GEOPHYSICS, VOL. 65, NO. 5 (SEPTEMBER-OCTOBER 2000); P. 1641–1653, 23 FIGS.

Noncausal spatial prediction filtering for random noise reduction on 3-D poststack data

Necati Gülünay*

ABSTRACT

A common practice in random noise reduction for 2-D data is to use pseudononcausal (PNC) 1-D prediction filters at each temporal frequency. A 1-D PNC filter is a filter that is forced to be two sided by placing a conjugate-reversed version of a 1-D causal filter in front of itself with a zero between the two. For 3-D data, a similar practice is to solve for two 2-D (causal) one-quadrant filters at each frequency slice. A 2-D PNC filter is formed by putting a conjugate flipped version of each quadrant filter in a quadrant opposite itself. The center sample of a 2-D PNC filter is zero. This paper suggests the use of 1-D and 2-D noncausal (NC) prediction filters instead of PNC filters for random noise attenuation, where an NC filter is a two-sided filter solved from one set of normal equations. The number of negative and positive lags in the NC filter is the same. The center sample of the filter is zero.

The NC prediction filters are more center loaded than PNC filters. They are conjugate symmetric as PNC filters. Also, NC filters are less sensitive than PNC filters to the size of the gate used in their derivation. They can handle amplitude variations along dip directions better than PNC filters. While a PNC prediction filter suppresses more random noise, it damages more signal. On the other hand, NC prediction filters preserve more of the signal and reject less noise for the same total filter length. For high S/N ratio data, a 2-D NC prediction filter preserves geologic features that do not vary in one of the spatial dimensions. In-line and cross-line vertical faults are also well preserved with such filters. When faults are obliquely oriented, the filter coefficients adapt to the fault. Spectral properties of PNC and NC filters are very similar.

INTRODUCTION

Certain prestack processes such as trace-by-trace spiking deconvolution or zero-phase spectral enhancement often increase the noise content of data. Generally, this noise is at the high end of the spectrum, and one might be tempted to apply a bandpass filter to remove it. However, at a given frequency such a filter would also attenuate the signal. This is because a bandpass filter operates on the summed amplitude of the signal and the noise and cannot determine how much of a given amplitude is signal and how much is noise. For this reason bandpass filters are often used only to view test data (such as deconvolution tests), and CMP stacking is used to filter out noise. However, when fold is low, the stacked sections are noisy and we need a tool that will separate signal from noise and attenuate random noise. Canales (1984) and Gülünay (1986) describe similar methods, known as f-x prediction filtering, that attenuate random noise on 2-D poststack data. These techniques use the (approximate) linearity of reflection times in a narrow space window and the finiteness of the number of linear events in a short temporal window of this space window. They operate in the temporal frequency domain (f). At each frequency, a spatial prediction filter is designed and applied to the input data along the spatial direction. Although the term f-x might imply a 2-D filter, f-x filters are one dimensional. Perhaps a better term would be x-domain filters (Hornbostel, 1991), but I use the former terminology since it refers to frequency as well as space.

Generally, f-x prediction techniques work in the forward direction (causal), where the present (noise-suppressed) sample is predicted from the previous (noisy) samples of the input data. Sometimes f-x prediction filtering is iterated to gain more noise suppression. That is, the filtered output of an iteration becomes input to the next one. A forward prediction filter smears and moves forward a spike (noise) by one sample. The filter interprets event truncations, e.g., faults, as noise. Therefore, there is danger of smearing faults toward the front (forward)

Manuscript received by the Editor September 8, 1997; revised manuscript received January 11, 2000.
*Western Geophysical, 10001 Richmond Ave., Houston, Texas 77042-4299. E-mail: necati.gulunay@westgeo.com.

direction, especially when many iterations of prediction filtering are used. To lessen this problem one can average forward and backward prediction outputs of raw data. A backward filter can be designed independently or it can be obtained by conjugating the forward filter if event amplitudes do not vary spatially for each dip.

Wang and West (1991) and Hornbostel (1991) use NC filters for random noise attenuation on stacked data. The NC filters are also known as interpolation operators (Claerbout, 1991) or smoothing operators (Robinson, 1967). If the sample at a constant temporal frequency that is being processed is called the present sample, then the samples to one side of it (spatially) could be considered past samples and the samples to the other side of it future samples. In causal filtering all data samples in the window are used to design the filter, and only past samples with respect to an output point are used to construct that output sample. In NC prediction filtering all data samples in the window are used to derive the filters, and past as well as future samples with respect to an output point are used to construct that output sample. Soubaras (1994) explains why an NC prediction filter used on poststack data should preserve signal better than a causal prediction filter.

There are some issues that noise elimination algorithms must address: Can the algorithm separate signal from noise accurately? Can the algorithm leave pure signal unaltered? After all, if a process cannot handle pure signal, we cannot expect it to handle noisy cases either. In the case of f-x prediction filtering, the answer to the first question is no. The f-x algorithms are known to distort signal levels significantly if the noise level is high (Spitz and Deschizeaux, 1994; Harris and White, 1997), and they are considered accurate only for low to moderate noise levels.

Signal preservation quality of a prediction filter applied to pure signal depends on the nature of the data as well as the algorithm. Factors affecting the performance of the process are the number of dipping events in the data window, the type of prediction filtering (forward, backward, and noncausal), and the filter length. As in image processing, there is no preferred direction (along the space axis) in poststack seismic data. Hence, prediction should be symmetrical (forward–backward). This argument leads to my use of NC filters in this paper. The type of data windowing used in the autocorrelation calculations is another factor affecting accuracy (Kay and Marple, 1981). One such windowing method, autocorrelation, assumes data to be zero outside the window of interest. Therefore, the number of contributors to a given autocorrelation lag decreases as the lag increases, introducing a bias to the autocorrelation because of the presence of a data window. Such autocorrelations are sometimes referred as pre- and postwindowed autocorrelations or windowed autocorrelations. I use biased autocorrelations because they provide stable matrix inversions (full rank) and are cheaper to invert (they are Hermitian–Toeplitz or block Toeplitz matrices). The same type of autocorrelations are used in FXDECON (Gülünay, 1986), and such autocorrelations are known to produce front-loaded, compact filters (Abma and Claerbout, 1995), which are useful in preserving faults. However, the bias in such autocorrelations can distort the signal estimate. Burg's (1975) maximum spectral entropy technique and its modification (Ulrych and Clayton, 1976; Nuttal, 1976; Kay and Marple, 1981; Tufts and Kumaresan, 1982; Spitz, 1991) are two well accepted algorithms which reduce distortions from windowing. Although the modified Burg algorithm (also known as the modified covariance method) has less distortion, filter design is more costly (because of the non-Toeplitz matrix structure). Furthermore, these filters are not front loaded and can smear faults if filter order is large. I show that NC prediction filtering, such as the modified covariance method, has no windowing effect when data consist of a single dipping event (with little or no noise). A single dipping event model is a good approximation for most seismic data when the data window is sufficiently small and NC prediction filters are compact (center loaded).

