# FWI with wave equation migration: well validation vs data validation vs well-and-data validation

Sergio Romahn and Kris Innanen


## ABSTRACT

Conventional acoustic full waveform inversion (FWI) involves the crosscorrelation of the back-propagated data residuals with the forward-propagated source to produce the gradient. This process can be seen as the reverse time migration (RTM) of the data residuals. The gradient then is scaled to create a velocity perturbation. This step is achieved by applying a line search of the step length in a typical gradient descent scheme. We used PSPI, a wave equation migration method, to obtain the gradient, and we compared three different ways to produce the velocity perturbation. Firstly, we used a line-search method to scale the gradient, a process called data validation. Secondly, we applied well calibration, a technique that is called well validation. Finally, we used a combination of well and data validation. We applied these techniques to two different models, one with moderate lateral velocity changes, and the other one to the more complex Marmousi model. For a simple geological setting the three techniques provided similar results. Well and data validation produced the best result in the presence of more complex geological settings.


## INTRODUCTION

FWI is a procedure that extracts the model parameter from seismic data by fitting observed and synthetic shots generated by wavefield modelling. Lailly (1983) and Tarantola (1984) provided the fundamentals of FWI. The inversion is described as a sequence of prestack migrations of the residuals. The solution is an iterative methodology that consists of a forward propagation of the sources in the current model and backward propagation of the data residuals. The correlation of these two wavefields, that can be seen as the application of RTM over the data residual, leads to a correction of the model parameters. The idea of using any type of depth migration algorithm to produce the gradient was introduced by Margrave et al. (2010). The authors used phase-shift and interpolation (PSPI) migration and highlighted its suitability of selecting the frequencies that we want to use in each iteration. Examples of FWI experiments based on the PSPI gradient are the works of Pan et al. (2014), Margrave (2014), Guarido et al. (2014), Arenrin and Margrave (2015), and Romahn and Innanen (2017). In order to obtain the velocity perturbation that will be used to update the model, Margrave et al. (2010) calibrated the PSPI gradient by matching it to the velocity residual (the difference between actual velocity and current velocity in the inversion process) at a well. We call this methodology well validation. However, the velocity perturbation obtained in this way does not ensure the minimization of the objective function shown in Equation 2. On the other hand, the line search method performs a data validation and explicitly aims for the reduction of the data residuals. In this report we compared the performance of data and well validation and propose the combination of both of them. We will show the application of these three techniques to scale the gradient in two different synthetic datasets.

## THEORY

### Full waveform inversion

FWI is a local optimization procedure. We start with an initial model and iteratively update it, so that after a number of iterations the model is able to produce synthetic data in reasonable agreement to the observed data. Assuming that we are in the $n^{th}$ update to the model parameter $s_0^{(n)} = \frac{1}{(c_o^{(n)})^2}$, where $s_0^{(n)}$ is the squared slowness model and $c_o^{(n)}$ is the P-wave velocity, our problem is to find the model for the next iteration (Equation 1), so that the magnitude of the data residual is maximally reduced.

$$s_0^{(n+1)} = s_0^{(n)} + \delta s_0^{(n)} \tag{1}$$

The misfit function $\phi$ is given by:

$$\phi(s_0^{(n)}(r)) = \frac{1}{2} \int d\omega \left( \sum_{r_s} \sum_{r_g} \|\delta P(r_g, r_s, \omega | s_0^{(n)}(r))\|_2 \right) \tag{2}$$

where $r_s$ is the source coordinate, $r_g$ is the receiver coordinate, and $\delta P$ is the data residual defined as the difference between observed data $P$ and synthetic data $G$:

$$\delta P(r_g, r_s, \omega | s_0^{(n)}(r)) = P(r_g, r_s, \omega) - G(r_g, r_s, \omega | s_0^{(n)}(r)) \tag{3}$$

The procedure to find the model perturbation $\delta s_0^{(n)}$ that miminized $\phi$ is as follows. We begin by truncating a Taylor's series approximation for $\phi(s_0^{(n+1)}) = \phi(s_0^{(n)} + \delta s_0^{(n)})$:

$$\phi(s_0^{(n)} + \delta s_0^{(n)}) \approx \phi(s_0^{(n)}) + \frac{\partial \phi}{\partial s_0^{(n)}} \delta s_0^{(n)} \tag{4}$$

then we compute the derivative with respect to the model parameters of equation 4 to identify the extrema of $\phi$:

$$\frac{\partial \phi(s_0^{(n)} + \delta s_0^{(n)})}{\partial s_0^{(n)}} \approx \frac{\partial \phi}{\partial s_0^{(n)}} + \frac{\partial^2 \phi}{\partial^2 s_0^{(n)}} \delta s_0^{(n)} \tag{5}$$

By solving for $\delta s_0^{(n)}$ when the left-hand side of Equation 5 equals zero, we obtain

$$\delta s_0^{(n)} = - \left( \frac{\partial^2 \phi}{\partial^2 s_0^{(n)}} \right)^{-1} \frac{\partial \phi}{\partial s_0^{(n)}} = -H^{n-} g^{(n)} \tag{6}$$

where $H$ is the Hessian matrix and $g$ is the gradient vector. The inverse Hessian in equation 6 can be replaced by a step length $\mu$, leading to the gradient-based step form:

$$s_0^{(n+1)} = s_0^{(n)} + \mu^{(n)} g^{(n)} \tag{7}$$

## Data validation

When we apply data validation we want to find the updated model in Equation 7 that minimizes the objective function in Equation 2. The goal is to find a value of $\mu$ that does so at each iteration.

