*Stanford Exploration Project, Report 73, November 18, 1997, pages 171–196*

# Dealiasing bandlimited data using a spectral continuity constraint

*Dave Nichols*[1]

***keywords:*** *slant stack, Fourier transform, amplitudes, aliasing, least squares*


## ABSTRACT

When aliased data is slant stacked it sums coherently at multiple locations in the slowness-frequency $(p - \omega)$ domain. Only one of these locations is the "true" location. If the true location can be distinguished from the aliased locations an unaliased representation of the data can be computed. The criterion I use to distinguish between the true dip and the aliased dip is that, for bandlimited data, the amplitude spectrum at the true dip should be a locally continuous function of frequency. Once an unaliased representation of the data has been obtained any slowness domain processing can be performed before transformation back to the space-time$(x - t)$ domain. The data can be transformed back to the $x - t$ domain at any trace spacing to give an dealiased version of the original data.


## INTRODUCTION

Recent work (Spitz, 1991) has demonstrated the use of prediction filters in the $x - \omega$ domain to interpolate aliased data. Spitz's method uses the predictability of the filters as a function of frequency to overcome the problems of aliasing. The filters from low frequencies at one trace spacing are used to predict the filters at a higher frequency for a finer trace spacing.

This method has two major limitations; first the use of prediction in the x-direction requires that the input traces be sampled on a regular grid; second, if the dips at low frequencies are not the same as the dips at high frequencies the low frequency filter from the original spacing is not a good estimate of the high frequency filter at a finer spacing.

I have chosen to use a simpler property of the spectrum of dipping events to remove the aliased energy from sampled data. In the $p - \omega$ domain a band limited dipping event has a locally continuous amplitude spectrum at the true dip and a discontinuous

[1]**email:** not available

amplitude spectrum at the aliased dip. I perform a least-squares inverse transform (see e.g. (Kostov, 1990)) to the $p-\omega$ domain with constraints in the $p-\omega$ domain that are based on a measure of local continuity as a function of frequency. Once I have found an unaliased representation of the data in the slant stack domain I can transform back to the $x-t$ domain at an arbitrary trace spacing.

This method does not suffer from the same limitations as Spitz's method. The use of a least-squares inverse slant stack means that irregular trace geometries can be handled satisfactorily. By using a local continuity constraint I can handle events at different dips that have different bandwidths. My only requirement is that events be continuous in frequency over some (user definable) length. If this length is too long the method may reject valid dipping events with a narrower bandwidth. If it is too short the method may fail to reject aliased energy.

## ALIASING IN THE SLANT STACK DOMAIN

The $p-\omega$ domain can be viewed as a mapping of the $f-k$ domain along lines of constant slowness (radial lines). It should come as no surprise that aliasing in the $p-\omega$ has features in common with aliasing in the $f-k$ domain. Figure 1 shows the amplitude spectrum of a single dipping event in the and $p-\omega$ domain. The most important point to note is that in the $p-\omega$ domain all the unaliased energy for a single dip lies along a single trace of constant slowness. The aliased energy lies along trajectories given by,

$$p_{alias} = p_{true} + n\frac{2\pi}{\omega\Delta x}.$$

Where $\Delta x$ is the spatial sampling of the data and $p_{true}$ is the true dip.

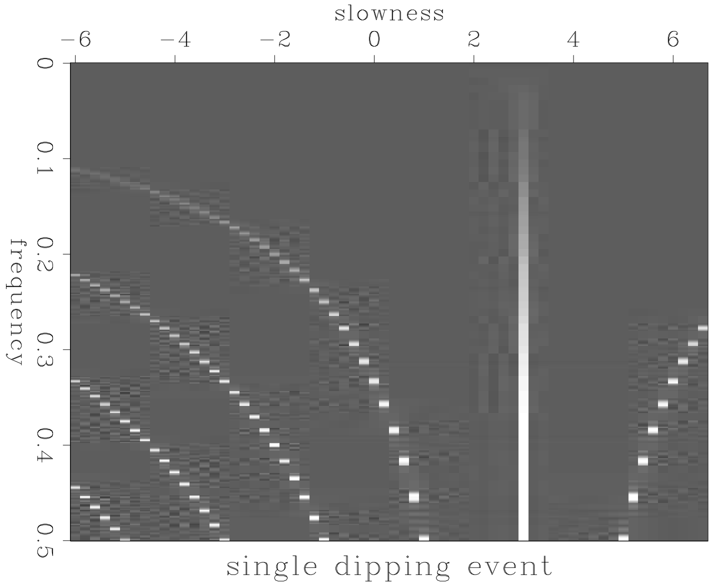



Figure 1: Amplitude spectrum of a single dipping event in the $p-\omega$ domain. The unaliased energy lies along a single "$p$ trace." The aliased energy is along curved lines that are the $p-\omega$ equivalent of the "wrapped lines" in the $f-k$ domain. dave1-single-pw [ER]

**Constructing an anti-alias mask function**

The simple representation of the unaliased energy in the $p - \omega$ space suggests a criterion for identifying aliased and unaliased energy. Energy is unaliased if the amplitude spectrum is continuous along a single slowness trace in the $p - \omega$ domain. If the amplitude spectrum changes rapidly along the trace the energy is assumed to be aliased. This is philosophically similar to the approach used by others (Zhang et al., 1992) where the zeros of a prediction filter are examined to see if they lie along a line through the origin in the $f - k$ domain.

This criterion will not be reliable if the amplitude spectrum of the wavelet for a particular dip changes rapidly. To overcome this problem I divide the amplitude spectrum for each $p$-trace by a local normalization factor. This factor is the maximum spectral amplitude over a local window in slowness. After this normalization the amplitude spectrum of the unaliased dip should be a slowly varying function of frequency while the aliased energy will still have a rapidly varying amplitude spectrum. This somewhat ad-hoc method is designed to remove spectral color whilst retaining local amplitude differences. It will fail when events with nearby dips have very different spectra.

I can now generate an "aliasing" measure by computing the continuity of the amplitude spectrum as a function of slowness. I calculate an energy normalized smoothness along each slowness trace given by,

$$\chi = \frac{(\sum_i a(i))^2}{(\sum_i a(i)^2 + \delta)}$$

The summation is over a running window over frequency. The length of this window should be longer than the bandwidth of the aliased energy on any one trace but shorter than the bandwidth of the unaliased energy. The variable $\delta$ is an energy related damping factor. If it is zero the method is similar to semblance, it will give a measure of smoothness that is independent of the energy along the trace. A high value of $\delta$ will bias the measure towards high amplitude events. I choose a value that is some fraction of the maximum smoothed energy in the whole $p - \omega$ spectrum. The fraction can be varied to allow the user to choose only high energy events or to allow events of all energies.

I generate a mask function by thresholding the smoothness measure. Any point with a value less than the threshold is assumed to be aliased energy and any point with a value greater than the threshold is unaliased energy. I have found that 70% of the maximum value is a reasonable threshold in most circumstances. Figure 2 shows the four stages in creating the mask function.

Figure 3 shows the spectrum of the event in the $p - \omega$ after the mask function has been applied. The aliased energy has been removed and the unaliased energy has been preserved.

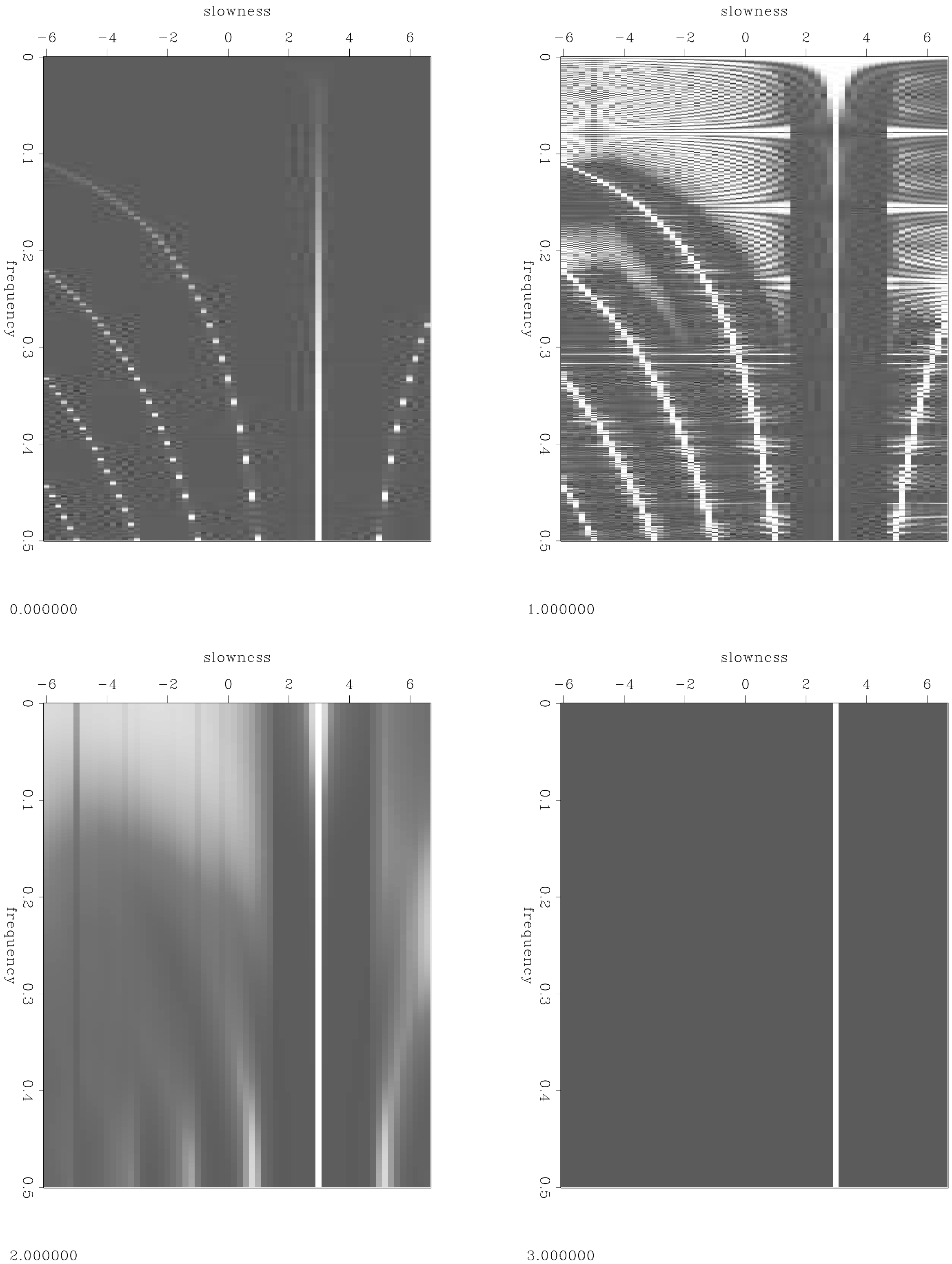



Figure 2: Four stages in creating the mask function. The input spectrum (top-left), the normalized spectrum (top-right), the coherence function (bottom-left), the mask function(bottom-right). $\boxed{\text{dave1-single-mask}}$ [ER]