PNC AND NC PREDICTION FILTERS FOR 2-D DATA

Because dips are handled properly during filter design, there is no harm in moving events laterally with causal prediction filters. However, residues from unpredictable components, such as amplitude anomalies, move forward by one trace with each application of the causal prediction filter. Since these filters are sometimes applied a few times successively to increase their effectiveness, discontinuities such as faults in the input will move forward a few traces. This is clearly undesirable. Therefore, it is important to emulate a symmetric process in the space direction. One way of doing this is to process the input data in both the forward and reverse directions and then average the results.

Another way is to design prediction filters that have conjugate symmetry. This can be achieved from the forward (or reverse) prediction filter. First, a causal prediction filter of length L,

$$(p_1, p_2, \ldots, p_L), \tag{1}$$

is calculated from the autoregressive Yule–Walker normal equations (Marple, 1987). In the geophysical industry real-valued versions of these equations, known as normal equations, are used to solve for spiking deconvolution operators. The right side of the equations contains a spike at the first sample. Then a new prediction filter is formed by placing a conjugate reversed version of the causal filter in front of itself with a zero between the two:

$$0.5\left(p_L^*, \ldots, p_1^*, 0, p_1, \ldots, p_L\right), \tag{2}$$

where * indicates complex conjugation (see Galbraith, 1991). The prediction point is the center of the filter. Each half of the filter predicts the signal; therefore, 0.5 is used to average their outputs. I refer to such operators as PNC f-x prediction filters. Since an operator designed in the reverse direction is approximately equal to the conjugate of the operator designed in the forward direction (except when amplitudes vary along each dip), PNC f-x prediction filter design is similar to (but not the same as) the averaging method mentioned above.

A conjugate symmetric prediction filter can also be obtained using two-sided (therefore NC) normal equations. Berkhout (1977) uses two-sided (real-valued) filters in wavelet deconvolution of seismic traces and proves desirable properties of the two-sided filters. Wang and West (1991) use such filters in f-x prediction. Soubaras (1990) shows the form of the normal equations needed for NC deconvolution filters and identifies it as a symmetrical Yule–Walker system. Hornbostel (1991) uses NC prediction filters for f-x prediction filtering of seismic data. Electrical engineers have been using 2-D NC prediction

filters in image processing for awhile (Jain, 1981; Xu and Azimi-Sadjadi, 1993).

For a 1-D input sequence containing L linear events, a causal filter of length L is necessary. This filter needs L autocorrelation lags (r_0 through r_{L-1}). An NC filter of length $2L+1$ is necessary to predict the same data. Such a filter needs $2L+1$ autocorrelation lags (r_0 to r_{2L}). Therefore, NC filter design involves nearly twice as many lags as causal filter design. The matrix in the NC filter equations (of size $2L+1$ by $2L+1$) is Hermitian–Toeplitz and positive definite. Furthermore, the matrix is main diagonal heavy (because of the windowed nature of the autocorrelation lags). The solution, the NC prediction error filter, has conjugate symmetry around the center element. This can be shown easily by

1) rewriting the original equations in reverse (row) order,
2) writing the original equations for the reversed filter,
3) conjugating step 2 to obtain a set of equations for the conjugate reverse filter, and
4) comparing the results of steps 1 and 3.

Knowing that the NC prediction error filter has conjugate symmetry, it is possible to reduce the order of the normal equations from $2L+1$ to $L+1$. It is difficult to derive general expressions of the reduced system of equations for large L. However, this can be done easily for small L. For example, when $L=1$, the normal equations (i.e., symmetric Yule–Walker equations) for the NC prediction error filter, e, are

$$\begin{bmatrix} r_0 & r_1^* & r_2^* \\ r_1 & r_0 & r_1^* \\ r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} e_{-1} \\ e_0 \\ e_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}. \tag{3}$$

They can be reduced to a simpler form, where e_0, e_1, and e_{-1} are given by

$$\begin{bmatrix} r_0 - \dfrac{|r_1|^2}{r_0} & r_1^* - \dfrac{r_1 r_2^*}{r_0} \\ r_1 - \dfrac{r_1^* r_2}{r_0} & r_0 - \dfrac{|r_2|^2}{r_0} \end{bmatrix} \begin{bmatrix} e_0 \\ e_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad e_{-1} = e_1^* \tag{4}$$

Therefore, the three-point NC prediction filter, p, is

$$\begin{aligned} p_{-1} &= p_1^* \\ p_0 &= 0 \\ p_1 &= -\frac{e_1}{e_0} = \frac{r_1 r_0 - r_2 r_1^*}{r_0^2 - |r_2^2|} \end{aligned} \tag{5}$$

For example, if the input sequence has N samples and is made of one event with a time stepup per trace of Δt, then

$$\begin{aligned} r_0 &= N \\ r_1 &= (N-1)z_0, \\ r_2 &= (N-2)z_0^2 \end{aligned} \tag{6}$$

where $z_0 = e^{j\omega\Delta t}$. This leads to the three-point NC prediction filter

$$\left(0.5z_0^*, 0, 0.5z_0\right). \tag{7}$$

This solution is different from the PNC filter of the same length. For a PNC filter of length 3, one needs to solve the normal equations for a causal filter—order one in this case. This leads to the causal filter

$$p_1 = \frac{r_1}{r_0}. \tag{8}$$

Therefore, the PNC filter is

$$\left(0.5\frac{r_1^*}{r_0}, 0, 0.5\frac{r_1}{r_0}\right). \tag{9}$$

Compare this result with equation (5). This result is close but not the same to that of the NC filter. For the numerical N-sample example given above, the PNC prediction filter is

$$\left(0.5\frac{N-1}{N}z_0^*, 0, 0.5\frac{N-1}{N}z_0\right). \tag{10}$$

This PNC prediction filter has an error $100/N\%$ on all output samples (except the edges), while the true NC filter in equation (7) has zero error everywhere (except at the edges). That is, the causal prediction filter or the PNC prediction filter obtained from it cannot predict a single event exactly even in the absence of noise. Furthermore, the error is larger for smaller space gates. This property is observed by others as well (Gülünay, 1986; Harris and White, 1997). Gülünay (1986) uses the complex Wiener filter method outlined by Treitel (1974) but keeps one extra sample for the crosscorrelation to eliminate the amplitude error caused by the biased autocorrelations. That remedy works only for a single dipping event. An NC design takes care of the windowing effect for a single dipping event.

