Wave-equation imaging and linearized inversion

Developing a methodology to correctly reposition recorded seismic reflection events is one of the most crucial aspects of exploration seismology. As a result, imaging algorithms have been an intensely researched topic over the last several decades. The underlying physics of seismic scattering, while complicated, are well understood; it is the computational feasibility of accurate wave simulation that remains a major bottleneck in image construction

Increases in computing power over the last two decades have brought forth the ability to, within certain assumptions, calculate direct solutions to the wave-equation. Accurate simulation of the two-way wave-equation is vital to capture many wavefield complexities, such as multiply scattered events, prismatic events, attenuation, overturned waves, elastic arrivals, mode conversions, anisotropy and steeply dipping arrivals (**?**; **?**; **?**). Earlier efforts to numerically solve the wave-equation, so called one-way extrapolators, could not reliably track any of these aforementioned complexities (**?**).

However, simplifications of the physics are still necessary, and many assumptions remain inherent. As computational power increases, so does the ability to capture more of the physics within a given imaging algorithm. Moreover, acquisition techniques are constantly improving, meaning datasets are increasing in size - both in terms of the number of shot points and the quantity of offsets/azimuths present. Most of this discussion will be limited to an acoustic, isotropic Earth, but solutions to the two-way wave equation will always be used. The reasons for these assumptions will be discussed at length in the forthcoming chapter.

## TWO WAY WAVE-EQUATION IMAGING

The acoustic wave equation can be expressed in equation 1.

$$\nabla^2 \mathbf{u} - \mathbf{h}\frac{\partial^2 \mathbf{u}}{\partial t^2} - \mathbf{g}\frac{\partial \mathbf{u}}{\partial t} = \mathbf{f} \tag{1}$$

$$\mathbf{u} = u(x, y, z, t); \mathbf{f} = f(x, y, z, t). \tag{2}$$

Here, $\mathbf{u}$ represents the wavefield, $t$ represents time, $\mathbf{h}$ represents the acceleration coefficients, $\mathbf{g}$ represents the diffusion coefficients and $\mathbf{f}$ is a forcing term. Depending on the situation, $\mathbf{h}$ and $\mathbf{g}$ may be scalar or vector terms. Table  details six different geophysical scenarios and the values of these differential coefficients under certain assumptions.

Of course, the wavefield of interest herein is the seismic reflection pressure field, and it will be assumed that the diffusion term is negligible ($\mathbf{g} = \mathbf{0}$). This assumption is valid for the scope of this thesis discussion.

| **Method** | **u** | h | g |
|---|---|---|---|
| Seismic Reflection | Pressure/displacement | $1/c^2$ | $2\alpha/c$ |
| Transient EM | Magnetic components | $\mu\epsilon$ | $\mu\sigma$ |
| Georadar | Electrical components | $\mu\epsilon$ | 0 |
| Low-frequency EM | Electromagnetic fields | 0 | $\mu\sigma$ |
| Geothermal | Thermal fields | 0 | $1/\kappa$ |
| Electical/Magnetic/Gravity | Potential fields | 0 | 0 |

Table 1: Wave equation coefficients for various geophysical methods

$$\frac{\partial^2 \mathbf{u}}{\partial t^2} = \mathbf{v}^2 \nabla^2 \mathbf{u} \tag{3}$$

$$u_{t+1} = -u_{t-1} + 2u_t + \Delta t^2 v^2 \nabla^2 u_t. \tag{4}$$

Numerical solutions to these various wave equations have been extensively researched, and implementations include frequency domain propagation (**?**), time domain propagation (**?**), mixed domain techniques (**?**), finite difference and finite element techniques (**?**), and many more. For 3D wave propagation, decomposed across a network of computers, a time-doman finite-difference approach can provide sufficiently accurate, computationally efficient solutions (**?**). The resultant equation to be solved is shown in equation 3.

The second equation represents how to practically update the solution. $u_{t+1}$ is the 3D wavefield, at time $t+1$, being solving for (one time increment in the future), $u_t$ is the current 3D wavefield, $u_{t-1}$ the previous wavefield, $\Delta t^2 v^2$ the time sampling time squared multiplied by the current velocity value squared and, finally, $\nabla^2 u_t$ is the second-order spatial derivative, or Laplacian, of the current 3D wavefield. It is clear, through comparison of the above equation, that these first two terms approximate the time derivative of the wave equation, and the final term approximates the spatial derivative.

Implementing the time derivative is relatively trivial, and a simple subtraction of the two wavefields is acceptable. Although, more accurate options exist (**?**). For the spatial derivative/Laplacian choosing an exact breakdown is more subtle. This depends on the desired accuracy and stability of the solution, and more subtly on the range of velocities present and the desired propagation bandwidth.

It is desirable to use a large stencil to approximate the spatial derivative - the larger the stencil, the more stable the numerical solution, since instabilities can be implicitly smoothed over (**?**). These short wavelength instabilities are of particular concern for exploration seismology, since velocity fields often feature sharp contrasts over short distances. These can be caused by geological scenarios such as salt bodies and anhydrous layers (**?**).

This larger stencil also means that the field can have a sparser representation, for a given level of derivative accuracy. A large stencil requires more computation;

current industry applications vary from using a 4th order stencil to a 40th order stencil. Generally, the choice depends on how many time-steps the wavefield will be propagated for, as each time step introduces new errors which are carried forward. An 8th order stencil introduces an error of approximately $10^{-7}$ per time step, this can be estimated by looking at the smallest stencil coefficient (**?**).

A representation of an 8th order stencil can be seen in Figure 1. It is comprized of 25 coefficients (nine in each dimension with a shared central value), only five of these coefficients are unique: the stencil is symmetric in all three dimensions about the central value. The Laplacian is applied by convolving the stencil about every point in the current wavefield, and multiplying by $\Delta t^2 v^2$, where $\Delta t^2$ is a constant and $v$ is the velocity field value for the central stencil point.

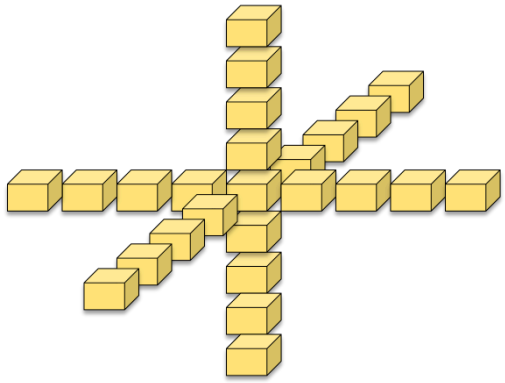

Figure 1: Representation of a centred, symmetric, 8th order finite-difference stencil. [**NR**]

## Reverse Time Migration

Simulating the wave-equation provides the core of seismic imaging, but is only part of the process. The act of repositioning reflection events to create a representative Earth image is called migration, and there are many approaches.

