Automatic selection of reference velocities for recursive depth migration by peak search method

Hugh D. Geiger and Gary F. Margrave, University of Calgary*

Summary

Wave equation depth migration methods such as phase-shift plus interpolation, extended split-step Fourier, or Fourier finite difference plus interpolation, require a limited set of reference velocities for efficient wavefield extrapolation through laterally inhomogeneous velocity models. In basic implementations, reference velocities are selected as either a linear or a geometric progression spanning the range of model velocities. However, it is unlikely that the model velocities are distributed linearly or geometrically. If the reference velocities can be distributed statistically to more closely approximate the actual distribution, the accuracy of the extrapolation step can be improved.

Key features of our automatic reference velocity selection algorithm are 1) division of the velocity distribution into clusters 2) entropy based statistical control to determine the minimal number of reference velocities required within a cluster, 3) selection of reference velocities within each cluster at or near peaks in the probability distribution, and 4) calculation of a single reference velocity if the velocity range within each cluster is less than a minimum threshold. The automatic velocity selection algorithm is implemented as part of a prestack PSPI depth migration of the 2D Marmousi model. The algorithm is suitable for 3D data, where the tradeoff between accuracy and efficiency is more pronounced.

Introduction

Wave equation depth migration methods such as phase-shift plus interpolation (Gazdag and Sguazzero, 1984), extended split-step Fourier (Stoffa et al., 1990), or Fourier finite difference plus interpolation (Biondi, 2002) require a limited set of reference velocities for efficient wavefield extrapolation through laterally inhomogeneous velocity models. All of these authors suggest that a geometric distribution can produce a good set of reference velocities. Kessinger (1992) suggests that, based on empirical testing, a percentage velocity increment of 15% provides a good compromise between accuracy and efficiency. Another standard method calculates a linear distribution of velocities (e.g. Ferguson and Margrave, 2002), which may have an advantage of increasing the accuracy of higher-angle wavefield propagation with depth. Bagaini et al. (1995) propose a statistical method based on statistical entropy and linear interpolation of the cumulative probability distribution for optimal selection of the number and location of reference velocities. We follow their basic philosophy: that the selection of reference velocities should be driven by the distributional statistics of the model velocities, but utilize statistical entropy and an algorithm similar to pairwise clustering (Velho et al. 1998) to select the number and location of reference velocities nearest to the peaks in the probability distribution.

The optimal choice of reference velocities will be a balance between accuracy and efficiency. In this study, we do not determine a quantitative method for evaluating optimality. Instead, we compare the selected reference velocities against the probability distribution of the velocities over the range for the profile. Figure 1 is a color contour of the Marmousi velocity model. The red lines indicate profiles used in this study.

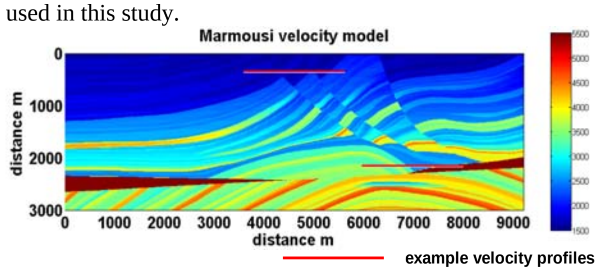

Figure 1: The Marmousi velocity model. Red lines mark velocity profiles used in this study.

Keeping in mind Kessinger's recommended empirical percentage increment of 15%, we can easily determine when we might have too few velocities (Figure 2a). However, a close fit to a velocity profile (as in Figure 2b)