Another benefit in using an NC filter is its ability to predict amplitude variations along space directions. The basic assumption of the causal f-x prediction is that each dipping event has constant amplitude along the space direction. Causal prediction filters can handle amplitude variations if they are in the form of decay in the forward direction (minimum phase). Similarly, causal filters can handle amplitude buildup if prediction is done in the reverse direction. NC f-x prediction filters, on the other hand, can predict such amplitude variations regardless of the direction of the decay (in the absence of noise). Therefore, for pure signal, NC f-x prediction filters preserve the amplitude levels of the events when the amplitude varies smoothly in the space direction. For example, a dipping event with exponential amplitude variation in the space direction can be exactly predictable (in the absence of noise) with an NC f-x filter but not with a PNC f-x filter (see Appendix A).

Figure 1 shows a dipping event with a spatially varying amplitude $e^{-0.05(n-1)}$ where $n = 1, 2, \ldots, 20$ is the trace count. The three-point NC f-x and PNC f-x filters are compared in Figure 2 by plotting the amplitudes of prediction filter outputs along the event and the corresponding errors. Curves at the top are the input, NC filter output, and PNC filter output. The bottom two curves are for the prediction errors. The NC f-x prediction filter predicts such data with almost no errors except at the edges. The PNC f-x filter, on the other hand, has an overall loss of amplitude (about 8%). Considering that a seismic section can have lateral amplitude variations, an NC f-x filter should be superior to a PNC filter.

PNC AND NC PREDICTION FILTERS FOR 3-D DATA

Multichannel least-squares filters have been used in the geophysical industry for awhile. In such applications the channels are generally seismic traces. Because a frequency slice of 3-D data is an image (a function of x, y-coordinates), image processing algorithms are readily applicable. A number of papers discuss 2-D prediction filters (Van Valkenburg, 1986; Jain, 1981). Such filters can be designed as causal, semicausal, or NC (see Figure 3). In image processing articles, the data points that the filter uses to construct an output sample are referred to as the region of support (Huang, 1981) or mask.

In the NC prediction filter design, full-plane finite support is used. The center lag of the prediction filter is zero, and every output point is predicted by using a mesh of points around it excluding itself. The filter dimensions can be chosen to be odd in each direction so the number of data points that the output is constructed from is equal in the forward and reverse directions. One set of normal equations is used to determine all forward and backward coefficients of the 2-D NC filter. The resultant 2-D NC prediction error filter and the prediction filter have conjugate symmetry around the sample being predicted. As in 1-D filters, it is possible to emulate a 2-D NC filter using 2-D causal filters (quarter-plane) by conjugate symmetric composition. For this process four quarter-plane filters or two half-plane filters are used. In the case of four quarter-planes, two of the planes can be obtained from the other two by conjugation if the data do not decay in the space direction. I call these 2-D PNC filters.

Chase (1992) extends f-x prediction filtering to three dimensions. In prediction filtering of poststack data volumes, he designs and applies 2-D prediction filters in the plane defined by the in-line (x) and cross-line (y) directions for each temporal frequency slice of the 3-D data volume—a PNC prediction filter obtained from half-planes. Following Gülünay et al. (1993), I propose using full-plane NC prediction filters to process each frequency slice of the 3-D data because NC prediction filters have some desirable qualities.

As in 1-D filter design, I use biased (windowed) 2-D autocorrelations in designing 2-D NC filters. Although this type of lag estimate, known as the autocorrelation method, biases operator calculation to some extent, it brings stability and robustness to the normal equations. Other methods that make different assumptions about data outside the data window while calculating the autocorrelation coefficients or while minimizing the error on the prediction are also possible. In the 1-D filter case, the autocorrelation method leads to a Hermitian–Toeplitz matrix structure, which allows fast solutions through Levinson

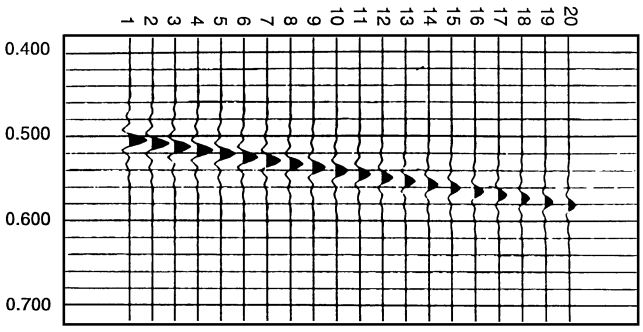

FIG. 1. A dipping event with an exponential decay in amplitudes. Wavelets on the traces are identical. The scale of the traces (peak wavelet amplitude) drops exponentially from left to right.

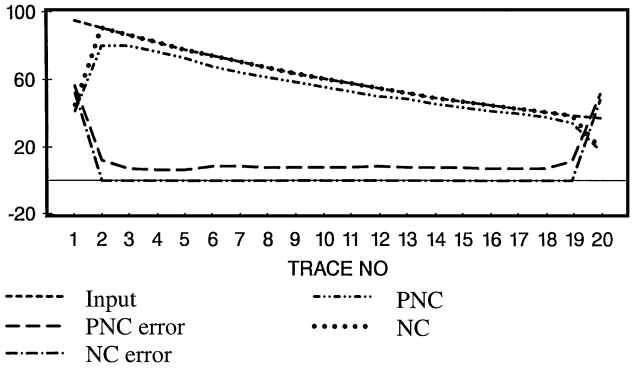

FIG. 2. Peak amplitudes on the exponentially varying data of Figure 1. Top three curves are for input, the PNC prediction filter output, and the NC prediction filter output. Bottom two are for difference amplitudes (prediction errors) for PNC and NC prediction filters.

