external (modelled by someone else), and our RTM of these data using the smooth velocity of the top left, is shown on the top right. Even when the original data were created also by a finite difference algorithm, there was no or very little inverse crime. The amplitudes, artifacts, source wavelet, noise and other characteristics of the input data are different from what the FD code in our program will generate during migration or inversion. On the bottom left, we show our RTM result with data modelled by ourselves (that is, with IC).

The results are much superior to the results on the no-inverse crime case. In addition to the difference between IC and no-IC, we modelled our data without multiples, so the RTM result can handle correctly all the seismic energy. The external data, on the other hand, seems to have surface and internal multiples. Although RTM could handle multiples, that would be only possible if the migration velocity model had all the sharp interfaces necessary to produce the multiples during the prediction (which is not a realistic case). On the bottom right of the figure, we perform a test generating the data ourselves with multiples, and using a different algorithm than the one used to migrate the data. This results on a partial IC case, but the presence of multiples on the data and not on the RTM shows by blurring the image in similar manner to the non-IC case.

In Figure 2, we illustrate realistic vs IC testing in a different scenario, a Foothills model. On Figure 2a, we see the velocity model after flooding the near surface. This could be thought of a datuming approach, putting the shots and receivers at the top of the model. However, a common IC case is to create the data on the flat datum and migrate from there (Figure2c). A more realistic test is to use the true velocity with topography (Figure 2b) and either create the data from topography using the proper boundary conditions (Cao and Chen, 2018, Sanchez, 2023) to generate near surface noise (Figure2e) or using external data modelled also from topography (Figure2f). It is interesting to compare with Kirchhoff tests (Figure 2c) which are usually done from topography because ray tracing is easier to accommodate to a variable surface than finite differences. Since Kirchhoff uses a very different algorithm than it is used to generate the data, Kirchhoff tests are often more realistic than RTM tests (unless the data were created with Kirchhoff modelling).

Similar considerations can be made for FWI. In fact, this is a more typical case since FWI is an inversion, whereas RTM is not unless is least squares RTM. For FWI we can do a similar comparison for the Foothills model. We can generate the data on the surface of a flooded model (Figure3, top left), which produce the FWI result on the top right. In this case, there was an attempt to remove the IC by using a multigrid approach with data and FWI results generated on different grids (Trad, 2021). However, a realistic test should be made by using external data from surface or generating the data from surface with the proper boundary conditions (Figure3, bottom left, initial velocity, bottom right after 14 iterations of FWI). We see results deteriorating as they get more realistic.

An even better way to make realistic tests is to work with real datasets for which we know the answer. In Figure 4 we present a test on physical modelled data generated in the CREWES physical model facility (Wong et al, 2009). The modelling in this case is done on a plate with two high density bodies on top of it. The acquisition geometry consists of shots and receivers located on rotating arms. A total of seventy shots were generated, with around the same number of receivers. In Figure4a we see one physical shot after muting events below the first arrivals. The goal of the FWI tests is, starting from the constant velocity model shown in Figure 4d, to detect the location of the bodies. To improve the sensitivity of FWI to the bodies, Wong et al. generated a second data set by removing the bodies from the plate (Figure 4b). The difference between the two datasets is shown in Figure 4c). We apply RTM and FWI (Figures 4e and 4f), and in just 4 iterations we can see the bodies with accuracy. This test can be compared to a time-lapse approach for detecting changes in a reservoir.

Although not discussed in this abstract, two key developments were required to perform these tests. a) Adapting open-source libraries to handle SEGy datasets (Trad and Sanchez, 2023) and b) implementing all finite difference algorithms with GPUs using Cuda (Trad, 2021).

Conclusions

In this abstract we have discussed the inverse crime problem and how we can design tests that either have no-inverse crim or when they do, we can understand the implications of the inverse crime by comparing results on different situations. We have look at RTM in marine, salt environments, with/without surface multiples, RTM and FWI in a Foothills model with data from datum/topography, and finally FWI for real data generated in a physical model facility. Understanding the inverse crime problem is a first and necessary step to produce meaningful research.

Acknowledgements

We thank Joe Wong by providing the physical model data and his help with preprocessing, Peliang Yang for RTM and FWI examples in Madagascar, Sam Gray for weekly fruitful discussions, and CREWES sponsors for their continued support over the years. This work was funded by CREWES industrial sponsors, NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 543578-19. We also gratefully acknowledge support from NSERC Discovery Grant and the Emerging Leaders in the Americas Program (ELAP) from Global Affairs Canada for funding the research internship.

References

Cao, J., and Chen, J.-B., 2018, A parameter-modified method for implementing surface topography in elastic- wave finite-difference modeling: Geophysics, **83**, No. 6, T313–T332.