We used a combination of backtracking and parabolic interpolation to search for the step length $\mu$. Backtracking is a way to adaptively choose the step size. Assuming a model perturbation of the form $\delta s_0^{(n)} = \mu^{(n)} g^{(n)}$, $\mu$ is defined as the multiplication of two small numbers, $\beta$ and $\alpha^{(n)}$, so that

$$s_0^{(n+1)} = s_0^{(n)} + \beta \alpha^{(n)} g^{(n)} \tag{8}$$

$\beta$ is initialized as a small number, for instance, 0.5, and a convenient $\alpha^{(n)}$ is obtained as

$$\alpha^{(n)} = 0.05 \sqrt{\frac{< s_0^{(n)} | s_0^{(n)} >}{< g^{(n)} | g^{(n)} >}} \tag{9}$$

where $<>$ denotes scalar product.

The data residual $res_1$ linked to the current value of $\beta$, called $\beta_1$, is compared to $res_0$, where $res_0$ is the data residual for the current model $s_0^{(n)}$. If $res_1$ is greater than $res_0$, $\beta$ is updated as $\beta = 0.5\beta$ and the process is iteratively repeated until $res_1$ is smaller than $res_0$ or $\beta$ is close to zero.

We need three points in order to apply parabolic interpolation. The first point is $(\beta_0, res_0)$ with $\beta_0 = 0$, the second point is $(\beta_1, res_1)$, and the third point $(\beta_2, res_2)$ is set up as the previous values of $\beta_1$ and $res_1$ in the backtracking process. The interpolation produces beta $\beta_{fin}$ that is expected to provide the minimum data residual for the three points.

The parabolic interpolation is given by

$$\beta_{fin} = \frac{1}{2} \frac{res_0(\beta_2^2 - \beta_1^2) + res_1(\beta_0^2 - \beta_2^2) + res_2(\beta_1^2 - \beta_0^2)}{res_0(\beta_2 - \beta_1) + res_1(\beta_0 - \beta_2) + res_2(\beta_1 - \beta_0)} \tag{10}$$

and the updated model for this iteration will be

$$s_0^{(n+1)} = s_0^{(n)} + \beta_{fin} \alpha^{(n)} g^{(n)} \tag{11}$$

In another scenario, if $res_1 < res_0$ for the first guess $\beta_1 = 0.5$, the third point is searched by using a larger value of $\beta$; for example, $\beta_2 = 2\beta_1$. In this scenario $res_2$ can be smaller or greater than $res_1$. Either way, we will have three points to apply the interpolation and obtain $\beta_{fin}$.

We assume that the three points fit the parabolic interpolation; however, there is a possibility that this is not true and $\beta_{fin}$ won't produce the smallest residual. In this case, we select from the values of beta already tested, the one that produces the smallest data residual.

### Well validation

Well validation refers to the use of well-log measurements of the model parameter to change the gradient into a velocity perturbation in a process that can be seen as a calibration of the gradient. The computational cost of this technique is negligible. It consists on matching the gradient of the objective function at the well $(g_w^{(n)})$ to the model parameter residual at the well $(\delta s_w^{(n)})$. In our acoustic example, the model parameter residual corresponds to the difference between the squared slowness measured in the well$(s_w)$ and the squared slowness of the current model $(s_w^{(n)})$ at the well location:

$$\delta s_w^{(n)} = s_w - s_w^{(n)} \tag{12}$$

We calculate the best (least-squares) scalar $a$ that matches the amplitude of the gradient trace to the amplitude of the model residual at the well. Then, we calculate a constant-phase rotation angle $\gamma$ that reduces the differences between the scaled-gradient trace and the model residual. After that, we recalculate the scalar $a$ and use it with $\gamma$ to produce a match filter $\lambda$. The model perturbation is derived when we applied the match filter to the whole gradient.

### Well and data validation

We analyzed the performance of the inversion when we combine the two techniques mentioned above. In this case, we firstly calibrated the gradient with well-log information and then we applied a line search to minimize the objective function.

## METHODOLOGY

The inversion process for data validation is shown in the diagram of figure 1. The input variables (initial model, wavelet, observed shots and frequency range) are shown inside green boxes. We firstly propagate the source wavelet through the initial model by applying finite-difference modelling to generate synthetic shots. Then we obtain the data residuals and migrate the selected frequencies by applying the PSPI algorithm. This step produces a pseudo-reflectivity gradient. In the next step, the reflectivity is converted from depth to time by using the current model. After that, we apply impedance inversion. As we are considering constant density such procedure provides a velocity gradient. After going back to depth, we apply line search to scale the gradient and obtain a velocity perturbation that will update the model.

The diagram in figure 2 shows the process when we apply well validation. In this case, the calibration with well-log information is done after converting the gradient from reflectivity to velocity while we are in time domain. We apply a band-pass filter to the well-log data with the same frequency range that we used in the migration of the data residuals. The calibrated gradient is then converted back to depth and finally is added to the current velocity model to update it.

The diagram in figure 3 shows the combination of well and data validation. We firstly calibrate the gradient with well-log information and then we apply a line search of the step length to scale the calibrated gradient.

