Extrapolation of low frequencies and application to physical modeling data

Scott Keating and Kris Innanen

ABSTRACT

A POCS based frequency extrapolation is tested on 1D synthetic and physical modeling data sets, with the goal of obtaining sufficient low frequencies for inversion. The effects of both noise and available frequency band are also investigated. The addition of noise is found to restrict the effectiveness of our inversion in recovering small impedance changes. Broad frequency ranges at high frequency are shown to produce similar results after extrapolation as small frequency ranges at low frequency. Application to physical modeling data met with moderate success, but offered substantive improvements in the inversion as compared to that of the original band-limited data.

INTRODUCTION

In seismic inversion, we attempt to recover some property of the subsurface based on our seismic measurements. While inversion is ideally performed on full bandwidth data, in reality constraints on seismic acquisition limit the frequency range of the data we have available. When we wish to recover the long scale trends, or absolute amplitudes of the property we invert for, it is the very low frequencies which are of the greatest importance to our inversion (Innanen (2015)). Unfortunately, these very low frequency data are some of the most difficult to obtain, and they are almost inevitably missing in real measurements. Commonly, this deficit is made up for with low frequency data from other sources, such as well logs. This ultimately means that the long scale behaviour in which we are interested is coming from these introduced sources, and the seismic data is not being used to recreate the data at all on these scales. An alternate approach is to approximate the data at our missing frequencies via extrapolation, using the data that we have measured, and to use these extrapolated frequencies in our inversion. In this paper, the use of a frequency extrapolation using the projection onto convex sets (POCS) algorithm, and its effect on inversion are studied.

POCS

In POCS, we view the data which we wish to extrapolate as the sum of two functions; the 'ideal', extrapolated result we hope to obtain, and some 'gapping function' which, when added to the ideal function, introduces the missing points we hope to extrapolate. The POCS method attempts to use discrimination between these two functions in order to extrapolate. At its core, POCS consists of applying a transform to the data, discriminating between the ideal data and the gapping function in the transform domain, and transforming back to the original domain (Abma and Kabir (2006)). By doing so, we obtain an estimate of the ideal function, which can then be used to replace the missing points in our measured data. There are two crucial assumptions behind POCS interpolation. The first of these is that the 'ideal' data are represented in the transform domain with relatively few, high amplitude coefficients. The second assumption is that the 'gapping function' is not special in the transform domain and so consists of randomly distributed and relatively low amplitude coefficients in this transform domain.

The POCS method consists of the following steps. First, the data are transformed to some transform domain. Provided our assumptions are satisfied, in the transform domain we will have a few high amplitude coefficients (from the ideal function), and an abundance of randomly scattered, low amplitude coefficients (from the gapping function). Due to the differences between the ideal function and the gapping function in this transform domain, we now have a means of distinguishing one from the other. The high amplitude coefficients should primarily represent the ideal function, with only small contributions from the gapping function (which is more evenly distributed). The low amplitude coefficients contain most of the information corresponding to the gapping function, and for the ideal function only represent whatever small fraction was not well represented as high amplitude in the transform domain. We then apply a threshold function, zeroing all coefficients in the transform domain below a certain amplitude. This should remove most of the data belonging to the gapping function, while preserving most of the data belonging to the ideal function due to their differences in amplitude. We then inverse transform, which gives us an approximation of the extrapolated, ideal function. By replacing our measured data with this approximation at the missing points only, we obtain an estimate of the extrapolated data. This estimate will not be perfect, as the application of the thresholding function removed those parts of the ideal function which fell below the threshold in the transform domain, as well as preserving the parts of the gapping function which were represented at points that laid above the threshold. Because our estimate is closer to the ideal function than our original data, the gapping function is now lower in amplitude than in the original data. This means that if we repeat the above procedure of forward transformation, threshold, inverse transform and replacement, we can safely lower the threshold (Abma and Kabir (2006)), as the gapping function will also be lower in amplitude. This allows us to recover more of the ideal signal, while still removing most of what remains of the gapping function. By iterating this procedure, we eventually obtain something very close to the ideal function.

The basic idea of the POCS method is illustrated in Figures 1 and 2. On the left, we see that our measured data is equal to the sum of an 'ideal function' (in this case a sine wave) and some 'gapping function' that introduces our missing points. On the right, we see the representations of each of these functions in the transform domain (in this case a Fourier transform is used). Here, we can easily distinguish between the ideal function, which by the nature of the transform chosen is high amplitude at a single point, and the gapping function, which is more evenly distributed. By inverting only the data above the threshold in the transform domain, we recover an estimate of the ideal function (Figure 2, top). We can insert this estimate into the missing portions of our measured data to obtain a first update to our measured data (Figure 2, middle). When we iterate this process, we can safely lower the threshold, as more of the gapping function is removed on each iteration (Figure 2, bottom).

Thus far, we have neglected the fact that any noise present in the signal will be attributable to neither to our ideal function, nor to our gapping function. On any given iteration, we expect the noise to be relatively low amplitude, and randomly distributed. This means that the ideal function will be easily distinguishable from the noise through the use

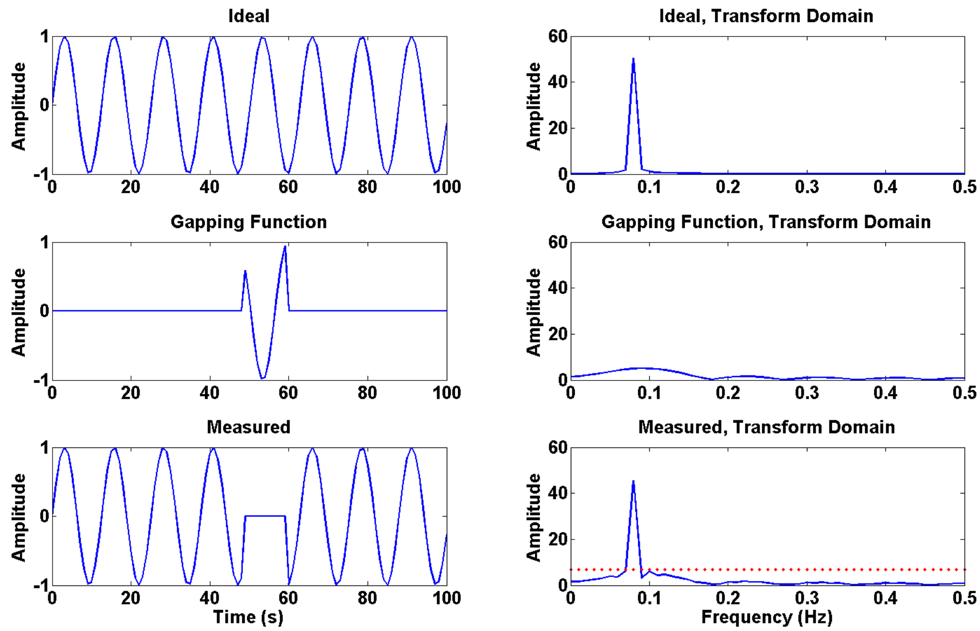

FIG. 1. Left: Measured data can be viewed as the sum of the ideal extrapolation we hope to achieve, and some gapping function. Right: Ideal extrapolation is easily distinguishable from the gapping function in the transform domain by means of amplitude difference. In this example, a Fourier transform is used. The red dotted line represents a possible threshold amplitude for this iteration.