**?**, formally introduced the concept of Reverse Time Migration for exploration seismology, referred to as RTM. It's construction is relatively intuitive: the algorithm propagates both a source wavefield and a receiver wavefield (these will be in 4D, $x, y, z$ and time.) The source wavefield is a forward simulation of the seismic experiment, the receiver wavefield is the back projected data. Events in these fields which are coincident in space and time will represent a true reflection event, all other events are non-physical. By applying an appropriate imaging criterion to these wavefields, a representative image of the Earth scattering field can be created. For RTM, this imaging condition is the cross-correlation of these wavefields, at zero time-lag. This is a representation of Claerbout's imaging condition (**?**), which was conceived by using the exploding reflector model (**?**).

$$m(\mathbf{x}) = \sum_{\mathbf{x}_s,\omega} f(\omega) G_0(\mathbf{x}, \mathbf{x}_s, \omega) \sum_{\mathbf{x}_r} G_0(\mathbf{x}, \mathbf{x}_r, \omega) d^*(\mathbf{x}_r, \mathbf{x}_s, \omega) \tag{5}$$

A frequency domain representation of RTM is more compact, and can be seen in equation 5. Here, $\mathbf{x}$ represents a generic coordinate system, $\mathbf{x}_s$ the current source wavefield coordinate, $\mathbf{x}_r$ the current receiver wavefield coordinate and $\omega$ the temporal frequency. $m(\mathbf{x})$ is the scattering model, $f(\omega)$ the injected source function and $G_0$ are Green's functions that represent the wavepaths between the relevant coordinates. Finally, the two summation operators approximate time and space integrals, and $d^*$ is the complex conjugate of the recorded data.

The aforementioned imaging condition is the implied multiplication of these two wavefields, and subsequent sum over $\omega$. This is equivalent to integrating over time, over the source coordinates, and over the receiver coordinates, resulting in an image created in all spatial dimensions. Only events coincident at the same point in time (and space) coherently interfere, to contribute to the estimate of $m(\mathbf{x})$. Other events are incoherent, and will correlate incoherently.

The data term features a complex conjugate, this reverses the sense of time of the receiver wavefield with respect to the source wavefield. This is clear when considering equation 5 expanded with the implied Fourier transforms; the conjugate changes the sign of the exponent which reverses the sense of time. Consequently, these wavefields cannot be estimated concurrently. When solving for $m(\mathbf{x})$, it is necessary to propagate and save these 4D wavefields separately. Typically, the source wavefield is computed and saved, the receiver wavefield is then back-propagated and the imaging condition is applied at each imaging time step. The result of this correlation is summed to the estimate of the scattering potential.

Multiple additional methods for dealing with this temporal incongruity have been postulated, and some are more appropriate for certain computational systems than others. These will be discussed in Chapter 6 and Section **??**.

While RTM can produce valuable images from few underlying physics assumptions, often, unphysical artifacts and imaging noise are prevalent. This is particularly true if acquisition was not uniform, or strong reflector contrasts were present. To improve the process RTM can be extended to linearized inversion.

## LINEARIZED INVERSION

RTM can be considered an adjoint procedure (**?**). RTM is the adjoint of an idealized modeling operator known as the first-order Born scattering operator. Applying this adjoint will serve to approximate the inverse of the operator, but ideally a better estimate of the inverse is desired (**?**).

$$d(\mathbf{x}_r, \mathbf{x}_s, \omega) = \omega^2 \sum_{\mathbf{x}} f(\omega) G_0(\mathbf{x}, \mathbf{x}_s, \omega) m(\mathbf{x}) \sum_{\mathbf{x}} G_0(\mathbf{x}, \mathbf{x}_r, \omega) \tag{6}$$

This idealized modeling procedure is represented as equation 6, where the terms have the same meanings as in equation 5. If this equation describes how the physics links the scattering potential, $m(\mathbf{x})$, to the recorded data, $d(\mathbf{x}_r, \mathbf{x}_s, \omega)$, then it is desirable to "un-do" the physics to recover the model. The exact inverse of equation 6 does not exist (the corresponding matrix is both singular and not square), but it is possible to achieve a better approximation than simply applying RTM. One solution is to pose the problem as an inversion.

This has been studied in detail over the last few decades, in reference to improving seismic imaging. **?** initially presented the theory and results for, what they referred to as, Least-Squares Migration. This was posed as a linearized inversion problem, and **?** expanded on this, in the particular focus of wave-equation migration and inversion. **?** demonstrated multiple regularization options for improving least-squares imaging, particularly around salt bodies, and **?** extended their orginal work to demonstrate the effectiveness of this technique with incomplete data. In accordance with the nomenclature of these previous studies, the extension of RTM to a linearized inverse process will be referred to as Least-Squares RTM, or LSRTM, from hereon.

$$\mathbf{d} = \mathbf{L}\mathbf{m} \tag{7}$$

$$\mathbf{m} \approx \mathbf{L}'\mathbf{d}. \tag{8}$$

Initially, equation 6 could be simplified by writing the full Born operator as $\mathbf{L}$, the data as $\mathbf{d}$ and the model as $\mathbf{m}$. Implicit in $\mathbf{L}$ is all knowledge of source coordinates, receiver coordinates, and the velocity field. The modeling procedure is now equation 7. Accordingly, RTM is now represented more simply as equation 8, where again $\mathbf{L}'$ is the adjoint of $\mathbf{L}$.

This can be considered as a discretized linear system, whereby the linear operator, $\mathbf{L}$, links our recorded observations to the model we are trying to obtain. In the case of seismic reflection imaging these data are recorded in a plane on the surface of the Earth. However, the model to be delineated, $\mathbf{m}$, exists in three dimensions. The resultant system is underdetermined: there is not enough information in these data to uniquely determine the model. From the view of linear algebra, $\mathbf{L}$ contains some zero eigenvalues, meaning it is rank deficient and consequently non-invertible.

$\mathbf{L}$ is not a square matrix, and exhibits a null space. A square matrix that can be estimated is $\mathbf{L}'\mathbf{L}$, and since it is square, $(\mathbf{L}'\mathbf{L})^{-1}$ can be constructed. The matrix $\mathbf{L}'\mathbf{L}$ is often referred to as the Hessian. Thus, it is possible to solve the system in equation 9.

$$\mathbf{m} = (\mathbf{L}'\mathbf{L})^{-1}\mathbf{L}'\mathbf{d} \tag{9}$$

However, computing the Hessian matrix, let alone the inverse, is not possible due to size constraints. Instead, the process can be posed in a different way. An algorithm can be used that attempts to reduce a measure of how closely the current model estimate matches the data. There are many ways to measure this misfit, and these are referred to as 'norms.' Least-sqaures imaging uses the L2 norm to measure this misfit, this is a scalar measure, and in this context would be the sum of the difference of the squares of the two datasets ($\Sigma(\mathbf{d}^2_{obs} - \mathbf{d}^2_{est})$).

Using this scalar measure of data similarity, an objective function can be constructed; this is a function that will aim to either minimize or maximize a numerical value. In this case, the objective will aim to minimise the L2 norm of these datasets, since this measurement will tend to zero as these datasets approach each other. Two such objective functions can be seen in equation 10.

$$J(\mathbf{m}) = \|\mathbf{d} - \mathbf{L}\mathbf{m}\|_2^2 \tag{10}$$
$$J(\mathbf{m}) = \|\mathbf{d} - \mathbf{L}\mathbf{m}\|_2^2 + \epsilon\|\mathbf{A}\mathbf{m}\|_2^2 \tag{11}$$

If these data are noise free and the operator exactly describes the physics, then it is possible to reduce $J(\mathbf{m})$ to zero.

## Data-space inversion

At this juncture it is necessary to consider two complementary approaches to the problem, these are data-space inversion and model-space inversion. In data-space inversion, forward modeling is used to estimate the misfit between the estimated data and the observed data (**?**). As a sequential process, consider algorithm 1.

**Algorithm 1** Linearized inversion

Calculate initial data-space residual $r = Lm_0 - d$
**while** iter < n_iter; iter++ **do**
  Create gradient $g = L'r$
  Create conjugate gradient $cg = Lg$
  Calculate step length
  Update $m$ and data residual
**end while**
Output model

The stopping criterion (when to output the final model) can either be designed to halt after a certain minimum residual is reached, or after a certain computational duration. For inverse imaging it is usually the latter.

