Stanford Exploration Project, Report 102, October 25, 1999, pages 1–128

Predictive signal/noise separation of ground-roll-contaminated data

Morgan Brown, Robert G. Clapp, and Kurt Marfurt[1]

keywords: *coherent noise, ground roll, nonstationary filtering*

ABSTRACT

Coherent noise contamination is a first order problem plaguing the imaging of seismic data acquired in both land and marine environments. We present a new method for the predictive separation of coherent noise from prestack data which operates in the $t-x$ domain. We apply the new method to real 2-D receiver lines coming from a 3-D shot gather, and nondestructively separate hyberbolic ground roll from primary reflections. This method performs favorably compared to other common techniques, even with an imperfect model of the ground roll, a fact which makes the method attractive in cases where the noise is difficult or expensive to model explicitly.

INTRODUCTION

Historically, high acquisition costs and data with low signal-to-noise ratio have handicapped land-based seismic acquisition efforts. Chief among the causes of low S/N ratio is *ground roll*, which to first order consists of direct-arrival and scattered energy from the shot-generated Rayleigh (SV) wave. Larner et al. (1983) recognized that coherent noise often add constructively after conventional imaging steps, producing spurious signatures that obscure real primary reflections, or worse, may masquerade as primaries themselves. The authors argued convincingly that coherent noise events should be removed in the prestack domain; this argument is not disputed today.

In (t, x, y) space, the ground roll falls nearly within a cone centered at the shot location. Traditional 2-D roll-along land surveys consist of a single receiver line with a shot in the center, such that the recorded ground roll lies at the apex of the cone and has nearly linear moveout. In 3-D surveys, the ground roll on a 2-D receiver line has hyperbolic moveout when there is crossline distance between the shot, as in the common "cross-spread" acquisition geometry, shown in Figure 1. A single 2-D receiver line from a 3-D shot gather from such a geometry is shown in Figure 2.

In many cases, regions containing ground roll are simply muted from the recorded data, since they are localized in $t-x$. Clearly, however, it is preferable to use both the near and far

[1]**email:** morgan@sep.stanford.edu, bob@sep.stanford.edu, kmarfurt@tank.agl.uh.edu

offsets, rather than trying to "outrun" the ground roll by recording at very far offsets. High acquisition cost or geological considerations may prevent the recording of far-offset reflections, especially for deep targets. Likewise, multicomponent seismic applications such as the determination of fracture orientation and shear wave extinction angles (Mueller, 1992; Li and Mueller, 1997), and also quantitative inversion techniques require both near and far offsets. Furthermore, as noted by Regone (1997), in regions of rugged topography, body waves sometimes convert into surface waves after scattering from near surface heterogeneities, creating a high velocity event that reaches into the far offsets. The need to adaptively identify and nondestructively filter ground roll is pressing, but the geographical variability of near surface geology gives ground roll wildly variable character, and prevents most current techniques of removal from achieving universally applicability.

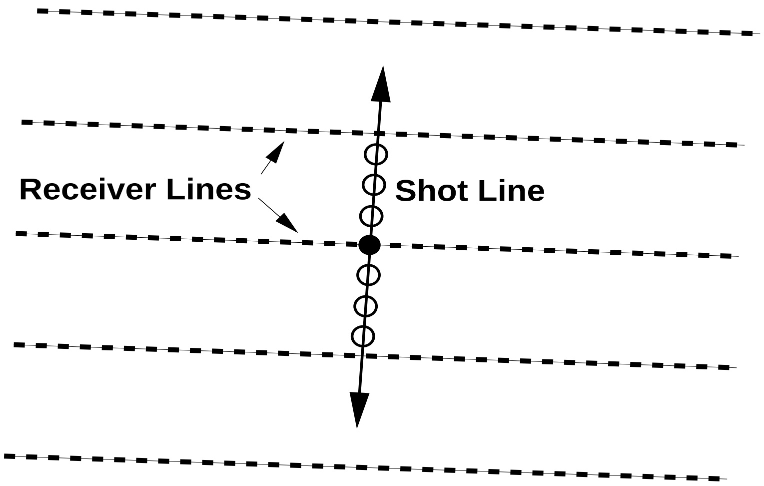

Figure 1: So-called "cross-spread" 3-D land acquisition geometry. The spacing between the parallel receiver lines is much greater than the inline receiver spacing. morgan1-cross-spread [NR]

In spite of a long and successful tradition in ground roll suppression, field arrays often damage the recorded signal. Averaging operations imply an inherent loss of information, which in this case is the frequency content, and hence the resolving power, of the recorded signal. Incident primary reflections are never perfectly vertical, nor is the near surface perfectly laterally invariant, so a field array stack will always blur the wavelet. When the static shifts caused by near surface heterogeneities are strong, the blurring is significant. Often, as with Ocean Bottom Cable (OBC) or smaller-scale seismic applications, placement of arrays with adequate areal extent is impractical for logistical and/or economic reasons. Furthermore, traditional 2-D field arrays are ineffective in removing out-of-plane backscattered energy, so in many cases a 3-D array is required (Regone, 1997), adding considerably to the acquisition cost.

Often, as seen in Figure 2, ground roll occupies a much lower frequency band than the primary reflections, in which case the ground roll is easily removable by (ω, x) mute, at the expense of the low frequency component of the signal, which is crucial for acoustic impedance inversion (Sbar and Blankenship, 1998). $f-k$ (or "fan") filtering (Treitel et al., 1967) is useful for removing linear events, but is sensitive to aliased data, and ground roll is almost always aliased at normal receiver spacing. The aliasing becomes even more pronounced in 3-D, since acquisition costs usually prevent adequate sample of the crossline.