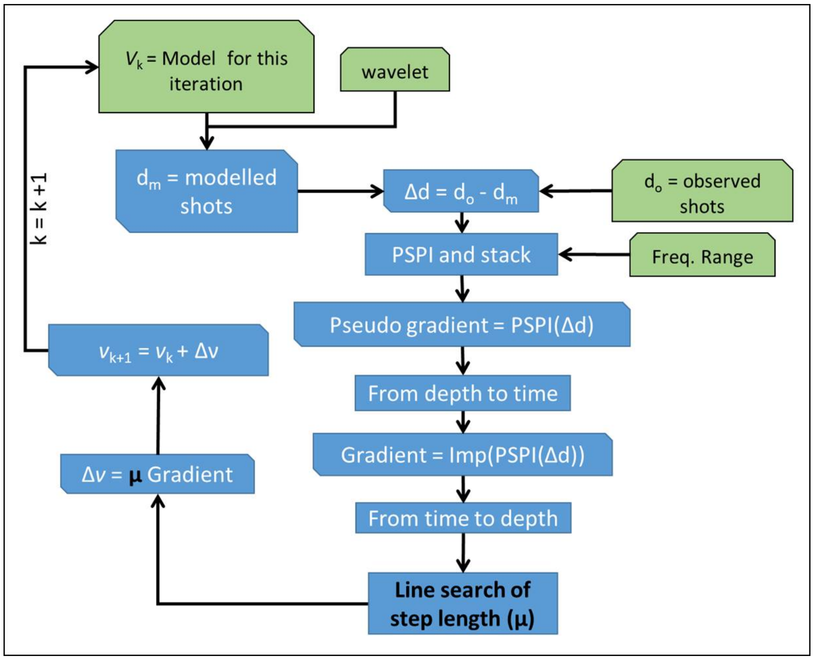


FIG. 1. Inversion process with data validation. A line-search method is applied to obtain $\mu$ that scales the gradient to produce a velocity perturbation.

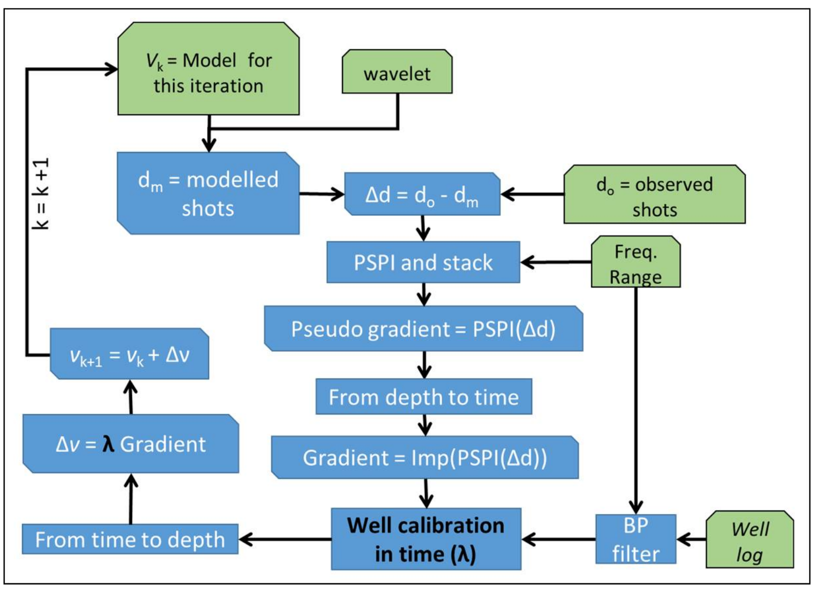


FIG. 2. Inversion process with well validation. The gradient is calibrated with well-log information to produce a velocity perturbation.

## NUMERICAL EXAMPLES

### Anticline model

For the first example we used a simple model characterized by a shallow anticline that constitutes a low velocity reservoir trap surrounded by a relatively high velocity medium

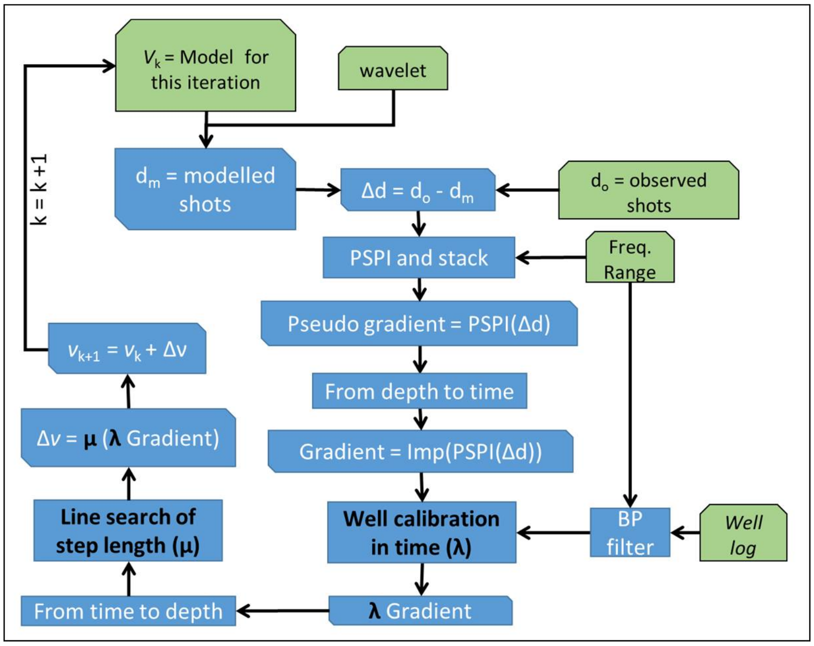


FIG. 3. Inversion process with well and data validation. The gradient is firstly calibrated with well-log data. After that, a line search is applied to find a scale that produces a velocity perturbation while minimizes the data residuals.