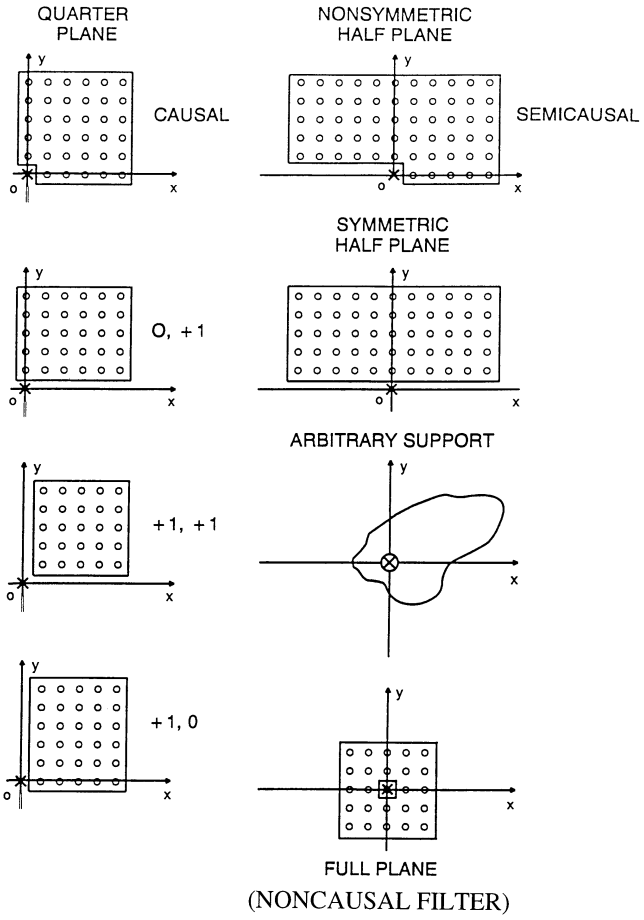

FIG. 3. Two-dimensional prediction filters. The location filter points in relation to the output point (i.e., x-y origin) determines the type of a filter.

recursion (Makhoul, 1975; Kay and Marple, 1981). For 2-D filters, a block matrix is obtained. The matrix is not Toeplitz, but each block is Toeplitz.

In 1-D prediction the one-step advance is unique. There is more than one way of obtaining a one-step advance in 2-D space: one step in $+x$ or $+y$ or one step in both $+x$ and $+y$. Therefore, it is simpler to devise the normal equations for a prediction error filter than for a prediction filter. As in causal prediction error filtering, the output of the NC prediction error filter must have minimum energy. This requirement leads to the normal equations (Jain, 1981) with a vector whose right side is all zeroes except at the middle sample. The matrix on the left side is formed from the autocorrelation coefficients.

The normal equations of the NC filter for 3-D data can be solved by standard means, such as the Gauss–Jordan technique for linear systems, after putting them into a real form as described by Treitel (1974). One may also use more efficient algorithms that use the block Toeplitz nature of the matrix being implicitly inverted. Once the normal equations are solved for the NC prediction error filter, the corresponding prediction filter can be obtained by normalizing the solution, reversing the sign, and zeroing the middle sample (zero lag). The 2-D NC prediction error filter and the 2-D NC prediction filter have conjugate symmetry around the origin. This same symmetry is exhibited by the complex 2-D autocorrelations. All the negative lags of the prediction filter can be obtained from its positive lags, but I use one set of normal equations to solve for all the lags simultaneously, as Jain (1981) does for the noncausal filters. An example of the 2-D NC prediction filters used in this paper is given in Appendix B.

BENEFITS OBTAINED FROM NC DESIGN

In 1-D prediction, L_x dipping events in the t-x (time–space) domain need L_x+1 prediction error filter coefficients with the causal approach and $2L_x+1$ prediction error filter coefficients with the NC approach. In 2-D prediction, if L_x+1 and L_y+1 are the dimensions of a one-quadrant causal prediction error filter, then $(L_x+1)(L_y+1)$ coefficients must be solved for a causal prediction error filter. On the other hand, the dimensions of the NC filter are $(2L_x+1)$ and $(2L_y+1)$. Appendix B gives a numerical example of a 3×3 filter for a single dipping event. The dimensions of a 2-D causal prediction error filter for the same event would have been 2×2. Therefore, a 2-D NC filter is costlier to design than a 2-D causal (or 2-D PNC) filter. One may naturally ask what benefits are obtained for such cost. The examples given by Chase (1992) show some smear across faults. I attribute this to the PNC design of the filters. A 2-D NC prediction filter, on the other hand, can be compact yet preserve many geological features well. For example, as long as geology varies along only one axis (x or y) and when an NC filter is used, even the shortest filter, such as 3×3 ($L_x=1$, $L_y=1$), can predict 2.5-D geology with zero prediction error in the absence of noise. This property of the surfaces may be termed exact predictability.

Structures that do not vary in one of the two space directions are exactly predictable; data correlate along the no-change direction perfectly, and the 2-D filter pulls most of its information from data in that direction. When the input data vary rapidly in one direction, the prediction will fail if the length of the 1-D filter is small and the prediction is made in the direction of change. On the other hand, a 3×3 NC filter predicts 2.5-D structures regardless of the amount of change in one of the directions.

A 3×3 NC filter passes vertical faults or lateral amplitude drops (oriented along the x- or y-direction) with no distortion. Figure 4 shows how such a filter would preserve an amplitude change, such as a drop from amplitude A to amplitude B (zero dip assumed for simplicity), oriented along one of the x- or y-directions. For this problem the resulting 3×3 prediction filter can be shown to be (−0.25, 0.50, −025, 0.5, 0, 0.5, −0.25, 0.5, −0.25), which is independent of A or B. This filter outputs an amplitude of A when it is centered on the left side of the fault and an amplitude of B when it is centered on the right side of the fault. When the filter is centered at the edge of the fault where the amplitude is A, its output is also A. Therefore, this filter preserves such lateral amplitude drops.

