

Full waveform inversion of Hussar synthetics

Babatunde Arenrin and Gary Margrave

ABSTRACT

In this study, we present the result obtained from incorporating well log information into a conjugate gradient optimization scheme in Full Waveform Inversion (FWI). We test this approach on synthetic datasets generated using the three sonic logs from Hussar. Using formation tops to guide the interpolation, the sonic logs are interpolated to form the 2D velocity model used in this study. The initial velocity model for the inversion is a linear $v(z)$ velocity model. We adopt the conjugate gradient algorithm as described by Magnus R. Hestenes and Eduard Stiefel. Our results show that combining well information with conjugate gradient directions in FWI can save computational time, as well as getting a good inverted model after a few iterations. The inverted model shows encouraging results and this proves that the algorithm works well and can resolve thin beds in the model.

INTRODUCTION

Full waveform inversion is an optimization technique that seeks to find a model of the subsurface that best matches the recorded field data at every receiver location. The method begins from a best guess of the true model, which is iteratively improved using linearized inversions methods although the FWI problem is non-linear (Warner et al, 2013). FWI is formulated as a generalised inverse problem with a numerical solver-a forward modelling code and its adjoint. FWI can be viewed as an iterative cycle involving modelling, pre-stack migration and velocity model updating in each iteration (Margrave et al, 2010).

Despite its success, FWI suffers from cycle skipping problems, and convergence problems when the starting model is far from the true model and in the absence of low frequencies. However different approaches have been developed to mitigate the problems with conventional FWI, such as incorporating well information to FWI (Margrave et al, 2011a). Well information can aid in (1) calculating the step-length (a scalar which multiplies the gradient for the model update), (2) constraining the line search calculation used in a steepest descent optimization scheme, and (3) improving the wavelet estimate which is essential for proper updates. Some other approaches that mitigate the problems with conventional FWI are Tomographic Full waveform Inversion (TFWI) which combines both FWI and WEMVA (Biondi and Almomin,2012), and Adaptive Waveform Inversion (AWI) in which the observed and predicted datasets are matched trace-by-trace using a least squares convolutional filter (Warner and Guasch, 2014).

In a previous paper (Arenrin et al, 2014), we compared using well log derived step length with a line search optimization scheme, and found that deriving step lengths based on well information produced good inversion results. In that paper, we also proposed that a combination of step lengths derived based on well information with any type of optimization scheme should produce desirable results than using either method only. In this report, we combine well information with conjugate gradient optimization scheme and obtain encouraging results within a few iterations.

FWI FUNDAMENTALS

The theory of FWI has been described in literature by Tarantola (1984), Lailly (1983). Pratt et al, (1998) used a frequency-space modelling formalism for FWI. A full mathematical derivation of the theory of FWI can be found in these papers. FWI compares observed and predicted data by subtracting the two datasets to obtain a residual, for real data we anticipate that this residual should be minimized in a least square sense. The FWI objective function is the L_2 norm of the residuals and can be represented mathematically as

$$\phi_k = \sum_{s,r} (\psi - \psi_k)^2, \quad (1)$$

where ϕ_k is the objective function we want to minimize, s, r are the sources and receivers over which the sum is taken, ψ is the observed data, and ψ_k is the predicted data for the k^{th} iteration (Margrave et al, 2010).

If we are interested in inverting for the velocity model of the subsurface, the model update can be expressed as the gradient of the objective function multiplied by a scalar expressed mathematically as

$$\delta v_k(x, z) = \lambda \int \sum_{s,r} \omega^2 \hat{\psi}_s(x, z, \omega) \delta \hat{\psi}_{r(s),k}^*(x, z, \omega) d\omega \quad (2)$$

where λ is a scalar, the hat (^) over a variable indicates its temporal Fourier transform, $\hat{\psi}_s(x, z, \omega)$ is a model of the source wavefield for source s propagated to all (x, z) , ω is temporal frequency, $\delta \hat{\psi}_{r(s),k}(x, z, \omega)$ is the k th data residual for source s back propagated to all (x, z) , and $*$ is complex conjugation. Specifically $\delta \hat{\psi}_{r(s),k}(x, z, \omega) = \hat{\psi}_{r(s)}(x, z, \omega) - \hat{\psi}_{r(s),k}(x, z, \omega)$ where $\hat{\psi}_{r(s)}(x, z, \omega)$ is the real data at receivers $r(s)$ as back propagated into the medium and $\hat{\psi}_{r(s),k}(x, z, \omega)$ is the k th data model for the same. (Margrave et al, 2010).