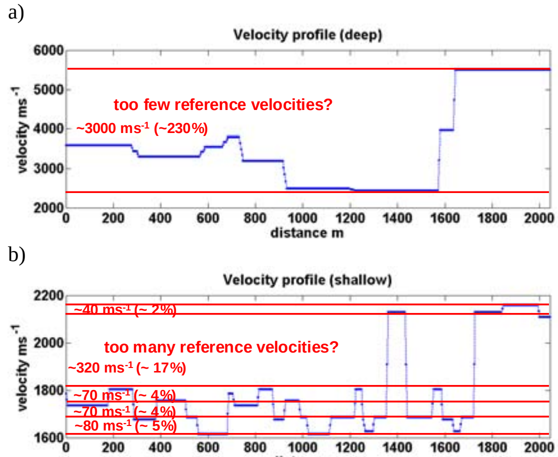

Figure 2: How should reference velocities be chosen for w-k_x domain extrapolators such as PSPI? In a) selection of the bounding velocities creates too large an increment (230%), whereas in b) a careful selection of velocities that closely matches steps in the velocity profile may produce increments that are too small (steps of 2%, 4%).

might not satisfy Kessinger's empirical increment: there might be too many reference velocities in some parts of the range, and too few reference velocities elsewhere. How then should we select velocities to create a best fit to the underlying probability disribution?

In this paper, we review the linear, geometric, and Bagaini et al. methods, show examples where Bagaini et al.'s choice of reference velocities are not necessarily near the peaks in the probability distribution of the velocities over a layer, and propose a new statistical method that produces a more optimal distribution of reference velocities. We test our method against Bagaini et al.'s method using velocity profiles from the Marmousi acoustic model (Versteed and Grau, 1991).

Linear and geometric methods

The linear method can be described as follows:
1) choose an approximate velocity increment $d\tilde{V}$,
2) calculate $nV = \text{int}\left(\left(V_{\max} - V_{\min}\right)/d\tilde{V}\right)$, and
3) determine references velocities as equal increments of $dV = \left(V_{\max} - V_{\min}\right)/nV$.

What is a good choice for $d\tilde{V}$ when using the linear method? One option is a value close to Kessinger's empirical percentage, but it will be difficult to choose one value that will be optimal for both high and low ends of the velocity range (see Figure 3a).

a)

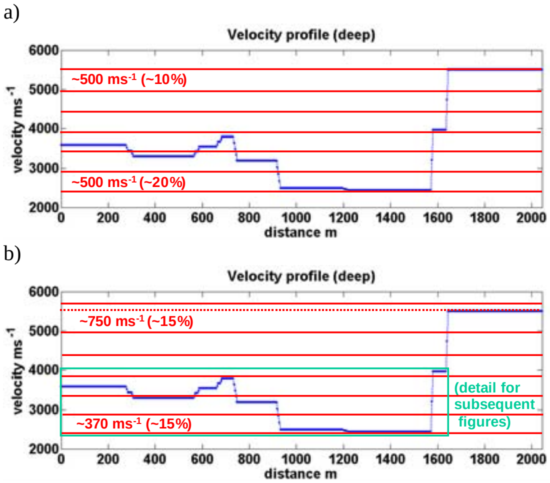

Figure 3: Standard methods for selecting reference velocities include a) a linear interpolation between the minimum and maximum velocities within a layer or cluster, and b) a geometric interpolation based on a fixed percentage increment from the minimum velocity. Empirical testing (Kessigner, 1992) suggests that an increment of approximately 15% provides a reasonable compromise between accuracy and efficiency. The green box in figure 3b) shows the velocity profile used in Figure 4.

The geometric method can be described as follows:
1) choose a percentage step $\%V_s$, and
2) calculate $V_i = \left(1 + \%V_s\right) * V_{i-1}$.

The percentage can be adjusted by solving a simple logarithmic equation if reference velocities are desired at both the minimum and maximum velocities. As mentioned above, Kessigner (1992) recommends $\%V_s = 0.15$. A geometric distribution may yield some reference velocities that are not close to actual model velocities (e.g. at 4300 ms^{-1} and 5000 ms^{-1} in Figure 3b).