This construction assumes that the background velocity is well constrained. Implicitly, Born modeling states that the velocity model can be divided into high-wavenumber and low-wavenumber components (**?**). The former of these controls the position and amplitude of all reflection events (a representation of the reflectivity), the latter controls the kinematics of the wavefield.

$$\mathbf{m} = \mathbf{b} + \mathbf{h} \tag{12}$$

Ordinarily, linearized inversion will be applied once a reasonable background velocity model has been estimated, otherwise events will be be misaligned when comparing estimated and observed data. Alternatively, linearized inversion can be augmented with a velocity update scheme, where imaging is the 'outer loop' of the process, and the 'inner loop' acts to update the velocity model (**?**). Velocity update schemes will not be discussed in detail in this thesis, but their importance should be mentioned while discussing seismic inversion.

## SEAM EARTH MODEL

Artificial Earth models can be invaluable for understanding how these imaging systems work. To understand and quantify how RTM and linearized inversion contrast it is necessary to use a representative example, where the physics can be controlled. To test these algorithms an imaging target that contains a wide variety of wavenumbers and velocity contrasts is desirable, in this section part of the SEAM (**?**) model is used. A constant density assumption will be implicit during these upcoming sections.

Figure 2 shows a section of the model which was chosen. Features include continuous sedimentary reflectors, rugose and discontinuous reflectors, and a high contrast, steeply dipping salt body. Within this salt body are sedimentary inclusions and a carbonate top. An effective imaging algorithm will be able to resolve these features and converge towards the correct image, within given acquisition restrictions. A high wavenumber representation is shown in Figure 3: these are the features that the imaging methodology should aim to recreate.

An example of a 3D shot simulated over the model is shown in Figure 4. The geometry was a relatively sparse, fixed receiver type survey to emulate that of Ocean Bottom Node (OBN) acquisition. Sources were at the surface in a regular grid, receivers were on a sparse grid below the surface with dense inline sampling (35 receivers, one every 200 m) and sparse crossline sampling (two inlines separated by one kilometre.) The Earth model itself is sampled every 10 m in the three cartesian directions and extends to a depth of 3600 m (almost 12,000 ft), has an average water depth of 800 m (2600ft) and extends 8km x 3km in the $x$ and $y$ directions respectively (26,000 ft x 10,000 ft). To mitigate spurious reflections from the domain boundaries during propagation each dimension is padded by an additional 40 model points (another 400 m, or 1300 ft.)

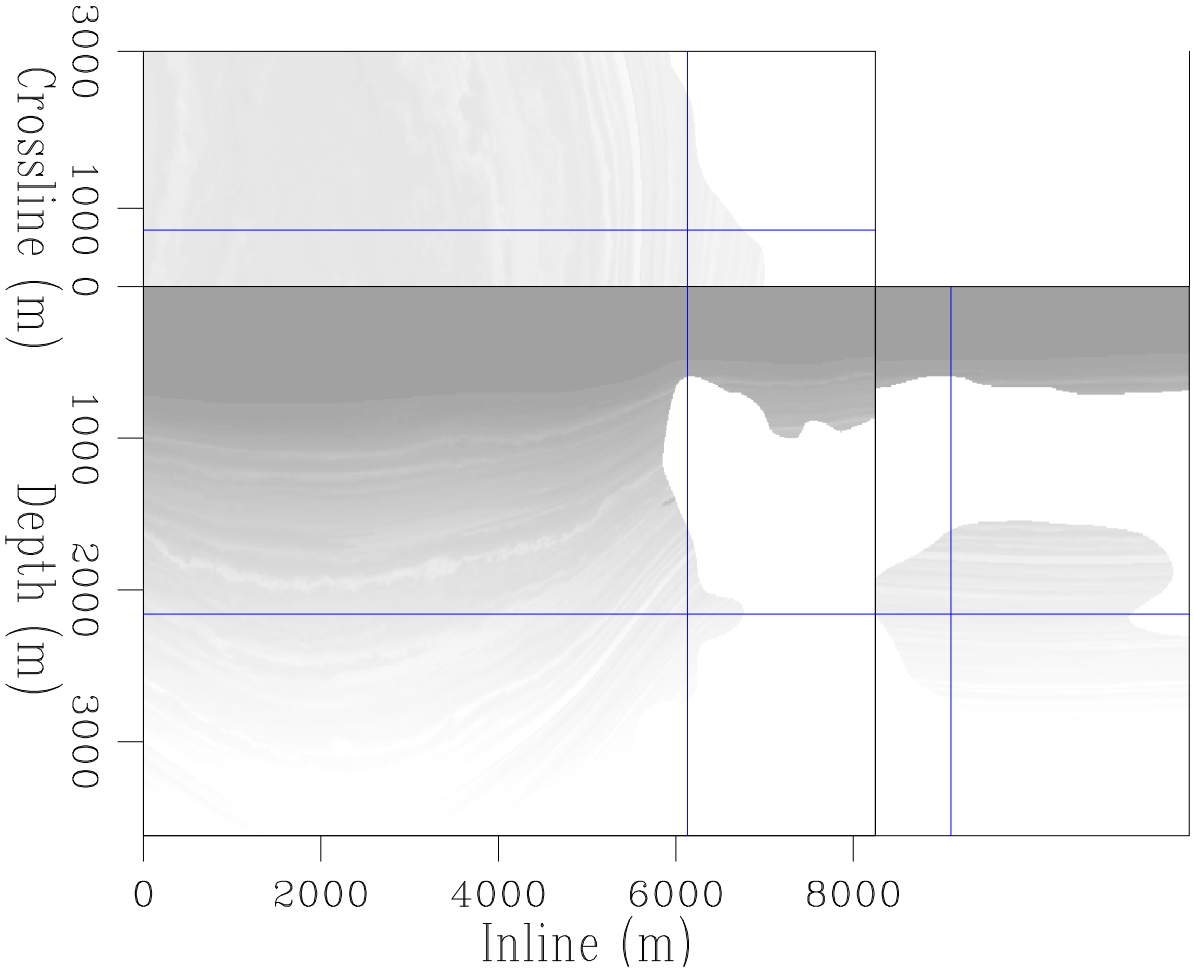


Figure 2: Section of the SEAM velocity model, used for modeling and imaging. [**ER**]