The scalar (step length) can be calculated using a line search algorithm or if there is well control, a method based on well information.

The scalar calculated from well compares the current velocity model to that of the known velocity at the well location. We define an objective function β which is the L_2 norm of the difference between the model update calculated from migrating the data residuals and the known velocity at the well and the background velocity model expressed by,

$$\beta = \left\| \lambda G_k - (V_{\text{well}} - V_{\text{BG}})_k \right\|^2 \quad (3)$$

where G_k is the migration of the data residuals stacked over all shots at the well location, V_{well} is the known velocity at the well location, V_{BG} is the background velocity (or the migration velocity) at the well location, and the L2 norm is taken over all the samples in the well. (With real data it is necessary to resample the well information to the same sample density as the velocity model).

The scalar λ is obtained by minimizing the objective function β in Equation 3 with respect to λ . Making λ the subject of the expression gives

$$\lambda = \frac{\sum_j \delta V_j G_j}{\sum_j G_j^2} \quad (4)$$

where $\delta V_j = (V_{well} - V_{BG})_j$ and j indicates sample number.

Conjugate gradient (CG) method based on Magnus R. Hestenes and Eduard Stiefel

The history and theory of conjugate gradient methods can be found in several literatures, however, we will just present one of the conjugate gradient algorithm described by Hestenes and Stiefel, 1952. We notice that this conjugate gradient algorithm in geophysical papers is often credited to Polak and Ribiere, and is known as the Polak-Ribiere method. However, we found the same algorithm in Hestenes and Stiefel 1952 paper which predates Polak and Ribiere's 1969 paper. The conjugate gradient is an iterative method that starts with an initial estimate of the solution, and one determines successively new estimates of the solution, each estimate being closer to the true solution (Hestenes and Stiefel, 1952). The algorithm can be summarised thus:

given a linear system (we try to keep the notations consistent throughout this paper), $Av = \psi$, that we wish to solve using the conjugate gradient method, where A (in the case of a seismic inverse problem) can be considered as the forward modelling operator or an operator that maps from the model space to the data space, v is the model vector, in this case the p-wave velocity, and ψ is the data vector. The solution to the inverse problem can be obtained iteratively by the Hestenes and Stiefel CG algorithm given below

$$\begin{aligned} \alpha_{k-1} &= \frac{|\nabla \phi_{k-1}|^2}{(p_{k-1}, Ap_{k-1})} \\ v_k &= v_{k-1} + \alpha_{k-1} p_{k-1} \\ \beta_k &= \frac{|\nabla \phi_k|^2 - (\nabla \phi_k, \nabla \phi_{k-1})}{|\nabla \phi_{k-1}|^2} \\ p_k &= \nabla \phi_k + \beta_k p_{k-1}, \end{aligned}$$

Figure 1. Hestenes and Stiefel CG algorithm

where $\nabla\phi$ and p_k are the gradient of the objective function and the conjugate direction respectively. β_k is designed to guarantee that p_k and p_{k-1} are conjugate, k is the iteration number, while α_{k-1} acts to scale the conjugate direction p_{k-1} .

Hestenes in 1990 introduced a scale factor of $(1 + \beta_k)^{-1}$ applied to the conjugate direction of the algorithm given above. The modified algorithm takes the form

$$\alpha_{k-1} = \frac{|\nabla\phi_{k-1}|^2}{(p_{k-1}, Ap_{k-1})}$$

$$v_k = v_{k-1} + \alpha_{k-1}p_{k-1}$$

$$\beta_k = \frac{|\nabla\phi_k|^2 - (\nabla\phi_k, \nabla\phi_{k-1})}{|\nabla\phi_{k-1}|^2}$$