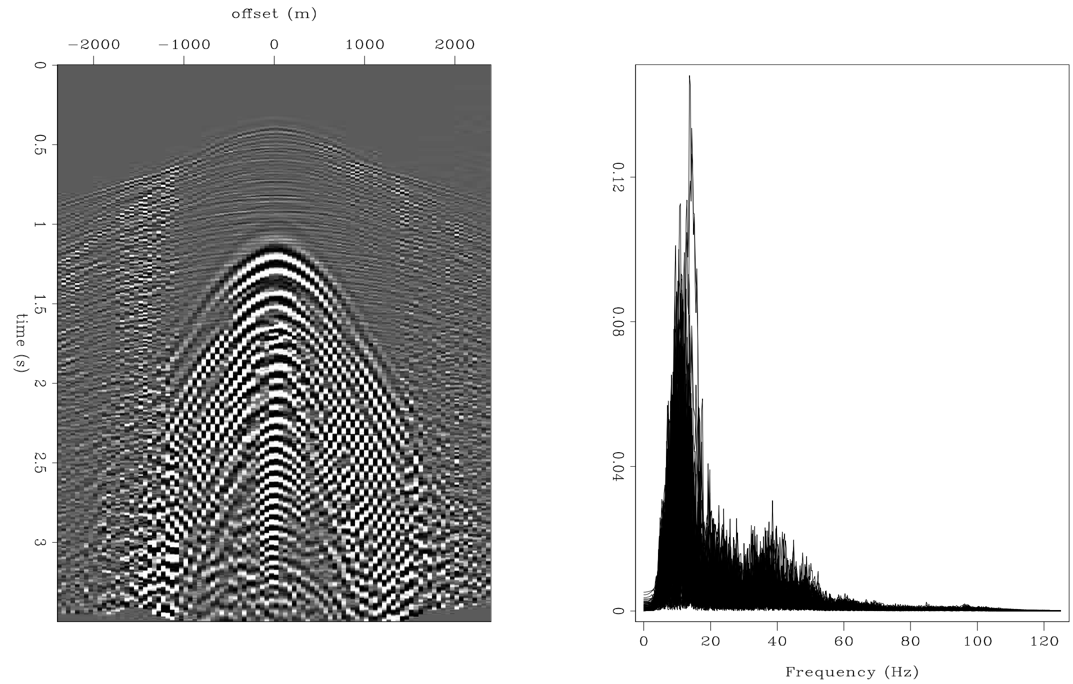

Figure 2: Left: Receiver line windowed from a 3-D shot gather of a Middle East cross-spread survey. Since the penetration depth of Rayleigh waves is a decreasing function of frequency, they are generally dispersive for $v(z)$ media. Strong, dispersive, aliased ground roll with hyperbolic moveout contaminates the lower portion of the section. The signal-to-noise ratio is quite low. Right: Data spectrum. Notice considerable separation of noise and signal spectra. morgan1-dune-spectrum [ER]

Ground roll and primaries normally differ greatly in moveout, so in that sense, the Discrete Radon Transform (DRT) can remove the ground roll in 2-D or 3-D. However, Marfurt et al. (1996) point out the extreme sensitivity of the DRT mute process to aliasing, so an a priori moveout correction to the ground roll is required, which in turn may cause aliasing of the primaries themselves. Additionally, the DRT is unable to remove backscattered energy from the crossline direction. Still, when the ground roll in a single 2-D receiver line is almost perfectly hyperbolic, as it is for instances of very well-behaved near-surface geology, the DRT may adequately separate it from the signal (Wang, 1997).

Canales (1984) introduced a fundamentally new view of the signal/noise separation problem. Rather than viewing noise as something to be explicitly subtracted from data, he exploited the fact that seismic signal is usually predictable relative to random noise. When a prediction error filter (PEF) is estimated from a given dataset, Gaussian noise components of the data are in the nullspace of the estimation. It is a known (Claerbout, 1976) that the frequency response of a PEF approximates the inverse spectrum of the data from which the PEF was estimated. By computing a PEF ($f - x$ domain) from the noise-contaminated data, this approach obtains an "in situ" estimate of the signal power spectrum. Application of Wiener reconstruction (Castleman, 1996) then yields an optimal (least squares) estimate of the signal.

Coherent noise suppression is a more difficult matter, since it requires an indirect method of estimating the signal PEF, like that of Spitz (1999), who presented a predictive $f - x$ technique to suppress multiple reflections. We present a similar predictive signal/noise separation scheme in the $t - x$ domain which utilizes recent SEP advances in nonstationary filtering, and apply it to the removal of complicated ground roll in 2-D receiver lines which come from a 3-D land survey. This approach requires only a rough model of the ground roll, obviating the need for explicit modeling techniques, which are generally complicated and very costly. Clapp and Brown (1999) have applied the same methodology to suppress multiple reflections in prestack marine data.

METHODOLOGY

Canales' (1984) implementation is in the $f - x$ domain, so he makes the implicit assumption that the data is time-stationary, i.e., that each trace is the convolution of a single time-invariant wavelet with the earth's random reflectivity sequence. Computationally, this approach is very efficient. However, since ground roll is highly dispersive, and thus nonstationary, we instead choose a $t - x$ method utilizing nonstationary PEF's (Claerbout, 1998a; Crawley, 1999; Clapp et al., 1999).

Consider the recorded data to be the linear superposition of coherent signal plus coherent noise:

$$\mathbf{d} = \mathbf{s} + \mathbf{n}. \tag{1}$$

Also assume that both the signal and noise are predictable, i.e., made up of of one or more local plane wave segments. The prediction error (residual) of the convolution of the signal and the noise with the corresponding nonstationary PEF's $\mathbf{S}$ and $\mathbf{N}$ is then uncorrelated. Writing these ideas as convolutional "fitting goals", we have

$$\begin{aligned} \mathbf{r_n} &\equiv \mathbf{Nn} \approx 0 \\ \mathbf{r_s} &\equiv \epsilon\mathbf{Ss} \approx 0 \end{aligned} \tag{2}$$

where $\mathbf{S}$ and $\mathbf{N}$ represent $t-x$ convolution with the nonstationary signal and noise PEF's, respectively, and the scaling factor ϵ balances the energies of the residuals. Equation (2) can be rewritten in an equivalent, but more familiar, notation as the minimization of a quadratic objective fuction:

$$Q(\mathbf{n},\mathbf{s}) \equiv \|\mathbf{r_n} + \epsilon\mathbf{r_s}\|^2 = \|\mathbf{Nn} + \epsilon\mathbf{Ss}\|^{\mathbf{2}},$$

but we will hereafter use the notation of equation (2).

Rewriting the constraint equation (1), $\mathbf{n} = \mathbf{d} - \mathbf{s}$, allows us to eliminate $\mathbf{n}$ from equation (2), and suggests a regularized least squares optimization problem for the unknown signal $\mathbf{s}$

$$\begin{aligned} \mathbf{Ns} &\approx \mathbf{Nd} \\ \epsilon\mathbf{Ss} &\approx 0, \end{aligned} \tag{3}$$

which is the approach used by Abma (1995). It can be easily shown (see Appendix A) that the predicted signal of equation (3) is the same as the optimal Wiener reconstruction (Castleman, 1996; Leon-Garcia, 1994) for the special case of uncorrelated signal and noise.

