Multiple attenuation via predictive deconvolution in the radial domain

Marco A. Perez and David C. Henley

ABSTRACT

Predictive deconvolution has been predominantly used as a method for attenuating multiples. The inconsistency of vertical spacing and amplitude of a primary and its corresponding multiples hinders the feasibility of applying this method to shot gathers. Algorithms used to overcome this problem may select a new operator and gap length for each trace with increasing offset. An alternative, less tedious method is presented which maps the shot gather into a more suitable domain. This approach utilizes the invariability between spacing of multiples and primary reflections for a larger range of offsets in the radial domain. The results indicate better attenuation of multiples for shot gathers in the radial domain and less sensitivity of the choice of operator length and stability factor in predictive deconvolution.

INTRODUCTION

This project was initiated to improve a method of deconvolution for removing multiples from a shot gather in an efficient way. In 1980, Taner introduced multiple attenuation by employing the radial transform. Recent work by Lamont et al. (1999), improved the transform to accommodate the removal of multiples from a dipping sea floor. In this paper, flat sea floor multiples of synthetic and real shot gathers were considered. Predictive deconvolution was applied to the shot gathers in both the x-t and radial domains. The findings are compared, and the improved quality obtained from applying the radial transform surmised. Parameters including operator length, stability factor and choice of origin for the radial domain transform are considered. The result of the radial domain multiple attenuation technique is then compared to predictive deconvolution on a synthetic NMO corrected shot gather.

THEORETICAL ASPECTS

Predictive Deconvolution

The purpose of predictive deconvolution is to estimate, or predict, the periodic portion of the input signal $x(t)$ at some future time, namely $x(t+\alpha)$, where α is the number of time samples ahead to be predicted. To do this, one convolves the input data, $x(t)$, with a filter, $f(t)$, to obtain the output, $g(t)$. Mathematically this is expressed as

$$\sum x(t)f(\tau - t) = g(\tau) = x'(t + \alpha)$$

$$(1)$$

where it is hoped that $g(t)$ is a time-advanced version of $x(t)$. The filter is the unknown operator which will output an estimate of $g(t)$. Since the estimate may not

be correct, an error series is defined as the difference between the true value at $x(t+\alpha)$ and the estimated value of $x'(t+\alpha)$. Hence for each predicted value, there is an associated error, expressed mathematically as

$$\varepsilon(t + \alpha) = x(t + \alpha) - x'(t + \alpha)$$

$$(2)$$

This error represents the unpredictable part of the signal $x(t)$.

To compute the time-advanced version of the input signal, the correct filter is required. It is assumed that the input signal has a definite periodicity. Therefore the output is approximately a time-advanced version of the input (Peacock and Treitel, 1969). Using least-squares filter theory, a set of normal equations are built to compute the filter. These equations have a general form of

$$\begin{bmatrix} a_0 & a_1 & . & a_{n-1} \\ a_1 & a_0 & . & . \\ a_2 & a_1 & . & . \\ . & . & . & a_0 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ . \\ f_{n-1} \end{bmatrix} = \begin{bmatrix} g_0 \\ g_1 \\ . \\ g_{n-1} \end{bmatrix},$$

$$(3)$$

where the n by n matrix is the autocorrelation of the input (Toeplitz matrix), the vector $g(t)$ is the crosscorrelation of the data and the desired output and the vector $f(t)$ is the filter. In order to solve for the filter in equation (3), the autocorrelation of the input, and the positive lag coefficients of the crosscorrelation between the desired output and the input must be known. However as mentioned previously, the output is simply the time-advanced input signal, and thus the crosscorrelation of output and input is simply the autocorrelation of the input. The normal equations can now be expressed as

$$\begin{bmatrix} a_0 & a_1 & . & a_{n-1} \\ a_1 & a_0 & . & a_{n-2} \\ . & . & . & . \\ a_{n-1} & a_{n-2} & . & a_0 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ . \\ f_{n-1} \end{bmatrix} = \begin{bmatrix} a_\alpha \\ a_{\alpha+1} \\ . \\ a_{\alpha+n-1} \end{bmatrix},$$

$$(4)$$

where the only unknown is the filter, $f(t)$. Now by a method of least squares, the filter is estimated, from which predictions can be made. The error series can be computed as a filter, or it can be obtained simply by subtracting the prediction from the input.

This theory is applied in geophysics to remove undesired periodic events like multiples from seismic data or to shorten event wavelets by removing the predictable part. Though one can use prediction filtering in both the time and offset direction within a seismic trace panel, multiples are basically periodic only in time. In marine data, reflection energy is often trapped in the water layer, giving rise to multiple reflections. Thus in a marine trace, one can see the same event occurring at approximately constant time lags repeating down the length of the trace. It is this part of the trace that is predictable. The primary reflection events are assumed to be

random and therefore unpredictable. Since multiples are to be attenuated, it is the error series that is the signal of interest.

There are details to be addressed in designing the correct prediction filter, especially when applied to a seismic signal. The first is the length of the input signal to be placed within the autocorrelation matrix. The seismic signal, a marine trace in this case, will contain a primary, the first occurrence of an actual event. It is copies of this primary which are to be predicted down the length of the trace. The length of the input signal should be long enough to contain all the information of the primary wavelet. This is referred to as the operator length since it is the data contained in this operator that will be predicted. The number of samples between the primary and the first copy determines the number of samples that the filter is to predict. This is referred to as the gap length. This method of determining operator and gap length can be used when the input signal is simple. When the signal is complicated, a larger number of points may be required to capture the information of the repeating energy. Complications arise in real data when the wavelet is not as well defined, as is generally the case with actual data. In the real data case, the wavelet may start at the time at which the primary occurs, yet still contain energy up to the onset of the first multiple. Here the wavelet will cover what would have been a null zone in a synthetic section. Other complications arise when prediction lag is not exact, and when multiple waveform differs from that of the primary.

Another subtlety to consider is the addition of a white noise factor. A complete discussion of this factor is given by Treitel and Lines (1982). Essentially, the method of least squares yields a solution, which is a function that minimizes the error, that is to say, the difference between the output and the expected output. However, the inversion is sensitive to the presence of uncorrelated noise in the input and to the relative size of the elements in the inverted matrix. The white noise factor, or stability factor, is added to the zero lag of the autocorrelation to stabilize the inversion and effectively suppress the noise in the data. Mathematically, the prewhitened autocorrelation matrix is expressed as

$$\underline{\underline{A}} + \lambda \underline{\underline{I}}, \tag{5}$$

where A is the autocorrelation matrix and I is the identity matrix. Now the filter will also attempt to minimize the λI factor. The factor of λ will stabilize the calculation of A^{-1} eliminating the presence of the small eigenvalues of A. It is these small eigenvalues, which could potentially inflate noisy values and cause instabilities. In addition this factor of λ also turns out to be a constraint in the output noise power. By making λ a small positive number, the output can have a decreased amount of noise power. However, it also degrades the resolution of the deconvolution. Thus if λ is too small, the noise could dominate the output signal due to instabilities in the inversion. If the input data set is clean however, the deconvolution will have a high resolution transforming the wavelet into a spike. If λ is a relatively large number, the noise will be attenuated at the expense of a less accurate approximation of the desired output. Therefore, the addition of a white noise factor is a trade off between approximation

accuracy and the ability to attenuate noise and control instabilities. It is desirable to choose as small a value for the stability factor as possible to maximize resolution.

The key to successful predictive deconvolution in shot gathers is that the predictability of trace events remains constant as the offset increases. Since the timing and amplitude relationships between the primary and multiples change with offset, the operators chosen for near zero offset predictive deconvolution become less appropriate with offset and the quality of the deconvolution falls off rapidly. The goal here is to map the shot gather into a more suitable environment in which the spacing between the primary and its corresponding multiples remains constant for a larger portion of the data.