$$p_k = (1 + \beta_k)^{-1} (\nabla\phi_k + \beta_k p_{k-1}).$$

Figure 2. Hestenes modified algorithm

Incorporating well information into Hestenes and Stiefel CG method

In this study, we incorporate well information into Hestenes and Stiefel CG algorithm by replacing α_{k-1} in the algorithm above with the value of λ from Equation 4, calculated at every iteration. Hence, the algorithm used in this study takes the form

$$\lambda = \frac{\sum_j \delta V_j G_j}{\sum_j G_j^2}$$

$$v_k = v_{k-1} + \lambda_{k-1} p_{k-1}$$

$$\beta_k = \frac{|\nabla\phi_k|^2 - (\nabla\phi_k, \nabla\phi_{k-1})}{|\nabla\phi_{k-1}|^2}$$

$$p_k = (1 + \beta_k)^{-1} (\nabla\phi_k + \beta_k p_{k-1}).$$

Figure 3. The algorithm used in this study

We test this algorithm on synthetic datasets generated from the three sonic logs from Hussar. By interpolating between the logs we obtain a 2-D velocity model that is used to generate synthetic datasets. The interpolation between the logs was guided by the formation tops. The datasets is generated using an acoustic finite difference forward modelling code form Acceleware. We test the algorithm using different starting models for the inversion: a linear $v(z)$ velocity model, and a slightly smooth version of the true model.

Modelling and Migration (inversion)

Figure 4 below is the 2-D velocity model obtained from the interpolation of the three logs from Hussar. Wells 14-35, 14-27, and 12-27 are superimposed on the velocity model. One can observe the fine stratigraphy and notice that the interpolation was guided by the formation tops as evident on the character of the logs. The 2-D velocity model is about 4km wide and 2Km deep, described by approximately 1.3×10^6 discrete parameters. A total of 61 shots is generated with the acoustic finite-difference code, with a shot spacing of about 67 meters, receiver spacing is 2.5 meters, sampling interval is 2ms, and a total of 1600 receivers. The source wavelet is minimum wavelet with a dominant frequency of 50 hertz. All boundaries for both the modelling and migration are absorbing except at the ground-air interface.

For the migration, we use a reverse time migration (RTM) algorithm provided by Acceleware. The RTM algorithm produces a cross correlation Imaging Condition (IC) reflectivity image, and the source illumination. With a proper pre-conditioning, the source illumination can be used to produce a reflectivity image that would have been obtained using a deconvolution imaging condition. Margrave et al, in their papers (2010, and 2011a) show a mathematical formulation of the deconvolution IC. However in this paper we use the cross correlation reflectivity image normalised by the source illumination for the gradient calculation. Since the FWI algorithm in this study is formulated in the time domain, a bandpass filter is applied to the data residuals prior to migration. The band pass filter is designed such that at every iteration, the lowpass corners of the filter is fixed but the highpass corners of the filter increases progressively at every iteration. We call this type of filtering the ‘fixed frequency window’. We imagine that there are several ways of implementing the filters, however we believe that since we are incorporating well information into the inversion, this is the best way to implement the filter.

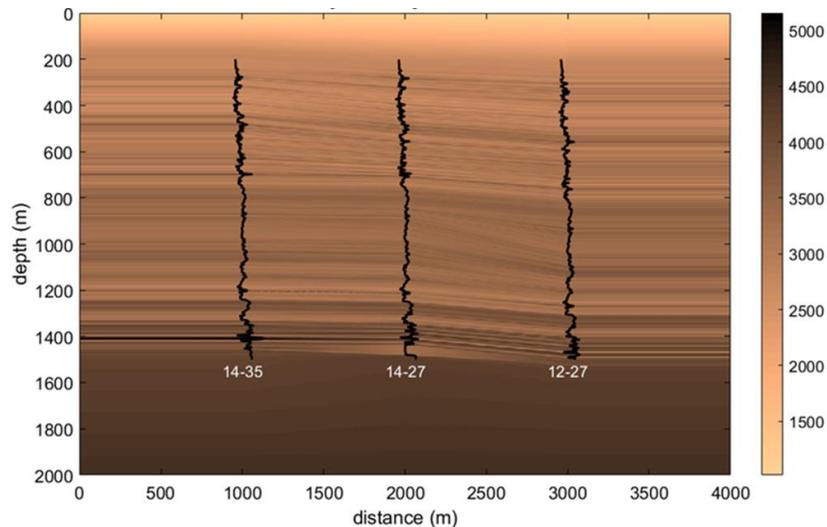


Figure 4. Hussar 2D velocity model interpolated from the three well logs shown in black.