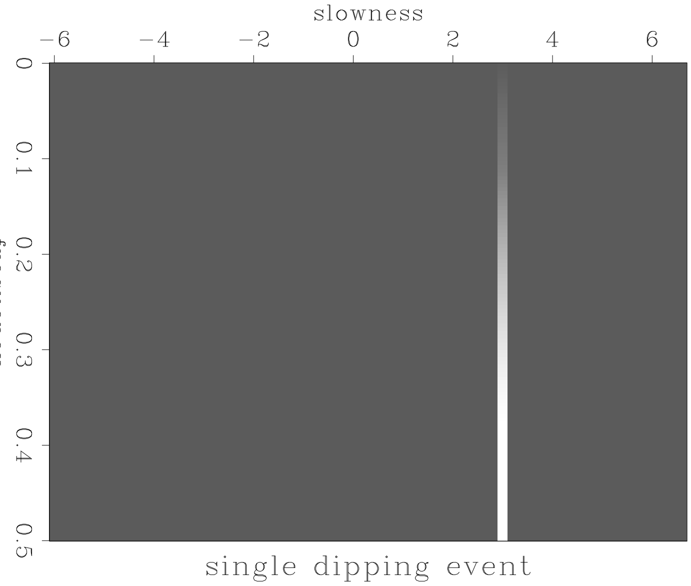



Figure 3: Amplitude spectrum of a single dipping event in the $p-\omega$ domain after application of the mask function to remove aliased events. dave1-single-wmask [ER]

**Mapping back to the $x - t$ domain**

The mapping from $p-\tau$ to $x-t$ is a simple superposition of linear events. Each sample in $p-\tau$ becomes a line in $x-t$. Figure 4 has three panels. The top panel is the original data in $x-t$; it is a single dipping event. The amplitude is constant along the event and there are no negative values in the data. The middle panel shows the $x-t$ data reconstructed from the aliased $p-\tau$ data. The high frequency energy scattered over the plot is the, mispositioned, aliased data. The bottom panel shows the $x-t$ data reconstructed from the slant stack data after application of the anti-alias mask. The dispersed aliased data is removed but there are still some artifacts. The frequency content of the original data has not been correctly recovered. This is despite the fact that a "rho" filter was applied to correct for the high frequencies lost in the slant stack transform. The rho filter is a correction derived under the assumption that the data is continuous and has infinite aperture. The artifacts result from the finite sampling and finite aperture of the original $x-t$ data.

## LEAST SQUARES INVERSE SLANT STACKS

The amplitude spectra displaying in the previous spectrum were created using a "direct" slant stack. The transformation to the slant stack domain was performed using the traditional "sum along lines" operator plus a rho filter.

$$U(p,\tau) = \rho(\tau) * \sum_{i=1}^{nx} U(x_i, t = \tau - px)$$

Where the rho filter is defined as $\rho(\omega) = |\omega|$.

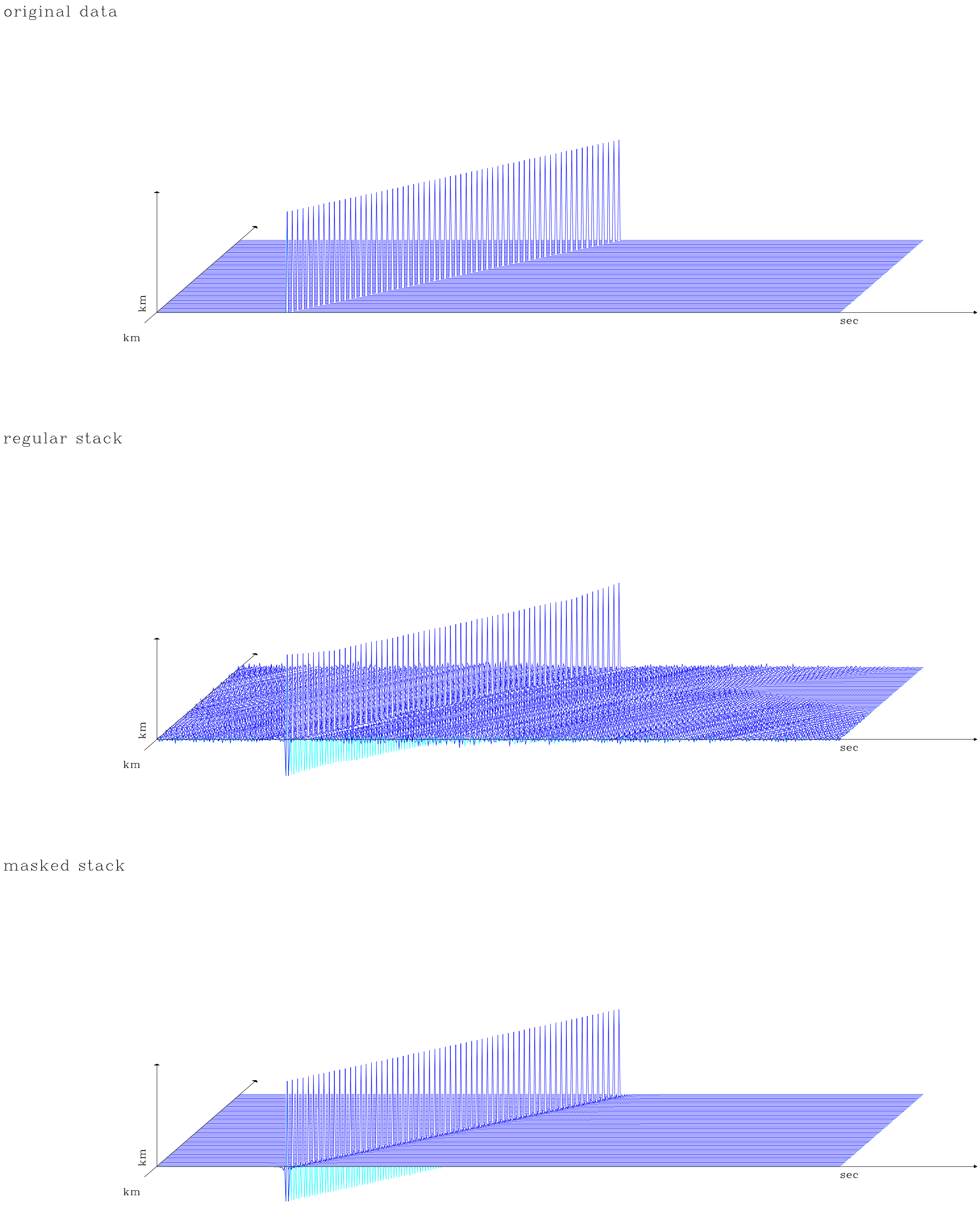



Figure 4: Top: original data, Middle: data reconstructed from the aliased slant stack, Bottom: data reconstructed from the slant stack with the aliased data masked out. $\boxed{\text{dave1-single-rec}}$ [ER]

The problems with using this approach have been well documented (Kostov, 1990). The limited aperture and sampling in the $x-t$ domain cause artifacts. Irregular sampling in x is not well handled by the direct stacking procedure. The alternative is to perform a least squares inverse slant stack. This method assumes that we can exactly perform the mapping from $p-\tau$ to $x-t$ and casts the mapping from $x-t$ to $p-\tau$ as an inverse problem. This is most conveniently calculated in the frequency domain.

The operator that maps from $p-\omega$ to $x-\omega$ is $L$,

$$L(p_i, x_j, \omega) = e^{-i\omega p_i x_j}$$

$$U(x,\omega) = L\,U(p,\omega)$$

Given the data in the $x-t$ domain we must construct a mapping to the $p-\tau$ domain that "best" reconstructs the $x-t$ data. The misfit function for the reconstruction is defined as a least-squares error.

$$||L\,U(p,\omega) \;-\; U(x,\omega)||^2$$

Minimization of this misfit is equivalent to solving the inverse problem, $L\,U(p,\omega) = U(x,\omega)$. The solution for the overdetermined problem is,

$$U(p,\omega) = (L^H L)^{-1} L^H U(x,\omega),$$

and for the underdetermined problem it is,

$$U(p,\omega) = L^H (L L^H)^{-1} U(x,\omega).$$

Where $L^H$ is the conjugate transpose of $L$.

The operator $(L^H L)$ has a Toeplitz form when the data is regularly sampled in $p$. This makes the inversion cheap to solve using Levinson recursion. A problem with this method is that the operator $L^H L$ becomes singular when the data is aliased. Kostov suggests two ways to overcome this problem, "Aliasing can be overcome by using prior information, either in the form of an interval of wavenumbers smaller than twice the Nyquist wavenumber, or by introducing "a priori" information via a model-covariance matrix." (I would suggest that the first statement can be modified to allow multiple wavenumber intervals as long as the total width does not exceed twice the Nyquist wavenumber and no two intervals occur at positions that are the alias of one another.)