Radial Domain

The radial domain is a simple mapping of the conventional x-t coordinate system into the radial (r-t) coordinate system. The transform follows the simple formulas

$$t' = t$$
(6a)

$$v = x/t$$
(6b)

where *t'* and *v* are the coordinates of the radial domain (Henley,1999). Essentially, data that lies on a specific slope (velocity) is mapped as a single trace in the radial domain as is shown in Figure 1.
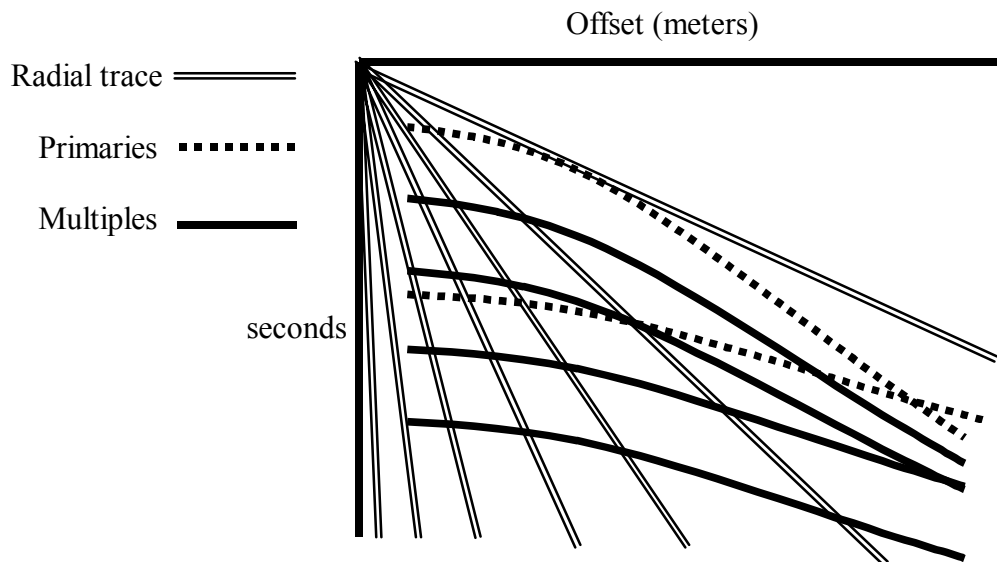

Figure 1. Shot gather with superimposed radial traces

It turns out that the spacing between the primary and its corresponding multiples is constant for larger offsets in the radial domain. This can be seen in Figure 2 below.
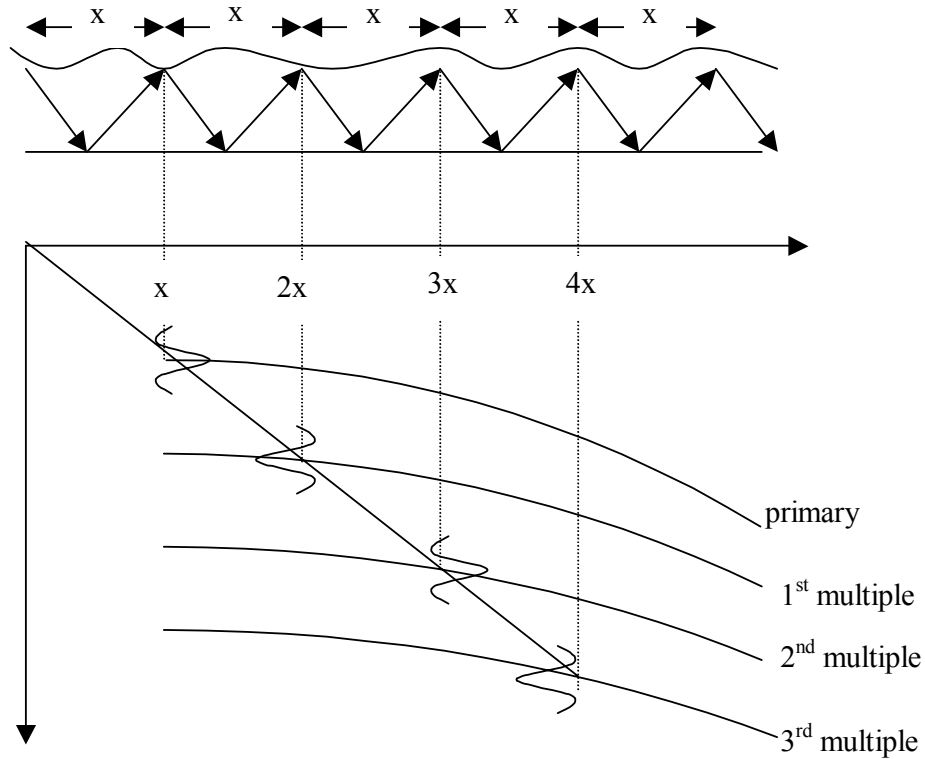
Figure 2. Heuristic diagram displaying constant primary to multiple spacing in the radial domain (Taner, 1980).

Through this diagram it is intuitively seen that a radial trace will contain a primary and its water borne multiples at constant time and distance increments. This corresponds to equal spacing in the radial domain. This is shown in Figure 3 below.
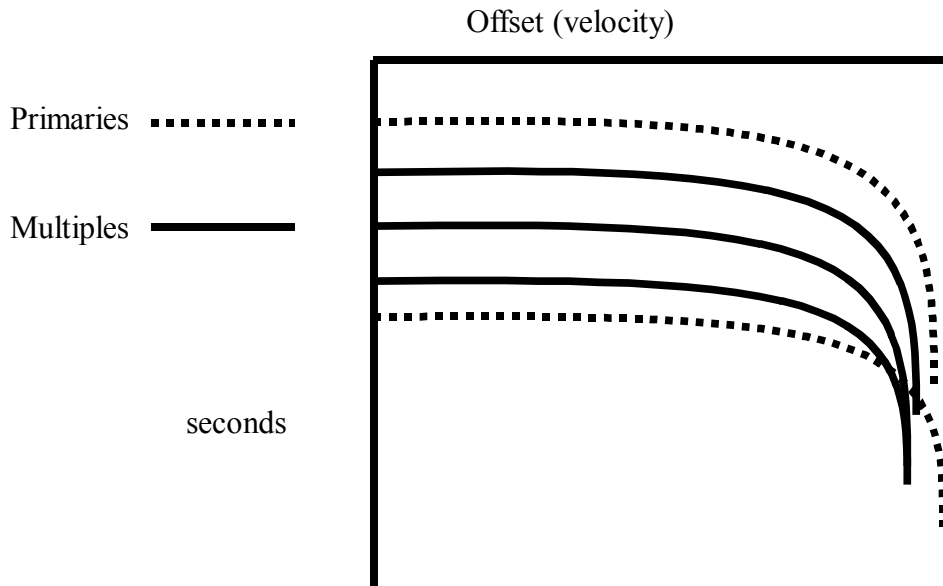

Figure 3. Radial transform of a shot gather

Though the transform is simple, there are key points to consider. One of the most important is the location of the origin of the radial traces. To have an accurate velocity transform, the origin should be at (0,0) in the x-t domain. The radial trace, which passes through the origin and is asymptotic to the water bottom reflection, determines the maximum velocity for the radial mapping. However, the point (0,0) is not always the optimal position for real data. The reason is that there is usually no data recorded for zero or small offsets on marine surveys. Thus for purposes of predictive deconvolution, some low velocity radial traces will not contain a primary event, leaving nothing for the filter to predict. Hence, the origin of the radial transform should be moved along the offset axis to the offset of the first trace with data. The origin should however still lie on the asymptote of the water bottom reflection hyperbola. The origin must be moved down the time axis, as well, until it intersects the said asymptote. This tactic will ensure that the predictive deconvolution will have all the information it needs. The price paid is that the spacing between primary and multiples is no longer quite as perfect as for the (0,0) origin. The ability to move the r-t origin is an important extension of Taner's application.