Sanchez-Galvis, I. J., 2023, Modeling and processing near-surface seismic scattering in multicomponent data: An approach based on elastic wave modeling and polarization filtering: Ph.D. thesis, School of Electrical, Electronics and Telecommunications Engineering - Universidad Industrial de Santander.

Schuster, G. T., 2017, Seismic inversion: Society of Exploration Geophysicists.

Trad, D., 2021, Computational aspects of FWI, *in* Geoconvention 2021, GeoConvention, Conference Abstracts.

Trad, D., 2022, Gpu applications for modelling, imaging, inversion and machine learning., *in* Geoconvention 2022.

Trad, D. and Sanchez I., Towards realistic modeling and inversion testing, CREWES reports, 2023.

Wong, J., Hall, K. W., Gallant, E. V., Bertram, M. B., and Lawton, D. C., 2009, Seismic physical modeling at the university of Calgary, *in* SEG Technical Program Expanded Abstracts 2009, Society of Exploration Geophysicists, 2642–2646.

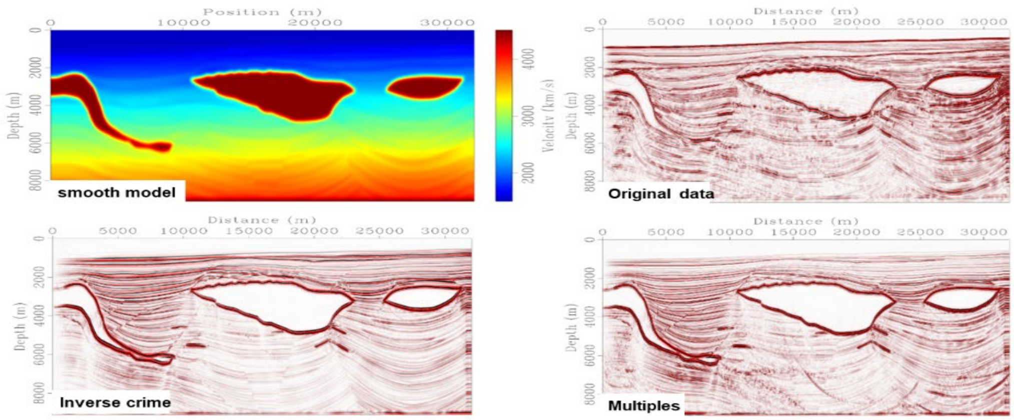

Figure 1-RTM for a Salt model (Pluto) without ('original'), with ('inverse crime') and partial ('multiples') inverse crime. Upper left is the migration velocity model which cannot generate internal multiples.

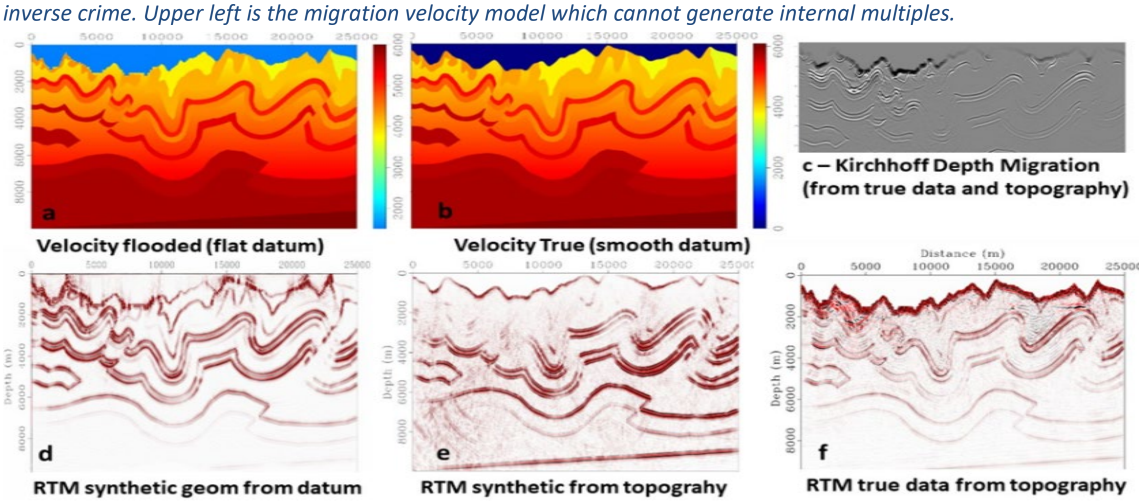

Figure 2- A Foothills model with a) flooded topography (flat datum), b) real topography. c) Kirchhof depth migration from real topography and original data. d) RTM from synthetic data on top of the flooded velocity, e) RTM from synthetic data modelled from the topography. f) External data from topography.