My first solution to the problem is to constrain the solution to be non-zero only in those places that I have defined as unaliased. I define a masking operator, $M$, that is one where the data is not aliased and zero where it is aliased. I then solve the least squares problem,

$$min||L\,M\,U(p,\omega) \;-\; U(x,\omega)||^2.$$

The equivalent inverse problem with the solutions,

$$U(p,\omega) = (M^H L^H L M)^{-1} M^H L^H U(x,\omega),$$

and

$$U(p,\omega) = M^H L^H (LMM^H L^H)^{-1} U(x,\omega),$$

does not have the same Toeplitz structure as the unmasked problem so a different solution method must be used. I use a simple conjugate gradient algorithm (Hestenes and Steifel, 1952) to solve the minimization problem. This problem is solved separately for each frequency so the size of each problem is only $np \times nx$ and the convergence is rapid. The masking accelerates convergence at high frequencies because the problem is now better determined, there is no ambiguity between aliased and unaliased dips at each frequency. Figure 5 shows the $p - \omega$ spectrum of the least squares inverse slant stack using the masking function. All the energy is concentrated at the correct dip and the amplitude spectrum is flat. Figure 6 has three frames. The top frame is the original $x - t$ domain data. The middle frame is the reconstruction of the $x - t$ data from the masked least squares inverse slant stack. The bottom frame is the difference between the two panels plotted at the same scale. The reconstruction has been completely successful. I am not limited to reconstructing the data at the original trace spacing, the data can be modeled in the $x - t$ domain at an arbitrary regular or irregular trace spacing.

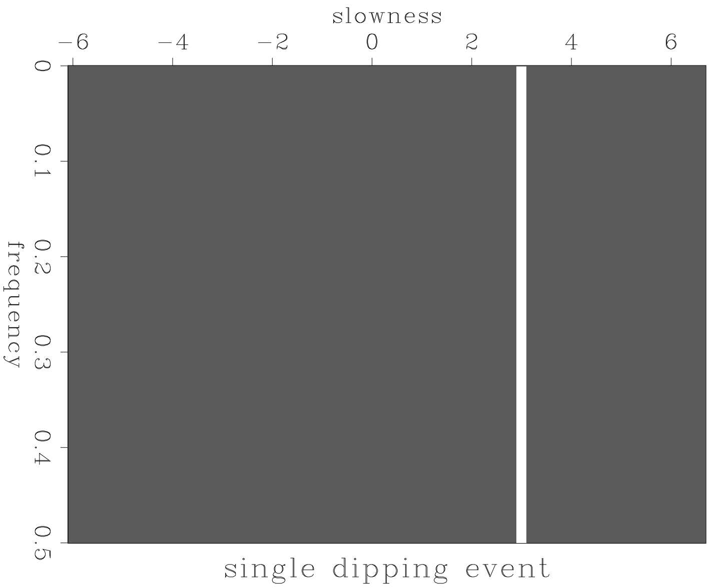
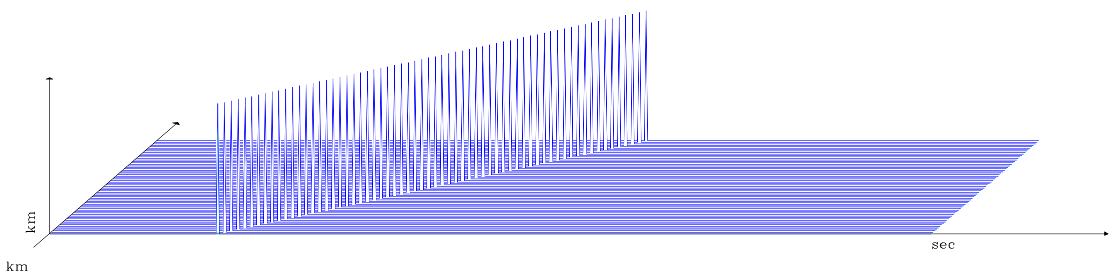


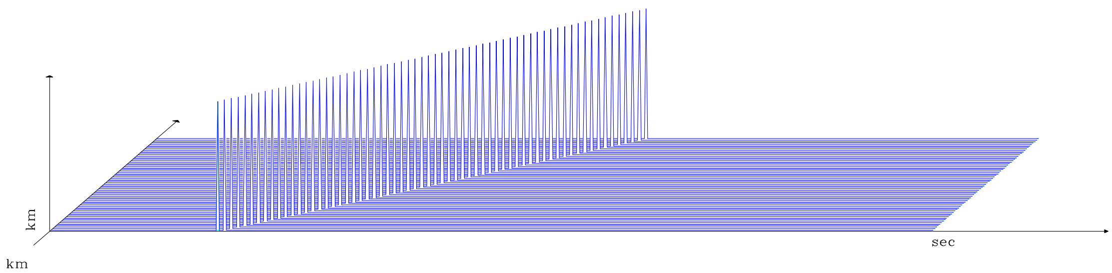

Figure 5: $p - \omega$ spectrum of the least squares transform using the masked operator. dave1-single-lsq [ER]

## BANDLIMITED DATA

This scheme can handle bandlimited data as long as the bandwidth is broad enough for the spectrum to be continuous over the length of "continuity detection" operator. Figure 7 shows the $f - k$ amplitude spectrum of three dipping events. The steepest event occurs only at high frequencies, the flattest event occurs only at low frequencies and the intermediate event occurs at medium frequencies. The original spectrum is on the left and the spectrum of the data after subsampling by a factor of four is on original data




masked LSQ stack




difference


Figure 6: Top: original data. Middle: reconstructed data using the masked least squares operator. Bottom: difference between the two plots. $\boxed{\text{dave1-single-lsqrec}}$ [ER]

the right. The data is severely aliased and the highest dip occurs only as an aliased event

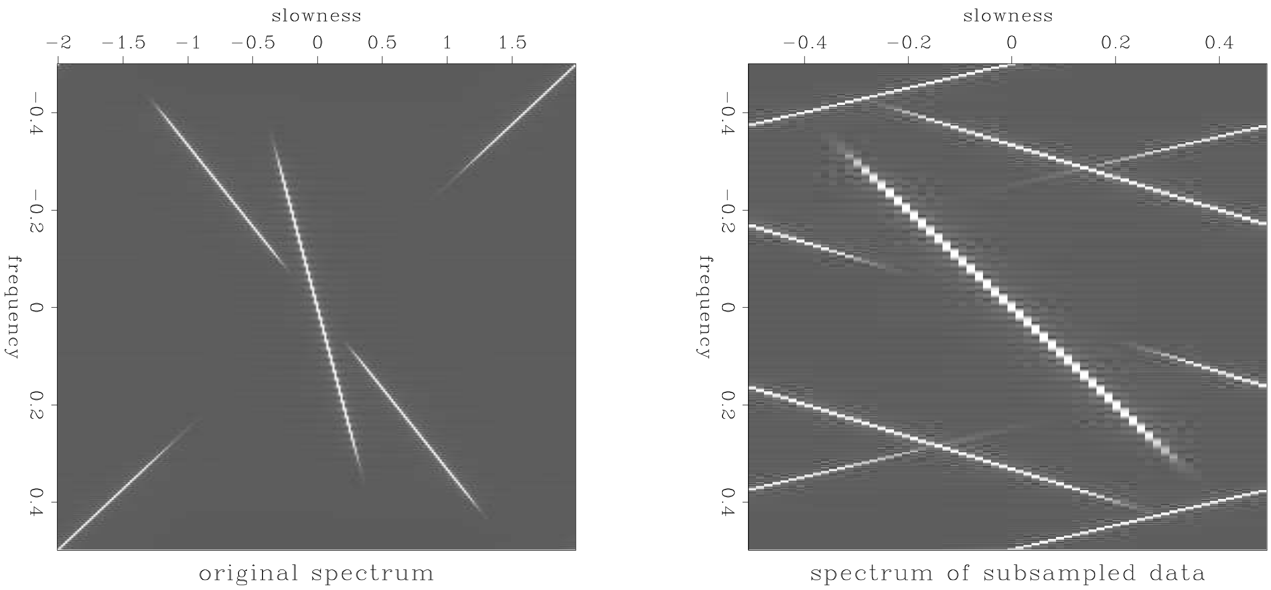



Figure 7: $f-k$ spectrum of the data, original data on the left, data at four times the sampling interval on the right. The steepest dip occurs only as an aliased event. dave1-three-fk [ER]

The $p-\omega$ domain amplitude spectrum of the subsampled data is shown in figure 8. The three events are clearly seen as events with continuity along single slowness traces whereas the aliased energy is not continuous in frequency.

Figure 9 shows the mask function derived from this spectrum. On this synthetic data the algorithm has worked excellently. The mask is an exact fit to the areas containing unaliased data.

The least squares inverse slant stack under the masking operator is shown in figure 10, the method has been very successful, the slant stack only has significant energy in the unaliased portions of the spectrum. Figure 11 shows the $f-k$ spectrum of the interpolated data, reconstructed at one quarter of the input trace spacing it is a perfect match to the original $f-k$ spectrum before interpolation, in figure 7.