Hussar 2D velocity model

In this subsection we will show results of using a linear $v(z)$ velocity model and a smooth version of the true model as the initial model for the inversion. However a slight modification is made in the case of the linear $v(z)$ velocity model. In order that the turning rays as observed in the observed data can be re-modelled correctly, the top of the true velocity model down to about a depth of 200 meters is left unaltered in the linear $v(z)$ velocity model. From 200 meters to the bottom of the model is a linear velocity function. Another way around this would be to use an F-K filter to get rid of the turning waves on the observed data prior to inversion.

In the case of the using a smooth velocity model of the true model as initial model, the true velocity model is smoothed with a Gaussian smoother to obtain the starting model. In both cases, well 14-35 was used in the algorithm, the other two wells were ignored.

Using a linear velocity function as the initial model.

The initial velocity model is shown in the Figure 5 below. The inverted velocity after 9 iterations is also shown. The inversion was stopped after 9 iterations because there was no appreciable change in the model and also in the norm of the data misfit function.

The frequency strategy for the inversion is presented in Table 1.

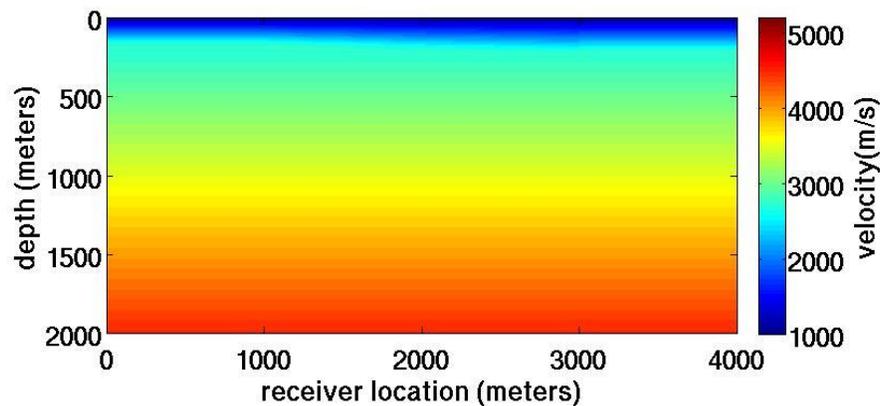


Figure 5. Initial velocity model (linear velocity function).

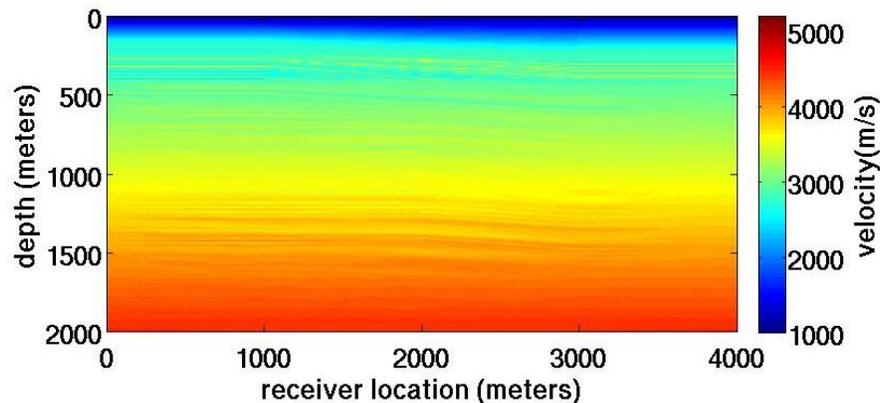


Figure 6. Inverted velocity model after 9 iterations.