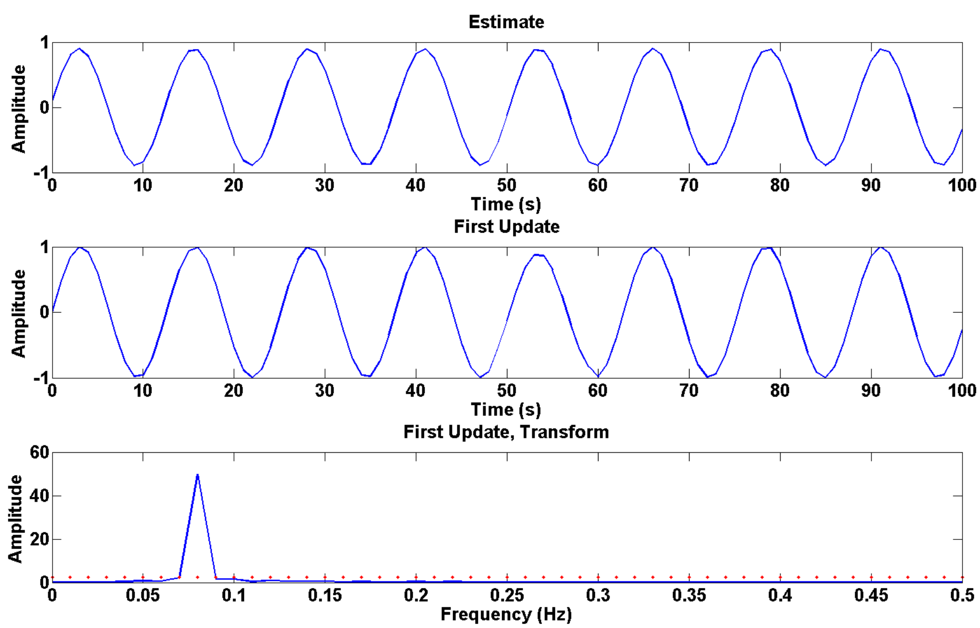

FIG. 2. Top: Estimate of the extrapolated data. Middle: For the update, missing data points in the measured data are replaced with our estimate. Bottom: Transform of the estimate. As our first iteration removed much of the gapping function, we can safely lower the threshold for further iterations. The red dotted line represents a possible threshold amplitude for this iteration.

of a threshold, in exactly the same way that the ideal function is distinguished from the gapping function. On each iteration, however, we will be replacing only those points in our measured data which were originally missing. This means that on every iteration, even though most noise will be eliminated in our estimate of the ideal function, only those point which were missing in the measured data will be replaced by this estimate. Consequently, as the noise is a property of our measured data, not the missing points, we will be keeping all the noise by preserving our measured points. So, while the gapping function will decrease in amplitude with each iteration, the noise will not. By remaining at the same amplitude on each iteration, the noise introduces a minimum amplitude we can lower the threshold to while still discarding most of the noise. This in turn limits the extent to which we can recover the ideal function, harming the quality of our extrapolation.

It is very important in POCS that we choose some transform which causes the ideal function to be sparse and high amplitude. In the case of extrapolation of frequencies of seismic data, the Fourier transform is an appealing choice. This is because the ideal function we hope to obtain in this case for seismic data should be the reflectivity series we expect in the time domain, specifically a series of delta functions. As delta functions are very high amplitude and sparse, the Fourier transform should fulfill the assumptions we make in POCS.

APPLICATION TO 1D SYNTHETIC DATA

As a test, a POCS based frequency extrapolation was applied to a simple 1-D, band-limited data set. This band-limited data was created by taking a synthetic series of reflectivities, and then applying a boxcar filter to this series. In Figure 3, a band-limited reflectivity series, using a 2-50 Hz boxcar filter is displayed. A POCS extrapolation of this band-limited series allowed for the recovery of the series shown in Figure 3. Importantly, while we have not introduced any new measurements to get this reflectivity, the impedance inversion we can perform using these extrapolated data shows massive improvements over that from our band-limited signal (Figure 4). Some small errors can be seen in this extrapolation. In the region of the two reflections near 0.8 s, the extrapolated reflectivity seems to have some errors. This is due to the fact that the two band-limited delta functions in this area interfere with one another, changing where the maximum amplitude is. In POCS, we assume that the highest amplitudes correspond to the ideal function, this interference can invalidate our assumption, and cause problems in our extrapolation. This also means that if we do not have the high frequencies needed to resolve very close reflections, we will not be able to extrapolate for these frequencies, as the resulting interference will render the assumptions we make in our POCS extrapolation incorrect. In Figure 4, we see that despite these high frequency errors, our inversion works well, as the high frequencies have little impact on the result.