If the faults are oriented in an oblique direction in the x–y-plane, a longer filter, such as 5×5, is needed. The

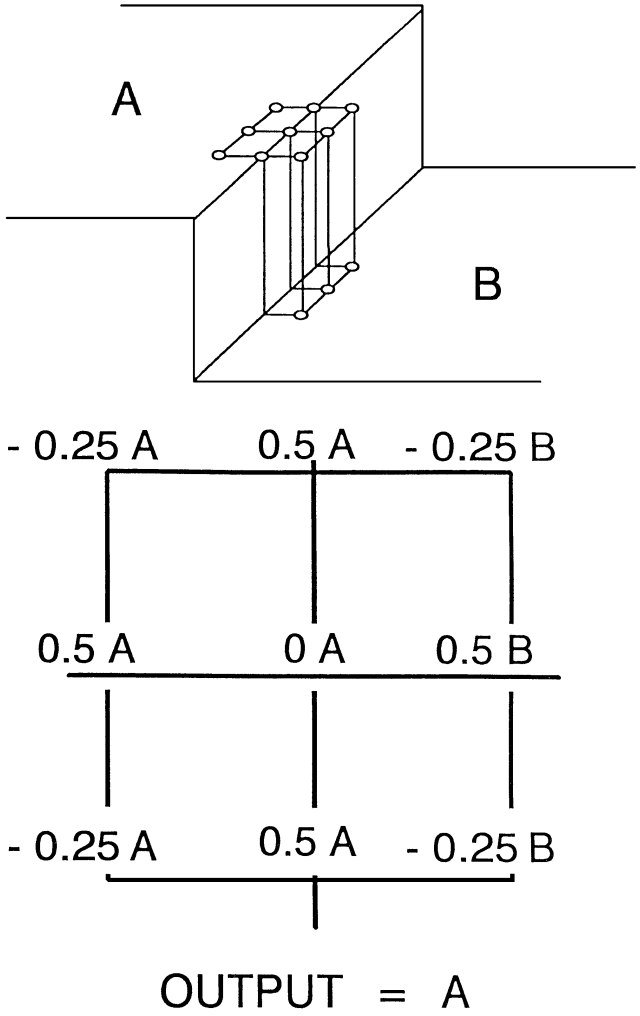

FIG. 4. Performance of a 3×3 NC f-xy prediction over an overamplitude drop from A to B in the frequency slice. The filter coefficients do not sense the presence of the fault. This filter produces amplitude A when the output point is to the left of the edge and when it is at the edge but produces amplitude B when the output point is to the right of the edge.

compactness of the NC f-xy filter is illustrated with such a fault in Figure 5. On the left side, an amplitude discontinuity runs parallel to the x-axis. The data have no noise, and a 7×7 filter was designed. The magnitude of the filter coefficients are plotted underneath using characters to indicate relative amplitudes (character 1 for amplitudes between 10% and 20%, character 2 for amplitudes between 20% and 30%, . . . , character 9 for amplitudes between 90% and 100%, and character * for peak amplitude). The significant coefficients of the filter occupy a 3×3 area instead of a 7×7 area, making the filter compact (center loaded) and making the output of higher spatial resolution than it would have been otherwise. In other words, specifying filter length larger than necessary does not alter the filter significantly. The right side of Figure 5 shows the same amplitude discontinuity after rotating it by 45°. The significant coefficients of the filter are now along the fault. Therefore, the filter will preserve the spatial resolution of the sections even for an oblique structure.

In all of the 2-D filter examples given below, I use equal length in the in-line and cross-line directions ($L_x = L_y$). In cases, the in-line and cross-line trace spacing may differ, and one may wish to have a square filter in geographical terms. This issue is not covered in this paper.

MODEL AND FIELD-DATA EXAMPLES FOR 2-D NC PREDICTION FILTERS FOR 3-D DATA

To study the behavior of the NC filter in a more complex setting, a 3-D model containing five diffractors at positions A, B, C, D, and E in Figure 6 was constructed. There are 120 lines, 120 CDPs in each line, and the data length is 2 s. Uniformly distributed random noise was added with noise amplitude 25% of the peak signal. A time slice at 1400 ms is shown in Figure 7. It shows four of the five circular wavefronts (the fifth one arrived earlier). Figure 8 is the NC f-xy prediction filter output. The size of the space gate was 20×20, and the size of the temporal gate was 500 ms. The operator size of the filter was chosen to be 5×5 ($L_x = 2$, $L_y = 2$). A small amount of noise still remains on the output, but the filtering has been very successful. The difference between the input and output of the process is shown in Figure 9 for the same time slice. The signal is only barely visible in the noise section. One can conclude from these figures that the NC f-xy filter preserves the wavefronts well. A corresponding difference section (input minus filtered output) after runs with a 5×5 PNC filter (constructed from four quadrants of 3×3 prediction error filters) is shown in Figure 10. The PNC prediction filter (not shown) suppresses more noise than the NC prediction filter does. However, Figures 9 and 10 show that the PNC filter rejects more signal than the NC filter does.

A land 3-D data cube is shown in Figure 11. The data cube of the f-xy NC filter output is shown in Figure 12, and the difference shown in Figure 13. The space gate is 64×64 CDPs. The size of the temporal gate is 1024 ms. The size of the NC filter is 5×5. Only a small amount of signal leaks to the noise cube of the f-xy process, giving confidence in the process. Also, steeply dipping weak coherent noise is well preserved by the process.

SPECTRAL PROPERTIES OF THE NOISE OUTPUT

The output of the prediction error filter represents the noise rejected from the input data and thus should contain no signal. In the model example above, the difference section of the NC prediction filters looks less coherent than the difference section of the PNC filter. For the NC prediction error filter, the noise output does not correlate with the input, but the noise spectrum is not white (Jain, 1989; Claerbout, 1991). The example to be given in this section suggests nonwhiteness is not a major issue. The spectrum of the NC filter is very much like the spectrum of the PNC filter for short filter lengths. Note also that the noise output of a finite-length causal prediction error filter will not

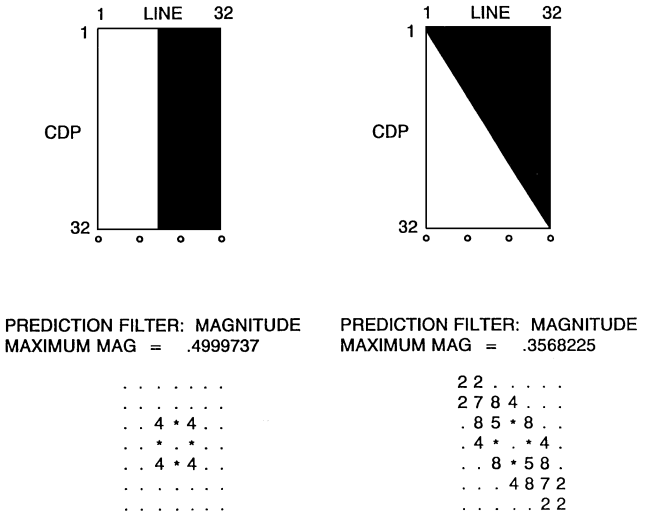

FIG. 5. A 7×7 NC f-xy prediction filter on two types of faults: a cross-line fault (left) and a 45° oriented fault (right).