(figure 4A). There is also a low-velocity stratigraphic trap located a depth of 800 m at the left-hand side of the model. A shallow high velocity body at a depth of 300 m is located at the right-hand side of the model. The shots, which will be considered the observed data, were generated by applying constant-density acoustic finite-difference modelling through the velocity model. We used a minimum phase wavelet with a dominant frequency of 10 Hz as a seismic source. The source and receiver intervals are 120 and 10 meters, respectively. The total number of shots is 83. An example of a shot located at the middle of the model is shown in figure 4B. The initial model for the inversion was generated by applying a Gaussian smoother with a half-width of 200 m to the true model (figure 5). The frequency range was from 1 to 6 Hz for the first iteration, then it was moved up by one Hz in each iteration. Well C, located at the middle of the model, was used to calibrate the gradient when well validation was applied. The top and base of well C are 400 and 900 m, respectively. For the data-validation case, we used only five shots distributed across the top of the model.

### *Results*

The performance of the data validation technique is shown figure 6. We display the error in the whole model, in the blind well B (which cuts the low velocity body at the left hand side of the model), in well C (calibration well), and the L2N of the data residuals. In general, the errors and data residual steadily decreases except in the location of well C where the error has a peak around iteration 8. After 20 iterations the inverted model recovers all features of the true model. In a previous work, Romahn et al. (2017) couldn't obtain a good result where applying line searh to scale the PSPI gradient and suggested it needed a phase correction that was provided by the well calibration. In that case, the gradient has not been subject to an impedance inversion as it's done in this report. We

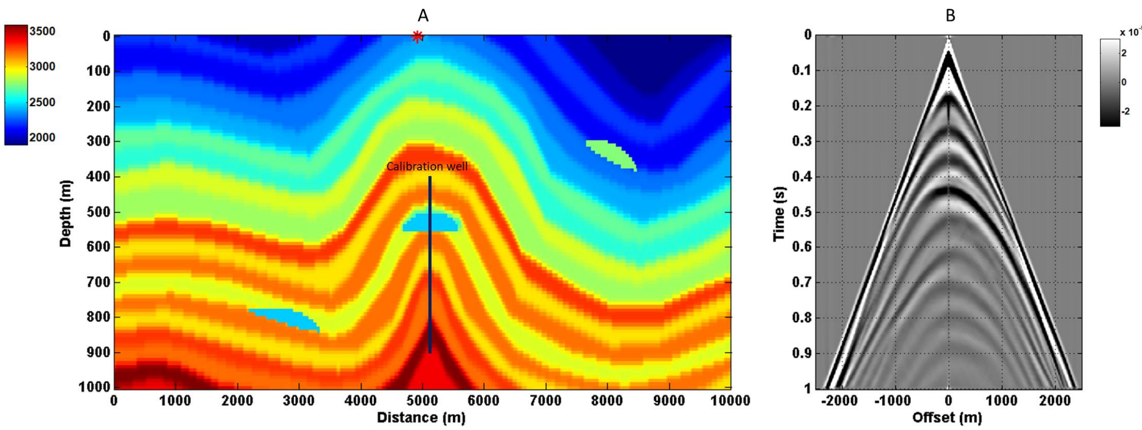


FIG. 4. True velocity model and observed shot located at the middle of the model. The calibration well is located in the middle of the model with a top and base of 400 and 900 m, respectively

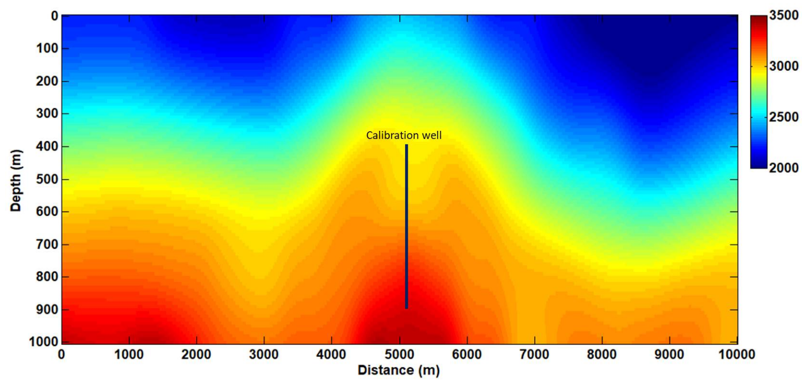


FIG. 5. Intial velocity model.

observe that the phase problem is diminished after converting the reflectivity gradient to impedance.

The result when we apply well validation is shown in figure 7. The errors and data residual are slightly smaller than in the case of data validation. This result suggests that there may be a residual phase problem in the gradient that is corrected with the well calibration.

The combination of well and data validation is shown in figure 8. This technique produces the best overall result because the well calibration accounts for any remnant phase issue in the gradient and the line search ensures that we reduce the data residual in each iteration.