Experimentation with the frequency band of our data can give us additional insight into the uses of this extrapolation. As we are using extrapolation to recover our missing frequency information, we may expect both the size of our frequency band, and the difference in frequency between our desired frequencies and our known ones to have an impact on the quality of the extrapolation. Specifically, we expect a decrease in extrapolation quality the further away our measurements are from the desired frequency, and an increase in

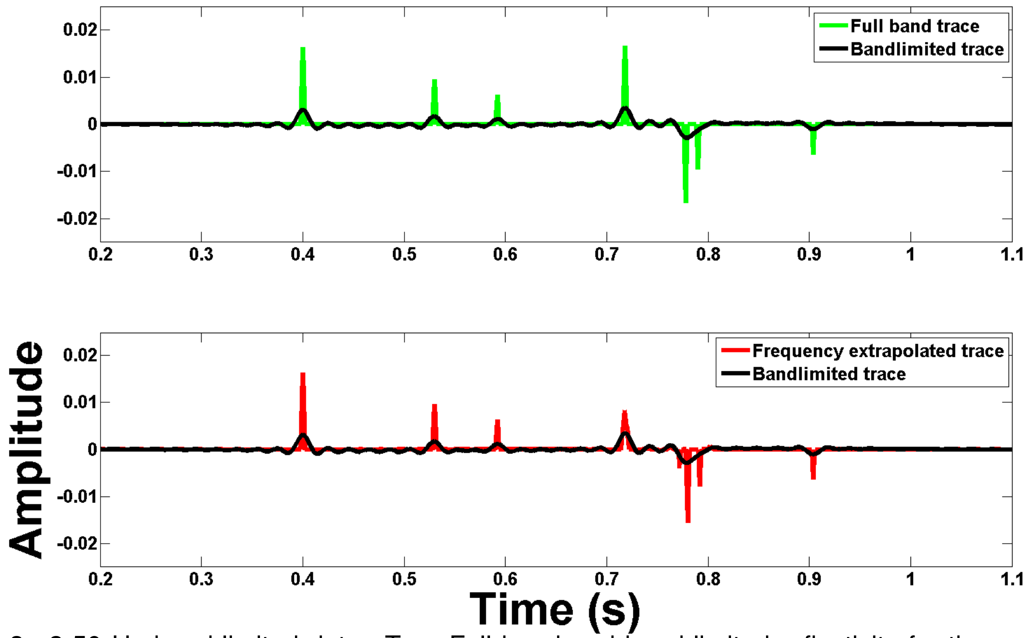

FIG. 3. 2-50 Hz band-limited data. Top: Full band and band-limited reflectivity for the synthetic model Bottom: Band-limited reflectivity and reflectivity after frequency extrapolation

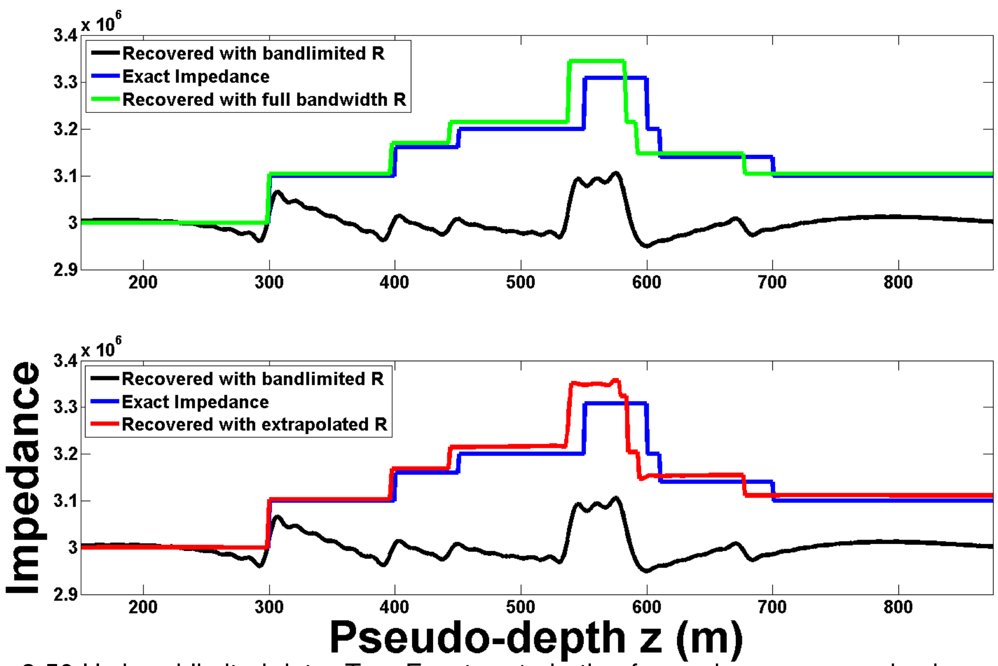

FIG. 4. 2-50 Hz band-limited data. Top: Exact perturbation from a homogeneous background, the result of an impedance inversion using the band-limited reflectivity, and the result using full band reflectivity. Bottom: Exact impedance from a homogeneous background, the result of an impedance inversion using the band-limited reflectivity, and the result using the frequency extrapolated reflectivity

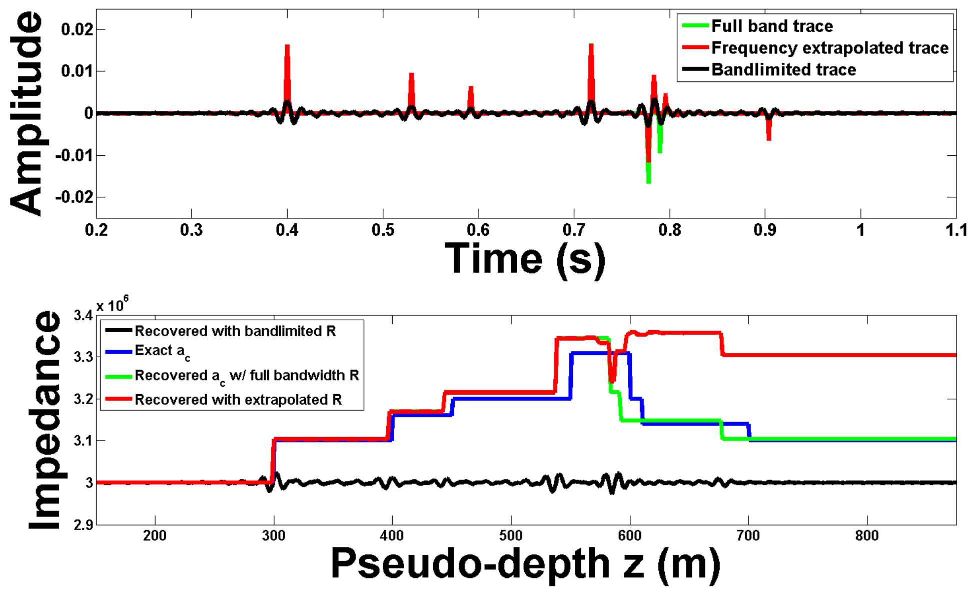

FIG. 5. 42-90 Hz band-limited data. Top: Full band reflectivity, band-limited reflectivity and reflectivity after frequency extrapolation. Interference between the two spikes near 0.8s has led to an incorrect extrapolation. Bottom: Exact impedance from a homogeneous background, the result of an impedance inversion using the band-limited reflectivity, the result using the frequency extrapolated reflectivity, and the result using full band reflectivity.