## USE OF MODEL WEIGHTING FUNCTIONS

While the masked operator method works well for synthetic data it is difficult to apply to real data. The masking operator is an "all or nothing" operator, data is classified either as "aliased" or "unaliased." It is hard to create a good mask function for real data because of the presence of noise, amplitude variation along events etc. If an incorrect mask function is used the masked operator may become singular or

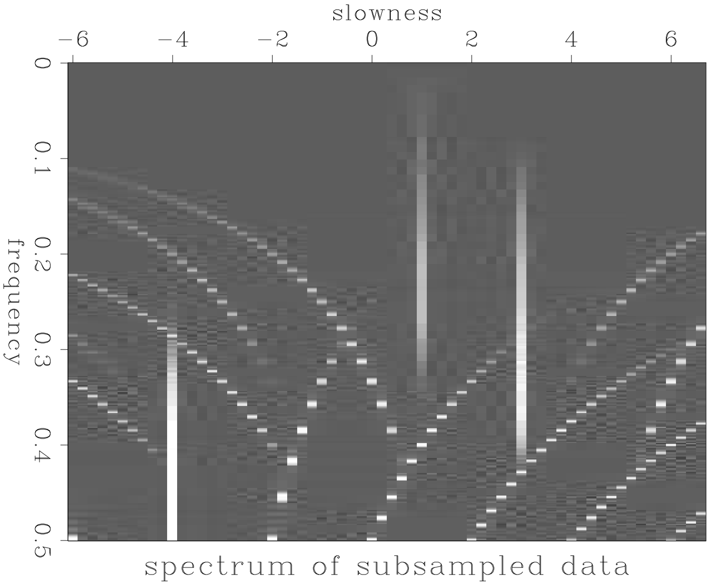



Figure 8: $p - \omega$ spectrum of the subsampled data. dave1-three-pw [ER]

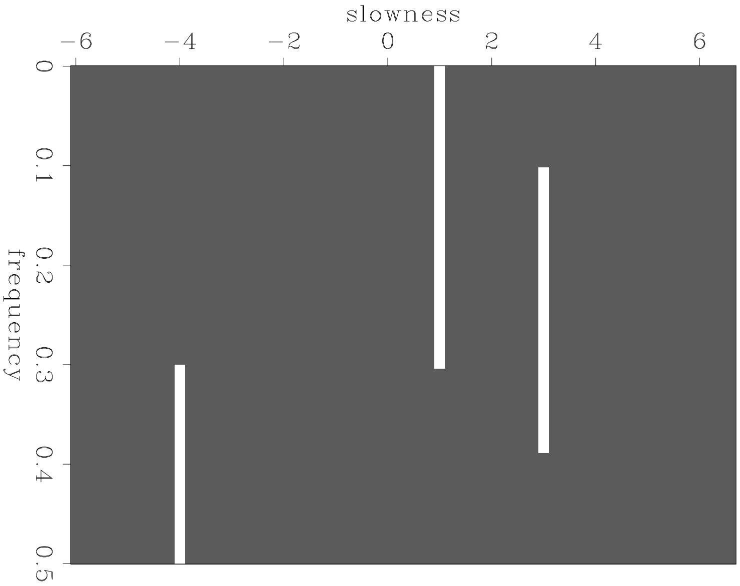



Figure 9: $p - \omega$ The mask function derived from the spectrum shown in figure 8. dave1-three-mask [ER]

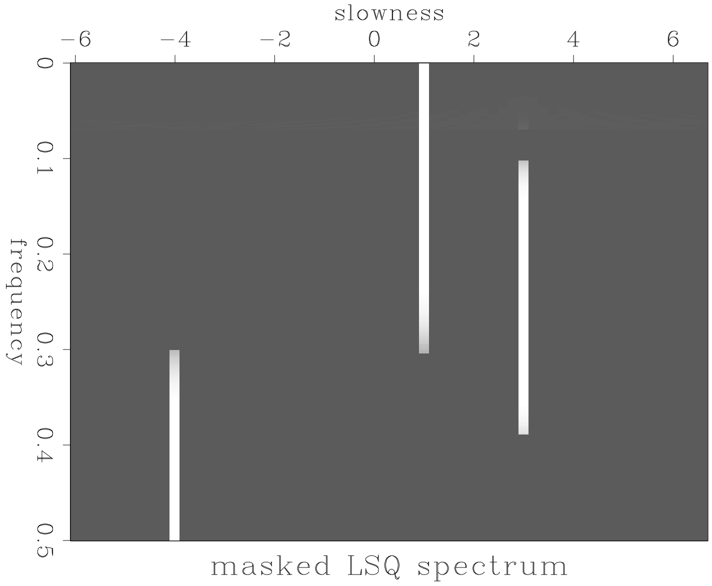



Figure 10: $p - \omega$ spectrum of the least squares inverse slant stack using the masked operator. dave1-three-masklsq [ER]

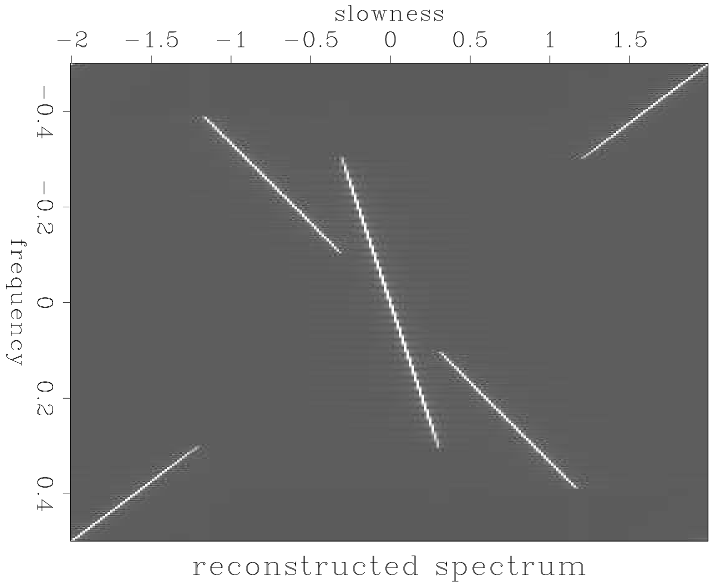



Figure 11: $f - k$ spectrum of the data reconstructed at four times the input trace spacing. Compare this with figure 7 dave1-three-rec4fk [ER]

have a large null space. The iterative solution process converges very slowly and it may produce spurious high amplitudes in the $p - \tau$ domain.

A more robust method is one that penalizes energy in the regions identified as "aliased" but does not force it to be zero. This method may be slower to converge than the mask method if the mask is chosen correctly but it will not fail completely if the weighting function is chosen poorly.

The new method minimizes the sum of the data mismatch and a weighted length measure.

$$min||L\,U(p,\omega) \;-\; U(x,\omega)||^2 + ||W\,U(p,\omega)||^2.$$

The weighting operator is a diagonal operator with weights that are high where the data is assumed to be aliased and low where it is assumed to be unaliased. If this problem is interpreted as a maximum likelihood scheme the operator $W$ is an estimate of the inverse model covariance matrix. The diagonal operator implies that model parameters are uncorrelated, with a variance that is low (where W is high) in the areas where the prior model (zero) is expected to be a good estimate, and high where the data is likely to be different from zero.

The new problem can be solved by casting it in the form of the linear system of equations,

$$\begin{pmatrix} A \\ W \end{pmatrix} U(p,\omega) = \begin{pmatrix} U(x,\omega) \\ 0 \end{pmatrix}$$

This system of equations is again solved using a simple conjugate gradient method.

I generate the weighting function in a manner similar to the masking function. Instead of using a threshold to convert the continuity function into a zero or one I scale the values to lie in the range zero to one. The inverse of this value is used as the weighting factor in calculating the weighted length of the model vector.

Figure 12 shows the inverse weighting function derived from the three-dip synthetic data and figure 13 shows the result of the inversion. As expected a little aliased energy has leaked through. This could be avoided by using a more severe weighting operator.

## A REAL DATA EXAMPLE

Figure 14 shows a real dataset. This is a crosswell dataset sampled at one foot receiver spacing. The upper frame is the data in the $x - t$ domain and the lower frame is its spectrum in the $p - \omega$ domain calculated with a "straight" slant-stack. There are two distinct regions with high energy. The large region around zero slowness corresponds to the events with low dip and some curvature. The broad dip range is due to the curvature along the events. The narrow region at a slowness of .00022(s/ft) corresponds to the steeply dipping events. These events have linear moveout and thus map to a narrow slowness band. There is no aliasing visible in this figure.

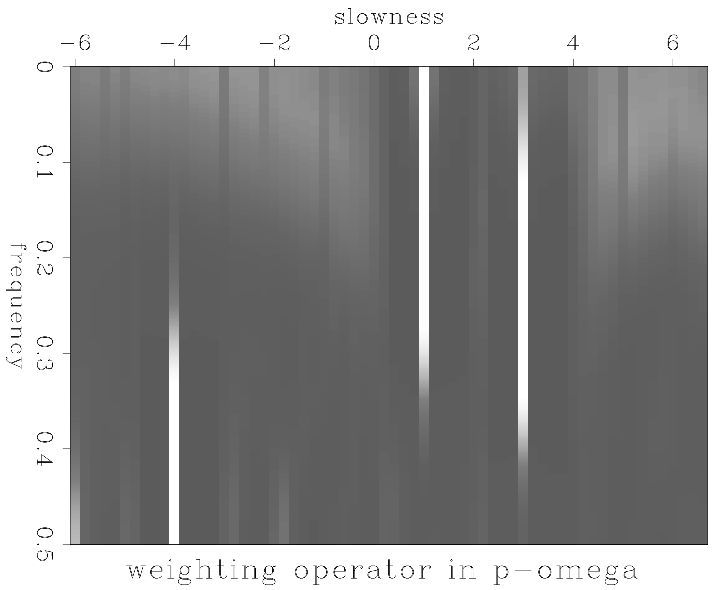



Figure 12: Inverse weighting operator displayed in the $p-\omega$ domain, the maximum value is 1.0 dave1-three-weightop [ER]

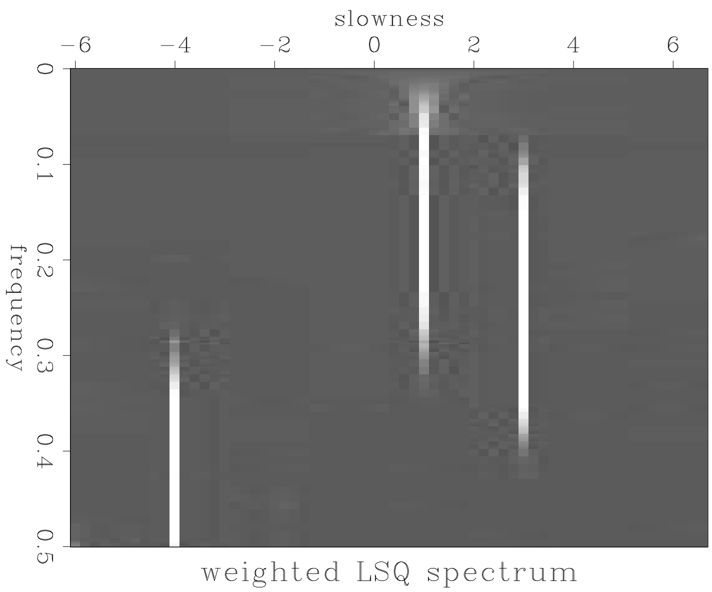



Figure 13: $p-\omega$ spectrum of the least squares inverse slant stack using the weighted operator. dave1-three-weightlsq [ER]

Figure 15 shows the same two figures after every other input trace is removed. At this input spacing the data is aliased at high frequencies. The alias of the tube wave that crosses zero slowness at about 2200Hz is particularly obvious. Figure 16 shows the same two frames when the data is sampled at three times the original trace spacing (three feet). At this trace spacing the data is aliased over the whole of the top half of the signal bandwidth.