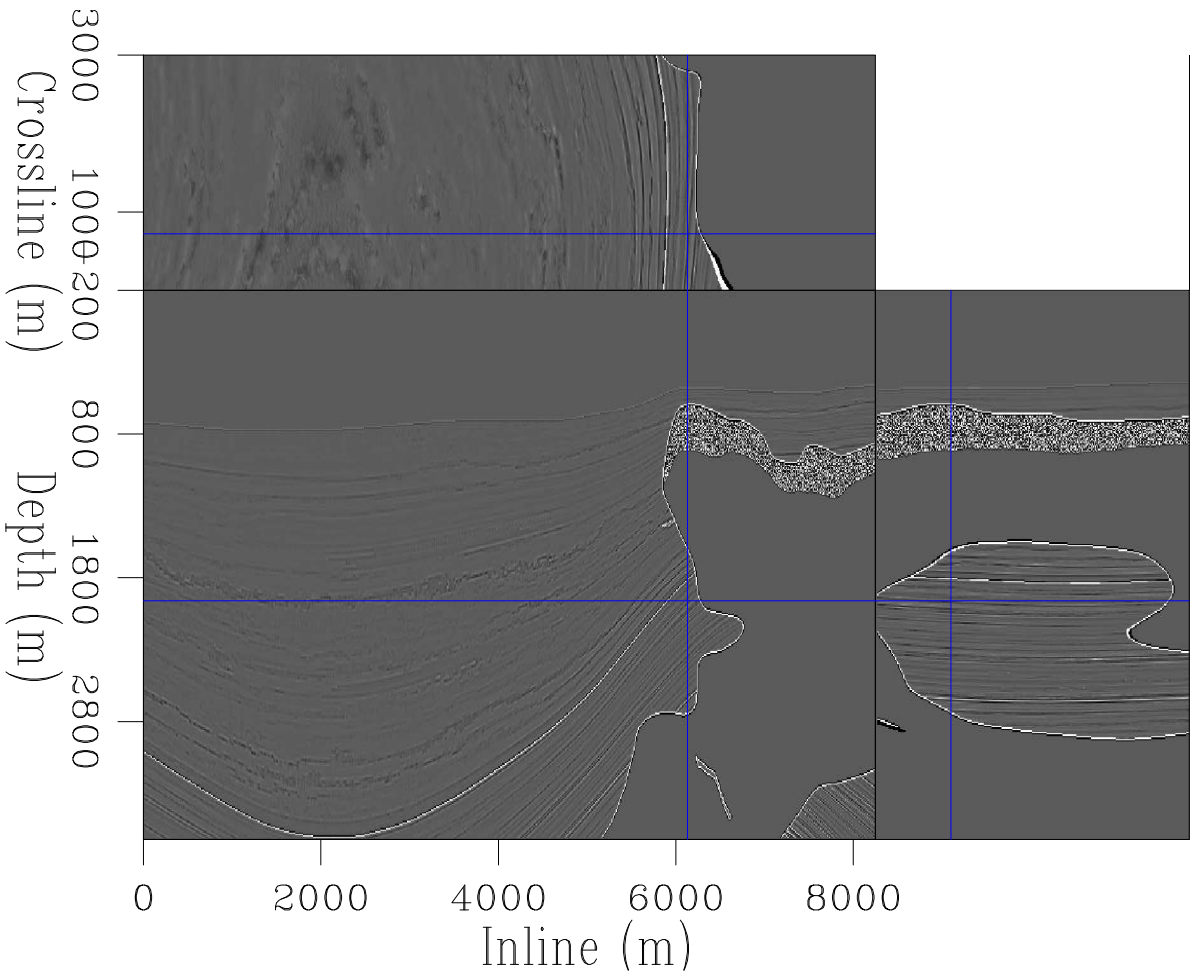


Figure 3: The scattering potential to be imaged, calculated by subtracting a lightly smooth version of the velocity model from itself. [**ER**]

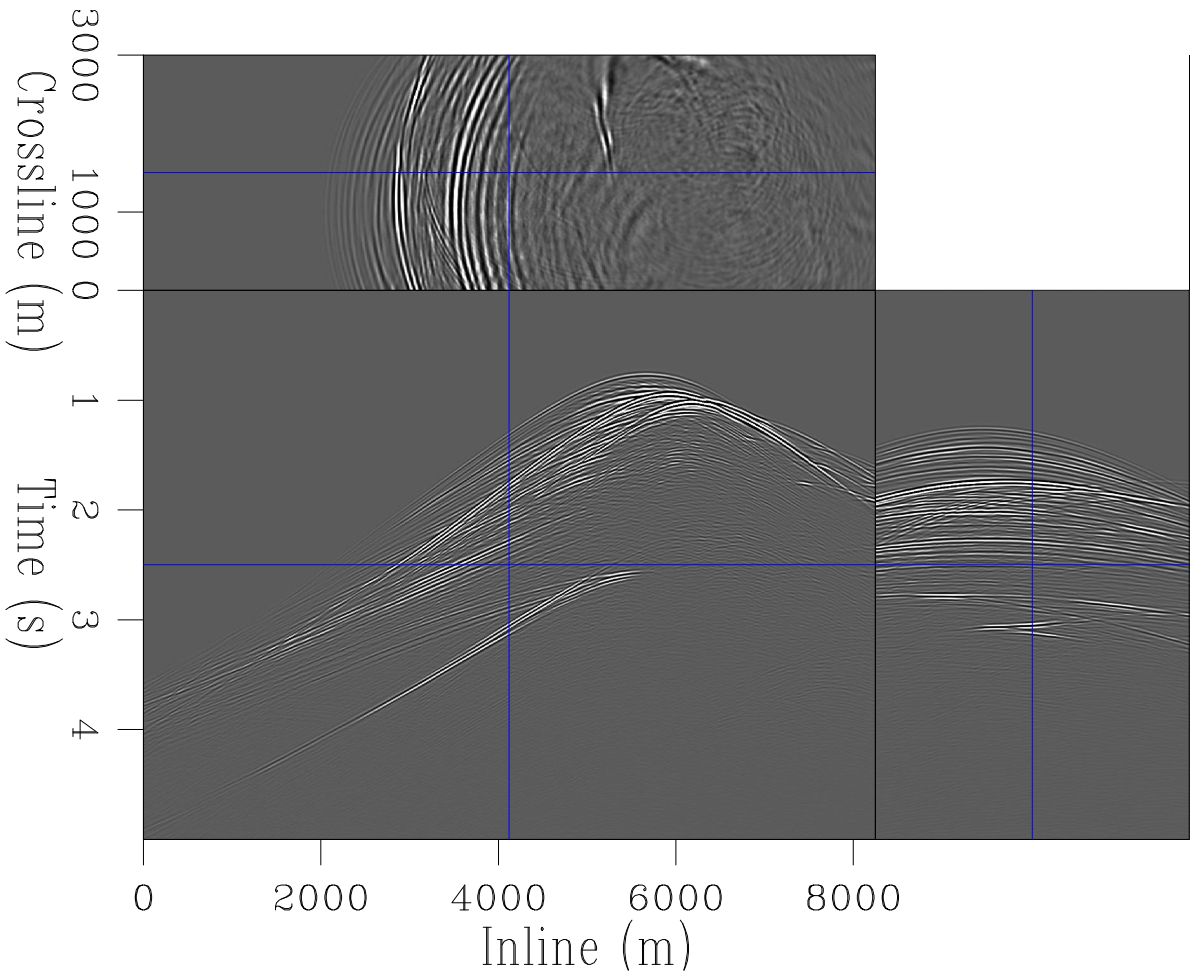


Figure 4: An example of a 3D shot simulated using linearized forward modeling over the proposed velocity model. [**CR**]