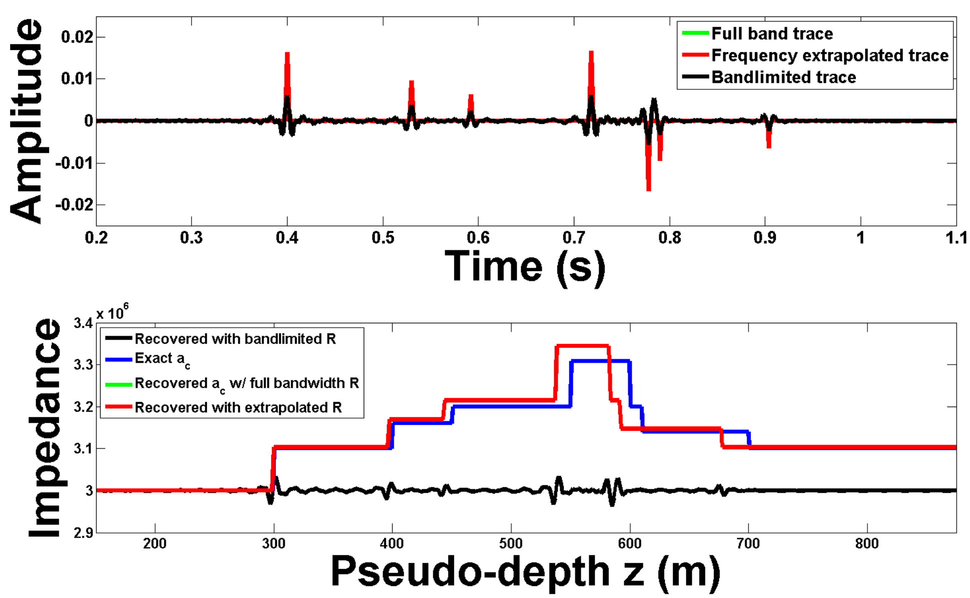

FIG. 6. 42-135 Hz band-limited data. Top: Full band reflectivity, band-limited reflectivity and reflectivity after frequency extrapolation. The full band and extrapolated reflectivity curves overlay one another. Bottom: Exact impedance from a homogeneous background, the result of an impedance inversion using the band-limited reflectivity, the result using the frequency extrapolated reflectivity, and the result using full band reflectivity.

quality the broader our measured frequency band is. Consequently, we might expect that a trace severely lacking in low frequency measurements might still be correctly extrapolated provided we measure a sufficiently broad range of frequencies. Suppose that instead of the 2-50 Hz band in Figure 3, we instead have a minimum measured frequency of 42 Hz. As our minimum frequency is much farther from the very low frequencies we are interested in, we may expect that a frequency band of the same size will yield poorer results. In Figure 5, the extrapolation result for a frequency band of 42-90 Hz is shown. As we expected, our extrapolation has not been successful in this case. Here, our extrapolation again lacks the high frequencies required to resolve the two reflections at 0.8s. Without the low frequencies we had in the 2-50 Hz case, this interference causes more severe problems in our extrapolation. If we were to increase the frequency band, we might expect to improve this result. With an increased frequency band of 42-135 Hz we can see that the frequency extrapolation, and subsequent inversion, are much more successful (Figure 6). We now have sufficient frequencies to distinguish the two nearby events, and correctly identify the locations of our full band delta functions. The success of a 42-135 Hz band means that even though we are most interested in very low frequencies in inversion, if it is easier to obtain a broad range of high frequencies, we can use them to arrive at the same result via frequency extrapolation.

APPLICATION TO NOISY 1D SYNTHETIC DATA

So far, this frequency extrapolation method has performed quite well on these simple synthetic data. As discussed in the POCS section however, noise can significantly impair the quality of the frequency extrapolation we can achieve. If we were to keep the same threshold as in the noise-free case, we would interpret much of the noise as being real reflectivity, and the resulting impedance inversion would suffer badly . We instead choose to limit the minimum value of the threshold based on the amplitude of the noise, which limits the extent to which we can recover the ideal function. In Figure 7, a frequency extrapolation is performed on the same 2 to 50 Hz band-limited trace as in Figure 3, but now with uniformly distributed noise added prior to band-limiting. The maximum amplitude of this noise is one tenth the maximum amplitude of the full band reflectivity. It is evident in Figure 7 that this noise has a negative effect on our extrapolation, and the resulting inversion. Specifically, the noise has caused the amplitudes of recovered reflectivities to be incorrect, and in the inversion the noise causes fairly large changes in the impedance values recovered.

Previously, the use of a higher frequency, but broader band signal was discussed as an alternate way of achieving the same frequency extrapolation. Interestingly, the addition of this uniformly distributed noise does not have an equal impact on these two extrapolations. Figure 8 shows the case where the 42-135 Hz band-limited data is extrapolated after introducing uniformly distributed noise, again with an average amplitude equal to one tenth the maximum amplitude of the full band reflectivity. In comparing Figures 7 and 8, we see an interesting result. In the low frequency case, the low frequency noise we have from our measurements has a very significant impact on the final result of our inversion, for exactly the same reason that low frequencies are important in our inversion. By contrast, in the broad band, high frequency case, we are left with high frequency noise, which has a relatively small impact on the inversion. In Figure 8, we also see a problem that is not specific