## DEALIASING BY "STRONG" ALIAS PREDICTION

The major problems in constructing model weighting functions occur in areas where strong aliased events cross weak unaliased events. The aliases of high energy events can be given low weights if they lie within the "pass window" of a lower energy event or alternatively a low energy event can lie within a "reject region" produced by a strong alias. One way to avoid this problem is to produce a initial estimate of the spectrum that has had the aliases of the high energy events predicted and removed.

I can use the "masked" slant stack method to create an estimate of the $p-\omega$ spectrum that contains only events that are continuous in frequency *and* high amplitude. I do this by designing a mask function using a high value for the energy damping factor, $\delta$, in the continuity estimator. Given this information I can subtract the aliases of this energy from the initial estimate of the spectrum to give an improved estimate.

The aliases are estimated by assuming that the mask function selects only unaliased energy. This data is then transformed to the $\omega - x$ domain at the original spacing and then back to the $\omega - p$ domain using the simple transpose operator; using the transpose operator will give a dataset with the aliased and unaliased events. The masking function is then used to remove the unaliased events leaving only the aliases of the high energy events.

$$
\begin{aligned}
U_{orig}(p,\omega) &= L^H U(x,\omega) \\
U_{unaliased}(p,\omega) &= M U_{orig}(p,\omega) \\
U_{alias}(p,\omega) &= (I-M) L^H L U_{unaliased}(p,\omega)
\end{aligned}
$$

I then remove the estimated aliases from the original data to give a better initial estimate of the slowness spectrum. The removal is done using a weighting factor that minimizes the energy in the dealiased data.

$$
\begin{aligned}
U_{improved}(p,\omega) &= U_{orig}(p,\omega) - \alpha U_{alias} \\
\alpha &= \frac{Real(< U_{orig}, U_{alias} >)}{< U_{alias}, U_{alias} >}
\end{aligned}
$$

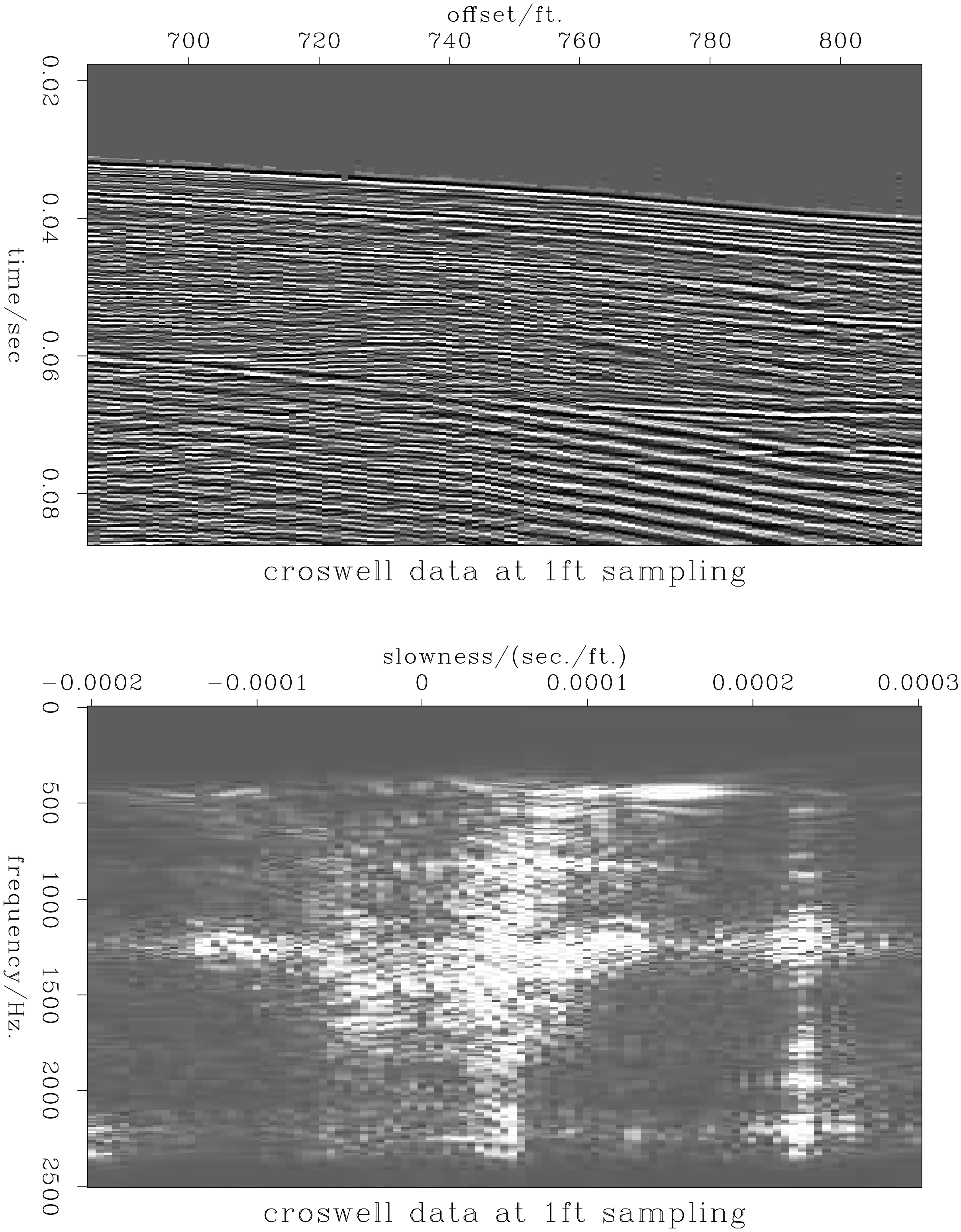



Figure 14: Crosswell data sampled at a one foot receiver interval. Top frame is the $x - t$ data, bottom frame is the $p - \omega$ data. dave1-real-orig [ER]

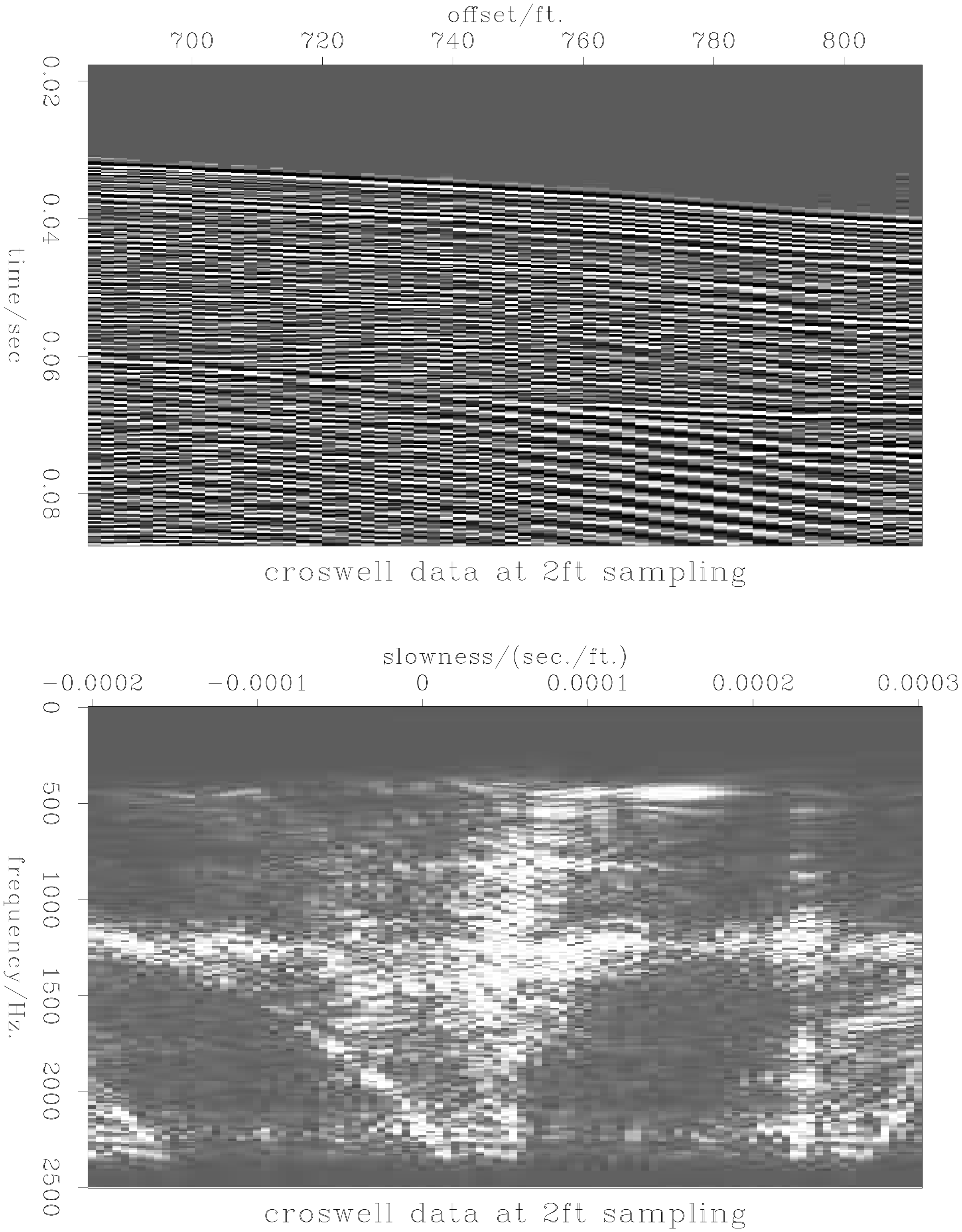



Figure 15: Crosswell data sampled at a two foot receiver interval. Top frame is the $x-t$ data, bottom frame is the $p-\omega$ data.Aliasing occurs above 1500Hz. dave1-real-sub2 [ER]