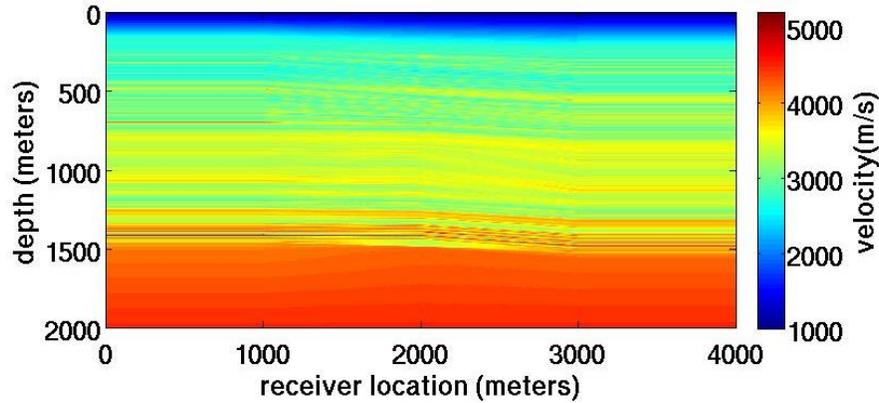


Figure 7. True velocity model.

The inverted model after 9 iterations using a linear $v(z)$ velocity function as the starting model reveals the thin beds in the stratigraphy between depths 1Km and 1.5Km. The area of interest for this study lies above the basement rock between 1.4Km and 1.5Km. We also observe that the area of interest has been resolved in the inverted model, although with little resolution. However, we can still make reasonable interpretation in terms of structure and stratigraphy.

Table 1. Bandpass strategy used in the inversion

Iteration 1	[4Hz, 6Hz, 8Hz, 15Hz]
Iteration 2	[4Hz, 6Hz, 13Hz, 20Hz]
Iteration 3	[4Hz, 6Hz, 18Hz, 25Hz]
Iteration 4	[4Hz, 6Hz, 23Hz, 30Hz]
Iteration 5	[4Hz, 6Hz, 28Hz, 35Hz]
Iteration 6	[4Hz, 6Hz, 33Hz, 40Hz]
Iteration 7	[4Hz, 6Hz, 38Hz, 45Hz]
Iteration 8	[4Hz, 6Hz, 43Hz, 50Hz]
Iteration 9	[4Hz, 6Hz, 48Hz, 55Hz]

Using a smooth version of the true model as the initial model.

The initial velocity model is shown in the Figure below. The inverted velocity after 8 iterations is also shown. Just like the case above, the inversion was stopped after 8 iterations because there was no appreciable change in the model and also in the norm of the data misfit function.

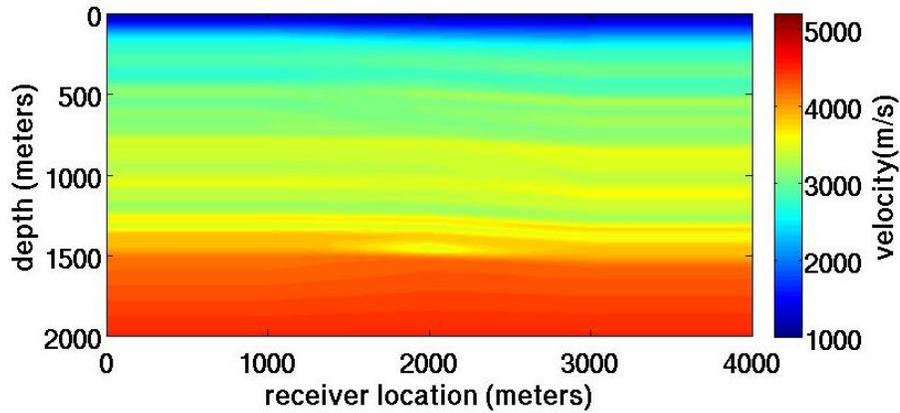


Figure 8. Initial velocity model (smooth velocity function).