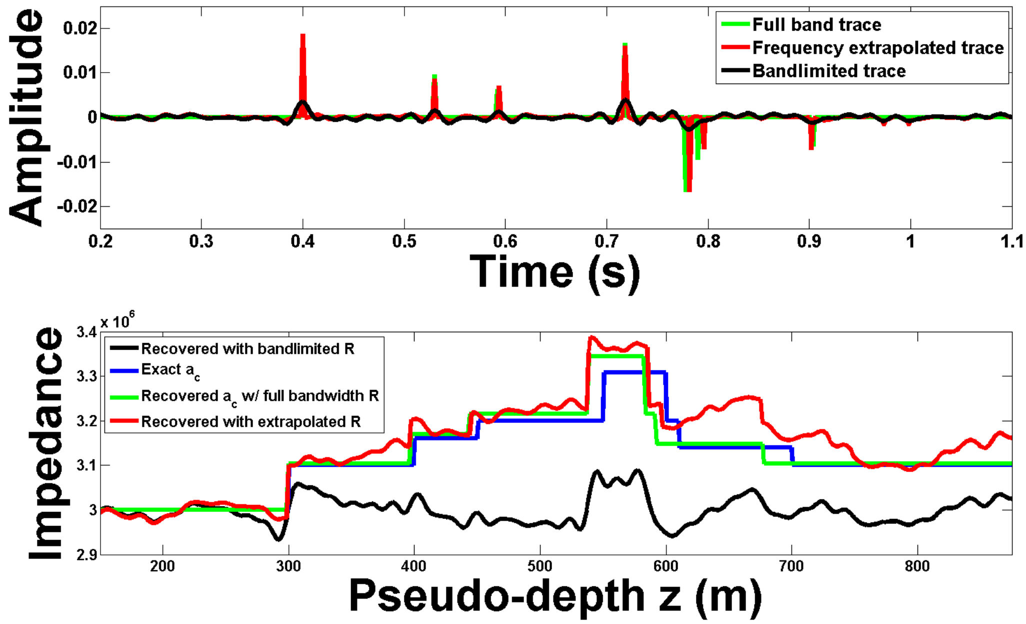

FIG. 7. 2-50 Hz band-limited data with noise. Top: Full band, band-limited and frequency extrapolated reflectivity. Bottom: Exact impedance from a homogeneous background, the result of an impedance inversion using the band-limited reflectivity, the result using the band-limited reflectivity, and the result using full band reflectivity.

to any one frequency range. Here, the reflection at 0.9 s has fallen below the minimum threshold used, as it lies below the maximum of the noise. Because of this, we are unable to recover this reflection.

There are also problems associated with the addition of noise in the high frequency case, however. When there are only high frequencies present, the measured signal varies very sharply, and near the peak associated with the delta spike in the full band reflectivity there are adjacent peaks of opposite sign and slightly smaller amplitude. In the absence of noise and interference with other band-limited spikes, the maximum amplitude peak will correspond to the location and sign of the delta function we hope to recover, and the POCS algorithm will perform as intended. When noise is present, however, it may happen to increase the amplitude of one of these adjacent peaks sufficiently to make one of these peaks the new maximum amplitude. In our extrapolation, we assume that the maximum amplitudes correspond to the ideal function, so this will cause our extrapolation to recover a delta function of the wrong sign at the position of these adjacent peaks. An error of this type is shown in Figure 9. This type of error is more likely to affect small amplitude reflections. The only obvious way to prevent errors of this type is by improving the signal to noise ratio.

TEST ON SYNTHETIC MODELING DATA

This method of frequency extrapolation was also applied to real data. Measurements from a physical modeling data set were used. The model setup is shown in Figure 10. The data were gathered using a source receiver pair at the same location, with one measurement at each of the 400 CMP points. Deconvolution was applied to the raw data to result in

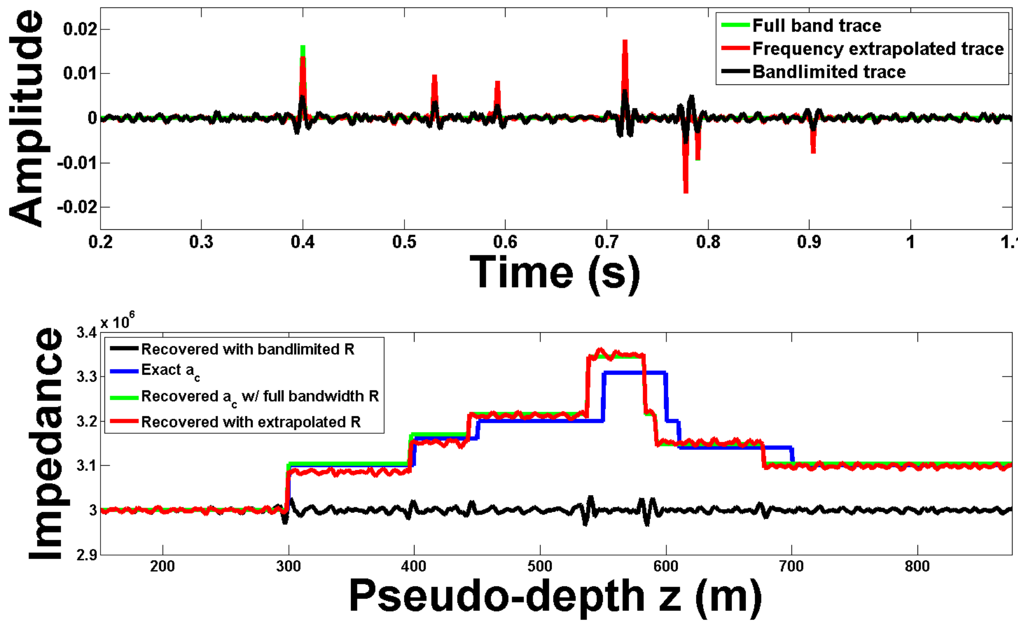

FIG. 8. 42-135 Hz band-limited data with noise. Top: Full band, band-limited and frequency extrapolated reflectivity. Bottom: Exact impedance from a homogeneous background, the result of an impedance inversion using the band-limited reflectivity, the result using the band-limited reflectivity, and the result using full band reflectivity.

