Mounting an offense against poor-quality shear data

Michael J. O'Brien, Allied Geophysics

Summary

This is an overview of an ongoing finite-difference model study aimed at improving shear wave data quality. Inspired by data quality challenges seen in an attempt to do 9-C monitoring of a CO_2 flood, the study started as a feasibility analysis to assess the major barriers to completion: developing the ability to build a meaningful model; defining the magnitude of the finite-difference calculation; gaining access to a code to calculate data with the appropriate physics; and gaining access to computers to generate synthetic data. I show a process that allows model frameworks to be built and manipulated quickly. The process uses distinct and realistically complex geologic elements that become part of the model's fingerprint. Example synthetic data from one trial model show that model data have many characteristics of field data. I also estimate that, for a carefully chosen problem, a 3D elastic model with heterogeneous HTI can be calculated on an average-sized commodity cluster in under one month.

Introduction

If the full potential of shear waves can be realized, there are many geologic settings where they provide valuable information that P-waves cannot provide. However, the general industry experience is that shear-wave data are lower quality than P-wave data for the same targets (Figure 1). This often remains true despite heroic efforts of anisotropists and shear data analysts. To address the many issues that degrade shear data quality, I have begun a long-term study that uses finite differences to simulate shear data. The finite-difference method has long been used to study and understand issues related to P-wave data quality and to advance data processing techniques useful for P-wave data. Similar studies to enhance the quality of shear data have not been undertaken, partly due to perceived cost and partly due to the difficulty in defining the physics of a meaningful problem. To guide the choice of physics used for this study, I have adopted a current study area of the Colorado School of Mines Reservoir Characterization Project (RCP) - the Postle field in western Oklahoma. At Postle field, shear wave data quality issues are said to be large statics, near-surface anisotropy, and scattered noise. In addition to using the Postle field as a basis for the model problem, RCP's seismic acquisition parameters have been adopted as model data acquisition parameters: multiple acquisitions of 1920 3-C source points into an orthogonal spread of 3-C receivers covering a 4 x 4 km patch.

Because this project is basically a model study, it requires competence in three general areas: building physically

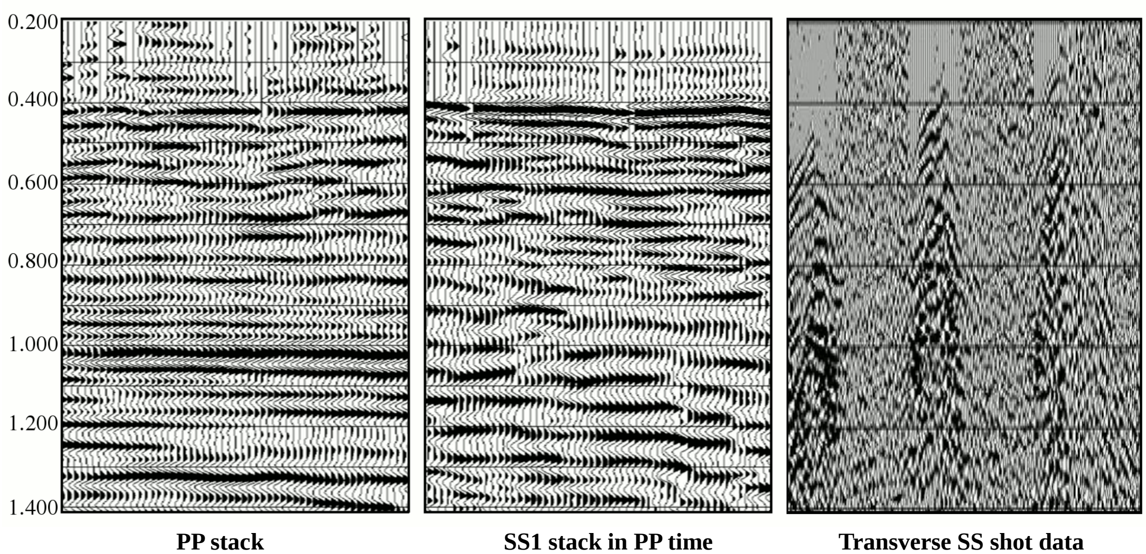


Figure 1. Data examples from Postle Field. The producing reservoir is at 1.05 seconds on stacked data. Continuity on PP data is generally good. The SS1 stack is much lower in quality, leading to higher risk shear data interpretations. Some of the shear data quality issues are evident on shot records where static breaks, surface scattering and aliased ground roll are evident.

meaningful elastic models; calculating synthetic seismic that honors the physics defined for the problem; and processing the data with the intent to learn the limits of current practices and to develop new methods that will improve shear data quality.