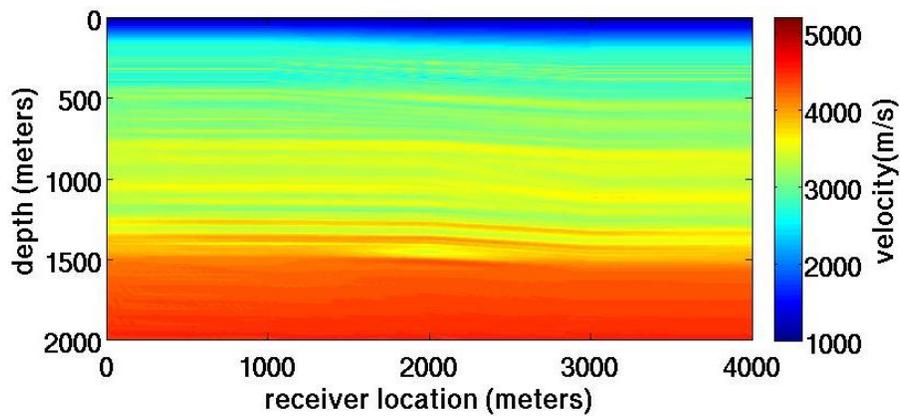


Figure 9. Inverted velocity model after 8 iterations

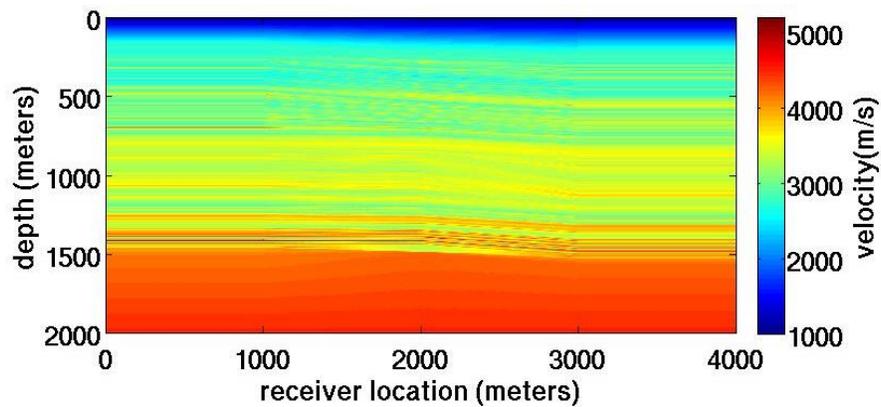


Figure 10. True velocity model

The inverted model after 8 iterations using a smooth version of the true velocity model also reveals the thin beds in the stratigraphy between depths 1Km and 1.5Km. We observe that the area of interest has been resolved in the inverted model, with much better resolution. This doesn't come as a surprise because the starting model is a good one. The frequency strategy in this case is the same as Table 1, except that convergence was at the 8th iteration.

Figure 11 shows a vertical profile of the true, starting and inverted models at the well location at well 14-35. This is the well that was included in the inversion. Figure 12 and Figures 13 is a vertical profile similar to Figure 11, but from wells 14-27 and 12-27 respectively. In all three figures, we observe a good match between the true and inverted model.