We now consider some issues involving the calculation of the nonstationary signal and noise PEF's, $\mathbf{S}$ and $\mathbf{N}$.

- **Signal PEF**
  Equation (3) treats the signal PEF $\mathbf{S}$ as known, but this assumption appears circular since the unknown signal $\mathbf{s}$ is itself required in order to estimate $\mathbf{S}$. Spitz (1999) proposes an estimate of the signal PEF in terms of a PEF, $\mathbf{D}$, estimated from the *data*, and a PEF, $\mathbf{N}$, estimated from a model of the noise.

  $$\mathbf{S} = \mathbf{DN}^{-1} \tag{4}$$

  $\mathbf{N}^{-1}$ is nonstationary *de*convolution, now better-defined in more than one dimension since the advent of the helix transform (Claerbout, 1998b). By conceptualizing equation (4) as a division of inverse spectra, Spitz' conjecture makes sense intuitively, that the unknown signal spectrum is the data spectrum scaled by the inverse of the noise spectrum. In the $t-x$ domain, we have seen heuristic proof of equation (4), and Sergey Fomel reports that a general proof is trivial in the $f-x$ domain.

- **Noise PEF**
  The assumption that the noise PEF $\mathbf{N}$ is known is also circular, since the noise is implicitly part of the unknown model. In many circumstances it is possible to explicitly model the noise well enough to estimate $\mathbf{N}$, like a convolutional model for multiples (i.e., the Delft SRME method of Verschuur et al. (1992)) Unforunately, variation in near surface earth properties makes it nearly impossible to explicitly model ground roll to cover all

observed cases from around the world, so it necessary to obtain the noise model directly from the data.

The noise model should be a first-order estimate of the noise - if it was perfect, the job would already be done. Ideally the noise model should contain the spatial correlation of all noise events, which will realistically differ in amplitude from the actual noise by a small, arbitrary time- and space-variant scale factor. Of course, a small amount of signal will contaminate the noise model, but as the ground roll will have much higher amplitude than any embedded signal, we assume that the L_2 PEF estimation of $\mathbf{N}$ will effectively ignore the embedded signal, though this is not always the case.

If the ground roll and signal are sufficiently well-separated in temporal frequency, application of a lowpass filter to the data may produce a satisfactory noise model. The degree to which the noise and signal are separated is dependent on many variables - geography, depth to target, geology, and source type, to name a few. Similarly, if a hyperbolic Radon transform focuses primaries well, then it may be possible to mute the primaries in $\tau - p$ space, then inverse transform to obtain a noise model. We have tested both approaches, and find that lowpass filtering is the most reliable, although this assertion is data-dependant. Clapp and Brown (1999) applied an L_1 iterative hyperbolic Radon transform to a multiple-infested 2-D seismic line, and found that the primaries were focused far better than for the usual L_2 least squares DRT.

Using Spitz' (1999) choice for the signal PEF, $\mathbf{S} = \mathbf{D}\mathbf{N}^{-1}$, rewrite equation (3)

$$\begin{aligned} \mathbf{Ns} &\approx \mathbf{Nd} \\ \epsilon \mathbf{D}\mathbf{N}^{-1}\mathbf{s} &\approx 0. \end{aligned} \tag{5}$$

Following Fomel et al. (1997), we precondition this iterative problem to improve performance. Make the change of variables

$$\mathbf{s} = (\mathbf{D}\mathbf{N}^{-1})^{-1}\mathbf{p} = \mathbf{N}\mathbf{D}^{-1}\mathbf{p}, \tag{6}$$

and rewrite equation (5):

$$\begin{aligned} \mathbf{N}\mathbf{N}\mathbf{D}^{-1}\mathbf{p} &\approx \mathbf{Nd} \\ \epsilon \mathbf{p} &\approx 0. \end{aligned} \tag{7}$$

This unconstrained least squares optimization can be solved using any gradient-based iterative technique, as in Claerbout (1998a).

RESULTS

We applied the technique to four 2-D receiver lines from a Middle East 3-D shot gather. One of these 2-D lines was shown earlier in Figure 2. The results are shown in Figures 3, 5, 6, and 7. Considering the low S/N, the extraction result is quite impressive. Many primary events which were hopelessly masked in the raw data become apparent after signal/noise separation.

As proof of the power of this method of separation, compare Figure 3 to Figure 4, where the latter is simply the result of subtracting the noise model from the data, or equivalently, lowcut filtering. The lowcut filter very effectively suppresses the ground roll, but severely damages the low frequency portion of the signal. The difference between the two results reinforces the fact that the predictive method can and will succeed with an imperfect noise model. Although the predictive method places some signal energy in the noise panel, the character of the signal seems preserved for the most part. Likely, any degradation of the signal within the noise cone is embedded signal which was *not* ignored in the estimation of $\mathbf{N}$.

DISCUSSION/CONCLUSIONS

Extension to 3-D

On our 4-processor SGI Origin 200, one normal-sized 2-D receiver line requires between 1 to 2 minutes of processing time, not including human interaction. For a full 3-D prestack survey, such performance is not damning, but some optimization is still desirable.

Since the spatial correlation of the events in a 3-D dataset are truly three dimensional, the best approach is to use three dimensional PEF's. Schwab (1998) notes that true 3-D PEF's more effectively whiten actual 3-D data than do ensembles of rotated 2-D filters. Events which appear incoherent in a 2-D slice may actually have hidden coherence in 3-D.

In 3-D, the way in which the data is sorted becomes an important issue. Among many different ways of sorting, two simple 3-D gathers for the cross-spread geometry (Figure 1) are common shot gathers (CSG) and common receiver-line gathers (CRLG). Define a CSG as the collection of records from all receiver lines in the survey for one shot, and similarly, a CRLG as the collection of records from one receiver line for all shots. Normally the shot spacing is far less than the receiver line spacing, so CRLG's are better sampled crossline than CSG's. This fact allows the use of fairly compact 3-D filters to predict the data in a CRLG, since the time delay between events is not large in the shot line direction. Still, the details involved in the choice of gather are strongly data-dependent, so these assertions are not exhaustive.