Another practical consideration for the radial transform is that of aliasing. Because of the geometric distortion of the radial mapping, care must be taken to ensure sufficiently dense sampling in the radial domain to prevent aliasing from occurring. Techniques for doing this however are beyond the scope of the work presented here.

ANALYSIS AND DISCUSSION OF EXAMPLES

Synthetic Data

The analysis begins with a noise free synthetic shot gather provided by Brian Hoffe of CREWES. The data is seen below in Figure 4.
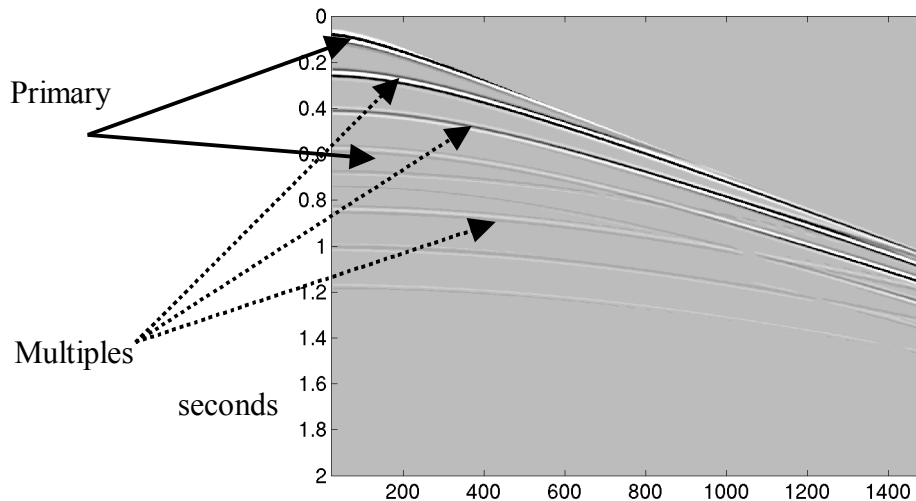

Figure 4. Input synthetic shot with multiples in the x-t domain. Horizontal axis is in meters.

The radial transform of this data is shown in Figure 5. Notice how the spacing between the primary and the multiples are constant for a larger range of offsets (velocity).
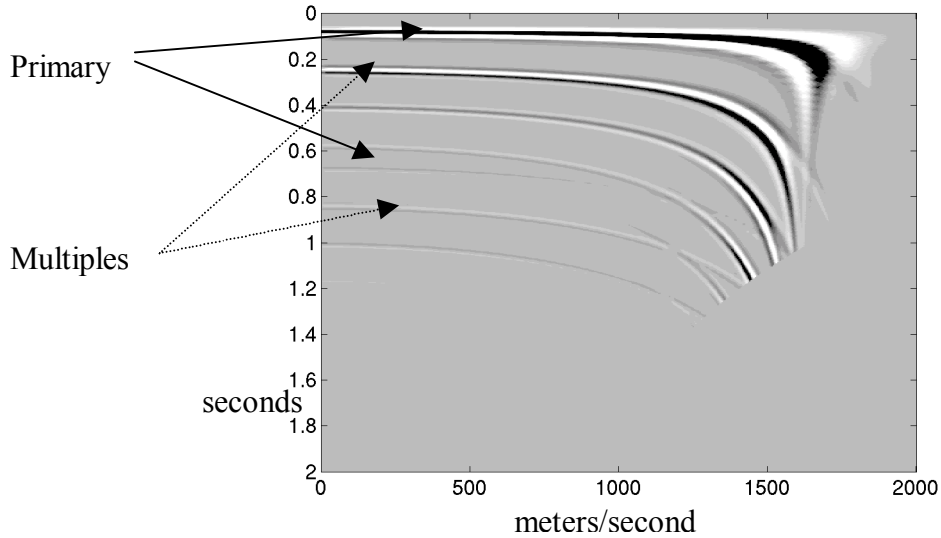

Figure 5. Input synthetic shot with multiples in the r-t domain

Notice that in Figures 4 and 5 there is a primary reflection at about 0.68 seconds. Its first water borne multiple occurs at approximately 0.82 seconds and with greater amplitude, a consequence of the modeling conditions. The simple predictive deconvolution assumptions are only appropriate for simple multiples like those generated by the sea floor. Thus for this report, this peg-leg multiple will be considered as a reflection event with the same properties as its primary. In the predictive deconvolution results we expect to see both events at 0.68 and 0.82 seconds.

Predictive Deconvolution in the x-t and r-t domain

The predictive deconvolution will use the same parameters in both domains. In fact the deconvolution for the first trace in each domain should be identical (zero velocity corresponding to a vertical trace in the x-t domain from the origin). The parameters used are displayed in Table 1 below.

Table 1. Parameters used in predictive deconvolution

Operator Length	Gap Length	Stability Factor
40 0.08 seconds	50 0.1 seconds	10^{-2}

It is expected that in the x-t domain, as the offset increases, the quality of the deconvolution decreases. In this case the quality decreases very rapidly as the majority of the traces still contain multiples. Figure 6 shows the result of the deconvolution in the x-t domain.
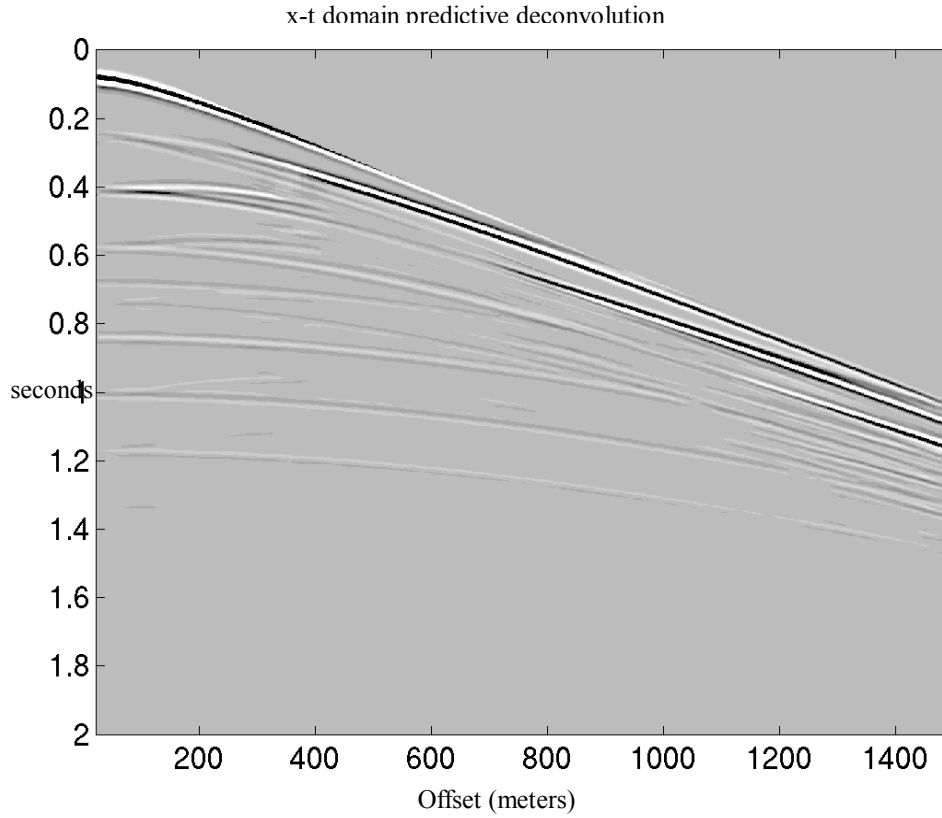
Figure 6. Predictive deconvolution result in the x-t domain.