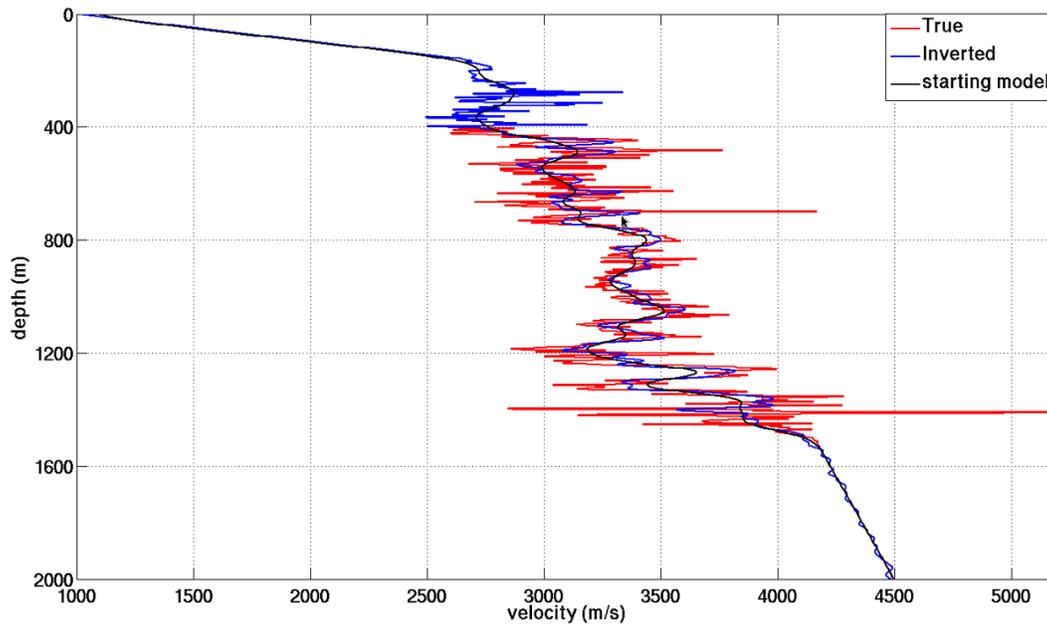


Figure 11. Vertical velocity profile at well 14-35. True velocity profile (red), smooth velocity profile (black), inverted velocity profile after 8 iterations (blue).

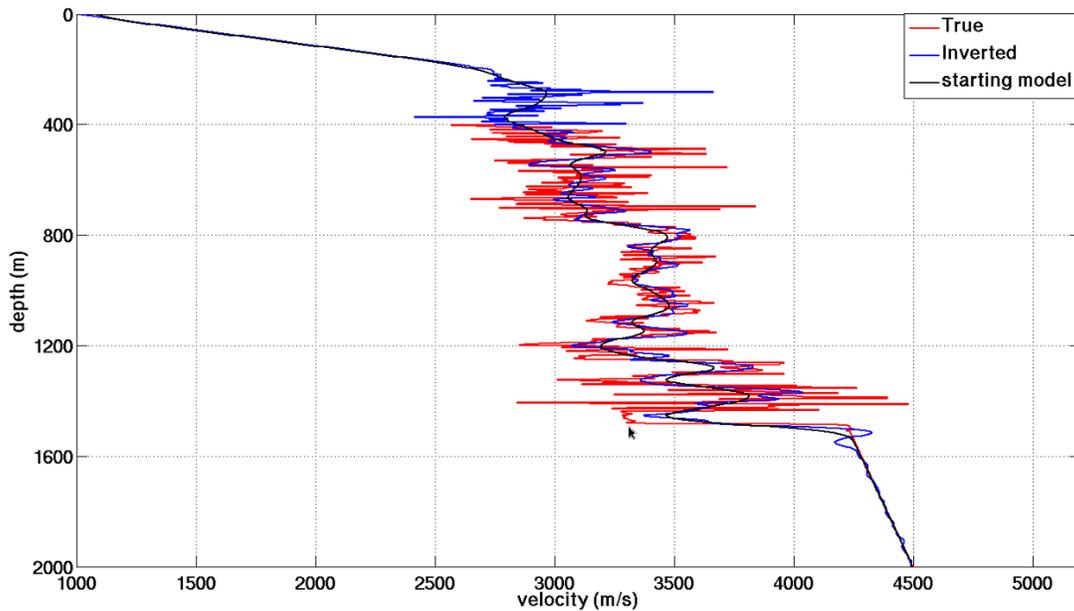


Figure 12. Vertical velocity profile at well 14-27. True velocity profile (red), smooth velocity profile (black), inverted velocity profile after 8 iterations (blue).