This experiments suggest that we can use any of these techniques and expect favorable results when the geology is as simple as the anticline model shown in this example.

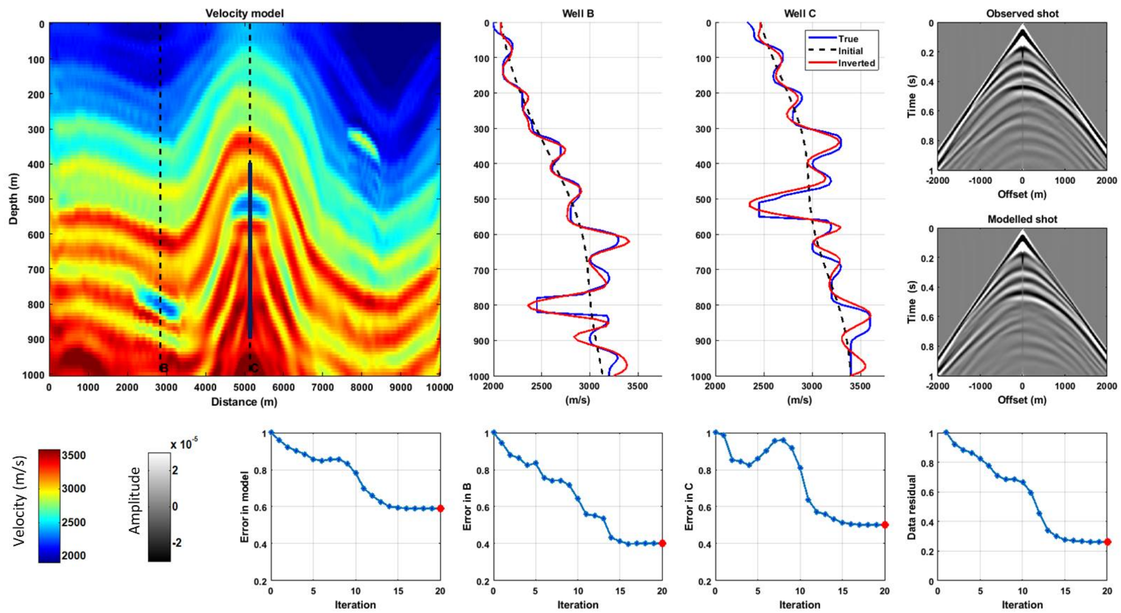


FIG. 6. Data validation applied to the anticline model.

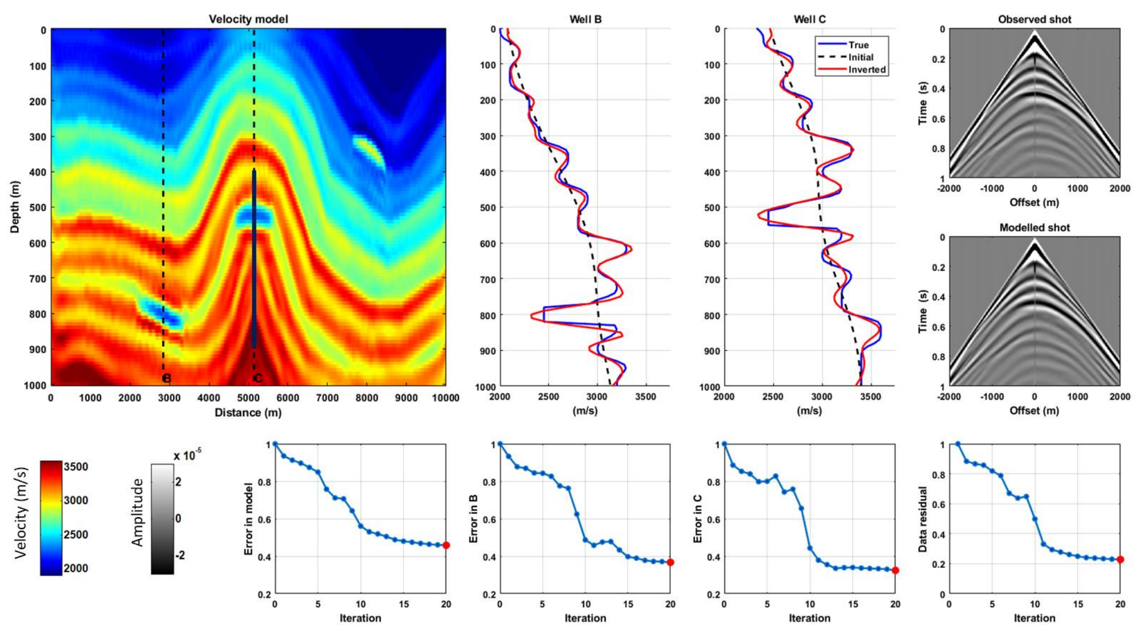


FIG. 7. Well validation applied to the anticline model.