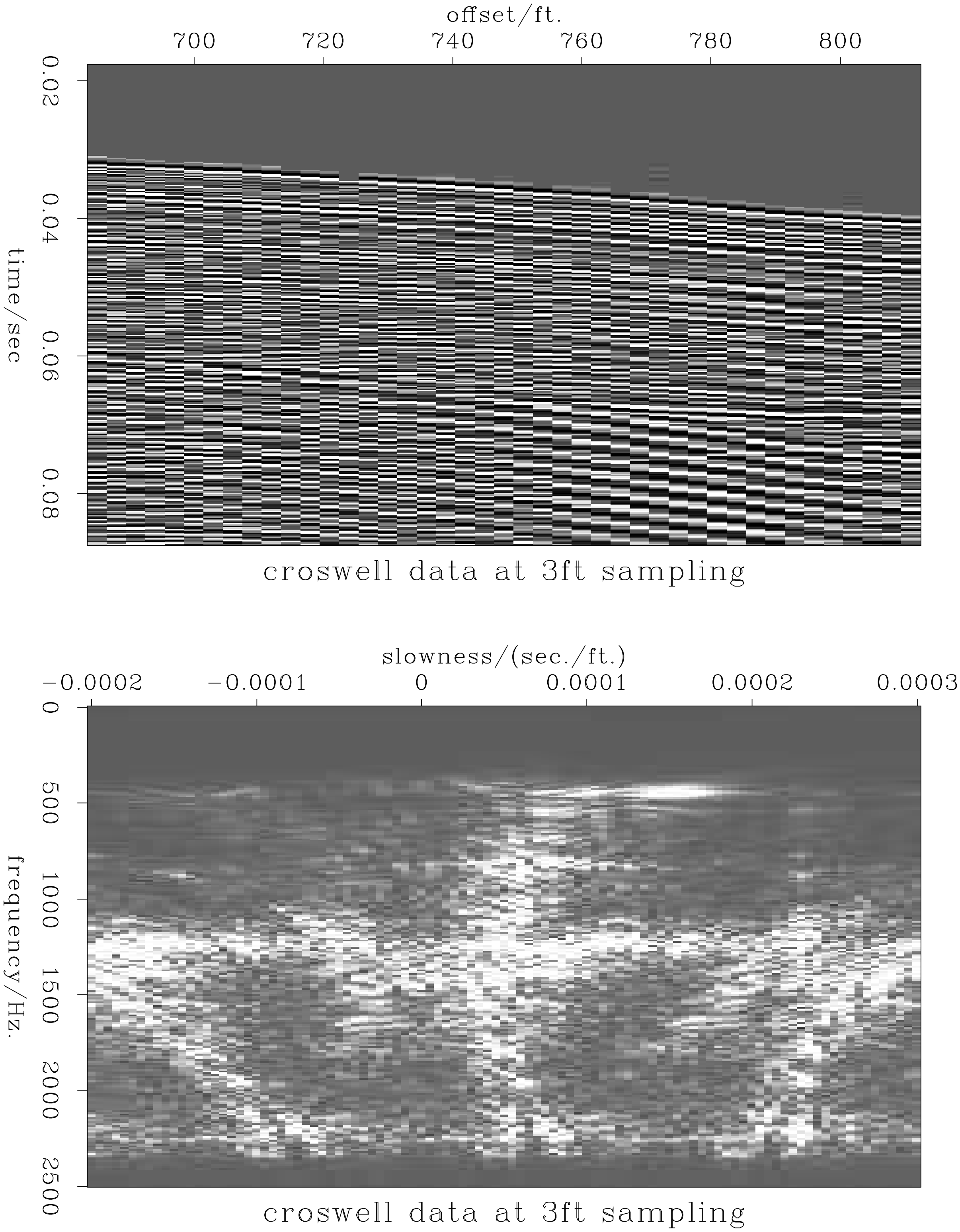



Figure 16: Crosswell data sampled at a three foot receiver interval. Top frame is the $x - t$ data, bottom frame is the $p - \omega$ data. Aliasing occurs above 750Hz. dave1-real-sub3 [ER]

This improved estimate is now used as the initial estimate of the solution in the iterative solver. Figure 17 shows four stages in the estimation of the improved initial estimate. The frames show the original data, the windowed data which is assumed to contain strong, unaliased events, the estimate of the aliases of the strong events, and finally the improved initial estimate of the $p - \omega$ spectrum.

The least-squares estimate of the $p - \omega$ spectrum generated from this improved estimate is displayed in figure 18 the equivalent panel for the data samples at a 3ft. interval is displayed in figure 19. Both are similar to the original spectrum calculated from the data at a 1ft. sample interval (figure 14). The detailed amplitudes are slightly different but on both plots almost all the aliased energy has been suppressed.

Figure 20 shows the data at 2ft. reconstructed at a 1ft spacing and the difference between that and the original data at a 1ft. spacing. Figure 21 shows the same figures for the data subsampled to a 3ft. spacing. The reconstruction from the 2ft data is satisfactory. The main differences appear to be trace-to-trace amplitude changes which were lost in the original subsampling and some energy dipping to the left which was outside the slowness range modeled in the slant stack data. The reconstruction from the 3ft. data is obviously worse but it still retains the major features of the original data and it does not contain any aliased data.

## LOCAL WINDOWS

The effectiveness of this method depends on the amount of overlap between the unaliased energy at one dip and the aliased energy of another dip. If all the aliased energy is in the same region as some unaliased energy it will not be rejected. If the unaliased energy has a broad dip range e.g. because of event curvature then overlap with aliased energy is much more likely. One way to reduce the dip-bandwidth of the $\omega - p$ spectrum is to process the data using local windows. The windows should be designed to include only a few distinguishable dips.

There are two tradeoffs to be considered when designing windows. As the windows become smaller in time the number of frequencies in the $\omega - p$ domain is reduced. When the number of frequencies is smaller the effectiveness of the continuity detection is reduced. As the windows become smaller in space the dip resolution of the slant stack is reduced. The real and aliased events become more spread-out in slowness and are more likely to interfere.

Figure 22 shows the result of reconstructing the original data using least squares transforms applied in windows on the data. The windows were 32msec long and 32 ft wide. Each window overlapped its neighbors on each side by 25%. I used the data patching algorithm described elsewhere in this report (**?**). This reconstruction is slightly superior to the one in figure 20 where the data was all processed together.

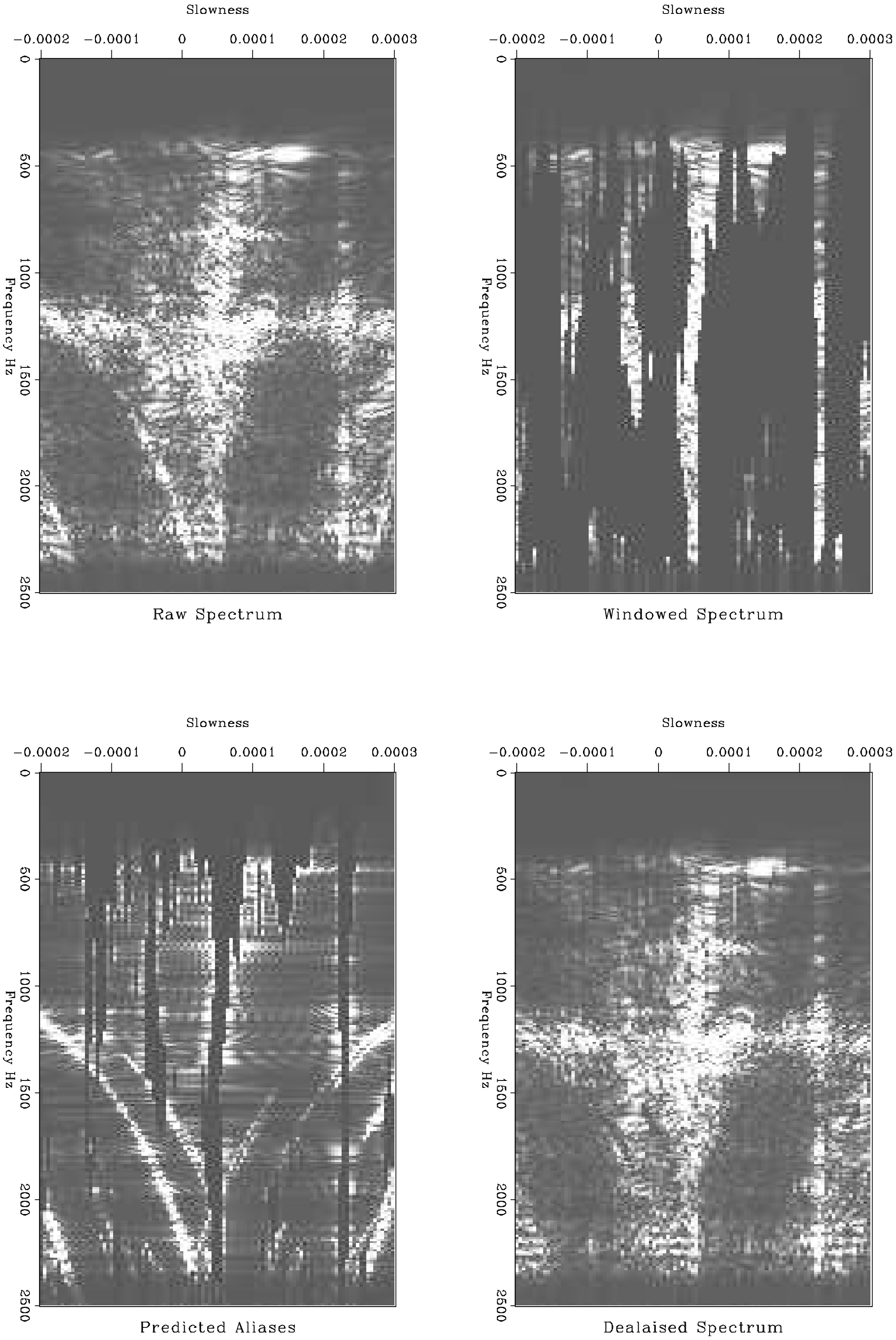



Figure 17: Four stages in the estimation of an improved initial estimate without the strong event aliases. The original data, the window of "unaliased events," the predicted aliases, the de-aliased data. dave1-real3-dealias [ER]