The method of Bagaini et al. (1995)

The method of Bagaini et al. (1995) can be described as follows:
1) choose a preliminary $d\tilde{V}$ (perhaps using linear or geometric methods),
2) determine $n\tilde{B}$ equally spaced temporary bins over $\left[V_{\min}, V_{\max}\right]$ (e.g. Figure 4a, where $n\tilde{B} = 6$),
3) bin velocities to give probability density P_i in each temporary bin such that $\sum P_i$,
4) calculate cumulative probability distribution Y_i at each temporary bin boundary, where $Y_i = \sum P_i$ (e.g. Figure 4b),
5) determine optimal number of bins nB by statistical entropy $S = \sum P_i \log P_i$, using $nB = \text{int}\left(\exp(S) + 0.5\right)$ (e.g. Figure 4b, where $nB = 5$),
6) each optimal bin will hold $1/nB$ of the cumulative probability distribution (e.g. $1/nB = 0.2$ in Figure 4b);
7) starting at $V_{\min}$, linearly interpolate position of reference velocities from velocities at temporary bin boundaries using increments $1/nB$ of cumulative probabilities Y_i (e.g. at 0, 0.2, 0.4, 0.6, 0.8 in Figure 4b).

The peak search method

We propose a new method that selects references velocities near peaks in the probability distribution of the model velocities. We use Bagaini's measure of statistical entropy to determine an optimal number of reference velocities. In all likelihood, however, the number of peaks in the probability distribution will not match the statistically optimal number of number of reference velocities. Our method uses a variety of techniques to position the reference velocities closest to the peaks, as follows:
1-5) optimal number of bins determined by statistical entropy following Bagaini et al's method (above).
6) bin model velocities using ~10 times the optimal number of bins, linearly spaced between $\left[V_{\min}, V_{\max}\right]$.
7) combine closest bins, starting with bins of lowest probability density. A new reference velocity is determined as a weighted distance between the two original peaks, except at the ends where the values of the minimum and maximum velocities are preserved as the bounding reference velocities (not necessary for split-step methods).

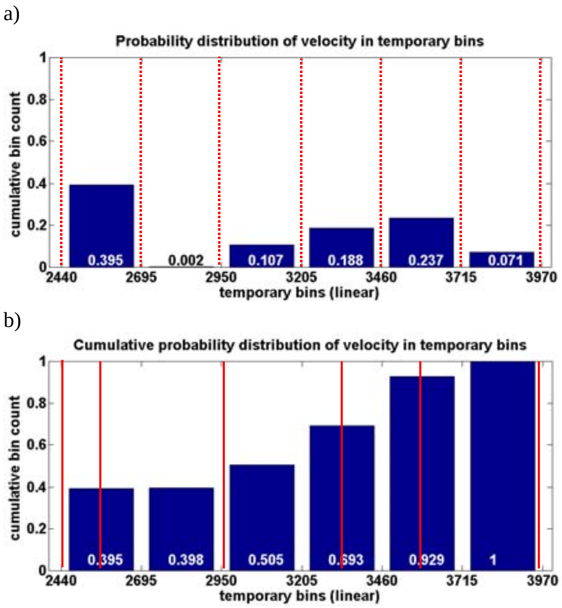

Figure 4: Bagaini et al's (1995) method for selecting reference velocities. a) the velocities are binned to create a discrete probability distribution. An optimal number of reference velocities nB_{opt} is calculated using a measure of statistical entropy. b) the discrete cumulative probability distribution is divided into nB_{opt} equal increments by linearly interpolating from the bin boundaries.

We also use a clustering algorithm to divide the velocity profile into clusters that might be separated by a jump of larger percentage difference than a desired value.

The general clustering method is described as follows

1) sort the velocities in the profile, select maximum percentage increment $c\%V_s$ ($c \geq 1$) and threshold increment $\%V_t \ll 1$,

2) select clusters where the percentage increment in sorted velocities exceeds some maximum value (e.g. 1.5x $\%V_s$),

3) each i^{th} cluster is now defined by a minimum velocity $V_{\min}^i$ and maximum velocity $V_{\max}^i$. Starting with the first cluster (lowest velocity), determine:

a) If $V_{\max}^i / V_{\min}^i \leq 1 + \%V_t$, then $V_{\min}^i \simeq V_{\max}^i$ and this cluster can safely be extrapolated at the average velocity.

b) If $V_{\max}^i / V_{\min}^i \leq 1 + c\%V_s$, then bounding velocities $V_{\min}^i$ and $V_{\max}^i$ are valid without any intermediate velocities;

c) If $V_{\max}^i / V_{\min}^i > 1 + c\%V_s$ there will be at least three velocities in the cluster (including the bounding velocities). Use the linear, the geometric, or statistical entropy to determine an optimal number of bins in the cluster, then apply the peak search method to locate reference velocities.