In the radial domain we can see that the deconvolution works for a greater number of radial offsets. This output is transformed back into the x-t domain to accurately compare the results. This is shown in the Figure 7.
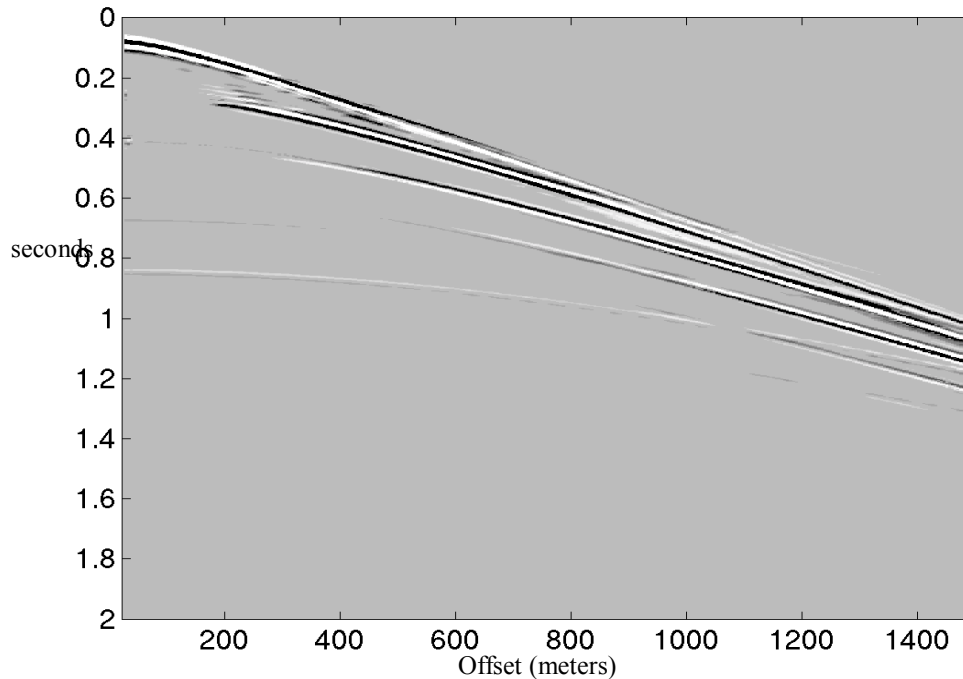
r-t domain predictive deconvolution

Figure 7. Predictive deconvolution in the r-t domain (vertical axis is in seconds)

Notice that the radial domain deconvolution does a significantly better job at farther offset, particularly on higher order multiples. Also, the two primaries are still seen on the figure at approximately 0.68 and 0.82 seconds. These results are very encouraging.

PRACTICAL CONSIDERATIONS ON SYNTHETIC DATA

Stability factor

As mentioned before, the white noise factor affects the quality of predictive deconvolution yet it is also necessary to prevent instabilities. In this section the sensitivity of the stability factor is investigated on noise free synthetic data. For this data, the best results shown previously used a safe stability factor of 10^{-2}. This factor will usually give adequate results for all data, yet not the best on any data. Figure 8 shows the effects that a smaller stability factor $(10^{-6)}$ has on the synthetic data.
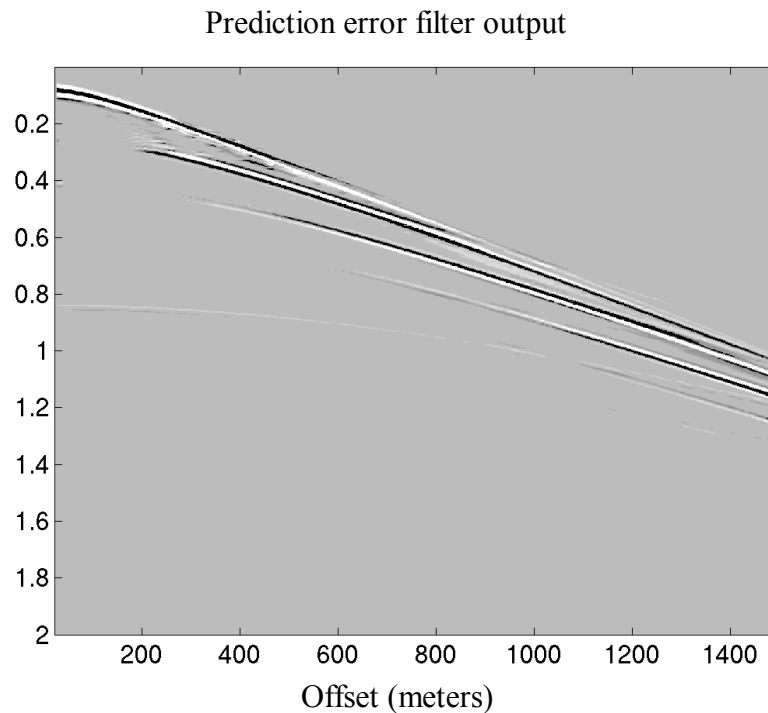
Prediction error filter output

Figure 8. r-t domain predictive deconvolution using white noise factor of 10^{-6}

The differences between the two white noise factors are not that noticeable. This is expected since the data being used is synthetic. However, it does seem that the resolution of the events is greater with a smaller stability factor, as was expected. The necessity for a larger white noise factor is usually only needed for data contaminated with noise.

Change of radial origin

The choice of origin in the radial domain affects the spacing between the primary and their corresponding multiples. This in turn affects the quality of the predictive deconvolution. The following figures are shown to illustrate this point. Figure 9, below, shows the input synthetic data with origins of (0,0) and (0,200) as opposed to the origin used to obtain the best results of radial domain predictive deconvolution seen previously in Figure 4.

Radial transform using origin of (0,0)   Radial transform using origin of (0,200)
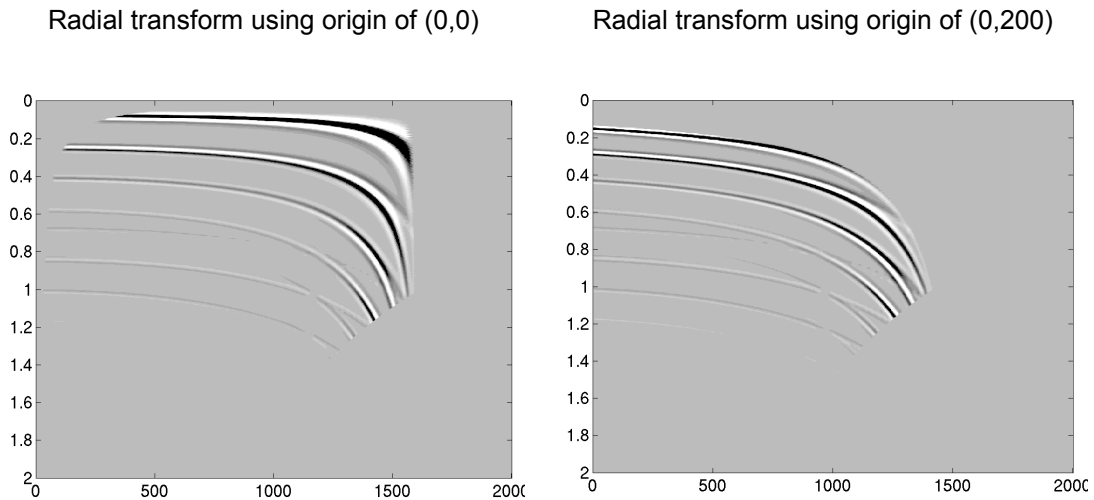
Figure 9. Radial transform using different origins. Axes are seconds in the vertical axis and offset (meters/second) in the horizontal axis.