3-D prestack seismic data is five-dimensional, and redundant over one or more of these, depending on the survey design. A tantalizing possibility is to exploit this redundancy by reusing PEF's and predicted signal panels from nearby gathers as starting guesses.

Pros

- The predictive signal/noise separation method operates successfully, even when the noise is severely spatially aliased. The data need not be sampled finely in space to remove ground roll, a fact which is profound, considering the high cost of land-based seismic acquisition. Crawley (1998) shows another example of the feasibility of using PEF's to estimate aliased, coherent seismic signals.

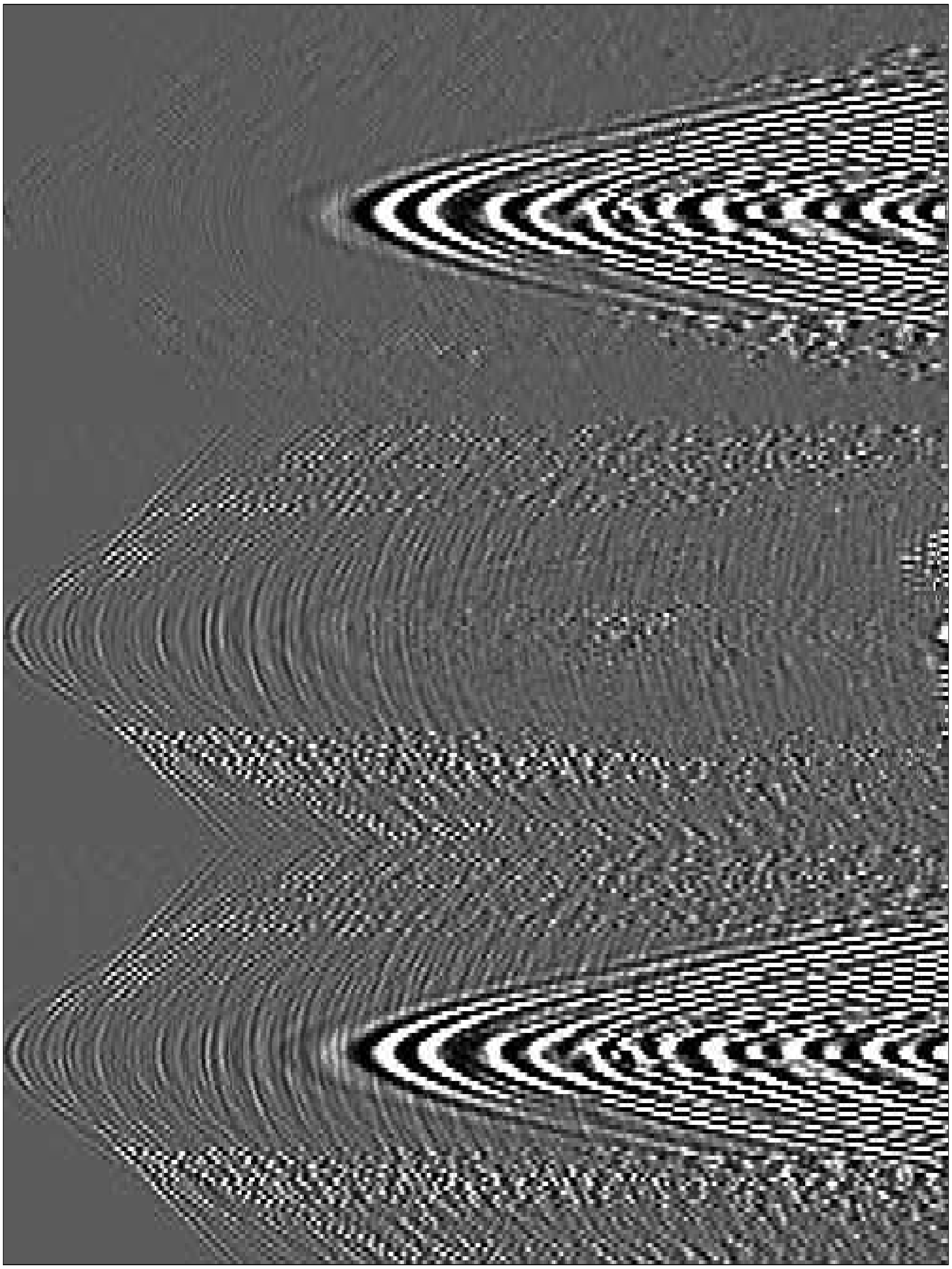

Figure 3: Receiver line #1 - predictive method results. From left to right: 1) Input data. 2) Predicted signal. 3) Difference, i.e., predicted noise. morgan1-dune-pred-daterr [ER]