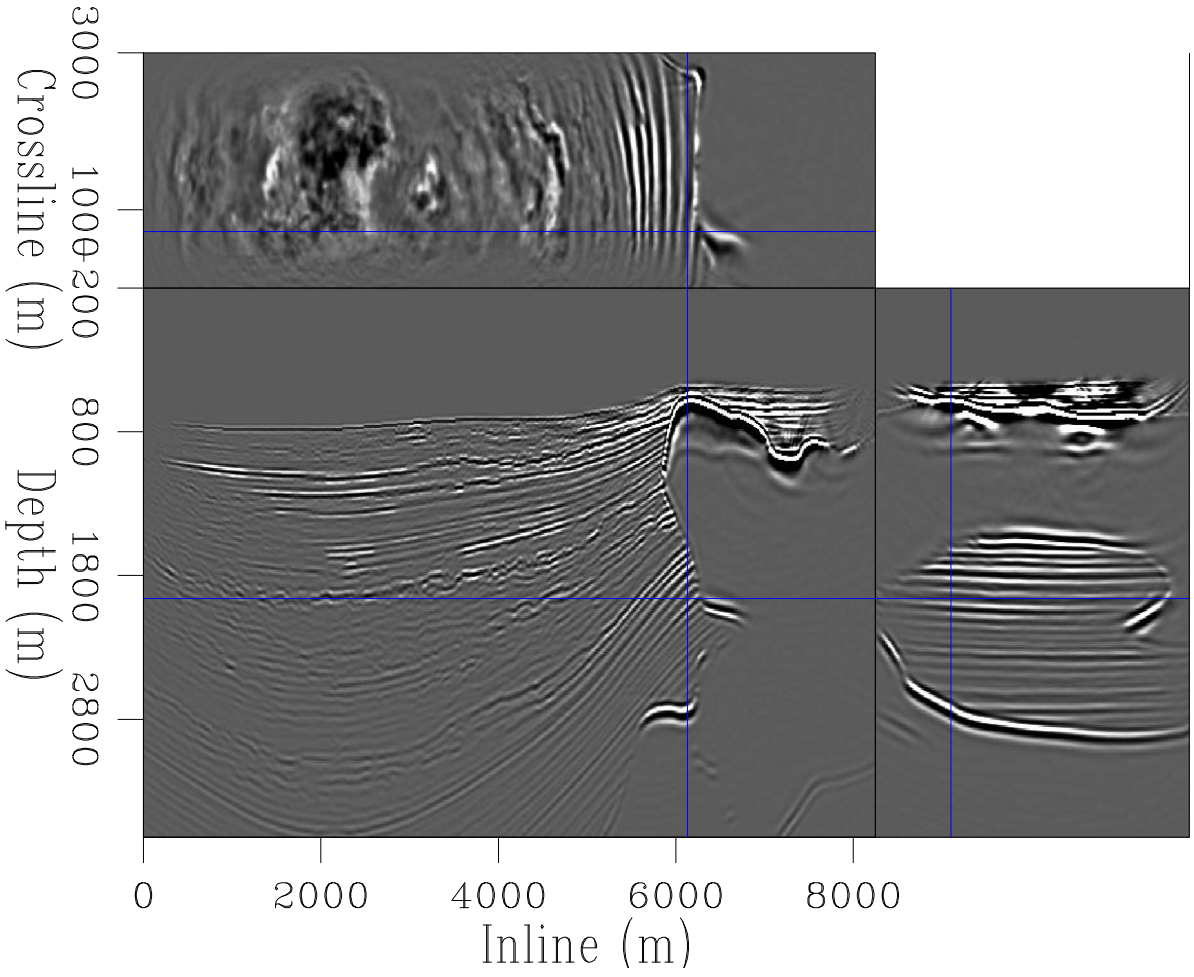


Figure 5: The unfiltered RTM image produced from the synthetic data shown above. For this image the source domain was well sampled with 800 inline shots and 30 crossline shots. [**CR**]

To create these data Born modeling was used. For imaging RTM is used first; this was a full 3D algorithm using random boundaries for the source wavefield modeling, for more details see Chapter 6. For the random boundaries the same additional padding was used (40 model points.) The image obtained from RTM is shown in Figure 5. Noticeably, many nuances of the high wavenumber model have been successfully recreated. Upon some scrutiny, several shortcomings are identifiable. Amplitudes are incorrectly balanced, the steep edges of the salt body are not well focused, the general frequency content of the image is lower than the model (not well balanced through the wavenumber domain) and low frequency artifacts are prevalent throughout. There is also an imprint of the wavelet along the reflectors.

These image imperfections can be attributed to two origins: the imaging procedure was not a true representation of the underlying physics (**?**); the acquisition was limited to a surface plane (**?**). Image noise is induced from limited acquisition, wavelet effects, random correlation noise and from coherent correlation noise (**?**). These will be briefly elaborated upon, in the next few paragraphs.

The imaging condition (equation 5) will result in a squaring of the source wavelet, which is implicit in both $f$ and $d$. Thus, reflectors will appear more spread out, as their spectra will be shifted; this is the wavelet effect alluded to. High frequency noise is present in the image (largely) because of random boundary source wavefield propagation, although this noise source is not significant. More noticeably, low frequency, unphysical events are seen as a result of coherent correlation noise. The imaging condition only holds for wavefields traveling in opposite directions. In a complex velocity model, occasionally the source and data wavefields may travel in the same direction. When correlated, low-frequency artifacts will be created (**?**). In particular, this is seen around the edges of the salt body (**?**). It is possible to harness this effect for velocity updates, since this same direction correlation can provide tomographic information (**?**).

Figure 6 shows the same image crosss-section after a single iteration of linearized inversion, then Figure 7 shows the same section after ten iterations of linearized inversion. Each iteration of linearized inversion is twice the cost of a single pass of RTM imaging. A conjugate direction algorithm, (**?**), was used at each iteration.

It is clear that this process improves upon all these imaging imperfections. Frequency content is higher and more consistently balanced, low-wavenumber artifacts are entirely gone and reflector amplitudes (and illumination) are representatively balanced throughout the section. In particular, inclusions in the salt body are more pronounced and the low-contrast carbonate atop the salt has been resolved. Additionally, the rugose nature of many of the shallow reflectors has been well resolved.

To measure how the procedure is performing the objective function can be plotted as a function of iteration number. Figure 8 shows this, by analysing at the normalized data-space residual (taking the difference of the squares of the estimated data and the observed data.) If the residual is zero then the estimated data exactly matches the observed data. After ten iterations the system has converged to within a few percent.

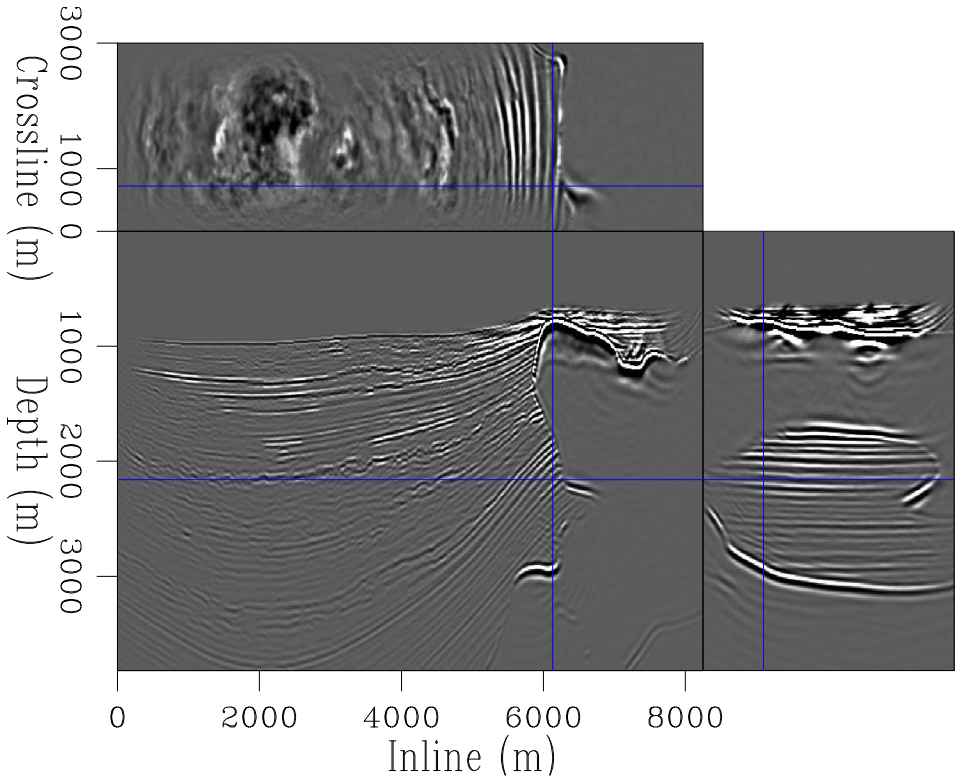


Figure 6: The recovered scattering model after a single iteration of conventional linearized inversion. [**CR**]