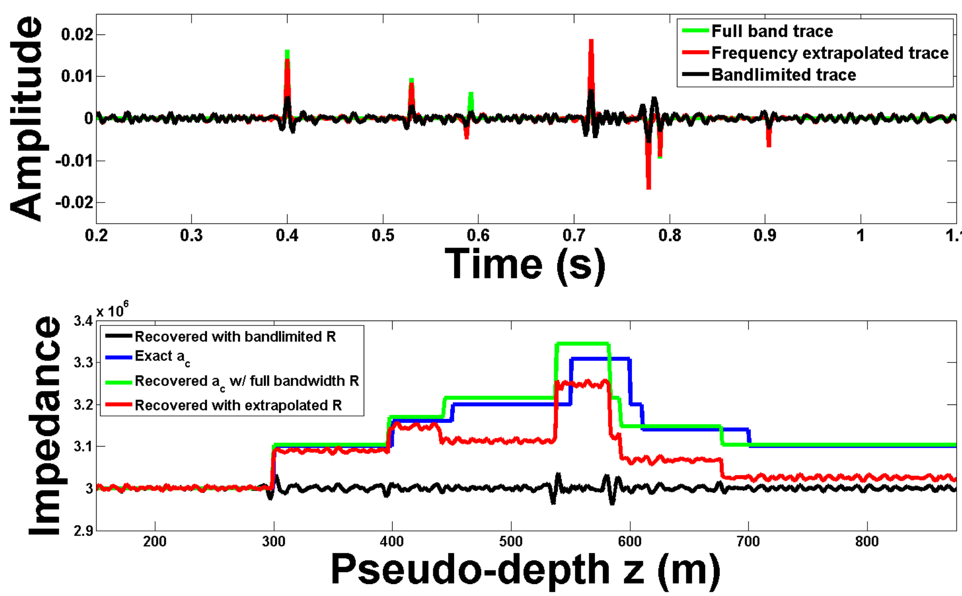

FIG. 9. 42-135 Hz band-limited data with noise. This Figure was created with identical conditions to Figure 8, but with different random noise. Top: Full band, band-limited and frequency extrapolated reflectivity. Bottom: Exact impedance from a homogeneous background, the result of an impedance inversion using the band-limited reflectivity, the result using the band-limited reflectivity, and the result using full band reflectivity.

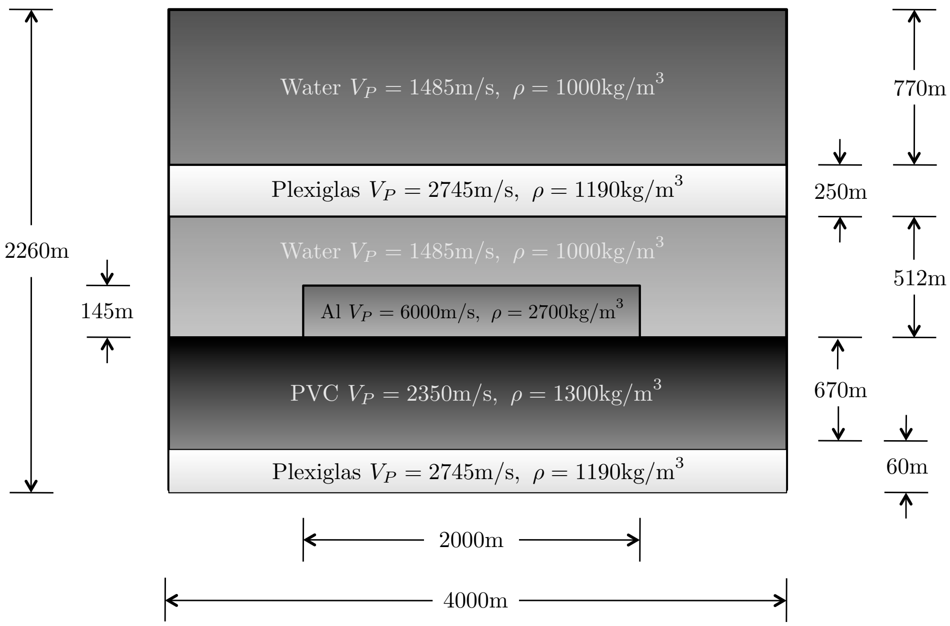

FIG. 10. Schematic of the physical model measured.

the measurements shown in Figure 11. Data gathered after 2.1s were removed due to the presence of free surface multiples. These data are band-limited, and so the result of an impedance inversion is trivial when immediately applied (Figure 12).

The frequency extrapolation method described above was then applied to these data on a trace by trace basis, and another impedance inversion performed. The result, shown in Figure 13 shows significant improvement over our starting result. Comparison with the inversion that would have been obtained using the exact reflectivities from the model (Figure 14), however, shows that there are still problems that remain. In the inversion from the model, we can see that we expect two very closely spaced reflections corresponding to the upper and lower boundaries of the aluminum layer. In our inversion, we see only one change in the perturbation in this area, and rather than a small region with a very large perturbation, we see a sustained increase in the perturbation measured at depths below this interface. Given the very small time separation of these two reflections, this is almost certainly a problem caused by interference between the two band-limited reflections. As such, it is likely that additional high frequency data would help to solve this problem.

Other than the interference problem at the aluminum layer, it is clear that there is one other obvious problem in the inversion in Figure 13. In the reflections at a pseudodepth of 1550 m, we see that the sign of the impedance change flips at certain positions. By looking at the traces corresponding to these locations, we can identify the problem. From the traces