From the plot on the left, it is immediately apparent that at small radial offsets, the primary does not exist and thus the filter will have nothing to predict. The plot on the right contains all the primaries, but the spacing between the primaries and the multiples is not constant. Figure 10a and 10b show the prediction error filter applied to the input using both origins seen in Figure 9.

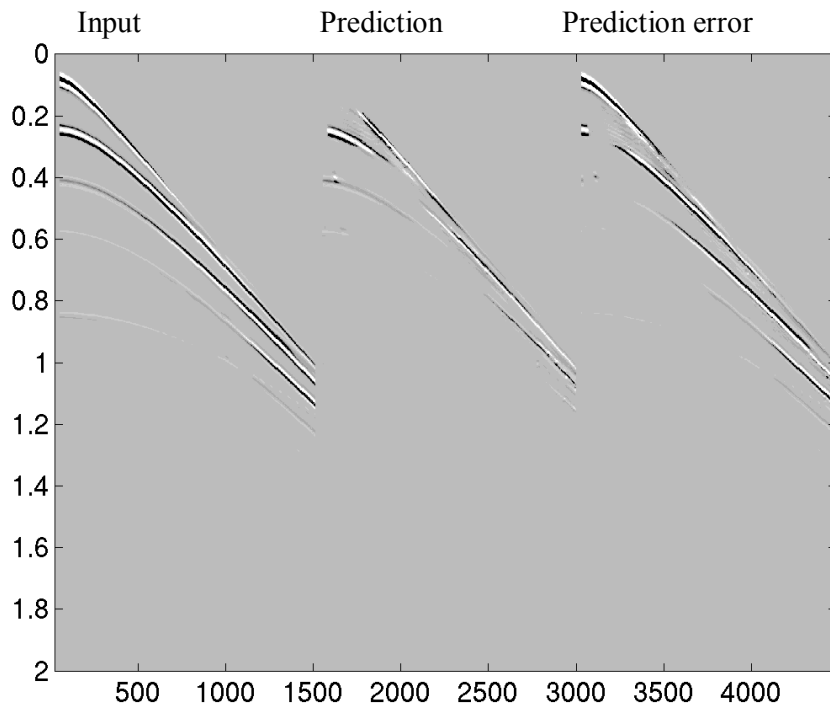Input   Prediction   Prediction error

Figure 10a. Predictive error filter applied to radial transforms with inappropriate origin of (0,0). Axes are in seconds along the vertical axis and meters along the horizontal axis.
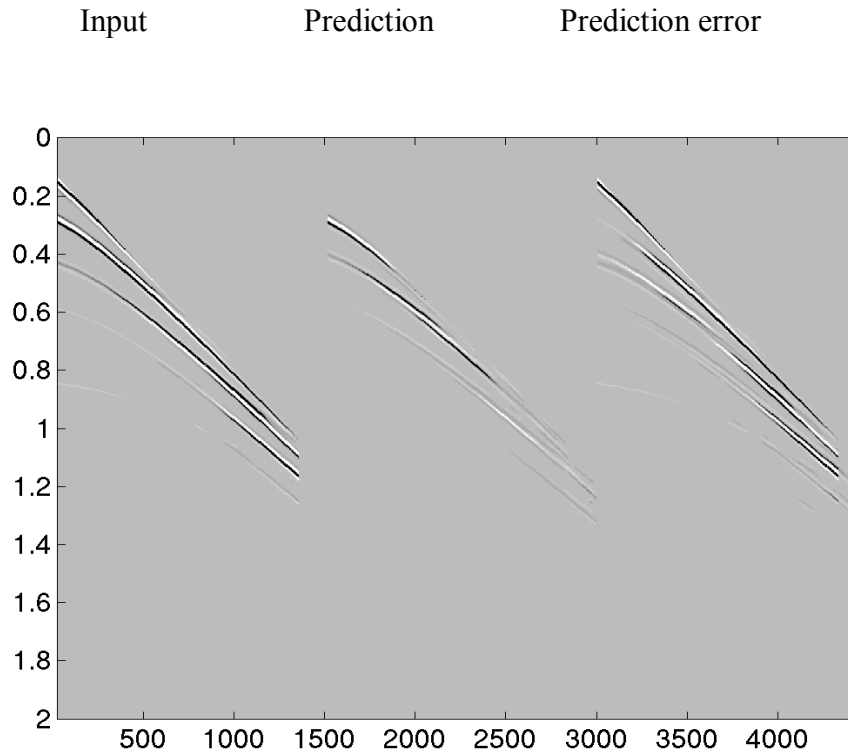
Input   Prediction   Prediction error

Figure 10b. Predictive error filter applied to radial transforms with inappropriate origins of (0,200). Axes are seconds along the vertical axis and meters along the horizontal axis.

The results given in the above figure clearly show that the quality of the filtering is sub par. The origin of (0,0), as expected, does poorly at near offsets though it does quite well at farther offsets. The choice of origin of (0,200) has incorrect spacing as the offset increases and this is reflected in the output section.

COMPARISON TO NMO OF X-T DATA

Another way to improve the predictive deconvolution of a shot gather is to apply normal moveout (NMO) to the traces. This correction essentially flattens hyperbolic events described by a specific velocity. Unfortunately, hyperbolas that are not characterized by this velocity will not be flattened. Also, as the offset increases, with the application of NMO, an event will stretch, distorting its spectrum. Since the wavelet is no longer as compact at larger offsets as it is at small offsets, the operator length will not contain all of the wavelet that needs to be predicted at some time in the future. In this case, the prediction error will reflect the fact that although the events are equally spaced and the gap length can remain constant, it is the operator length that must change. Also to compensate for the modification of the wavelet, the stability factor may change due to the lower frequency content at larger offsets. Figure 11 shows the effects of the application of NMO.
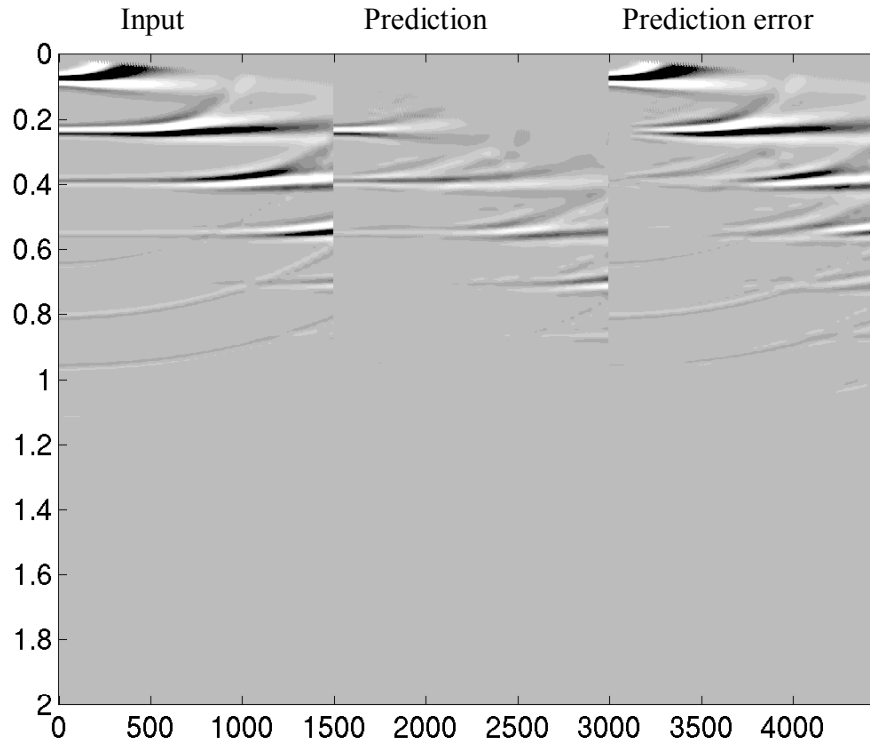
Figure 11. Application of NMO to x-t data with no stretch mute applied. Axes are time in seconds along the vertical axis and offset in meters on the horizontal axis.