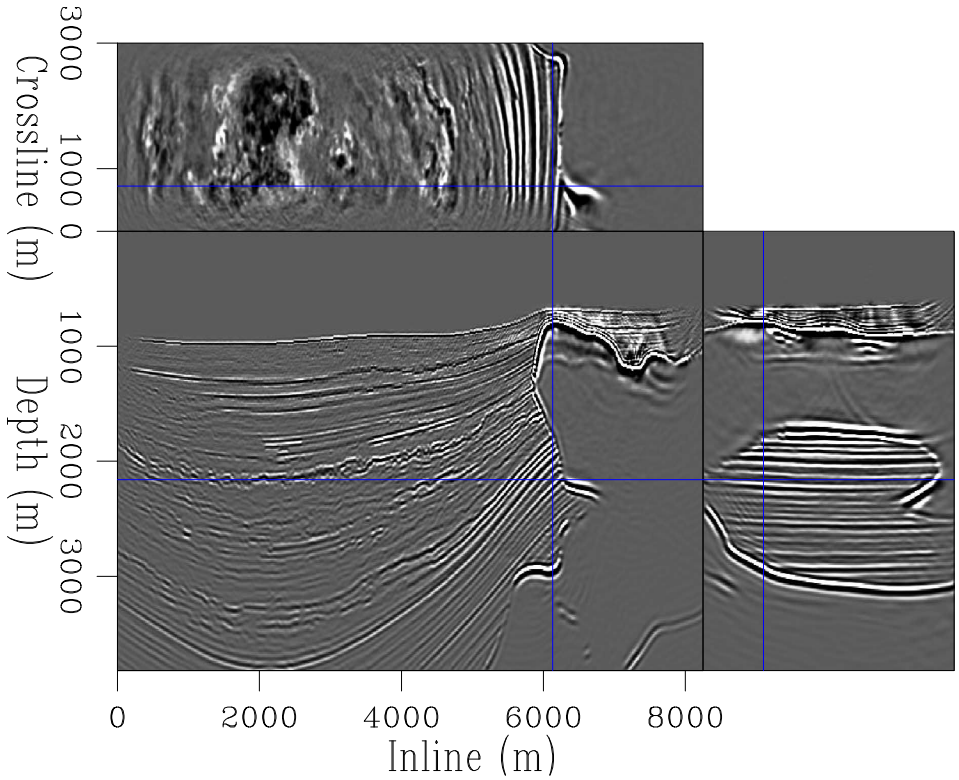


Figure 7: The recovered scattering model after ten iterations of conventional linearized inversion, clipped identically to Figure 6. [**CR**]

This is as expected, since noise-free, Born modeled data were used. With field data convergence to within ten, or even twenty, percent is regarded as successful.

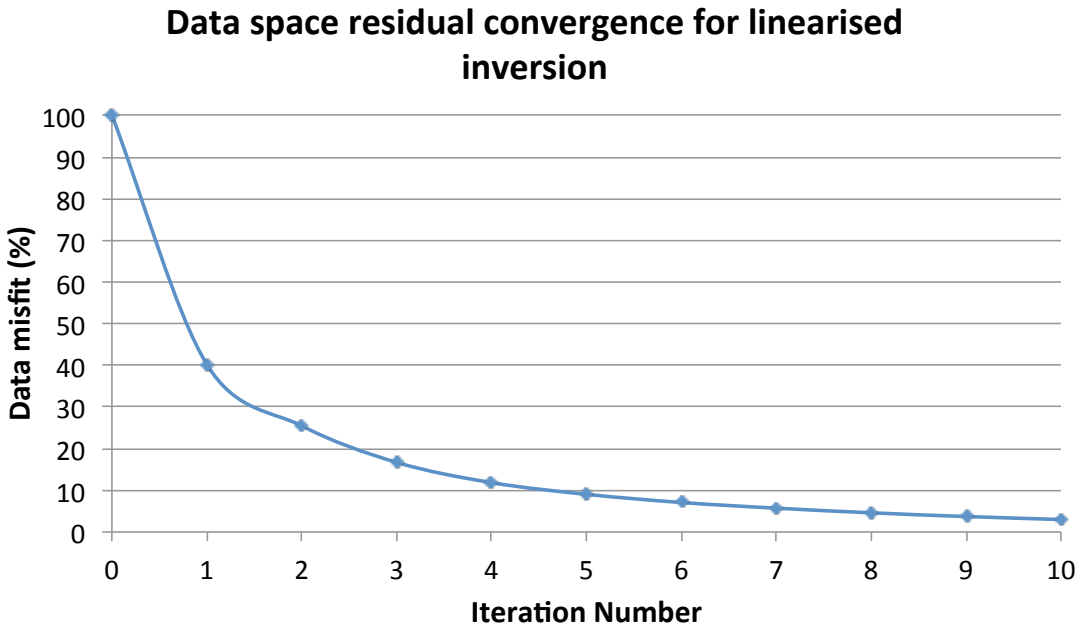


Figure 8: Normalized data misfit as a function of iteration number. Measured from the SEAM model above. [**NR**]

# PRECONDITIONING

For many applications, even ten iterations can become prohibitively expensive. The results in the previous section took the most naive approach to linearized inversion, with no attempt to influence the system with any prior knowledge. There are two main categories of influence that can be applied in seismic inversion, known as regularization and preconditioning (**?**). The former acts to steer the model space towards some preconceived notion or statistic that the model should exhibit (**?**), the latter aims to create a change of model variable that may improve convergence characteristics (**?**). This section will focus on preconditioning, whereas Chapter 4 will address some regularization options.

$$J(\mathbf{m}) = \|\mathbf{d} - \mathbf{L}\mathbf{m}\|_2^2 \quad (13)$$

$$J(\mathbf{p}) = \|\mathbf{d} - \mathbf{L}\mathbf{S}\mathbf{p}\|_2^2 \quad (14)$$

By making the change of variable $\mathbf{m} = \mathbf{S}\mathbf{p}$, equation 10 can be rewritten as equation 13. During each new iteration this operator, $\mathbf{S}$, and it's adjoint must be applied.

$$\begin{aligned} \mathbf{r} &= \mathbf{L}\mathbf{m}_0 - \mathbf{d} \\ \mathbf{gg} &= \mathbf{L}'\mathbf{d} \\ \mathbf{rr} &= \mathbf{L}\mathbf{g}\mathbf{g} \\ \mathbf{r} &= \mathbf{r} + \alpha\mathbf{r}\mathbf{r} \\ \mathbf{m} &= \mathbf{m} + \alpha\mathbf{g}\mathbf{g} \end{aligned} \tag{15}$$

The system will act to recover the most representative $\mathbf{p}$; during the final step this preconditioner can be reapplied to recover $\mathbf{m} = \mathbf{S}\mathbf{p}$.