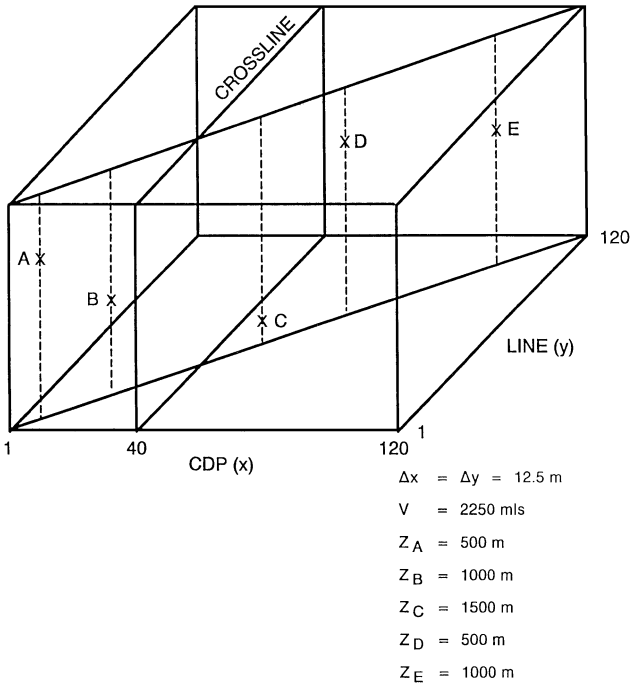

FIG. 6. The hyperboloid model with five point diffractors: A, B, C, D, and E, which are on CDPs 6, 26, 56, 76, and 106, respectively.

be white either since the input sequence of the f-x prediction process is not autoregressive (Kay and Marple, 1981; Soubaras, 1994).

In this section, I compare the average k-spectra (magnitude of the Fourier transform samples averaged over all temporal frequencies) for the noise output (i.e., the difference section) of the NC and PNC f-x prediction filters for flat events, buried in random noise. Figure 14 shows the flat events after adding random noise. The flat events had a white temporal spectrum before noise addition. Figure 15 is the result of PNC f-x prediction filtering with $L=1$ (length of PNC filter is 3). Spatial windows with some overlaps were used in processing the 100 CMPs shown (the space window size was 30 traces and the time window size was equal to the data length). The difference section for the PNC filter is shown in Figure 16. This section has horizontal streaks, indicating a significant amount of signal was rejected. Figures 17 and 18 are the corresponding output and difference sections of the NC filter, with $L=1$ (length of NC filter is 3). It is hard to identify visually any rejected signal in the difference section of the NC prediction filter (it is present, however, as will be seen).

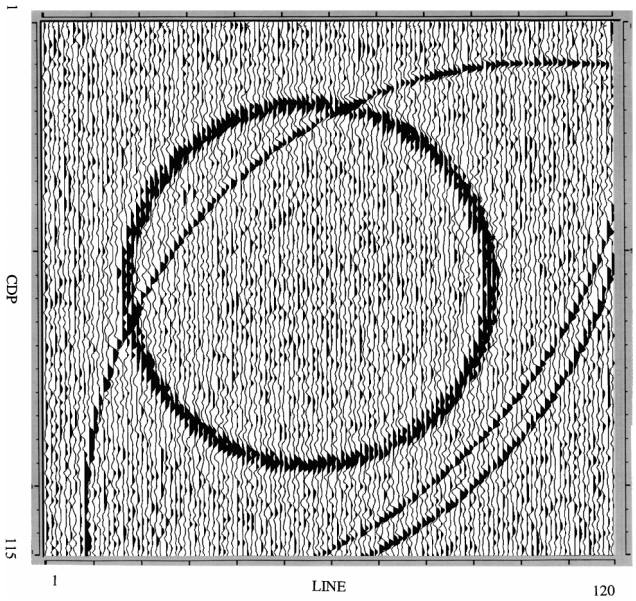

FIG. 7. The time slice at 1400 ms from the hyperboloid model after noise addition.

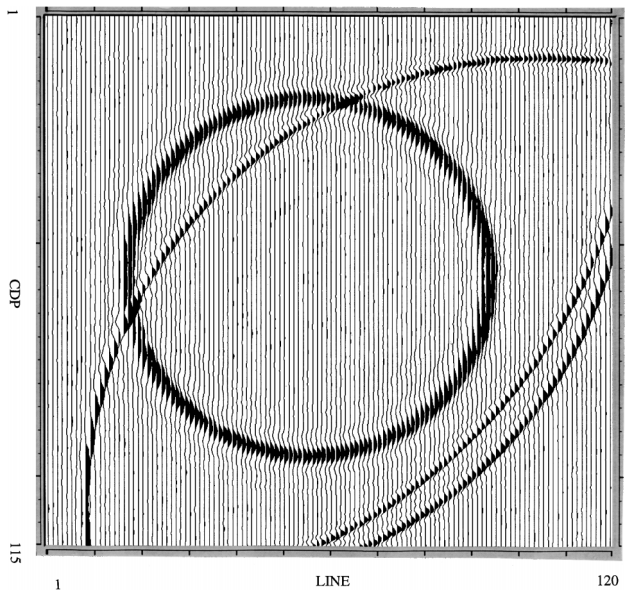

FIG. 8. The time slice from the output of an NC f-xy prediction filter.

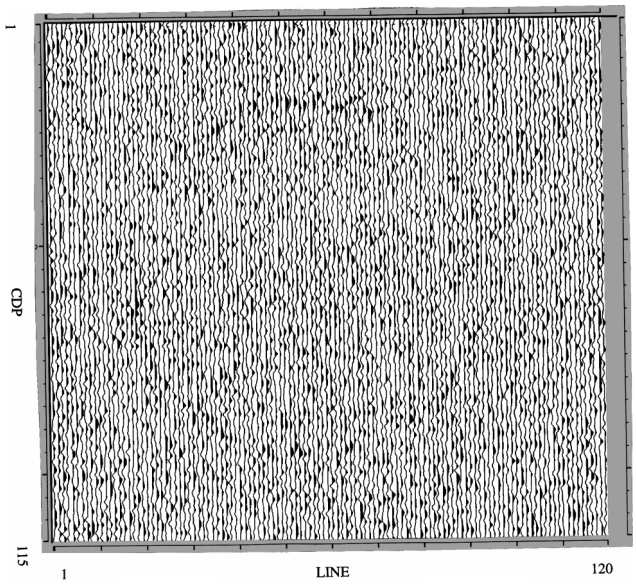

FIG. 9. The time slice from the difference section for the NC f-xy prediction filter.