In Figure 11 it is clearly seen how applying NMO would adversely affect predictive deconvolution with an increase in offset. Still, the results should be better than simple x-t predictive deconvolution without NMO applied. Figure 12 shows the results of the application of predictive deconvolution on the NMO'd data and consequently having the NMO removed so that a comparison can be made.

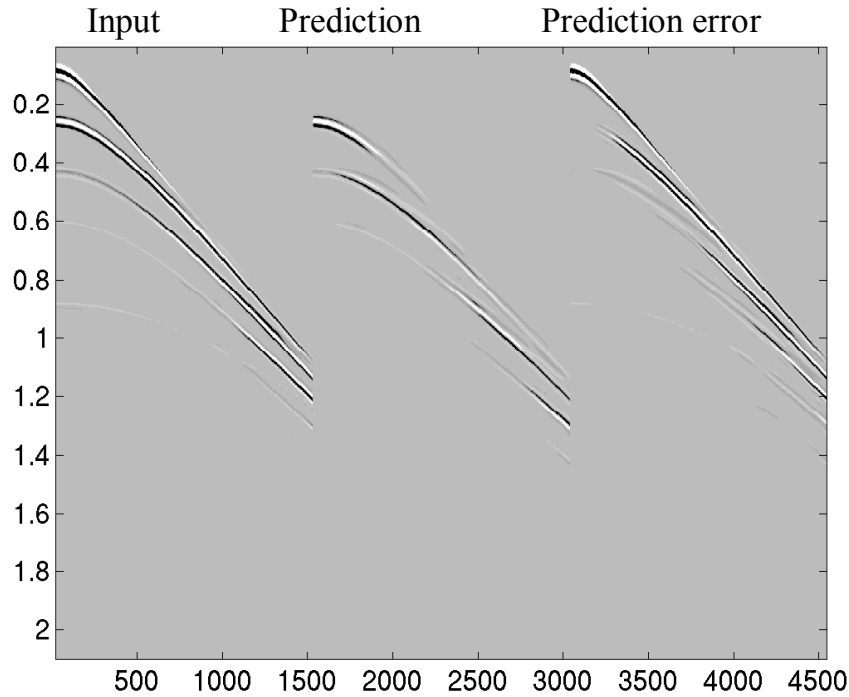Input   Prediction   Prediction error

Figure 12. Predictive deconvolution result to NMO'd data with no stretch mute applied. Axes are time along the vertical axis and meters along the horizontal axis.

REAL DATA

The testing is now extended to real data provided by Brian Hoffe of CREWES. The shot gather that will be tested is seen in Figure 13 in both the x-t and r-t domain.
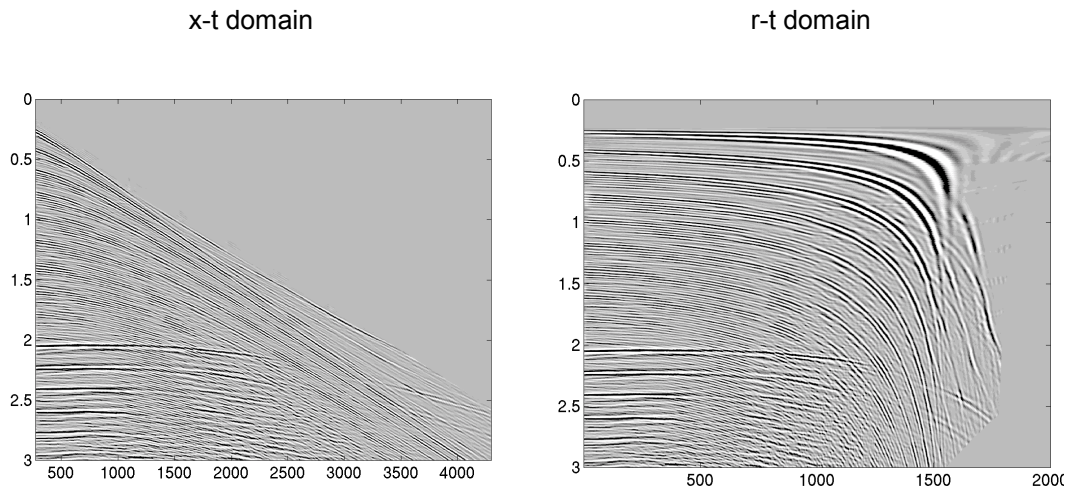
x-t domain   r-t domain

Figure 13. Real data represented in both the x-t and r-t domains. Figure on the left has a vertical axis of time in seconds and a horizontal axis of offset in meters. Figure on the right has a vertical axis of time in seconds and a horizontal axis of offset in meters/second.

Here we can see that the spacing between primary and multiples is constant at longer offsets in the radial domain.

Predictive deconvolution in the x-t and r-t domain

The parameters used in performing the predictive deconvolution are displayed in Table 2.

Table 2. Parameters used on real data to perform predictive deconvolution.

Operator length	Gap length	Stability factor
100 0.2 seconds	1 0.002 seconds	10^{-2}

Looking at the results of the predictive deconvolution in Figures 14a and 14b, it is seen that the filtering works better in the radial domain. From the two figures, it is noticeable that the multiples are more highly attenuated in the r-t domain. An additional benefit is the attenuation of long-offset linear arrivals.
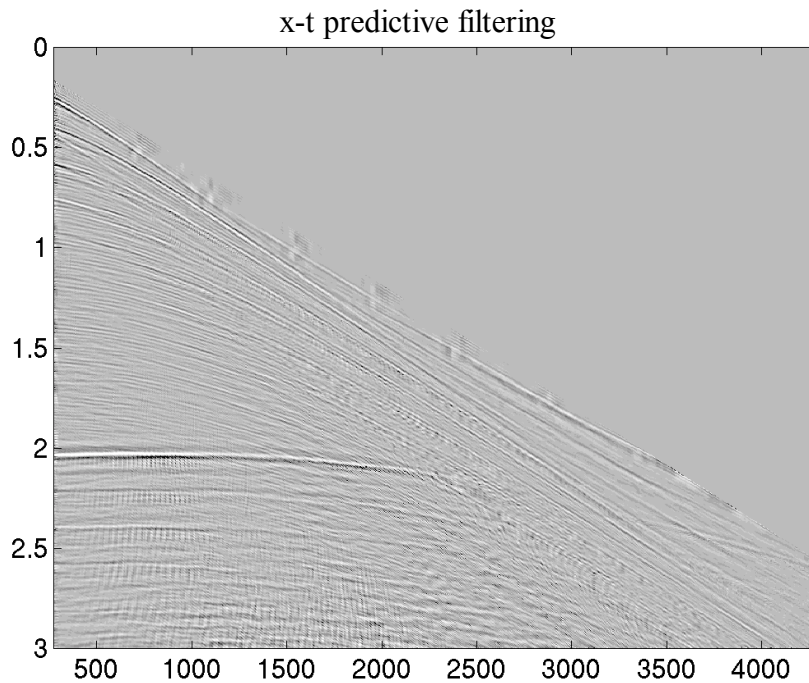
Figure 14a. Result of predictive deconvolution in the x-t domain. Figure has axes of time in seconds along the vertical axis and meters along the horizontal axis.
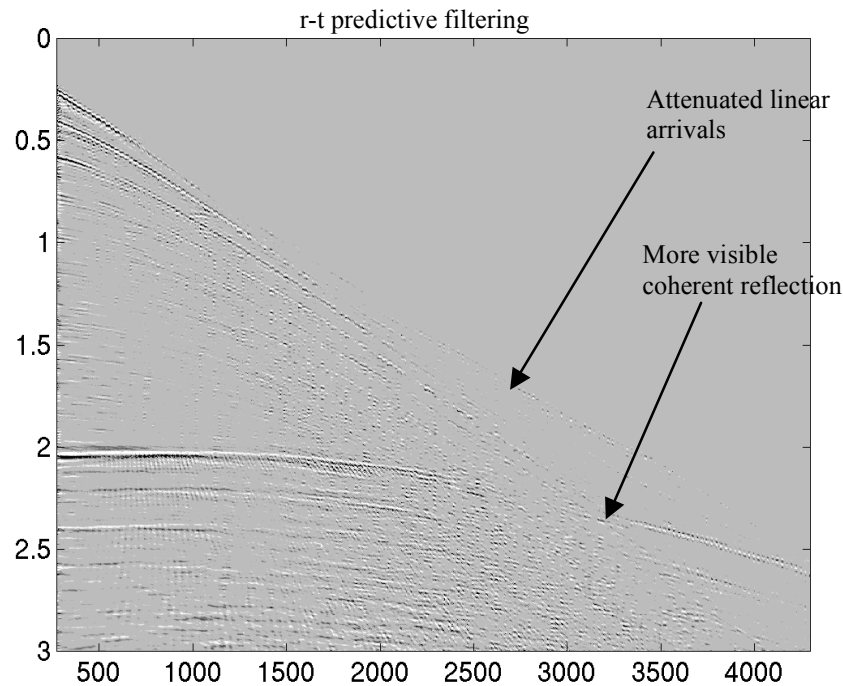
Figure 14b. Results of predictive deconvolution in the r-t domain. Figure has axes of time in seconds along the vertical axis and a horizontal axis of offset in meters.