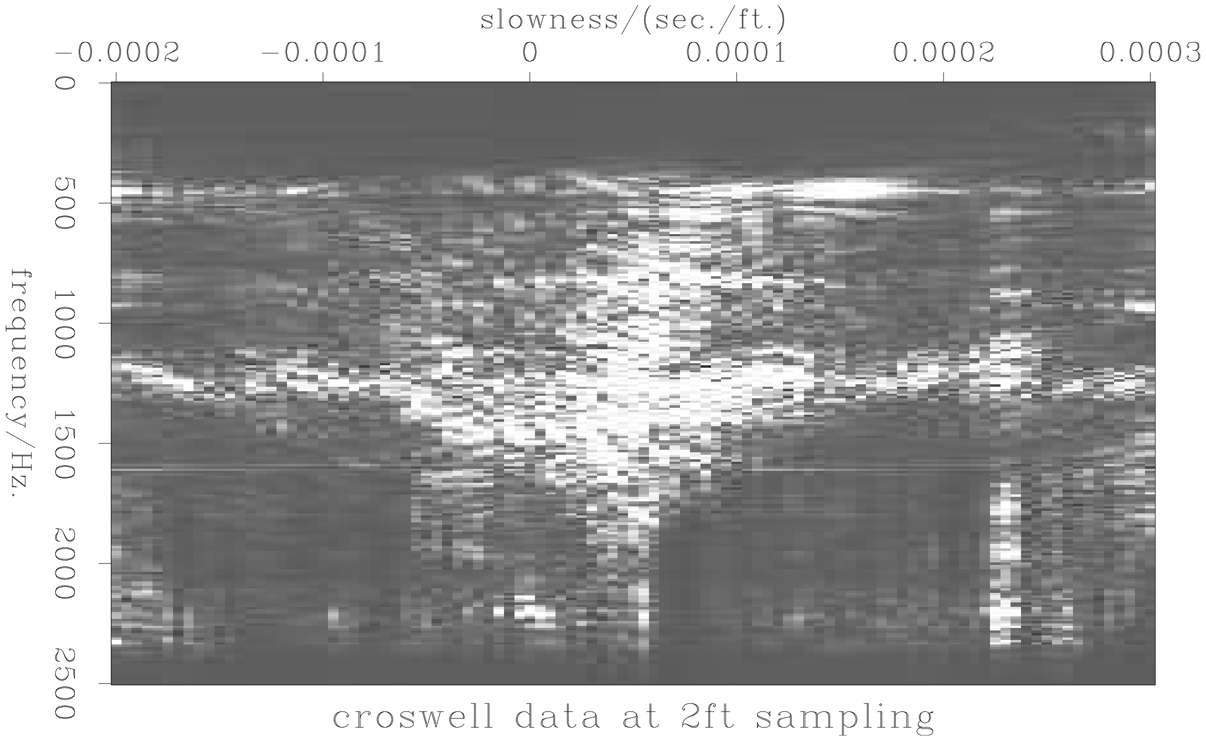



Figure 18: Crosswell data sampled at a two foot receiver interval. This is the weighted least-squares estimate of the $p - \omega$ spectrum after initial removal of strong aliases. dave1-real3-ps [ER]

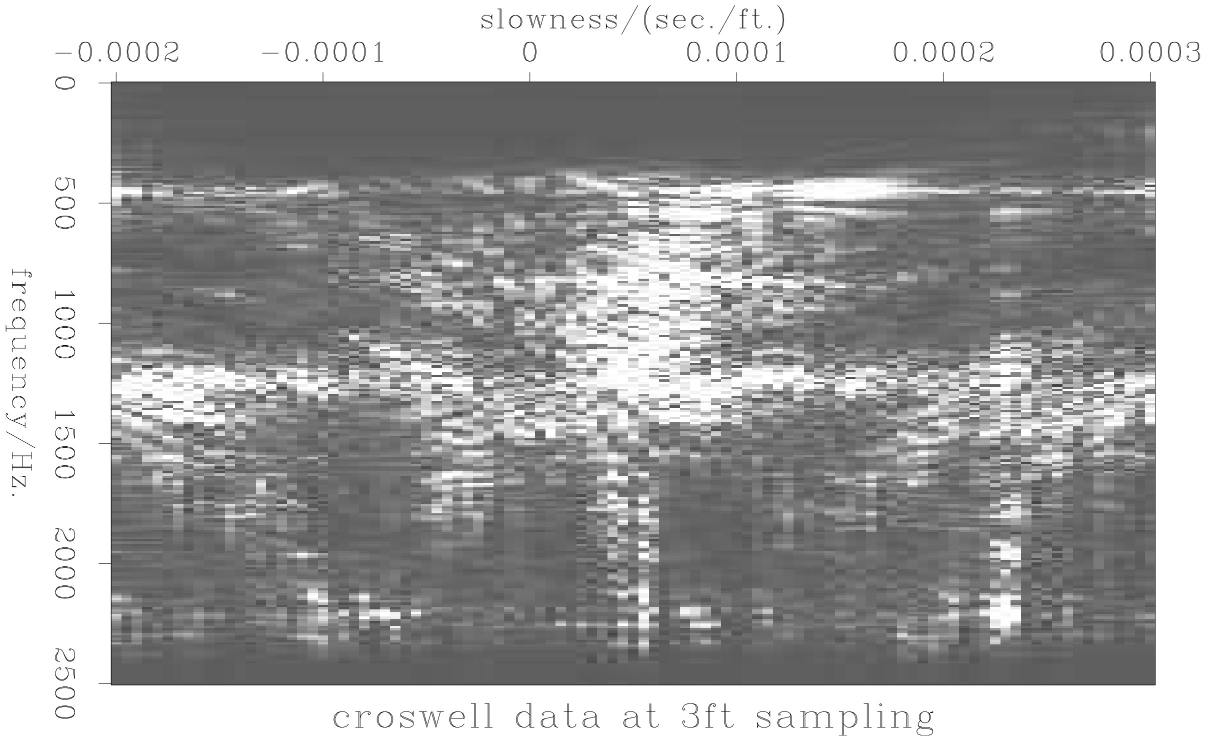



Figure 19: Crosswell data sampled at a three foot receiver interval. This is the weighted least-squares estimate of the $p - \omega$ spectrum after initial removal of strong aliases. dave1-real4-ps [ER]

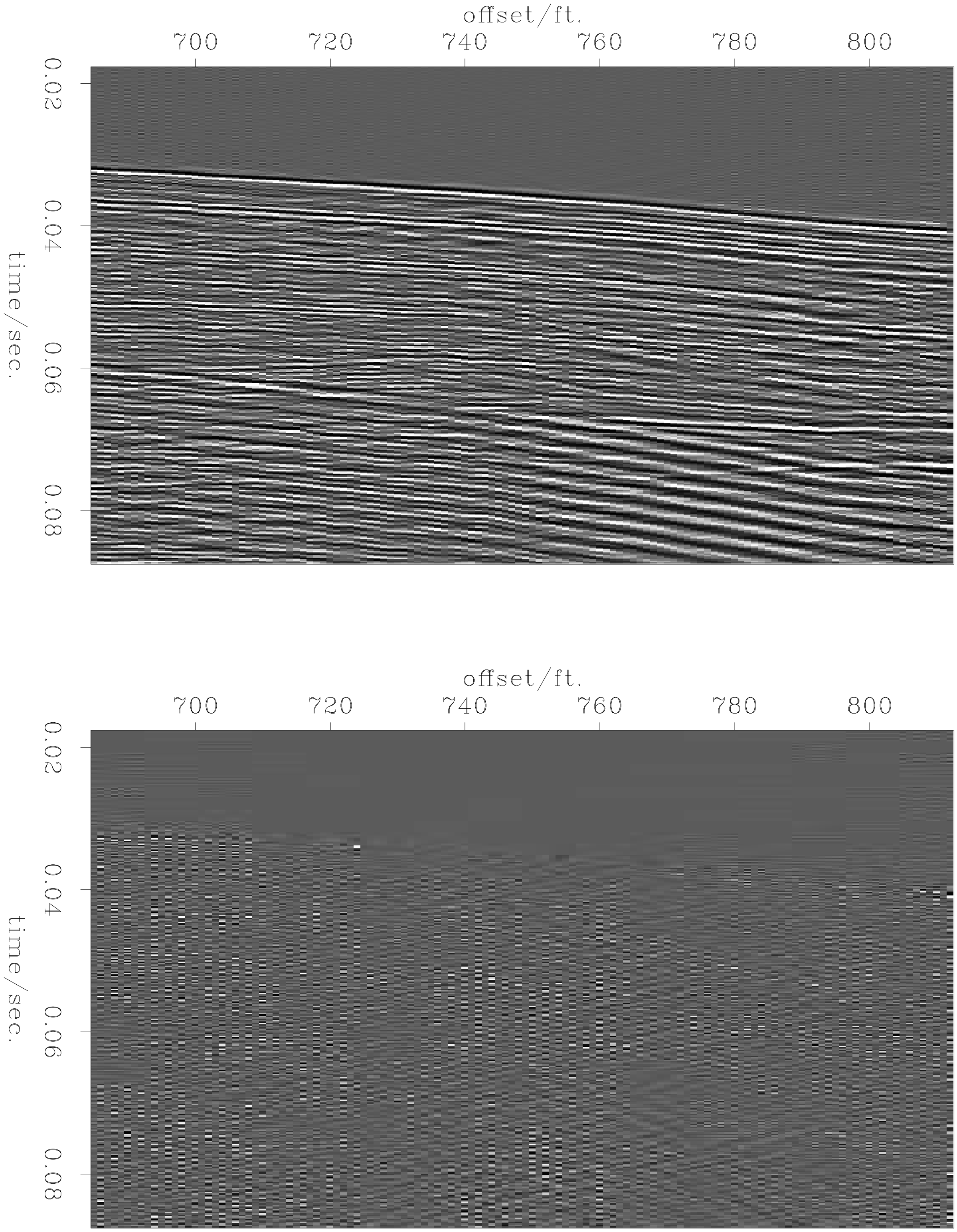



Figure 20: Results of reconstructing data at the original 1ft spacing from data at a 2ft spacing. Top panel is the reconstructed data, the bottom panel is the difference from the original data. The main difference is some back-scattered energy that was not within the dip range modeled in the slowness domain. dave1-real3-diff [ER]