Equation(s) 15 demonstrates the sequence of operations for linearized inversion. By including a preconditioner, the sequence subtly changes to that seen in equation 16.

$$\begin{aligned} \mathbf{q} &= \mathbf{L}\mathbf{S}\mathbf{p}_0 - \mathbf{d} \\ \mathbf{ss} &= \mathbf{S}'\mathbf{L}'\mathbf{d} \\ \mathbf{qq} &= \mathbf{L}\mathbf{S}\mathbf{s}\mathbf{s} \\ \mathbf{q} &= \mathbf{q} + \alpha\mathbf{q}\mathbf{q} \\ \mathbf{p} &= \mathbf{p} + \alpha\mathbf{s}\mathbf{s} \end{aligned} \tag{16}$$

It is necessary to consider what the preconditioning operator, $\mathbf{S}$, could be to improve convergence. The mathematics of inversion instructs that the gradient can be multiplied by any positive definite matrix and convergence will be observed. It is, perhaps, most instructive to consider how the first gradient could be improved/cleaned to better represent the system. In this case the initial gradient is the RTM image. The subsequent step of the inversion is to model a dataset using this gradient, to construct the conjugate gradient. One could conjecture that a preconditioner that could improve the frequency content of the image and remove the low amplitude artifacts would improve convergence. Options for this could include a vertical derivative, or a spatial Laplacian, applied to the image. Both of these would remove low-wavenumber artifacts and shift the frequency content.

The appearance of the initial gradient after application of these operators can be seen in Figure 9, Figure 10 and Figure 11. Both operators perform as expected, however, it is important to notice that the vertical derivative sets all zero-wavenumber events to zero, meaning it will destroy any vertical events. This is noticeable on the edges of the salt body.

Convergence curves for three preconditioners (also the identity matrix is used, meaning no preconditioning) can be seen in Figure 12. The application of these preconditioners (in terms of computation) is entirely negligible, so the cost of these systems is the same as unpreconditioned inversion. Consequently, there is no need to distinguish between iteration number and cost. In Chapter 3 a wider of inversion

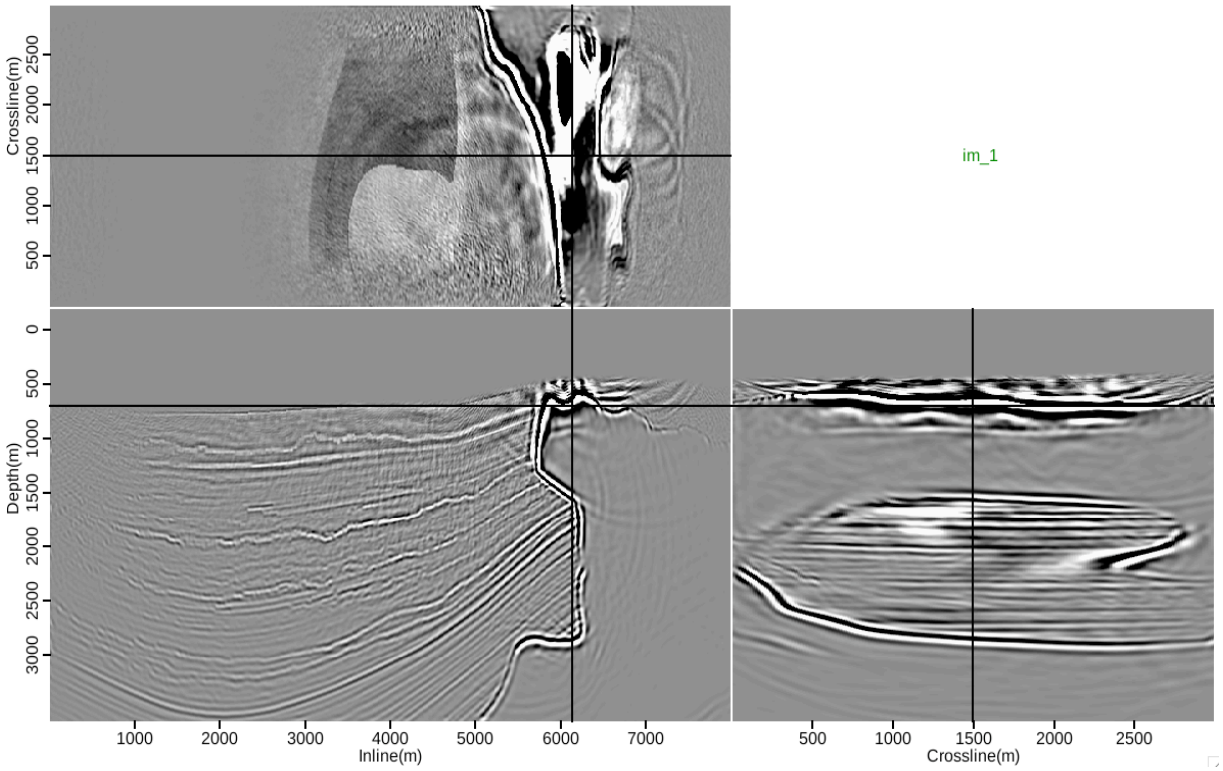


Figure 9: The first gradient of linearized inversion without any preconditioning. [**CR**]

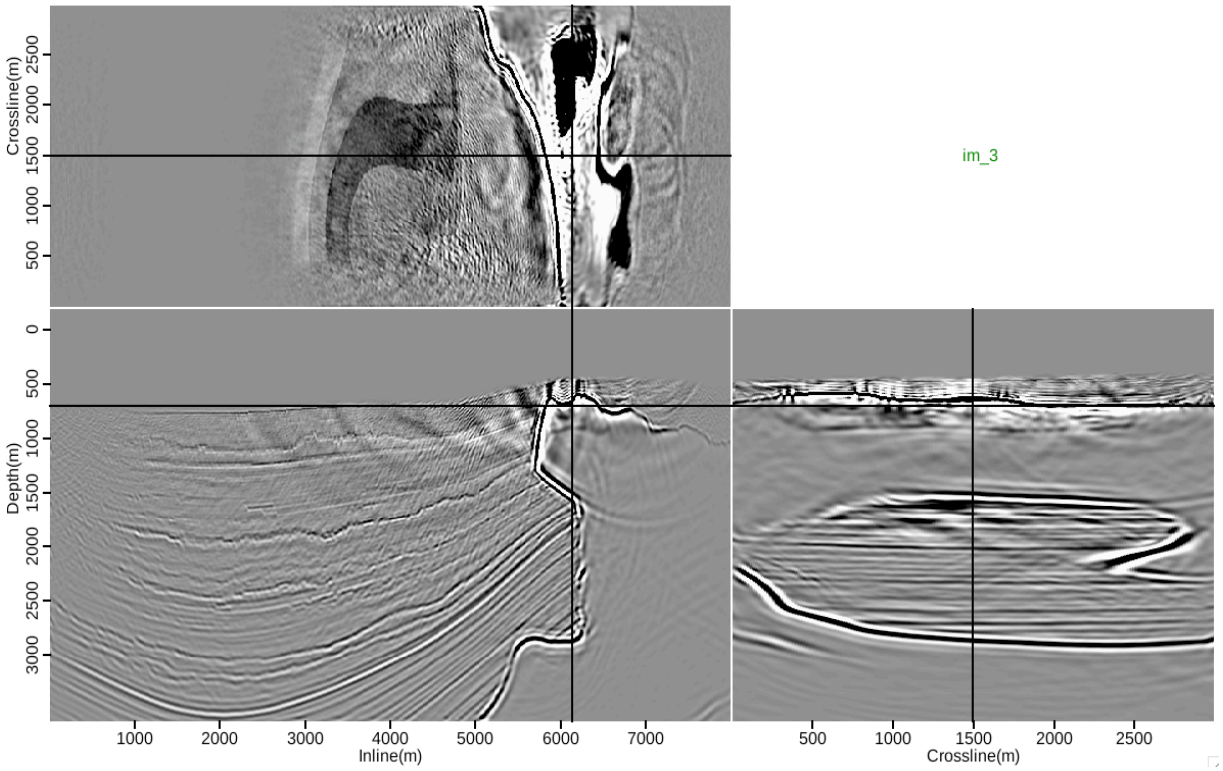


Figure 10: The first gradient of linearized inversion preconditioned with a vertical derivative. [**CR**]