PRACTICAL CONSIDERATIONS ON REAL DATA

Operator length

The choice of operator length is expected to be more important for real data than for synthetics. In this case, there seems to be a large amount of reverberation energy immediately after the first primary occurs which is at approximately 0.25 seconds. What is considered to be the first multiple occurs at around 0.45 seconds. Therefore when attempting to remove the reverberating energy the operator length will include both length operator and the gap as was suggested earlier. The results of using this operator were already shown in Figures 14a and 14b. Suppose now that the operator length did not take into account this reverberating energy between the primary and what is considered to be the first multiple. Using an operator length of 20 and a gap length of 80, the following results were achieved using predictive deconvolution in both the x-t and r-t domain.
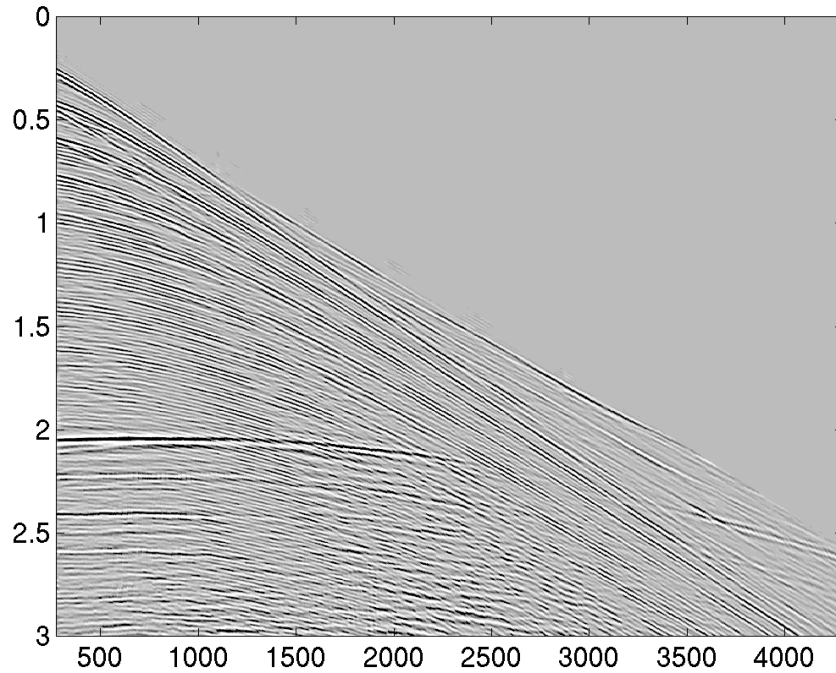
Figure 15a. Predictive deconvolution results when using an operator length of 20 and a gap of 80 in the x-t domain. Axes are time in seconds along the vertical axis and offset in meters in the horizontal axis.
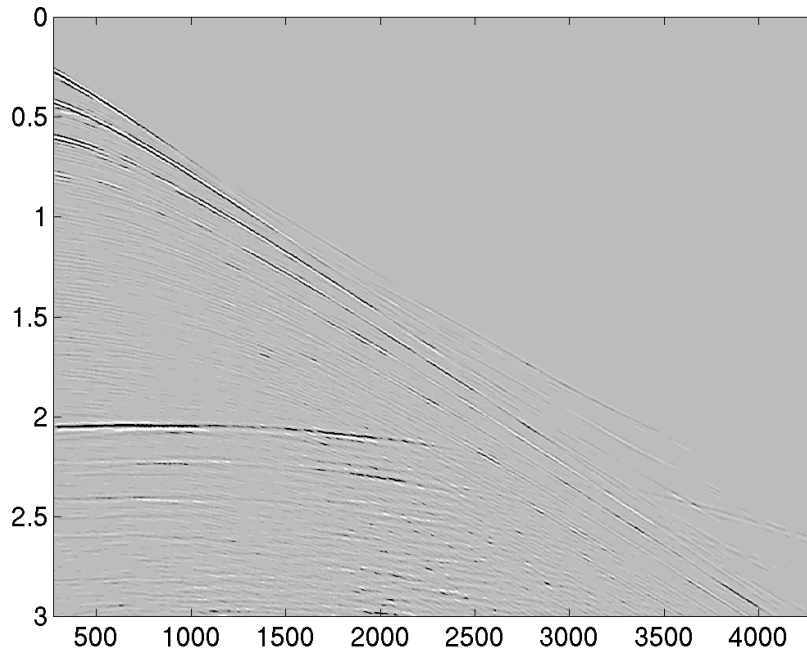

Figure 15b. Predictive deconvolution results when using an operator length of 20 and a gap of 80 in the r-t domain. Axes are time in seconds in the vertical axis and offset in meters in the horizontal axis.

As expected the filtering in the x-t domain is poor and though it appears the multiples are removed for short offsets, the reverberating energy remains. Interestingly, in the

filtering in the radial domain the reverberating energy is attenuated much better. This suggests that the choice of operator length is not as sensitive in the radial domain. This is not to say that the result is better than it was when using an operator length of 100, but it is an encouraging and unexpected result that the predictive deconvolution in the radial domain is not as sensitive in the r-t domain as it is in the x-t domain. A possible reason is that everything within the radial domain becomes much more predictable, not just the multiples. Thus events, which are relatively random on the x-t, become predictable in the r-t domain leading to better deconvolution with the ability to use shorter operators. Since the predictive deconvolution will not have to deal with as much 'noise', the operator length can simply be the number of samples describing the primary wavelet.