Given that no one has yet published results of a 3D elastic model study of this magnitude, there is some uncertainty about whether such an undertaking is within the scope of an academic consortium. At the time of this abstract submission, this question remains unanswered.

Method

First, the technical resources needed to conduct the study were identified, as were gaps in RCP's ability to meet these resource needs. Required resources include the following: an ability to build models with geologically meaningful frameworks; valid parameters with which to populate such models that will generate problems observed in field data; access to a modeling program that will honor the desired physical problem; and the computers capable of calculating synthetic data in a practical time frame. All of these resources exist somewhere among the sponsors of the consortium, but the ability to bring them together to conduct a collaborative project is still not evident.

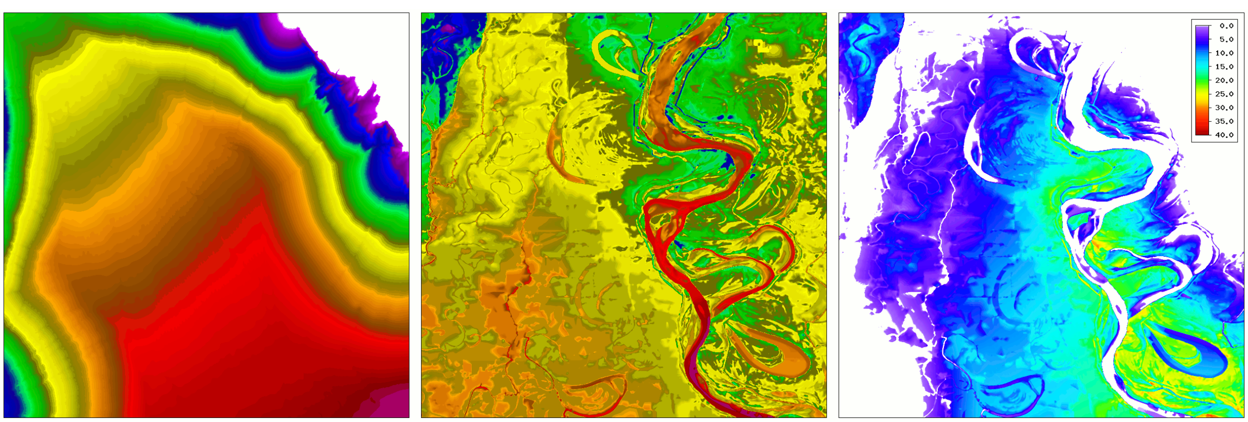
With the primary obstacles defined, several ideas are being explored to determine how they can be overcome. To this end, an attempt to quickly build and rebuild the geologic framework of trial models by using readily available digital information has been investigated. The computational resources needed to synthesize a full 3D elastic dataset have also been assessed. Physical parameters that create the observed problems have not yet been defined. Other resources, as yet to be acquired, include access to a program to calculate data and access to computers on which the calculations will be performed.

When feasibility is established, the physical problems needed for the study will be clearly defined. The final model(s) will be built, the model and modeling code will be sent to the computing center(s), and calculated data will be distributed to consortium members interested in using the data to better understand ways of improving data quality.

Building a meaningful model

Model building started with a few parameters from the RCP field trials at Postle: 4 x 4 km orthogonal source-receiver spread imaging a reservoir at 1860 m. These few parameters, and a desire to build a model with room to grow, led to a model region covering 20 x 20 x 2.5 km. The decision was made to make the model generic - not an exact analogy to the known field - but to use Postle field characteristics when convenient. This, in large part, is because the time frame needed for a 3D elastic simulation appears much longer than the rate at which new field data can be introduced.