5) cycle through all clusters determined in 2).

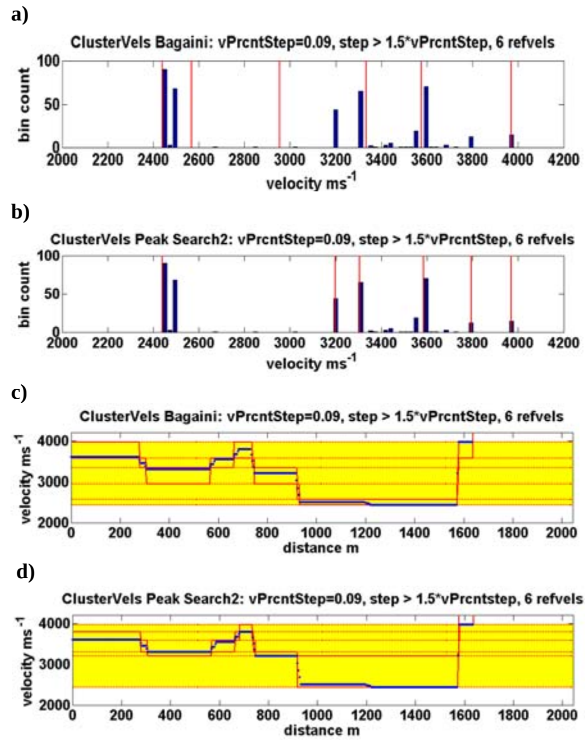

Figure 5:a-b) Probability distribution of velocities (blue bars) for portion of deep profile (green box in Figure 4b) and selected reference velocities using various methods (red lines). c-d) model velocity profile with distance (blue line), selected reference velocities (red dotted lines) and bounding velocities (red lines). Note that in a) and c) the reference velocities by the method of Bagaini et al (1995) are not necessarily at the peaks in the distribution. In b) and d), the peak search method selects reference velocities at the peaks in the probability distribution, providing a more optimal selection of velocities.

Marmousi example

We applied the clustering and peak search techniques to select reference velocities for PSPI migration of the Marmousi data set (Figure 6). Figure 6a is the bandlimited reflectivity calculated using gradients of the velocity and density model. The peak search method produces an excellent image of Marmousi (Figure 6b). The image of the target zone at depth (Figure 6d) compares well with the reflectivity (Figure 6d).

Conclusions

An optimal selection of reference velocities within a depth layer can maximize the accuracy and efficiency of wavenumber/spatial domain wavefield extrapolators such as PSPI.. A linear or geometric distribution does not take into account the underlying distribution of the velocities.