Stability factor

Now the results are compared to the predictive deconvolution results of 10^{-6}. It is expected that the noise contaminates the signal in such a way that the prediction error filter will perform poorly. The results are shown in Figures 16a and 16b.
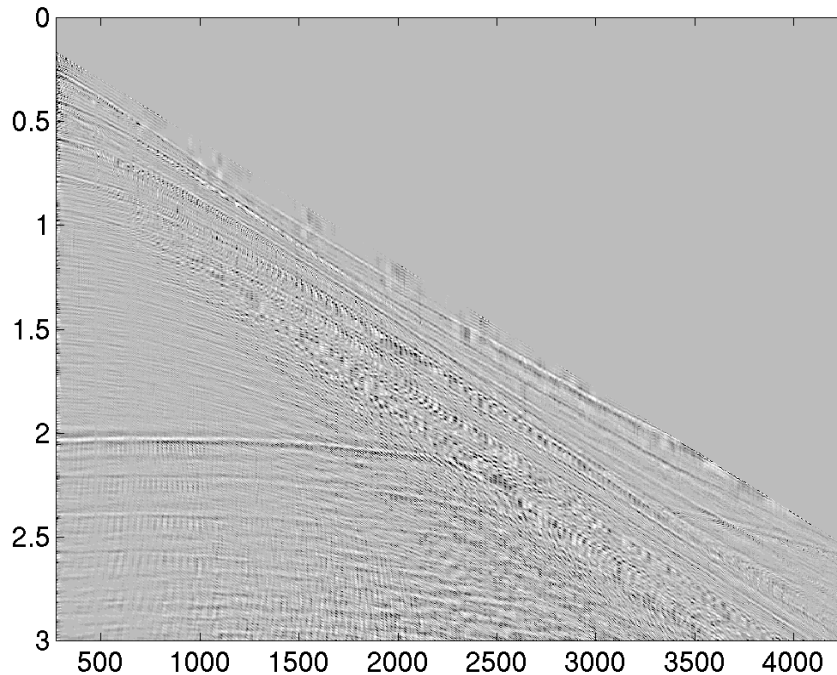

Figure 16a. Predictive deconvolution results using a stability factor of 10^{-6} in the x-t domain. Axes are time in seconds in the vertical axis and offset in meters along the horizontal axis.
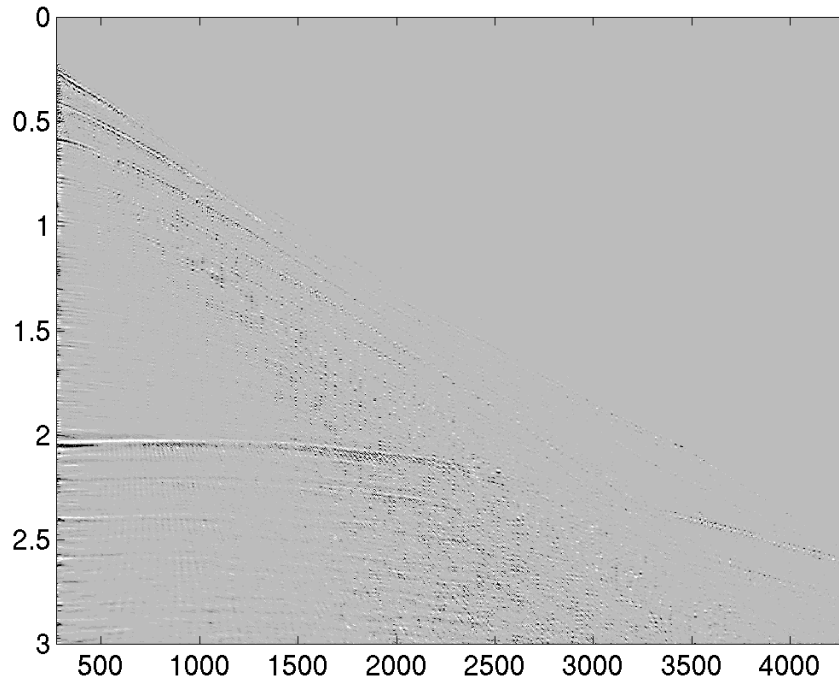
Figure 16b. Predictive deconvolution results using a stability factor of 10^{-6} in the r-t domain. Axes are time in seconds in the vertical axis and offset in meters along the horizontal axis.

As expected the x-t domain result shows a section mostly dominated by noise. However, in the r-t domain, the section looks comparable to using a stability factor of 10^{-2} as seen in Figure 14b. Again this is an unexpected result that may prove to be very useful. Since the stability factor is smaller, the recovered bandwidth of each event is much greater than for a larger factor. This is not ideal in the x-t domain because much of the extra bandwidth is dominated by noise. Hence, it can be concluded that the radial domain also allows for a smaller stability factor thus resulting in a more interpretable section. A possible reason for this follows the same line of reasoning as to why a smaller operator length can be used in the radial domain. The manner in which the radial domain organizes the data concentrates much of the coherent noise on a few specific radial traces allowing the predictive deconvolution to work mostly with cleaner traces. It is convenient to work in such a domain since the deconvolution will resolve the events better with a smaller stability factor. Though initially the motivation for using the radial domain was to keep the primary and multiple spacing constant, initial testing also shows that in the radial domain, predictive deconvolution is less sensitive to the choice of operator length and stability factor.

CONCLUSION

Predictive deconvolution can be a powerful tool, if used properly, in attenuating multiples. In a shot gather in the x-t domain, the increment spacing between the primary and its corresponding multiples is not constant. This means that more than one operator length must be used when applying predictive deconvolution and that predictive deconvolution itself fails as offset increases. Applying NMO to the section

does alleviate the problem of primary-multiple spacing yet introduces another problem. With the application of NMO, a stretch will occur, increasingly distorting the spectrum of the event with offset thus destroying the effectiveness of predictive deconvolution. Transforming the data into the radial domain reduces the extent of the problems encountered in the x-t domain. In the r-t domain, the primary-multiple timing is constant for a much larger portion of the data. The predictive deconvolution works much better and attenuates multiples for a larger range of effective offsets. In the radial domain, the wavelet describing the primary event does change but only significantly at large radial offsets. Were this the only benefit to predictive deconvolution in the radial domain, it would justify the use of predictive deconvolution solely in the r-t domain, but there are other unexpected benefits to working in the radial domain. Initial testing suggests that the choice of operator length and of stability factor is less sensitive in the radial domain than in the x-t domain and that both can be smaller than those used in the x-t domain. In terms of stability factor, the choice of a smaller factor in the r-t domain can lead to greater resolution of deconvolved reflections. For the choice of operator length, shorter operators can be used to more effectively predict multiples. A side benefit is the attenuation of linear noise that can occur simultaneously with multiple attenuation in the r-t domain. These unexpected results still require further testing yet at this point in time they reinforce the reasons for using predictive deconvolution in the radial domain.

ACKNOWLEDGEMENTS

REFERENCES

Claerbout, J.F., 1992, Earth soundings analysis: processing versus inversion: Stanford University Press.

Henley, D.C., 1999, Coherent noise attenuation in the radial domain: introduction and demonstration: CREWES 1999 research report, 11, 455-491.

Lamont, M.G., 1999, Multiple attenuation using the MMO and ISR preconditioning transforms: The Leading Edge, 18, 110-114.

Peacock, K.L. and Treitel, S., 1969, Predictive deconvolution: theory and practice: Geophysics, 63, 155-169.

Taner, M.T., 1980, Long-period sea-floor multiples and their suppression: Geophys. Prosp., 28, 30-48.

Treitel, S. and Lines, L.R., 1982, Linear inverse theory and deconvolution: Geophysics, 47, 1153-1159.

Wadsworth, G.P., Robinson, E.A., Bryan, J.G., and Hurley, P.M., 1953, Detection of reflection on seismic records by linear operators: Geophysics, 18, 539-586.