Finding geologically reasonable and interesting structures to make up the model's framework is among the challenges identified early on. The decision to create a generic version of the known problem gave the flexibility to seek high quality sources of geologic information from outside the field area. One such source is the USGS digital elevation archive (http://seamless.usgs.gov). Since elevations show relief that is largely controlled by geology, they are used as a proxy for



Depositional Surface **Erosional Surface** **Bed Thickness**

Figure 2. Interesting geologic features are added to the model by combining elements of digital topography. Layers are built up by defining a depositional surface, filling the space above with a desired lithology, and then eroding into that lithology with some other geologic form. The remaining bed thickness contains the fingerprint of the bounding surfaces. The use of readily available digital data has allowed model construction and modification to proceed rapidly.

geologic boundaries. Many different shapes that make up familiar geologic architectures are available. Figure 2 shows how elevation maps can be combined to create sedimentary packages with recognizable features. By repeatedly adding layers into a 3D grid, causing them to erode or truncate against underlying layers, an interesting geologic framework can be built up with relative ease. Some parts of that framework are constrained by real field geology, but other parts are designed simply to produce the desired data quality problems.

A focus when building the geologic framework is to use readily available digital inputs and to enable rapid replication of the process. Model building processes were restricted to relatively simple 1D and 2D operations on maps imported to a 3D grid. A desire to avoid complicated functions (e.g. manipulating 3D geobodies) has allowed rapid progress but has not compromised the desire to construct a model with detailed and interesting stratigraphic features.

For the first trial models, practical considerations forced some compromise. For example, various field data show shear wave frequencies up to 40 Hz and ground-roll velocities of nearly 200 m/s. Taken together, these require a grid spacing of 2.5 m to properly sample the complete wavefield. Such a small grid interval leads to an excessively large number of grid points. To reduce the size of the problem, the minimum velocity was redefined to be 500 m/s, allowing larger grid spacing. Other decisions made to move the feasibility study along were to limit the computational domain to use +/-4000 m offsets in X and Y, to model from a flat surface, to exclude attenuation, and to reduce the maximum velocity from the observed value of 5800 m/s to 5030 m/s. All of these decisions are subject to change if changes allow improvements in the project's goals, although some further compromise may be needed to enable the final calculation.

Physical parameters

Building a geologic framework is only part of the model-building process. Defining physical parameters that combine with the structure is also needed. The combination should create a realistically complex seismic response. Postle field has many sources of information that can be used to guide parameter selection. These include dipole sonic and density logs, as well as core and rock lab studies. For this project, an attempt is being made to replicate the amplitude and AVO response of the reservoir for both incident P- and S-waves. But because using exact overburden and reservoir properties is impractical, the model properties are scaled to give approximately the same response seen when modeling AVO from log data.

In constructing the near-surface region of the model, much artistic license is needed. Structure and velocities need to be combined in a way that generates problems observed in field seismic, but because logs are not run to surface and the well spacing is vastly greater than the distance between model grid nodes, parameters and structure in the near surface are interpretive inventions.

Model QC

Quality control starts at the point where structure and parameters begin to be combined. If the intent is to create a near surface with quantifiable statics, then transit times through the model should be checked to verify that the statics problem is reasonable (Figure 3). Additionally, trial shots are needed to verify that synthetic data contain other characteristics seen in field data and that those characteristics give a reasonable match to data-quality problems that prevent field data from resulting in high quality images (Figure 4). A goal is to create synthetic data that is challenging but not impossible to image. A challenge to meeting this goal is that one never really knows if he's succeeded until after model data have been acquired and imaged under ideal conditions.

Model QC may indicate a need to stage our way from a relatively simple model to a more realistic one. It may be valuable to separate the effects of heterogeneous HTI from the statics problem, for example. While the scattering and statics

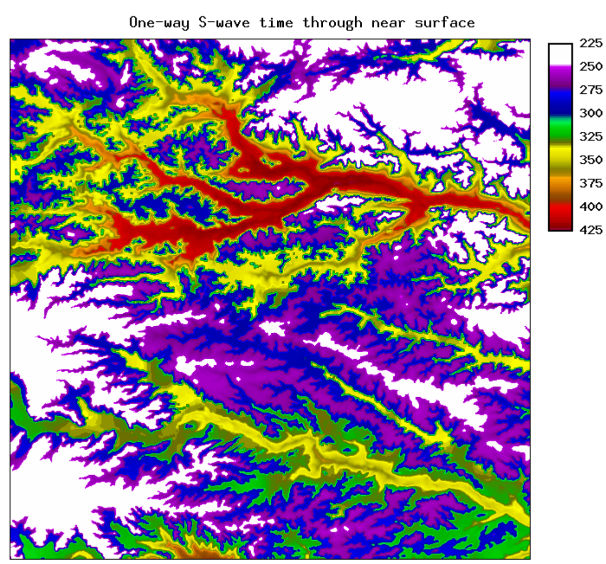


Figure 3. Much QC is needed to ensure the model will produce data with the problems seen in field data. This image shows differential shear wave statics on the order of 200 ms, similar to the magnitude of statics derived from field data.

problems appear to be coupled, there may be an advantage in studying those problems in the absence of shear wave splitting. Attenuation is also thought to be part of the problem at Postle field, but it may be better to create elastic data first, as a control dataset to compare with anelastic data generated later.