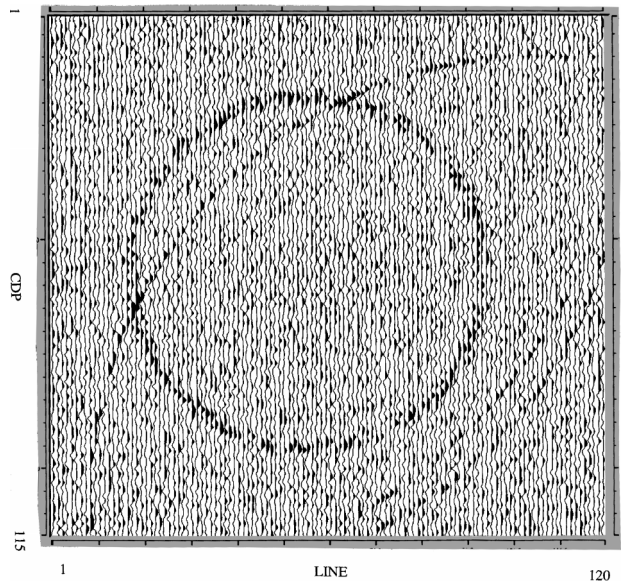

FIG. 10. The same time slice from the difference section for the PNC f-xy prediction filter.

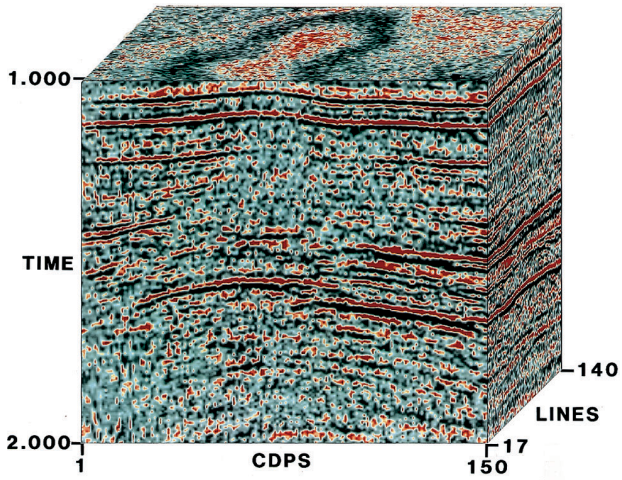

FIG. 11. A data cube from a noisy 3-D land survey.

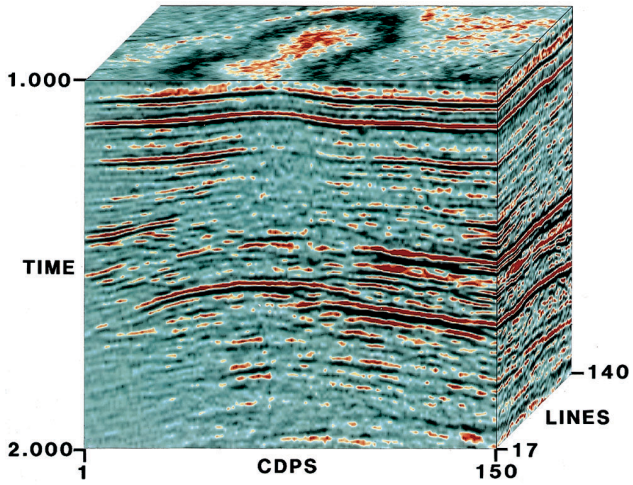

FIG. 12. The same cube in Figure 11 after NC f-xy prediction filtering. Steeply dipping coherent energy (noise in this case) is preserved.

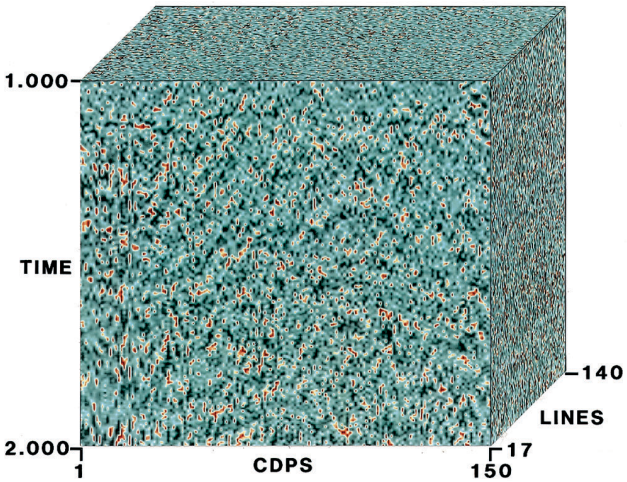

FIG. 13. The difference cube for the NC f-xy prediction filtering. No significant amount of coherent energy exists in this output.

Figure 19 shows the average k-spectra (magnitudes averaged over temporal frequencies) for the true input noise and for the noise found by the PNC and the NC prediction filters (Figures 16 and 18, respectively). A higher value (at $k=0$) than the actual noise in the data means that signal damage has occurred. The signal's average value at $k=0$ was about 7000. While both filters damage a signal, the NC prediction filter

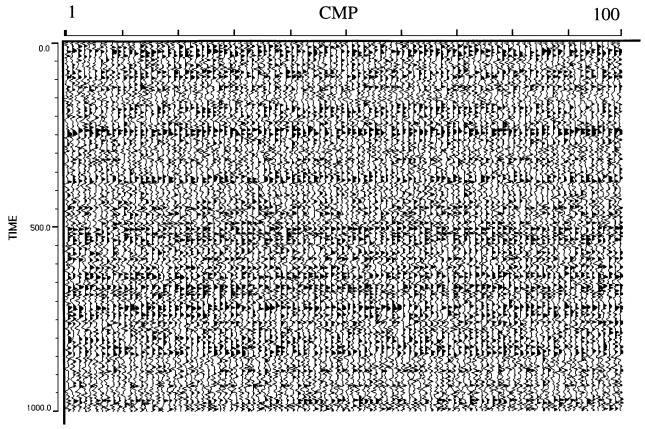

FIG. 14. A noisy synthetic containing flat events.