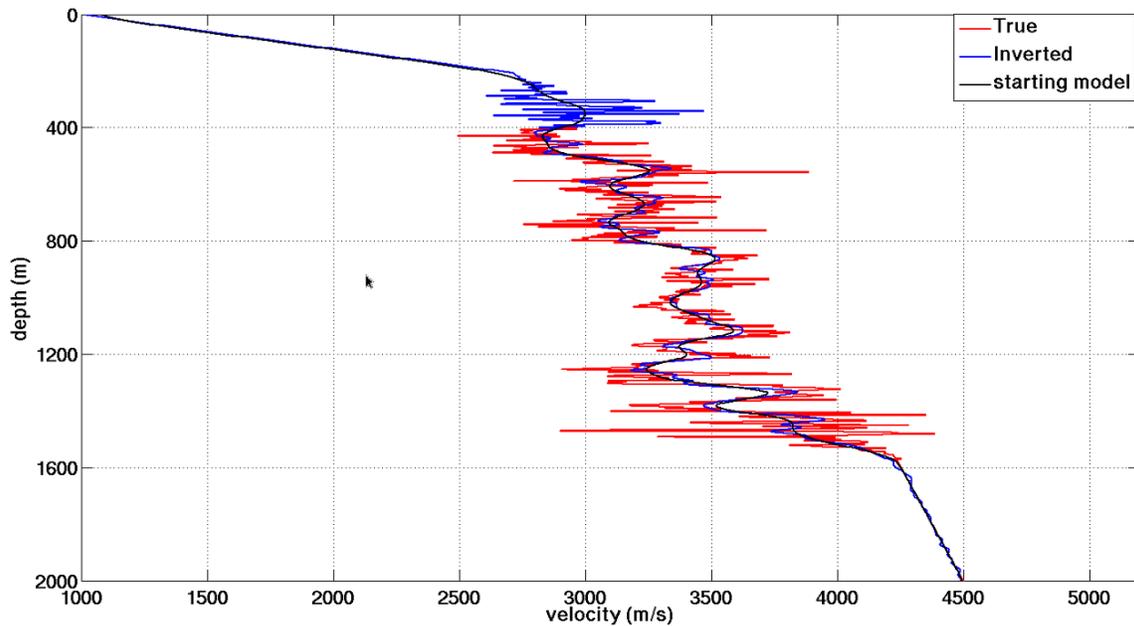


Figure 13. Vertical velocity profile at well 12-27. True velocity profile (red), smooth velocity profile (black), inverted velocity profile after 8 iterations (blue).

CONCLUSIONS

We have been able to incorporate well information into FWI using a conjugate gradient optimization algorithm developed by Hestenes and Stiefel. The algorithm works well and we have been able to bring down the number of iterations needed for convergence. In the case of using a linear $v(z)$ velocity function, we were able to reach convergence in 9 iterations. A linear $v(z)$ velocity function may not be the best starting model for FWI, however, we observe that the algorithm was able to resolve the thin beds and also resolve the area of interest just above the basement rock.

In the case of using a smooth version of the velocity model as the starting model, we obtain a high resolution inverted model after 8 iterations. The thin beds are resolved, and the area of interest can be clearly seen.

The algorithm will be tested on some more synthetic models and real data in the future.

ACKNOWLEDGEMENTS

We thank the sponsors of CREWES for their support. We also gratefully acknowledge support from NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 461179-13. We thank Acceleware for their contribution.

REFERENCES

- Arenrin, B., Margrave, G.F., and Bancroft, J., 2015, Comparing step length calculation using well logs with a line-search method, CSEG GeoConvention 2015
- Arenrin, B., Margrave, G.F., and Bancroft, J., 2014, FWI with source illumination: A synthetic case study in the 26th Annual Research Report of the CREWES Project.
- Biondi, B., Almomin, A., 2012; Tomographic full waveform inversion (TFWI) by combining full waveform inversion with wave equation migration velocity analysis, SEG Technical Program Extended Abstracts, 1-5.
- Hestenes, M. R., Conjugacy and Gradients, 1990, "A History of Scientific Computing," Stephen G. Nash, editor, pp. 167-179.
- Hestenes, M. R., and Stiefel, E., 1952, Methods of Conjugate Gradients for Solving Linear Systems, Journal of Research of the National Bureau of Standard, Vol. 49, 2379,409-436.
- Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Conference on Inverse Scattering, Theory and Application, Society of Industrial and Applied Mathematics, Expanded Abstracts, 206-220.
- Margrave, G. F., Ferguson, R. J., and Hogan, C. M., 2010, Full waveform inversion with wave equation migration and well control: in the 22nd Annual Research Report of the CREWES Project.
- Margrave, G. F., K. Innanen, and M. Yedlin, 2011 a, Full waveform inversion and the inverse Hessian: in the 23rd Annual Research Report of the CREWES Project.
- Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, Part I: Theory and verification in a physical scale model: GEOPHYSICS, 64, 888-901.
- Pratt, R.G., Shipp, M., 1999; Seismic waveform inversion in the frequency domain-II: Fault delineation in sediments using crosshole data, Geophysics 64, 901-913.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: GEOPHYSICS, 49, 1259-1256.
- Warner, M., Guasch, L., 2014; Adaptive waveform inversion, SEG Technical Program Extended Abstracts, 1089-1093