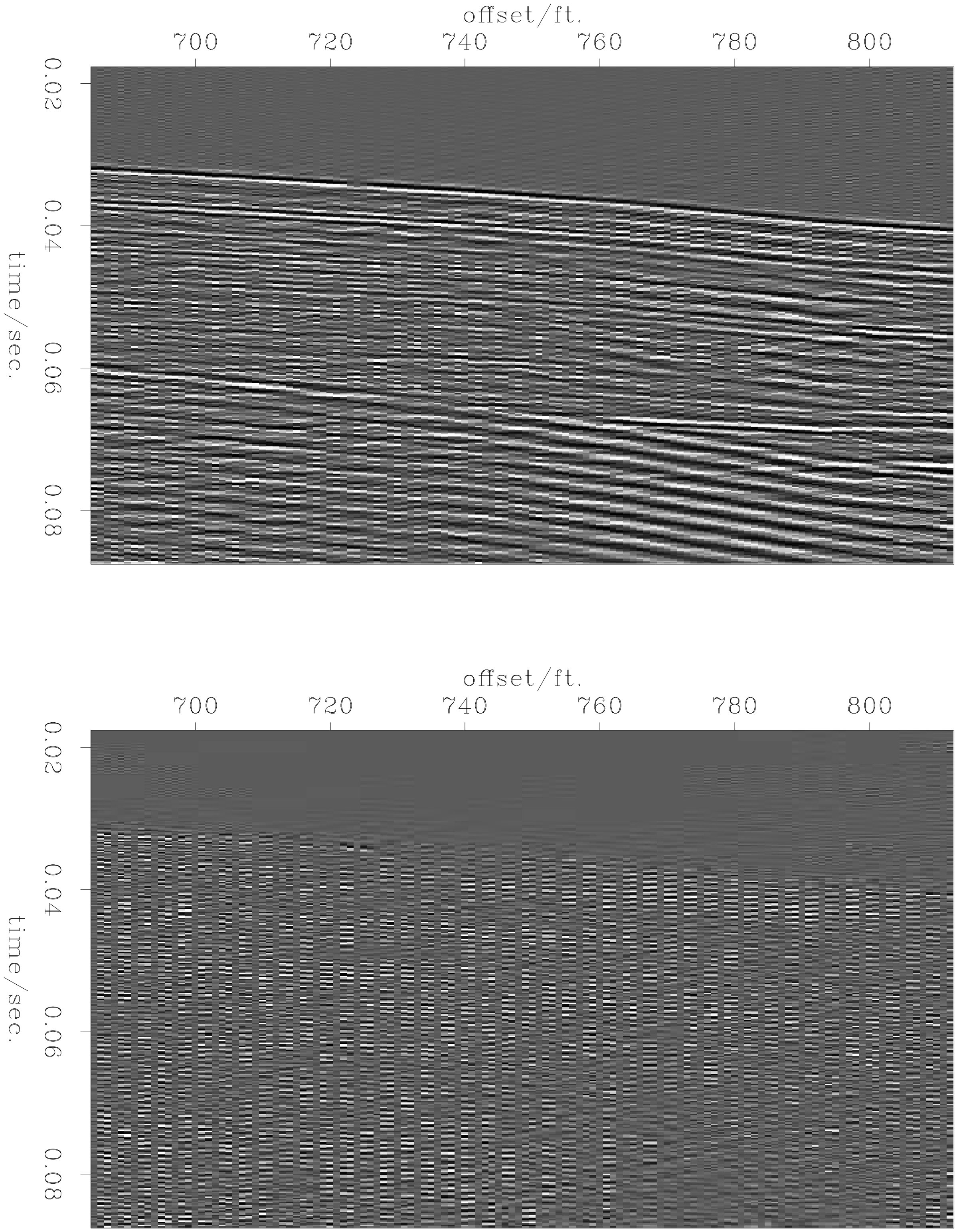



Figure 21: Results of reconstructing data at the original 1ft spacing from data at a 3ft spacing. Top panel is the reconstructed data, the bottom panel is the difference from the original data. These results are clearly worse that the results from the 2ft. data but they are not a total failure. $\boxed{\text{dave1-real4-diff}}$ [ER]

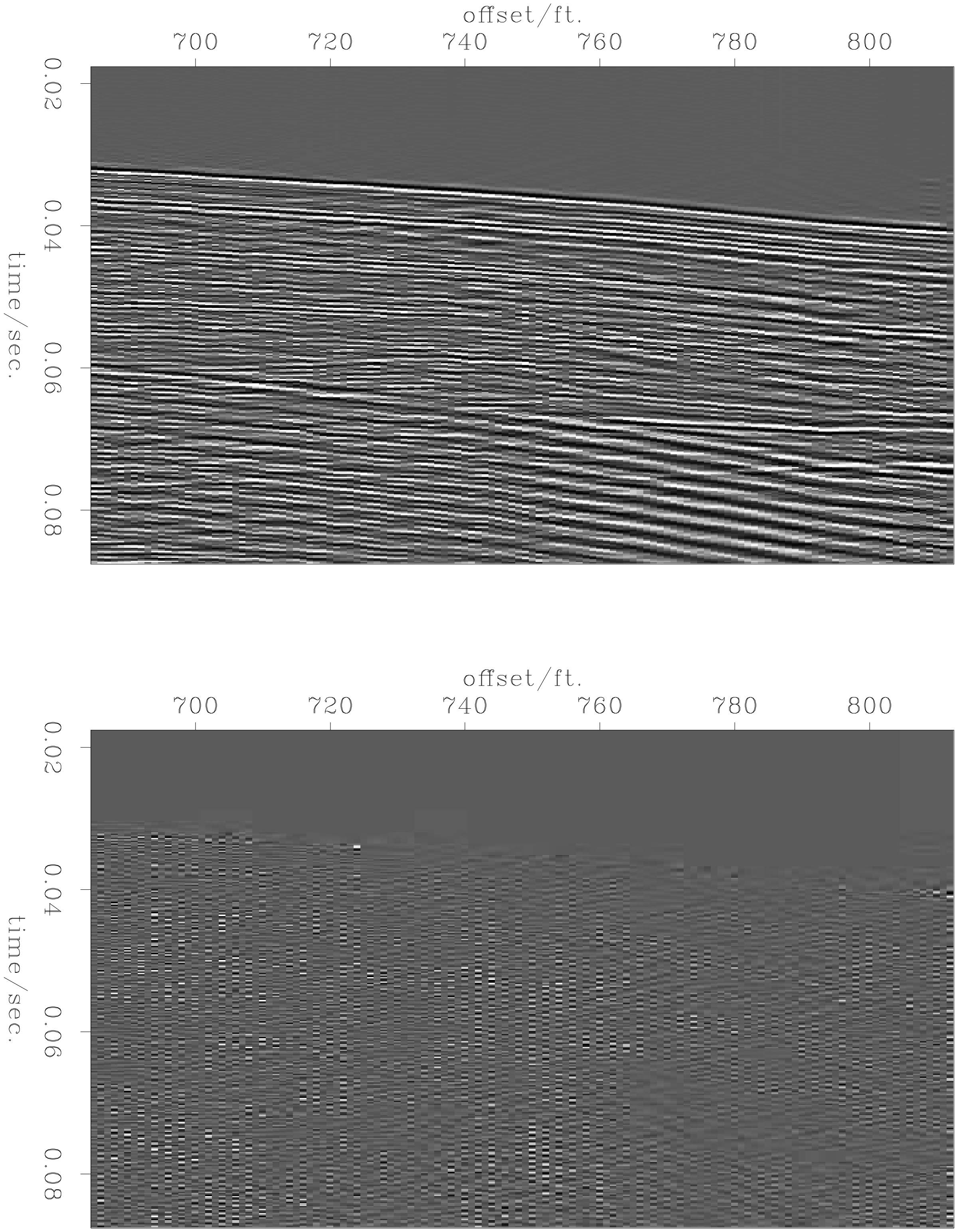



Figure 22: Results of reconstructing data at the original 1ft spacing from data at a 2ft spacing. The least squares slant stack was applied in windows that were 32msec long and 32 ft wide. Top panel is the reconstructed data, the bottom panel is the difference from the original data. This result should be compared with 20 dave1-real3-win-diff [ER]

## CONCLUSIONS

I have successfully used a spectral continuity constraint to dealias the least-squares slant stack operation. Two methods were used, the first applies a windowing constraint to the slant stack transform. This method works well for synthetic data but was too harsh for real data. The second method applies a weighting operator in the slant stack transform. The windowed slant stack can be used to predict and remove strong aliases before application of the weighted inverse slant stack. This method was effective in dealiasing a real dataset.

The method can be applied to a complete dataset or in local windows on the data. When choosing the window size there is a tradeoff between the effectiveness of the dealiasing and the ability to dealias curved events.

## REFERENCES

Hestenes, M., and Steifel, E., 1952, Methods of conjugate gradients for solving linear systems: J. Res. N.B.S., **49**, 409–436.

Kostov, C., 1990, Seismic signals from a drillbit source: Ph.D. thesis, Stanford University.

Spitz, S., 1991, Seismic trace interpolation in the f-x domain: Geophysics, **56**, no. 6, 785–794.

Zhang, L., Claerbout, J. F., and Ji, J., 1992, Seismic trace interpolation beyond spatial aliasing: SEP–**73**, 185–202.