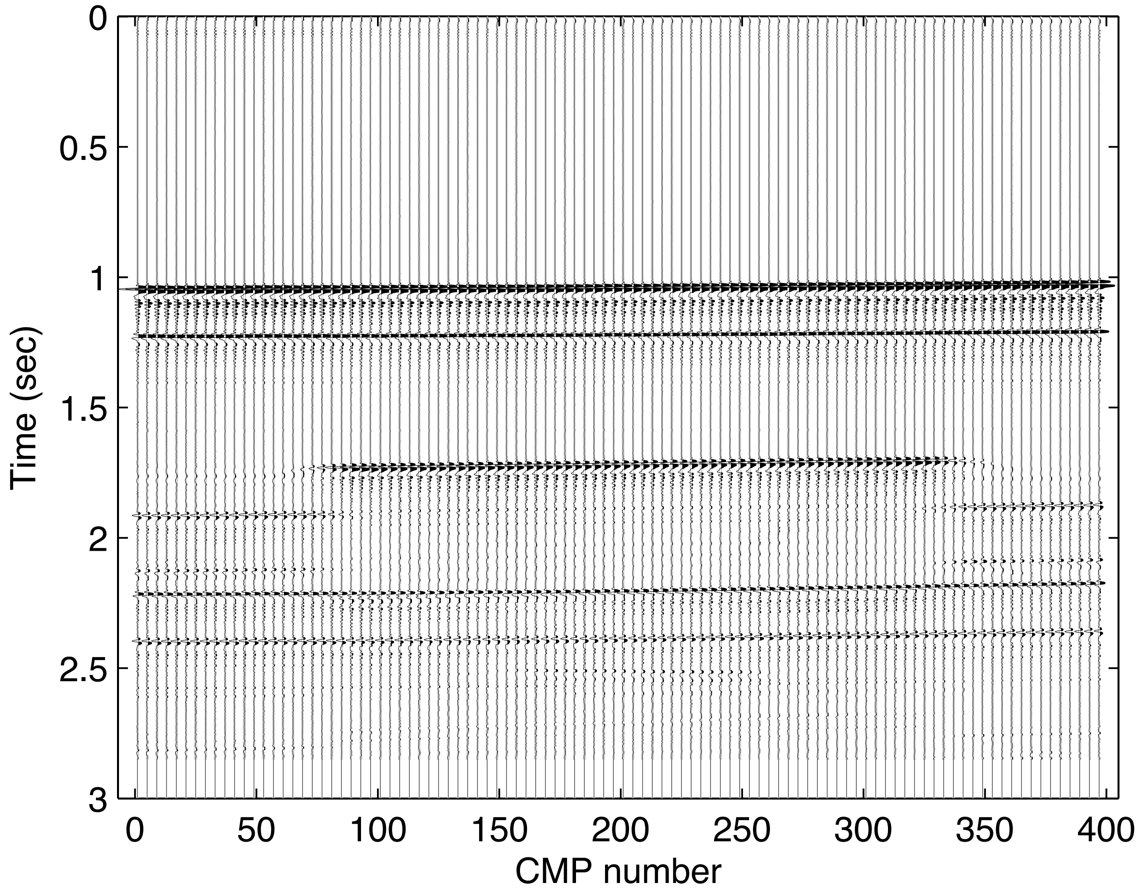

FIG. 11. Measured data from the physical model.

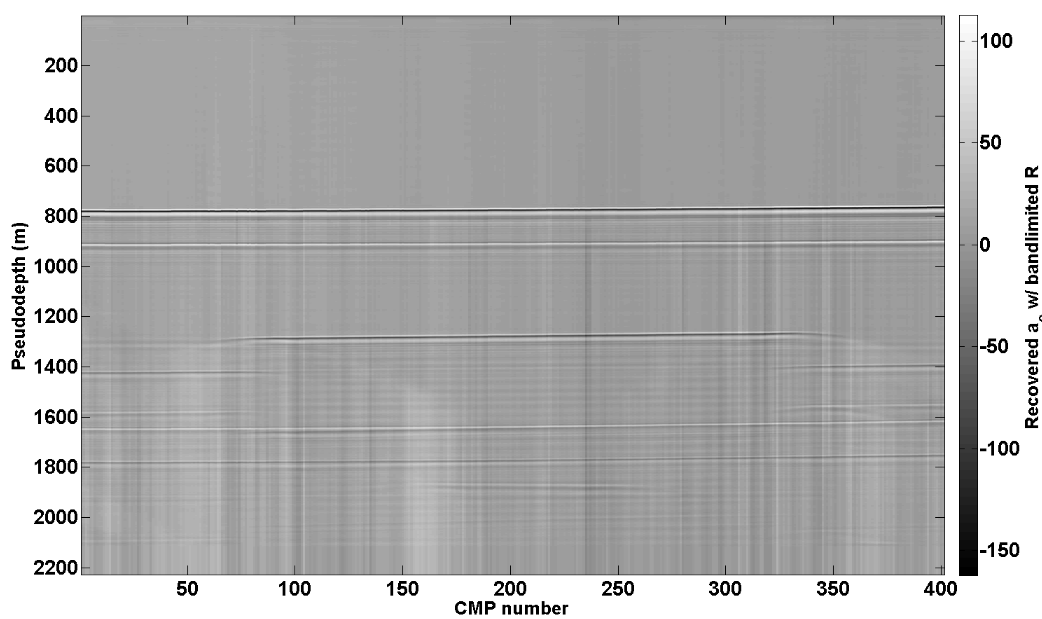

FIG. 12. Impedance inversion of the physical modeling data.

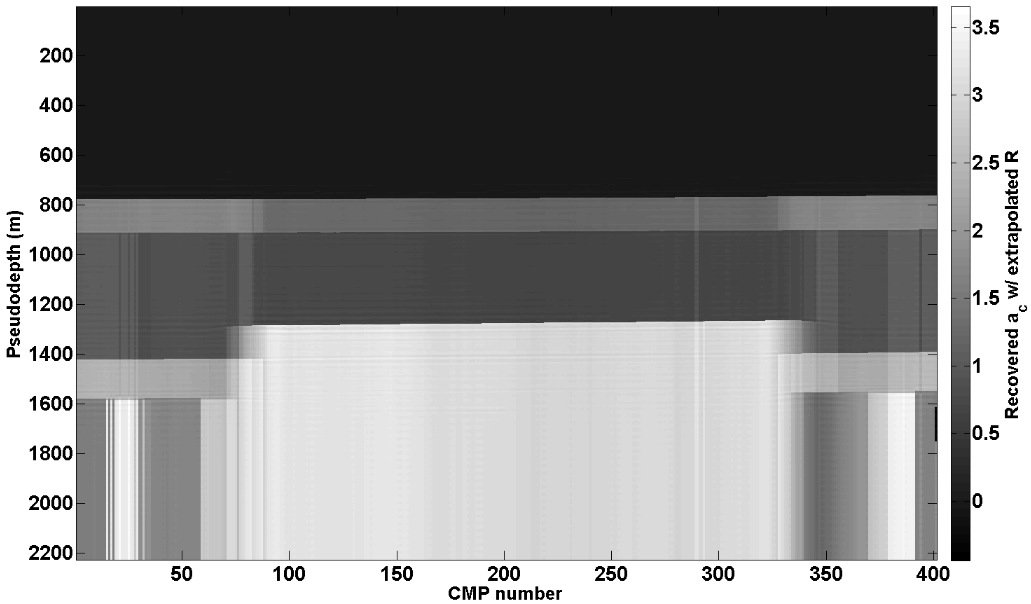

FIG. 13. Impedance inversion of the physical modeling data after frequency extrapolation.