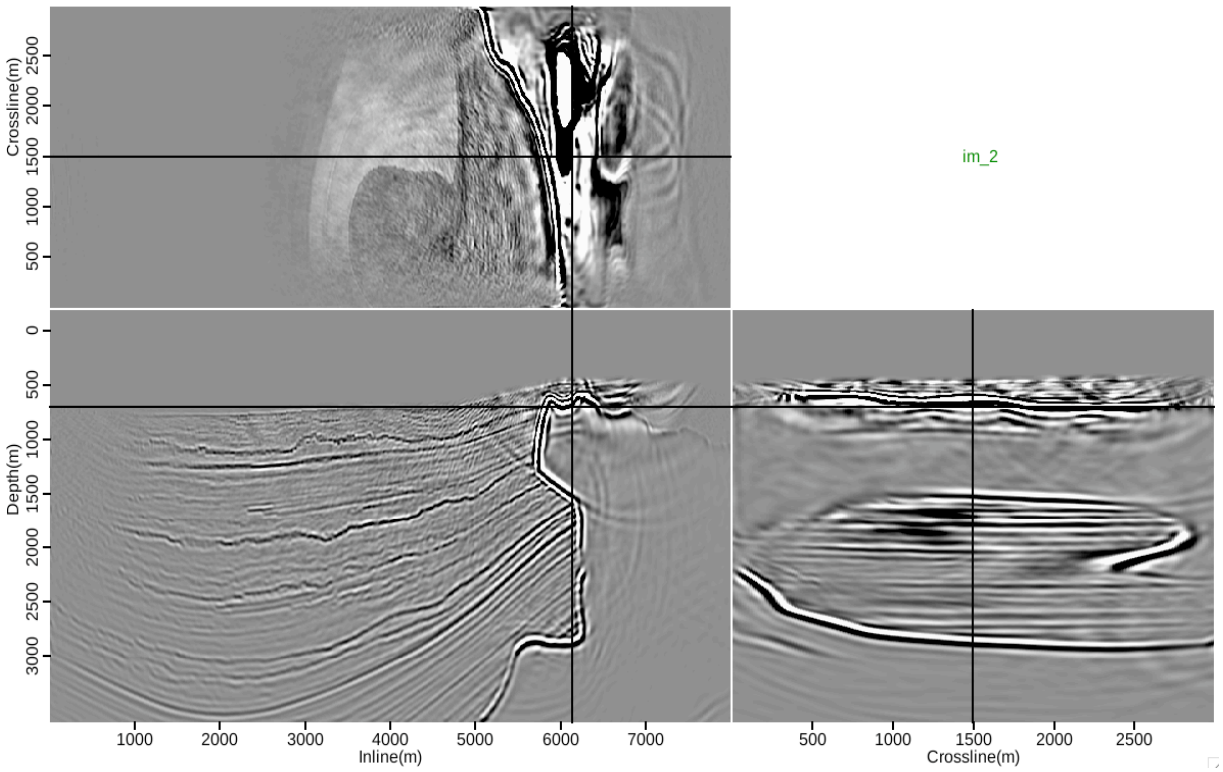


Figure 11: The first gradient of linearized inversion preconditioned with a Laplacian. [**CR**]

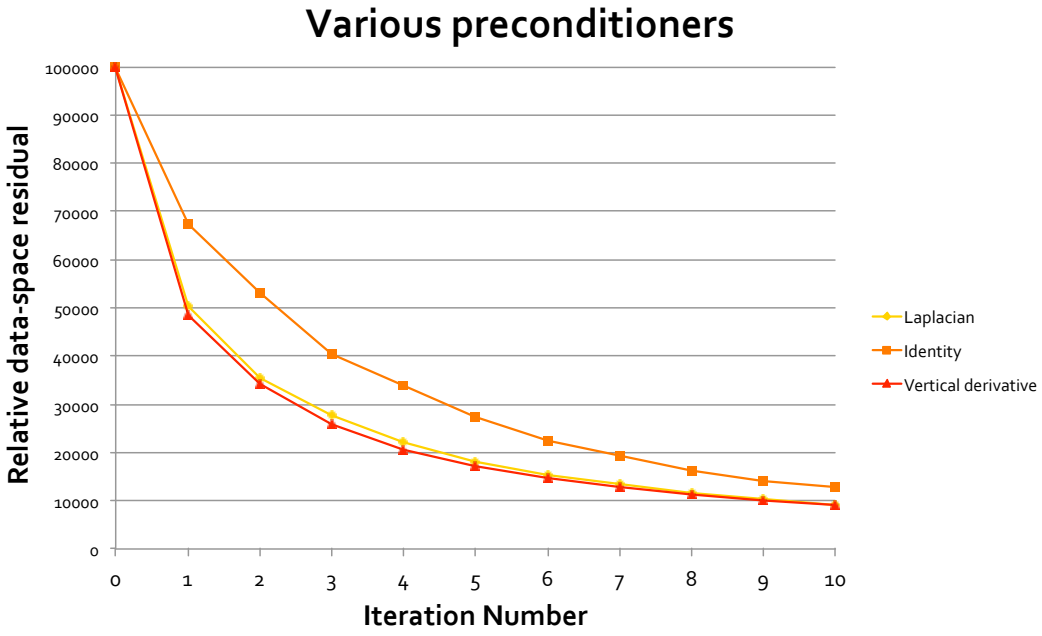


Figure 12: How each preconditioned system performs as a function of iteration, measured by L2 data misfit. [**NR**]

techniques will be discussed, and here a distinction between cost and iteration number with become imperative.

Predictably, the preconditioned systems converges to a lower data-space residual faster than the unconditioned system. Particularly for low iteration results the difference is considerable, giving much fast data-space convergence, at no extra cost. In this case preconditioning is only necessary for the first iteration, as after these artifacts have been removed the inverse system can work on amplitude balancing. One way to think about this is simply providing the system with a more accurate initial model.

Whilst convergence has been accelerated, the procedure it is still very expensive. The subsequent chapter will discuss methods of reducing the data-space in order to greatly reduce convergence as a function of cost.

## CONCLUSIONS

Methods of creates seismic images using wave-equation based approaches were discussed. By assuming data were created from an acoustic, isotropic, Earth, the two-way wave-equation can be reduced to three terms - a forcing term, a spatial derivative, and a temporal derivative. If knowledge of the subsurface velocity exists then the seismic experiment can be simulated.

Through simulating the source wavefield, and back projecting the receiver wavefield using the recorded data, RTM can be applied. RTM gives accurate, representative images of the subsurface, within these restrictions. These images can be greatly improved by extending this linear system to an inversion, and by using an update scheme to iteratively improve the model estimate.

This linearized inversion concept, as a function of convergence, can be accelerated by using preconditioners which can remove RTM induced artifacts faster than the inverse scheme.