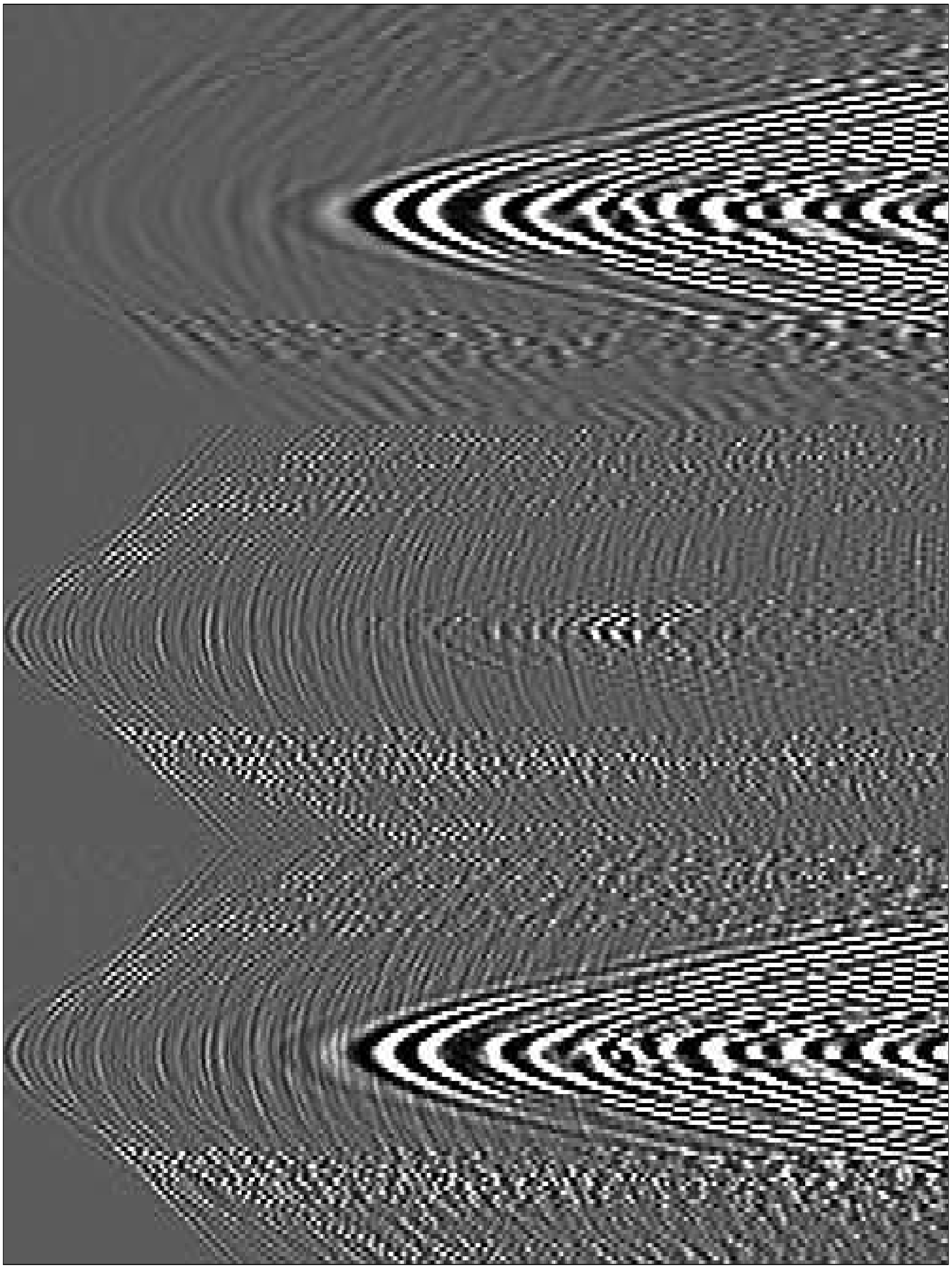

Figure 4: Receiver line #1 - highpass filter results. From left to right: 1) Input data. 2) Highpass (20 Hz cutoff) filter applied to data 3) Difference, i.e., noise model. morgan1-dune-locut-daterr [ER]

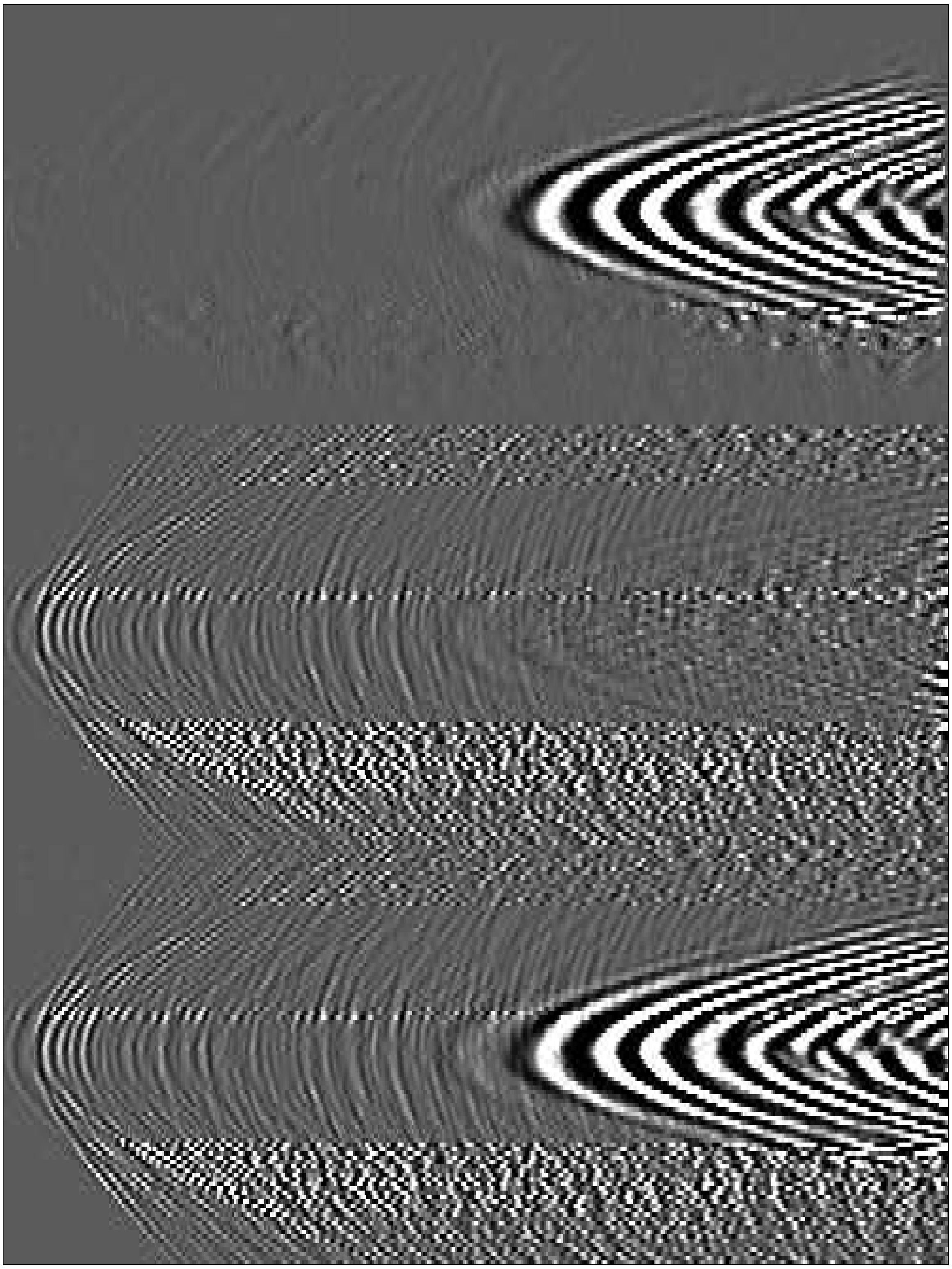

Figure 5: Receiver line #2 - predictive method results. From left to right: 1) Input data. 2) Predicted signal. 3) Difference, i.e., predicted noise. morgan1-dune1-pred-daterr [ER]