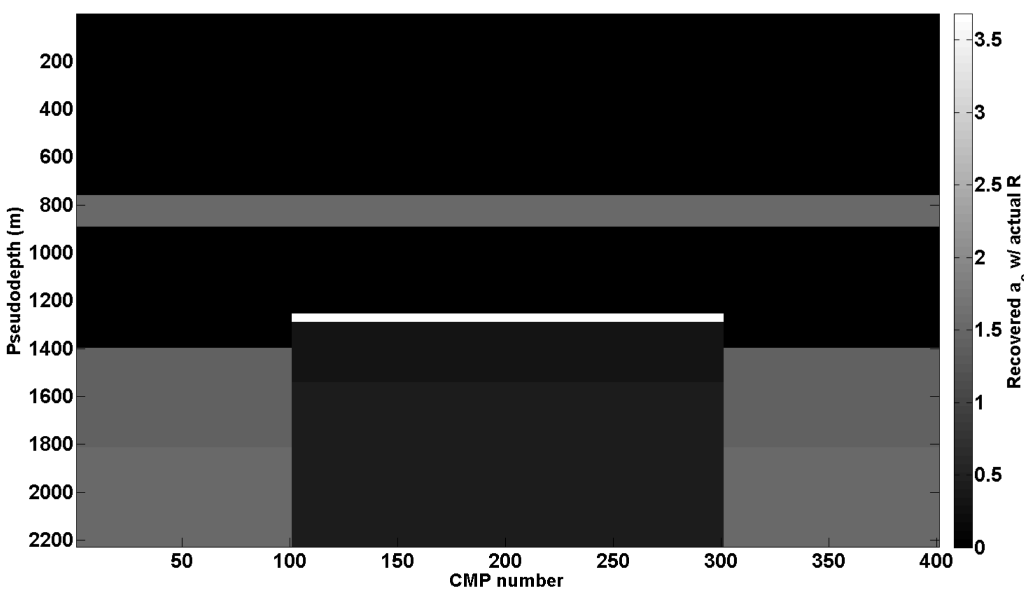

FIG. 14. Impedance inversion of the exact reflectivities as determined from the model setup

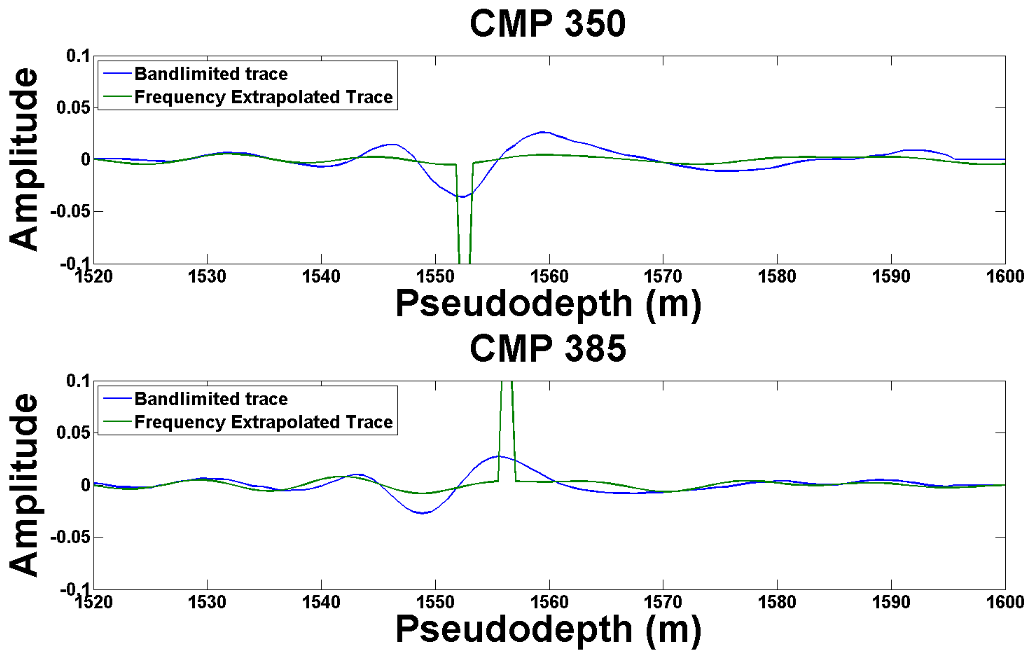

FIG. 15. Top: Band-limited trace and extrapolated trace for CMP 350. Bottom: Band-limited trace and extrapolated trace for CMP 385.

in Figure 15, we can see that the band-limited signal at these points is not zero phase. There are two peaks of opposite sign and nearly the same amplitude very close together at these points. As seen in the 1-D synthetic example, this could occur as a result of noise in the signal, but if this were the case, the changes in polarity we see should not have any consistent pattern from trace to trace, whereas we can clearly see that these changes occur in distinct regions. It is more likely that our deconvolution is consistently failing to output a zero phase signal at these reflections. As a result, the position and polarity of the delta function we retrieve for the reflectivity at this location varies depending on the actual phase of our deconvolved data, which may cause one peak or the other to be the greatest amplitude.

DISCUSSION

While in the simplest synthetic cases, relatively low frequency, limited band data performs equivalently to a high frequency, broader band signal for the purposes of frequency extrapolation, the addition of random noise seems to favor the use of high frequency, broad band data. This is largely counterintuitive, as in inversion it is the lowest frequencies that are most important to recovering usable information. High frequencies help to remove the interference between closely spaced reflections that can cause problems in frequency extrapolation. Additionally, low frequency measurement necessitates that low frequency noise remains after the extrapolation, which can have a large impact on our inversion. If high frequency measurements are instead used, the remaining noise is high frequency, and has little impact on the final output of our inversion. While there are significant problems in the application to the physical modeling data, the most prominent of these, the single recorded aluminum interface, would be solved with sufficient addition of high frequency data. Overall, it seems that frequency extrapolation offers potential for development of a method for inversion which focuses on the measurement of high frequency seismic data, rather than the introduction of low frequency data from non-seismic sources.

CONCLUSIONS

In inversion, low frequency data are crucial to recovering the long scale changes in the properties we are inverting for. Because seismic acquisition is limited in the frequencies it can measure, these low frequencies are often drawn from other sources. Frequency extrapolation offers a way of using the measured frequencies to approximate the frequencies not measured. Although a small band of low frequencies and a large band of higher frequencies can both be used to recover the very low frequencies, practical considerations such as noise and resolution seem to favor using a larger band of high frequencies. Frequency extrapolation applied to physical modeling data met with moderate success, but certainly offered improvement over inversion of band-limited data.

ACKNOWLEDGEMENTS

The authors thank the sponsors of CREWES for continued support. This work was funded by CREWES industrial sponsors and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 461179-13. Scott Keating was also supported by the Queen Elizabeth II scholarship. We also thank Mostafa Naghizadeh for his initial help with the POCS algorithm.

REFERENCES

Abma, R., and Kabir, N., 2006, Minimum weighted norm interpolation of seismic records: Geophysics, **69**, No. 6, E91–E97.

Innanen, K., 2015, Goph 673 course notes.