## Marmousi model

We wanted to compare the three techniques to produce the velocity perturbation with a more complex model such as Marmousi's (figure 9A). We used a minimum phase wavelet with a dominant frequency of 10 Hz to generate the "observed" shots. The source and receiver intervals are 120 and 10 meters, respectively. The total number of shots is 88. An

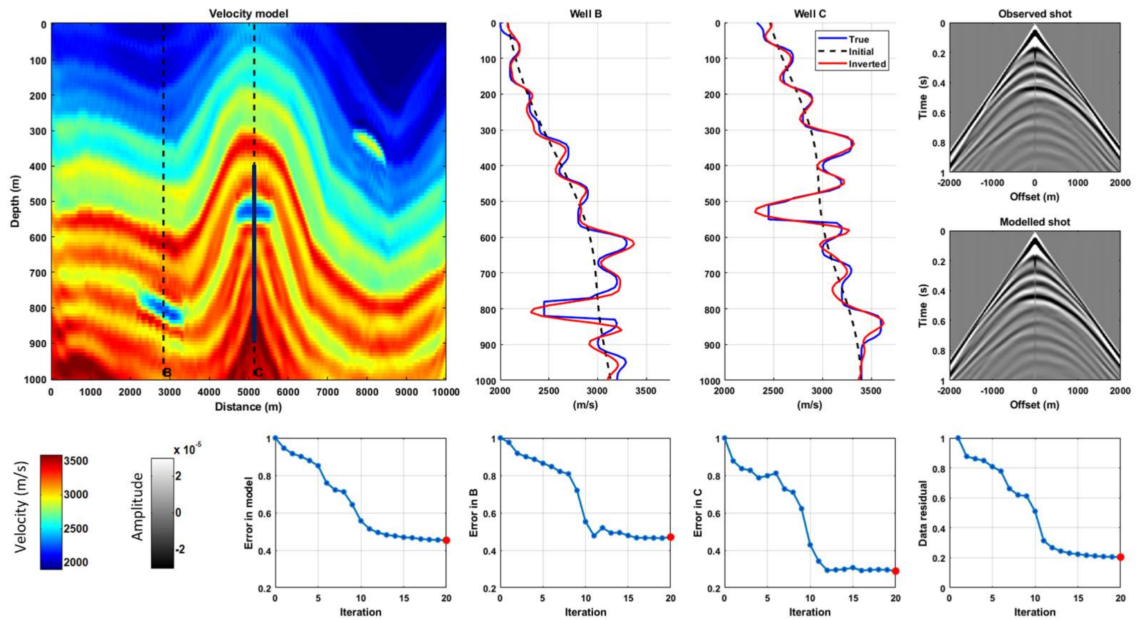


FIG. 8. Well-and-data validation applied to the anticline model.

axample of a shot located at the middle of the model is shown in figure 9B. The initial model for the inversion was generated by applying a Gaussian smoother with a half-width of 200 m to the true model (figure 10). The frequency range was from 1 to 5 Hz for the first iteration, then it was moved up by one Hz in each iteration. The calibration well C is located at the middle of the model cutting an anticline at a depth of 2500 m. The top and base of this well are 400 and 2700 m, respectively. A blind well B was simulated at the horizontal distance of 7100 m. We used five shots, distributed along the top of the model when we apply data validation.

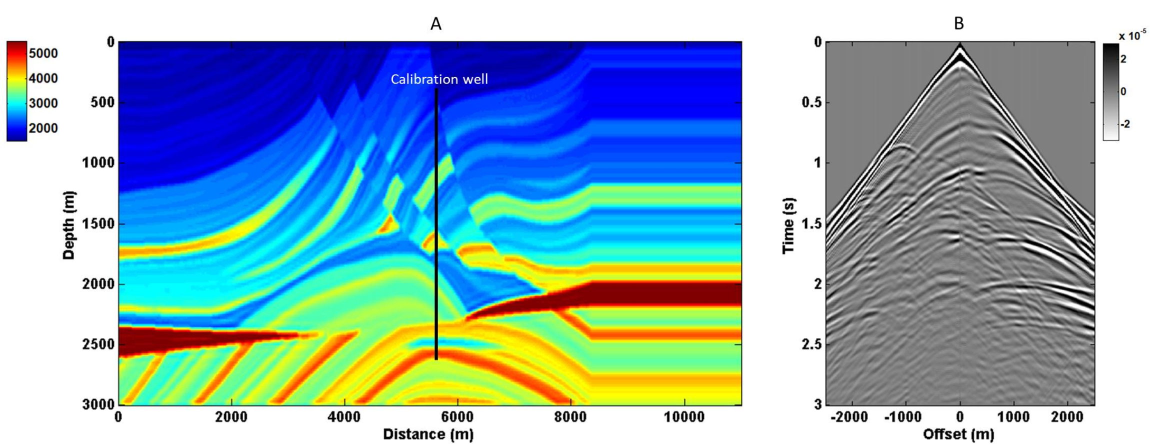


FIG. 9. True velocity model and observed shot located at the middle of the model. The calibration well C is located at the middle of the model with a top and base of 400 and 2700 m, respectively.

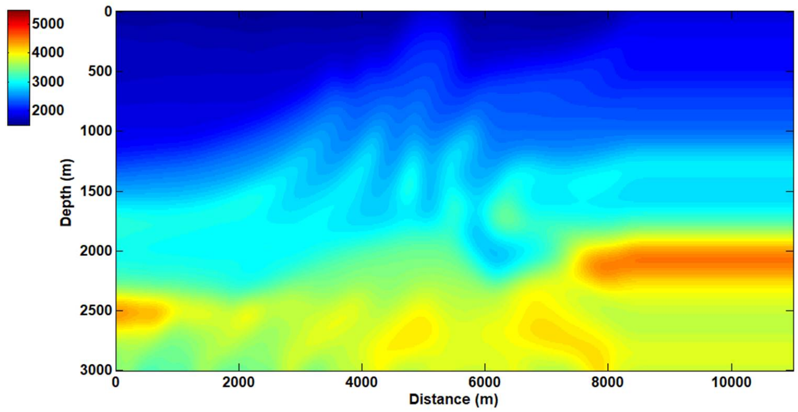


FIG. 10. Initial velocity model.