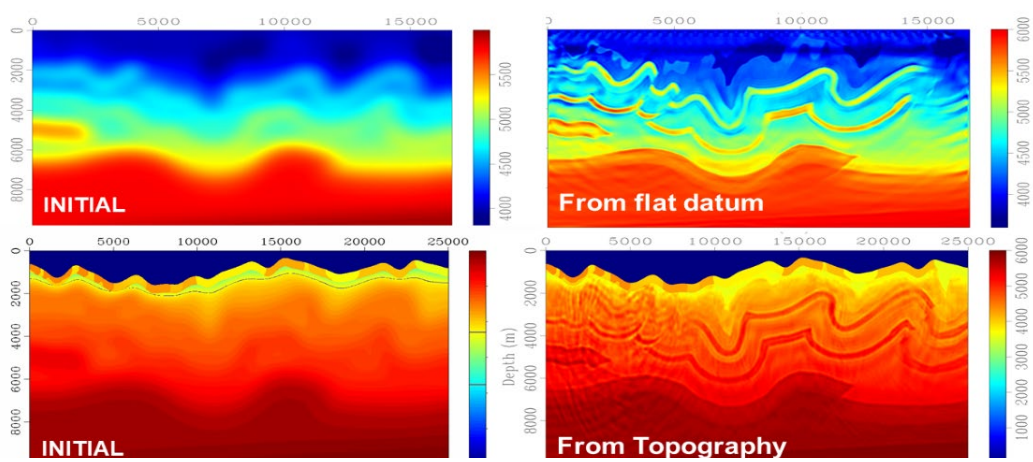

Figure 3 - Differences between FWI from flat datum (upper row) and topography (bottom row). Left column: initial model, Right column: FWI after 10 iterations. Color maps are different because the flooded model starts at 2000m/s, the model from topography starts from 0m/s (air).

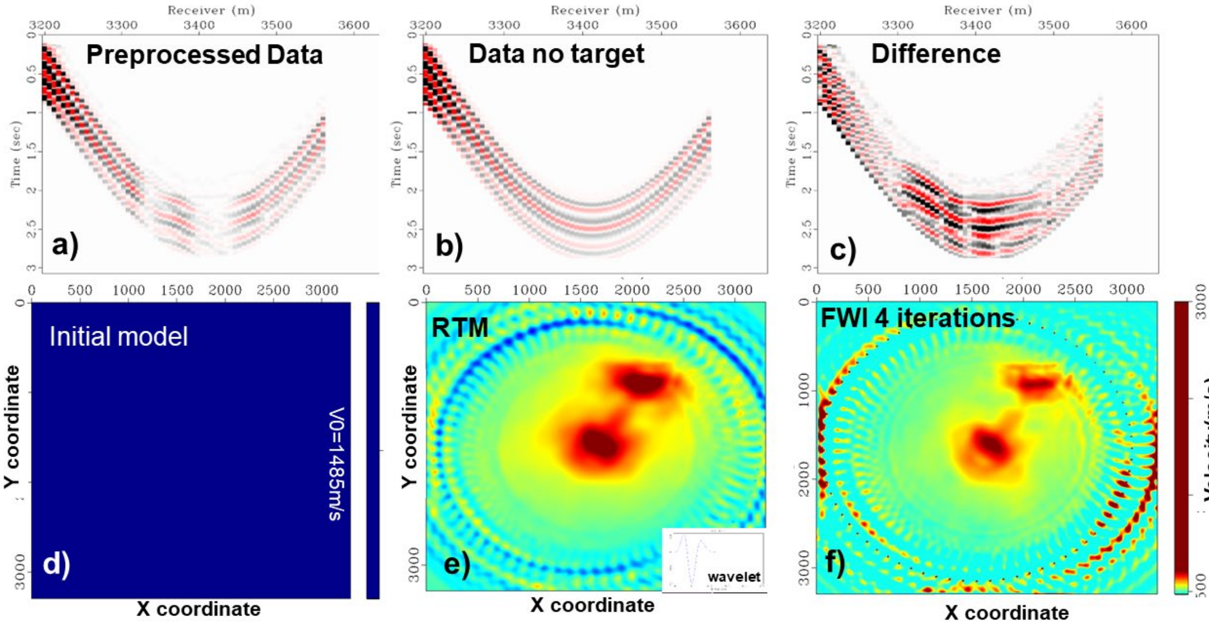

Figure 4-FWI for Physical model data: a) a single shot after muting later arrivals, b) Data acquired in a model without the target bodies, c) Difference between a and b, d) Initial model for FWI (constant velocity), e) RTM of c, f) FWI of c.