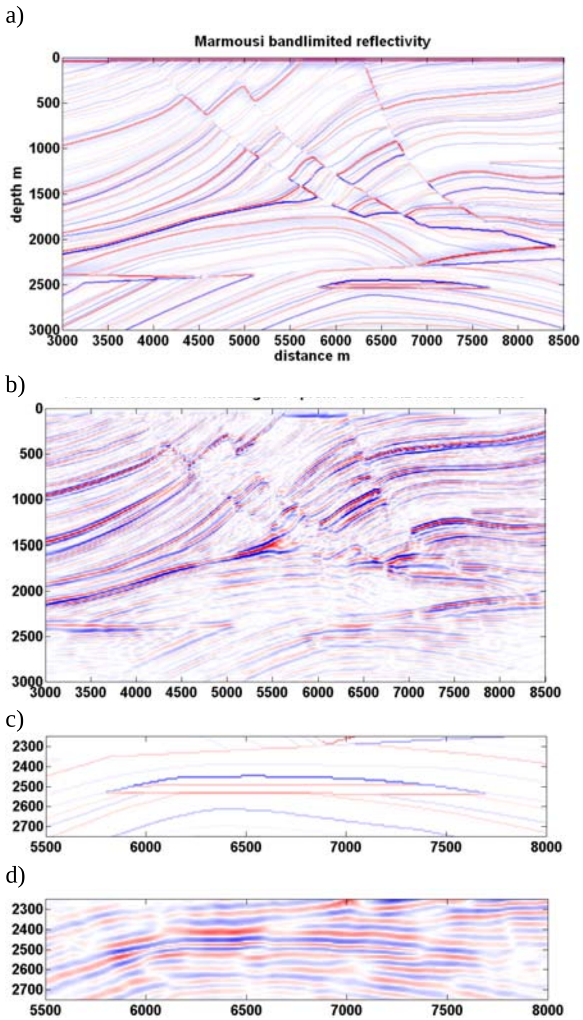

Figure 6:a) Bandlimited reflectivity for Marmousi acoustic model. b) Depth image by PSPI migration with cross-correlation imaging condition using peak search algorithm with clustering, where the optimal number of reference velocities within a cluster are determined by statistical entropy (Bagaini et al., 1995). c) and d) Zoom of full bandwidth reflectivity and PSPI image at the target zone. The target is accurately imaged, although the limited bandwidth of the data results in some ringing.

Hence, large interpolations between reference wavefields may be required, which can result in reduced accuracy and a poor image. The statistical method of Bagaini et al. (1995) distributes the reference velocities based on equal intervals of the cumulative probability distribution of the velocities within a layer. On average, this will reduce the size of the interpolations, but may still produce a reference velocity distribution that requires some large interpolations. We propose a modification to the method of Bagaini et al., whereby the velocities within a layer are first separated into clusters divided by large gaps in the distribution, an optimal number of velocities within a cluster is chosen based on Bagaini et al's method of statistical entropy, and a peak search algorithm is used to select reference velocities within each cluster as close as possible to the peaks in the probability distribution or as weighted averages between smaller pairs of peaks if there are more peaks than the optimal number of velocities. A single reference velocity is selected for clusters or layers where the spread of velocities is less than a specified threshold.

References

Bagaini, C., Bonomi, E., Pieroni, E., 1995, Data parallel implementation of 3-D PSPI: 65nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 188-191.

Biondi, B., 2002, Stable wide-angle Fourier finite-difference downward extrapolation of 3-D wavefields: Geophysics, 67, 872-882.

Ferguson R. and Margrave, G., 2002, Prestack depth migration by symmetric nonstationary phase shift:Geophysics, 67, 594-603.

Gazdag, J. and Sguazzero, P., 1984, Migration of seismic data by phase shift plus interpolation: Geophysics, 49, 124-131.

Kessinger, W., 1992, Extended split-step Fourier migration: 62nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 917-920.

Stoffa, P., Fokkema, J., de Luna Friere, R., and Kessinger, W., 1990, Split-step Fourier migration: Geophysics, 55, 410-421

Versteeg, R., and Grau, G., Eds., 1991, The Marmousi experience: Proc. 1990 Eur. Assn. Expl. Geophys. Workshop.

Velho, L., Gomes, J. and Sobriero, M., 1998, Visualization of color image quantization using pairwise clustering, Proc. 14th Ann. Symp. Comp. Geom., 40-408.

Acknowledgments

We thank the industrial sponsors of both the Consortium for Research in Elastic Wave Exploration Seismology (CREWES) and the Pseudodifferential Operator Theory in Seismic Imaging (POTSI) Project. We also thank the Canadian Organizations: Natural Sciences and Engineering Research Council (NSERC), the Mathematics of Information Technology and Complex Systems network of centers of excellence (MITACS), and the Pacific Institute for the Mathematical Sciences (PIMS) for their funding assistance. Hugh Geiger is a grateful recipient of a PIMS postdoctoral fellowship.