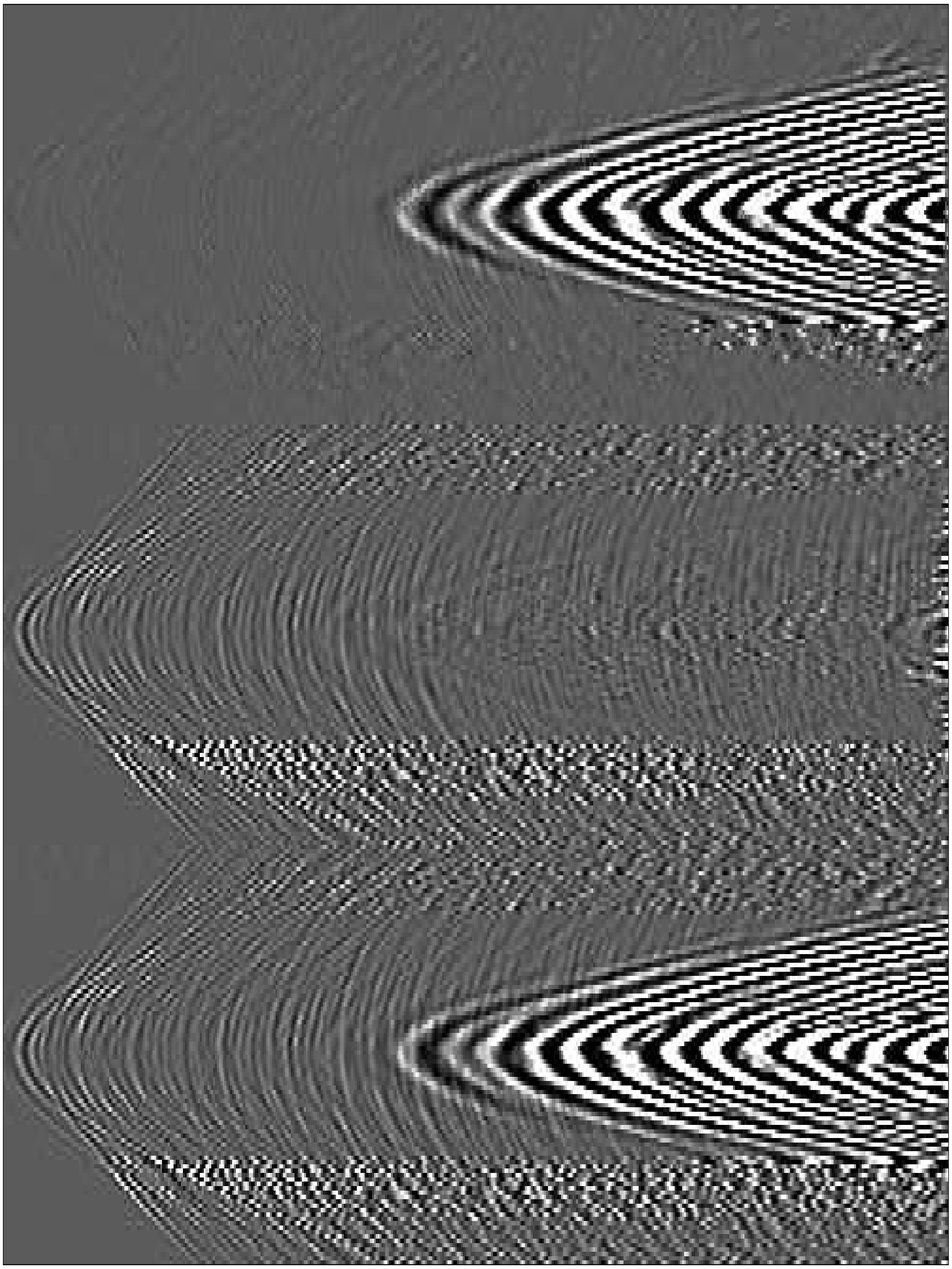

Figure 6: Receiver line #3 - predictive method results. From left to right: 1) Input data. 2) Predicted signal. 3) Difference, i.e., predicted noise. morgan1-dune2-pred-daterr [ER]