*Results*

Figure 11 shows the inverted model for data validation. The inverted model is able to recover all features of the true model; however, it has a blurry character that may be related to the residual phase issue inherent to the PSPI gradient.

Figure 12 shows the inverted model for well validation. This technique addresses the phase issue of the PSPI gradient, but does not ensure the decrement of the data residuals; therefore, the size of the gradient is clearly overestimated with no control specially below 2500 m of depth where high amplitude events appear.

The use of the combination of well-and data validation produces the best result as is shown in figure 13. All features of the model are clearly defined, including the high velocity bodies. The overall errors and data residual are significantly smaller than in the previous results when we only applied well or data validation. We would recommend the use of well and data validation in the case of complex geology.

## CONCLUSIONS

We compared three ways to scale the PSPI gradient in the process of FWI. The first technique, called data validation, is based on the widely applied line search that scale the gradient aiming for a direct minimization of the objective function which involves the difference between observed and modelled shots. The second technique, called well validation, is focused in matching the gradient trace at a well location to the velocity residual in that well. The third technique, that combines well and data validation, looks for a good match in the well location and also minimizes the data residuals. For a simple model, the three techniques produce good results. We would recommend the use of well validation when we are dealing with this kind of geological setting because it is the less computational expensive method. We also compared the performance of these scaling technique with the Marmousi model, which is a more complex model. We found significant difference in the inversion result in this case. We conclude that the well-and-data validation technique produces the best result for complex geological settings.

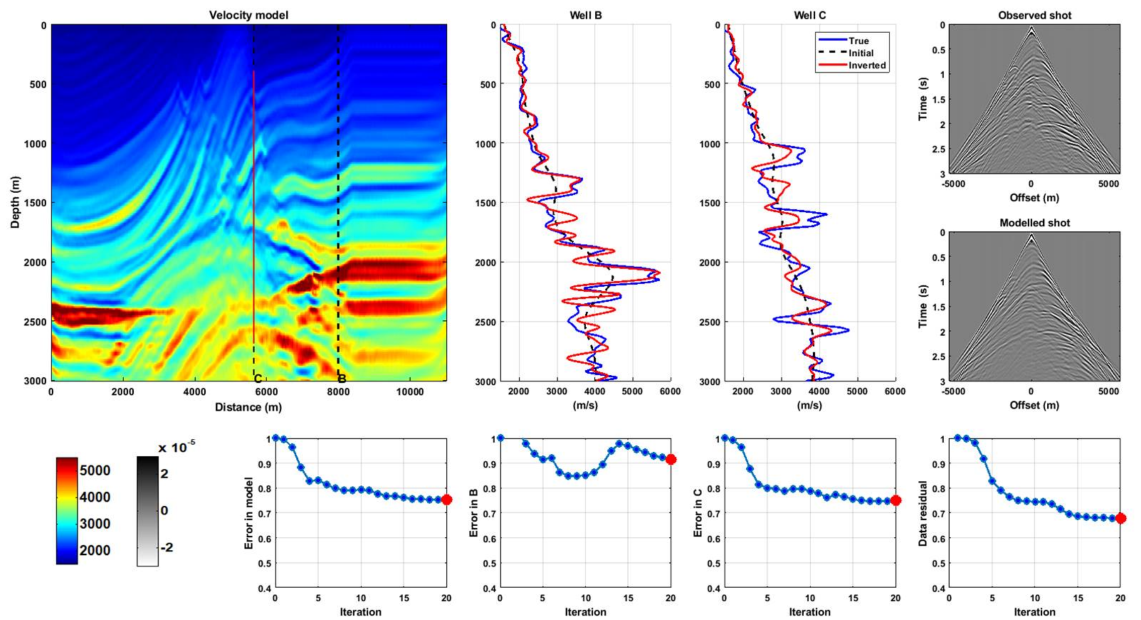


FIG. 11. Data valitation applied to the Marmousi model.