Timing experiments and accelerations

An early goal for this model was to be able to calculate a single 4 second shot record in 25 hours using 4 hosts (100 host hours) in a modern commodity computer cluster. Modeling parameters were initially defined as 40 Hz upper frequency, a velocity range of 500 to 5030 m/s, a computational domain of 8 x 8 x 2.1 km, and 3 grid points per minimum wavelength. These parameters led to a grid spacing of $4^1/_6$ m. Timing results reported here were computed on a single host using the staggered-grid, finite-difference algorithm described by Etgen (1987). This off-the-shelf technology from the 1980s is capable of calculating models with any symmetry, but the implementation in the test program is restricted to isotropic, VTI, and orthorhombic media. The extension to heterogeneous HTI is not available, but the additional computational load can easily be factored into results obtained with more symmetric models. Given the parameters and sizes listed above, an HTI shot should complete in ~220 hours, somewhat longer than the original goal.

Several ideas are available to reduce time needed to calculate a shot. Some, but not all, require compromise in modeled results. For example, increasing the grid interval to 5 m reduces the upper frequency to 34 Hz but also reduces run times by over 50%. Reducing the XY size of the computational domain from 8 x 8 km to 6 x 6 km gives another 50% runtime reduction. Both ideas require compromise, but in combination they bring the time to calculate a single shot down to approximately 55 hours, well within the original target of 100 hours. Given an average-sized commodity cluster with 512 hosts and the paired down parameters, the calcution of 1920 3-C shots (5760 shots total) can be completed in a little under one month.

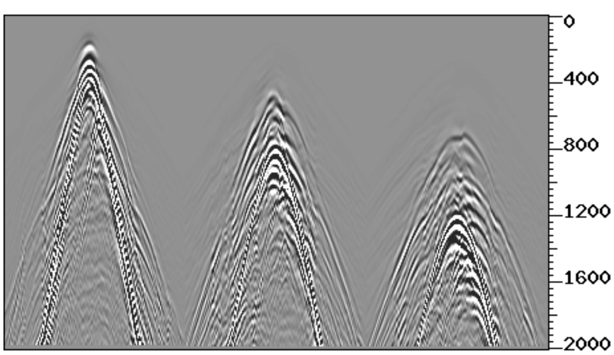


Figure 4. Part of the QC process is to learn whether problems seen in field data are also seen in calculated data. This image of model data shows discontinuities caused by statics, surface-wave scattering and ground roll energy that is highly aliased.

Not all ideas for acceleration require compromise, but they may require other good fortune. For example, if *really* fast FFTs can achieve N Log N performance, the reduction in the number of grid points enabled by using spectral space derivatives conservatively leads to a 75% reduction in run times with no loss in frequency content. Other ideas offer varying amounts of computational speed-up, but no decisions on direction or the amount of compromise have been made before the time of this review.

Conclusions

In areas where complex near-surface geology causes severe data-quality challenges, S-wave data are frequently more difficult to process than P-wave data. For this reason, creative ways to improve data quality are needed. An objective way to evaluate new ideas is to test them with synthetic data from a model study. I have shown that, for carefully chosen problems, 3D anisotropic elastic model studies are feasible. The barriers of building useful models and bringing the size of the calculation down to a managable size have been overcome. Readily available digital information enables rapid construction of meaningful models. Calculating data in those models can still be expensive, but with some compromise, I estimate an RCP-sized 9-C dataset with frequencies to 34 Hz can be calculated on an average-sized commodity cluster in under one month. Whether this calculation can be conducted within RCP is unknown. The challenges of finding a code that can be shared among consortium members and finding computers to do the calculation are not yet resolved.

Acknowledgments

I thank RCP for their technical support and the inspiration needed to approach this problem with a model study. I thank BP for use of their 'vintage' 3D elastic modeler, used to test trial models and assess computational needs.

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2010 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Etgen, J. T., 1987, Finite-difference elastic anisotropic wave propagation: Stanford Exploration Project Report, **56**, 23–58.