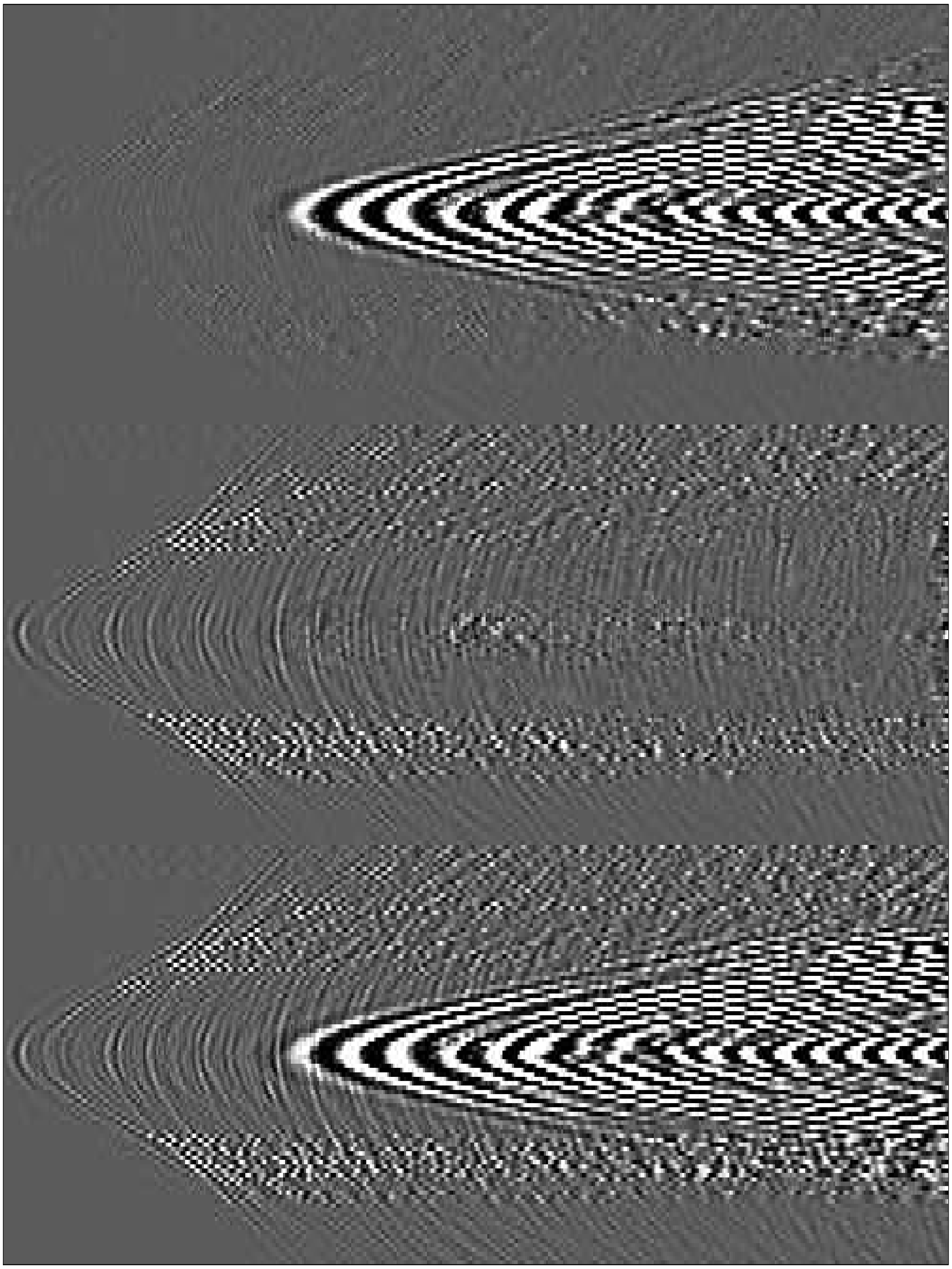

Figure 7: Receiver line #4 - predictive method results. From left to right: 1) Input data. 2) Predicted signal. 3) Difference, i.e., predicted noise. morgan1-dune3-pred-daterr [ER]

- Ground roll comes in many different flavors. A method which exploits the particular moveout patterns of one observed case of ground roll is sure to fail on others. The predictive method is blind to this; assuming a first order separation gives a viable noise model, the specific moveout patterns of signal and noise are irrelevant.

- Ground roll is almost always dispersive. The $f-x$-based algorithm of Spitz (1999) is computationally more efficient than our $t-x$ domain algorithm, but it requires a time-invariant seismic wavelet. Ground roll is usually a potpourri of different wave trains, all with different dispersion rates, strongly violating this temporal stationarity assumption. The nonstationary $t-x$ domain technique has no such limitation. Nonstationary filtering *is* possible in the frequency domain, but the distinct computational advantage over time domain methods is lost.

- An effective, general method for ground removal in 3-D may have profound effects on the future of seismic acquisition. If severe ground roll necessitates the recording of extremely long offsets or the use of large receiver arrays, the ability to remove the ground roll robustly from single-sensor data will cut acquisition costs considerably. Such a ground roll removal technique could greatly assist smaller-scale survey efforts (environmental or university research) which do not boast the same resources as a large, multinational oil company.

- Parenthetically, it should be emphasized that we have presented a method to do signal/noise *separation*, not simply noise removal. The output estimated noise may contain useful information, such as shear wave velocities. Multicomponent seismology is a promising specialty which would certainly benefit from any additional constraints provided by a robust "noise" extraction.

- Considerable effort is currently expended in pursuit of the perfect noise model, particularly in multiple suppression. However the results presented here show that predictive signal/noise separation produces good results with an imperfect noise model. An open question remains: might more primitive methods of obtaining a multiple model suffice if passed to a predictive signal/noise separation algorithm?

Cons:

- As discussed in Appendix A, we are concerned that the current approach may not fully account for correlation between signal and noise.

- While the nonstationary $t-x$ PEF estimation is more robust than stationary $f-x$ methods in terms of accurately predicting all the coherent events, the cost is considerably higher. As mentioned above, the method must be optimized before it becomes an industry standard.

- Our approach is parameter-intensive. Fortunately, in practical cases, the parameter choices are similar for most data gathers in the same survey.

ACKNOWLEDGEMENTS

Thanks to Sergey Fomel for writing much of the core nonstationary PEF code used for this algorithm, and also for insightful discussions.

REFERENCES

Abma, R., 1995, Least-squares separation of signal and noise with multidimensional filters: Ph.D. thesis, Stanford University.

Canales, L. L., 1984, Random noise reduction: 54th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, Session:S10.1.

Castleman, K. R., 1996, Digital image processing: Prentice-Hall.

Claerbout, J. F., 1976, Fundamentals of Geophysical Data Processing: Blackwell.

Claerbout, J. Geophysical estimation by example: Environmental soundings image enhancement:. http://sepwww.stanford.edu/sep/prof/, 1998.

Claerbout, J., 1998b, Multidimensional recursive filters via a helix: Geophysics, **63**, no. 5, 1532–1541.

Clapp, R. G., and Brown, M., 1999, Applying sep's latest tricks to the multiple suppression problem: SEP–**102**, 91–100.

Clapp, R. G., Fomel, S., Crawley, S., and Claerbout, J. F., 1999, Directional smoothing of non-stationary filters: SEP–**100**, 197–209.

Crawley, S., 1998, Shot interpolation for radon multiple suppression: SEP–**97**, 173–182.

Crawley, S., 1999, Interpolation with smoothly nonstationary prediction-error filters: SEP–**100**, 181–196.

Fomel, S., Clapp, R., and Claerbout, J., 1997, Missing data interpolation by recursive filter preconditioning: SEP–**95**, 15–25.

Larner, K., Chambers, R., Yang, M., Lynn, W., and Wai, W., 1983, Coherent noise in marine seismic data: Geophysics, **48**, no. 7, 854–886.

Leon-Garcia, A., 1994, Probability and random processes for electrical engineering: Addison-Wesley.

Li, X.-Y., and Mueller, M. C., 1997, Case studies of multicomponent seismic data for fracture characterization: Austin chalk examples, *in* Marfurt, K. J., Ed., Carbonate seismology: Soc. Expl. Geophys., 337–372.

Marfurt, K. J., Schneider, R. V., and Mueller, M. C., 1996, Pitfalls of using conventional and discrete Radon transforms on poorly sample data: Geophysics, **61**, no. 5, 1467–1482.