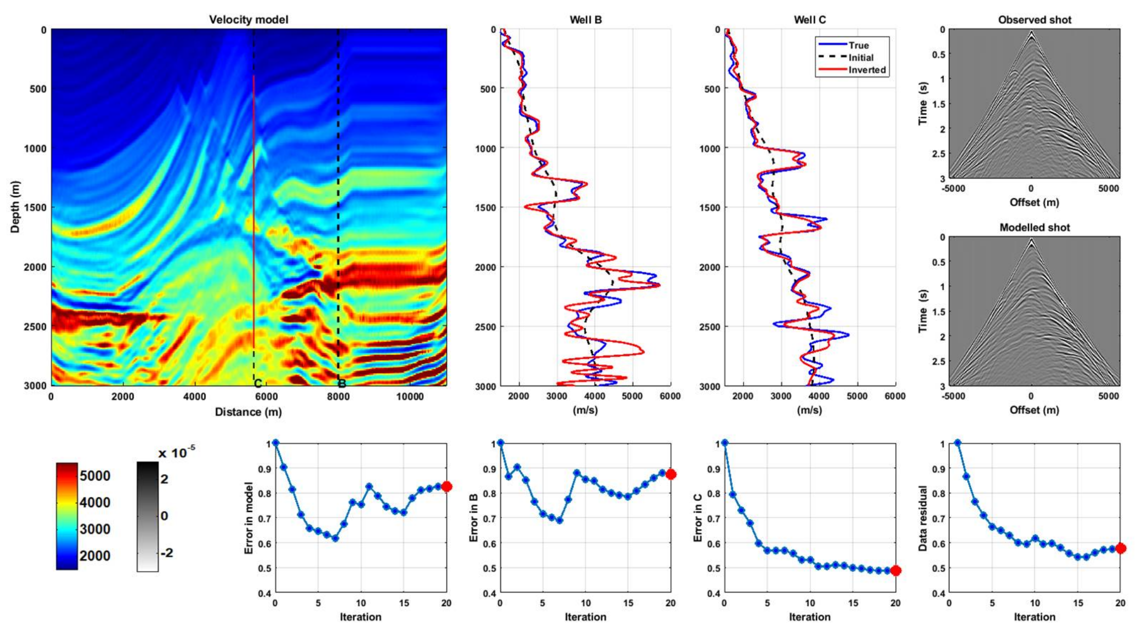


FIG. 12. Well valitation applied to the Marmousi model.

## ACKNOWLEDGEMENTS

We thank the sponsors of CREWES. We acknowledge support from NSERC through the grant CRDPJ 461179-13. Author 1 thanks PEMEX and the government of Mexico for founding his research.

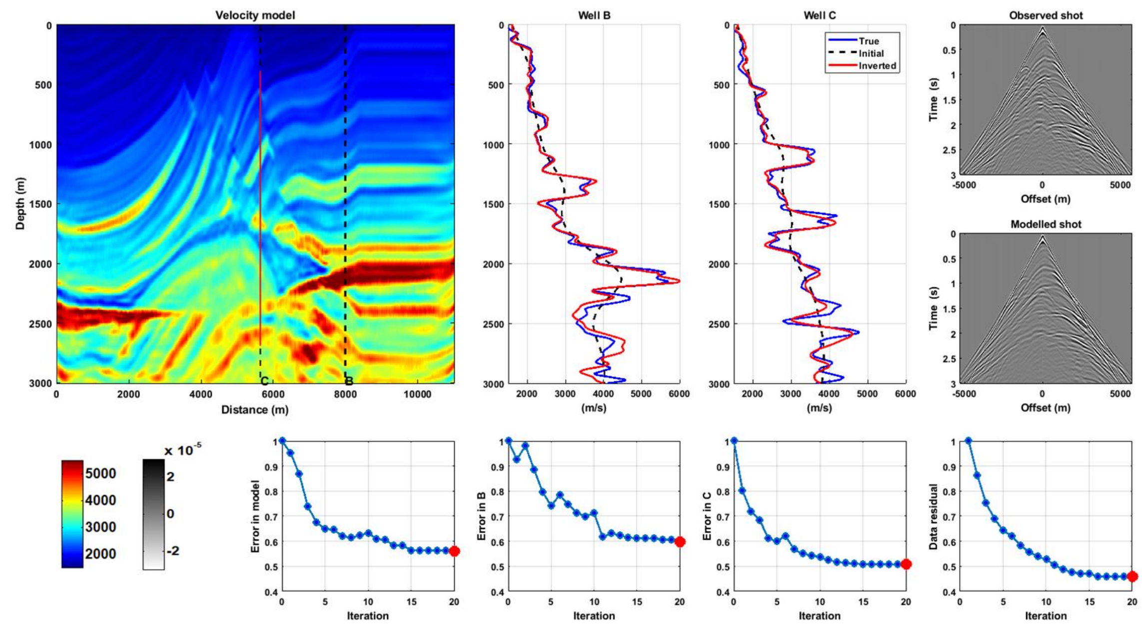


FIG. 13. Well-and-data valitation applied to the Marmousi model.

## REFERENCES


Arenrin, B., and Margrave, G. F., 2015, Full waveform inversion of hussar synthetics: CREWES Research Report, **27**.

Guarido, M., Lines, L., and Ferguson, R., 2014, Full waveform inversion - a synthetic test using the pspi migration: CREWES Research Report, **26**.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migration: SIAM, 206–220.

Margrave, G. F., 2014, Post-stack iterative modeling migration and inversion (immi): CREWES Research Report, **27**.

Margrave, G. F., Fergurson, R. J., and Hogan, C. M., 2010, Full-waveform inversion with wave equation migration and well control: CREWES Research Report, **22**.

Pan, W., Margrave, G. F., and Innanen, K. A., 2014, Iterative modeling migration and inversion (immi): Combining full waveform inversion with standard inversion methodology: 84th Ann. Internat. Mtg., SEG, Expanded Abastracts, 938–943.

Romahn, S., Guarido, M., and Innanen, K. A., 2017, Comparison between rtm gradient and pspi gradient in the process of fwi: CREWES Research Report, **29**.

Romahn, S., and Innanen, K. A., 2017, Iterative modeling, migration, and inversion: Evaluating the well-calibration technique to scale the gradient in the ful waveform inversion process: SEG Technical Program Expanded Abstracts 2017, 1583–1587.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.