Mueller, M. C., 1992, Using shear waves to predict lateral variability in vertical fracture intensity: The Leading Edge, **11**, no. 2, 29–35.

Nemeth, T., 1996, Imaging and Filtering by Least Squares Migration: Ph.D. Thesis, University of Utah.

Regone, C. J., 1997, Measurement and identification of 3-d coherent noise generated from irregular surface carbonates, *in* Marfurt, K. J., Ed., Carbonate seismology: Soc. Expl. Geophys., 281–305.

Sbar, M. L., and Blankenship, C., 1998, Recoverable volume prediction from absolute acoustic impedance with errors: An example from the Nirvana, MC162 no. 3 well, Gulf of Mexico: Recoverable volume prediction from absolute acoustic impedance with errors: An example from the Nirvana, MC162 no. 3 well, Gulf of Mexico:, 68th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 910–913.

Schwab, M., 1998, Enhancement of discontinuities in seismic 3-D images using a Java estimation library: Ph.D. thesis, Stanford University.

Spitz, S., 1999, Pattern recognition, spatial predictability, and subtraction of multiple events: The Leading Edge, **18**, no. 1, 55–58.

Treitel, S., Shanks, J. L., and Frasier, C. W., 1967, Some aspects of fan filtering: Geophysics, **32**, no. 5, 789–800.

Verschuur, D. J., Berkhout, A. J., and Wapenaar, C. P. A., 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, no. 9, 1166–1177.

Wang, C.-S., 1997, 3-D ground roll suppression: 3-D ground roll suppression:, 67th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1043–1046.

APPENDIX A

OPTIMAL SIGNAL EXTRACTION

The frequency domain representation of the Wiener optimal reconstruction filter for uncorrelated signal and noise is (Castleman, 1996; Leon-Garcia, 1994)

$$\mathbf{H} = \frac{\mathbf{P_s}}{\mathbf{P_s} + \mathbf{P_n}}, \tag{A-1}$$

where $\mathbf{P_s}$ and $\mathbf{P_n}$ are the power spectra of the unknown signal and noise, respectively. Multiplication of $\mathbf{H}$ with the data spectrum extracts the signal spectrum which is optimal in the least squares sense. Abma (1995) solved a constrained least squares problem to separate signal from spatially uncorrelated noise.

$$\begin{aligned} \mathbf{Nn} &\approx 0 \\ \epsilon\mathbf{Ss} &\approx 0 \\ \text{subject to} &\leftrightarrow \mathbf{d} = \mathbf{s} + \mathbf{n} \end{aligned} \qquad \text{(A-2)}$$

which reduces to equation (2). Explicitly minimizing the quadratic objective function suggested by equations (A-2) or (2) leads to the following expression for the predicted signal:

$$\mathbf{s} = \left[\frac{\mathbf{N'N}}{\mathbf{N'N} + \epsilon^2 \mathbf{S'S}} \right] \mathbf{d} \qquad \text{(A-3)}$$

Recalling that the frequency response of the PEF is a smoothed approximation to the inverse spectrum of the data from which it was estimated

$$\begin{aligned} \mathbf{P_s} &\approx \mathcal{F}\left\{ \frac{\mathbf{1}}{\mathbf{S'S}} \right\} \\ \mathbf{P_n} &\approx \mathcal{F}\left\{ \frac{\mathbf{1}}{\mathbf{N'N}} \right\} \end{aligned} \qquad \text{(A-4)}$$

it is easy to show that the Wiener reconstruction result is equivalent to Abma's. Claerbout (1998a) uses this approach, and we extended a variation of it to obtain the results obtained in this paper. When spatially coherent events cross, as they do with ground roll and primaries, they are not uncorrelated. We believe that in order to maintain a high degree of rigor in our future formulation of this problem, the correlation between signal and noise should be accounted for. A more general form of the Wiener optimal reconstructor (Castleman, 1996; Leon-Garcia, 1994) is

$$\mathbf{H} = \frac{\mathbf{P_{ds}}}{\mathbf{P_d}}, \qquad \text{(A-5)}$$

where $\mathbf{P_{ds}}$ is the Fourier transform of the cross-correlation of the data and the unknown signal. Such a formulation is considerably less intuitive than equation (A-1), and we currently have no concrete ideas as to a starting point. Nemeth (1996) presents a more rigorous formulation for the separation of coherent noise and signal. As it is collected, the data is composed of overlapping signal and noise events, so the goal is to map the data to a domain where the signal and noise are uncorrelated, and thus separable without crosstalk. Nemeth's model is a composite vector, $[\mathbf{m_s}\ \mathbf{m_n}]^T$, consisting of the independent signal and noise model in the transformed (migrated) domain. His composite modeling operator, $[\mathbf{L_s}\ \mathbf{L_n}]$, is adjoint to migration, so his method accounts for the arbitrary moveout of real data - not the idealized hyperbolic moveout assumed for Radon-family transforms. The least squares inverse for Nemeth's model is

$$\begin{bmatrix} \mathbf{m}_s \\ \mathbf{m}_n \end{bmatrix} = \begin{bmatrix} \mathbf{L}_s^T\mathbf{L}_s & \mathbf{L}_s^T\mathbf{L}_n \\ \mathbf{L}_n^T\mathbf{L}_s & \mathbf{L}_n^T\mathbf{L}_n \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{L}_s^T \\ \mathbf{L}_n^T \end{bmatrix} \mathbf{d} \qquad \text{(A-6)}$$

In the context of prediction-error filtering, the "model" would be the residual of some PEF convolved with the data, such that the signal and noise are separated in the model space. The cross terms in Nemeth's "inverse model covariance matrix," $\mathbf{L}^T\mathbf{L}^{-1}$, account for correlation between signal and noise. In practice, Nemeth's method is weakened by the need for an explicit ground roll model. However, if Nemeth's migration operators are cleverly replaced with prediction error filters, a similarly powerful formulation could probably be derived, one which is free from the need to model ground roll explicitly. In any case, Spitz' (1999) choice of signal predictor, $\mathbf{S} = \mathbf{D}\mathbf{N}^{-1}$ gives good results, and probably handles the correlation between signal and noise correctly. Though we don't show the result in this paper, we have found that when the *actual* noise model is used (synthetic data), the estimated signal is nearly perfect.