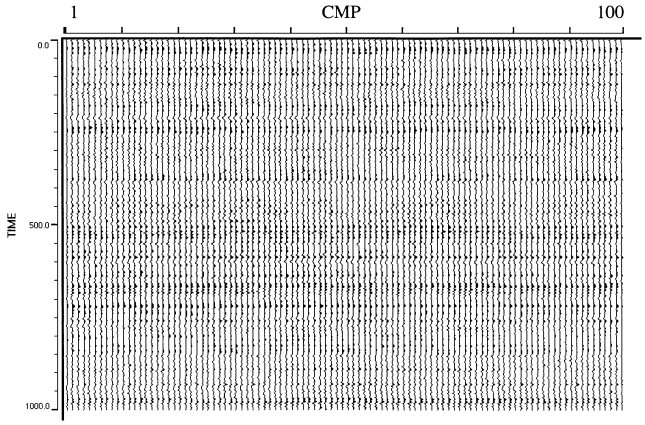

FIG. 15. The output of the three-point ($L=1$) PNC f-x prediction filtering. Low amplitude at the edges results from the two-sided filter not having enough live input samples.

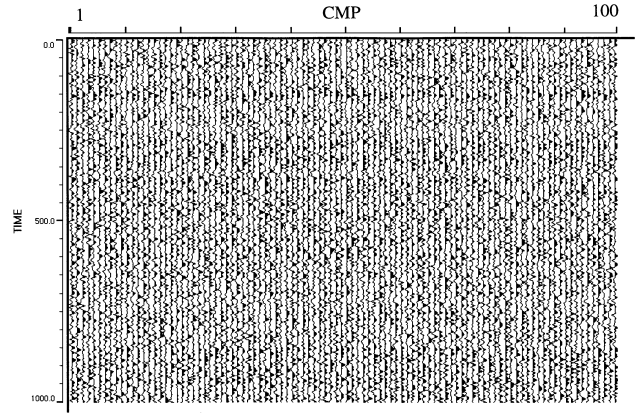

FIG. 16. The difference section of the three-point PNC f-x prediction filtering.

does less damage (about 2500/7000) than the PNC filter (about 3500/7000). Note the nonflat spectral shape of the noise outputs. Both filters seem to take out more energy at and around $k = k_N$ (Nyquist) than needed. The shape of this response is very much the same as the spectral response of a three-point filter $(-0.5, 1, -0.5)$. At low k-values where there is signal, both filters seems to reject less noise than they should have. The NC filter rejects less noise at $k = 0$ than the PNC filter does. It also rejects more noise around $k = k_N$ than the PNC filter does. Figures 20 and 21 show similar comparisons for longer filters ($L = 2$ and $L = 3$). The value $L = 2$ is still far from being spectrally correct, but $L = 3$ is getting close to the desired spectrum, suggesting that filter length is a more important factor than filter type. Also, signal leakage lessens as filter length increases. Figure 22 displays the average spectra at $k = 0$ versus L for both NC and PNC filters. The NC filter rejects less signal than the PNC filter for the same filter length. If one displays k-spectra for the filtered output rather than the rejected noise, one sees that the PNC filter does a better job suppressing noise than the NC filter does (for the same L). The noise level in the NC output is about 50% higher than the noise level in the PNC output. This is perhaps because of the fewer degrees of freedom the PNC filter has compared with the NC filter (the NC filter uses $2L + 1$ autocorrelations, while the PNC filter uses only $L + 1$). The relative performance of the two filters is reversed, however, in regard to the signal. Figure 23 plots $k = 0$ values, the plot of signal preservation. The desired level is about 7000 for the filter output. The NC filter does a better job than the PNC filter (for the same L) in preserving the signal.

The spectral comparisons made in Figures 19 through 23 are for 1-D filters used in f-x prediction filtering. The 2-D PNC filter design (obtained from four quarters) used in f-xy prediction filtering seem to behave in a similar way. For small space gates 2-D PNC filters reject more signal than 2-D NC filters do, whereas 2-D PNC filters suppress more random noise than 2-D NC filters do.

CONCLUSIONS

Noncausal f-x (or f-xy) prediction filters may be used for prediction filtering of 2-D (or 3-D) data at each temporal frequency. These filters have less prediction error than their PNC counterparts. These filters can be compact. For example, a 3×3 NC filter ($L_x = 1$, $L_y = 1$) can predict geology that does not vary in one direction. Slightly longer filters, such as 5×5 ($L_x = 2$, $L_y = 2$), are needed when faults lie in an oblique direction. The windowed nature of the autocorrelations used in NC filter design leads to stable filters. For data with high S/N

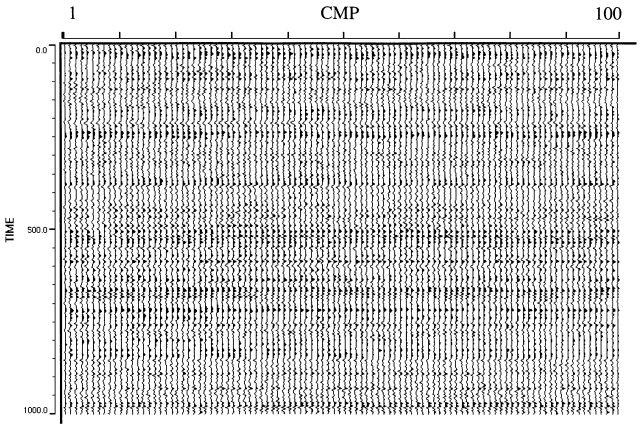

FIG. 17. The output of the three-point ($L = 1$) NC f-x prediction filtering. Low amplitude at the edges results from the two-sided filter not having enough live input samples.

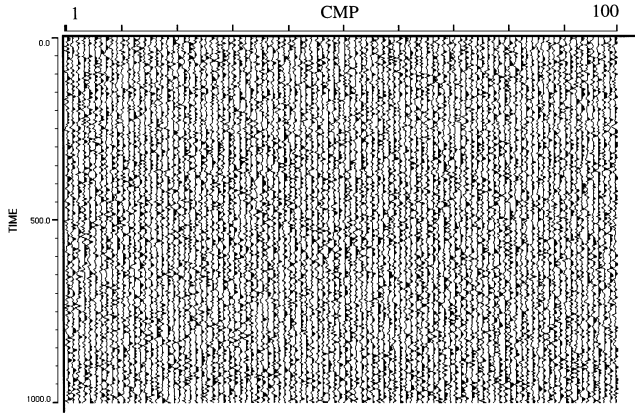

FIG. 18. The difference section of the three-point NC f-x prediction filtering.

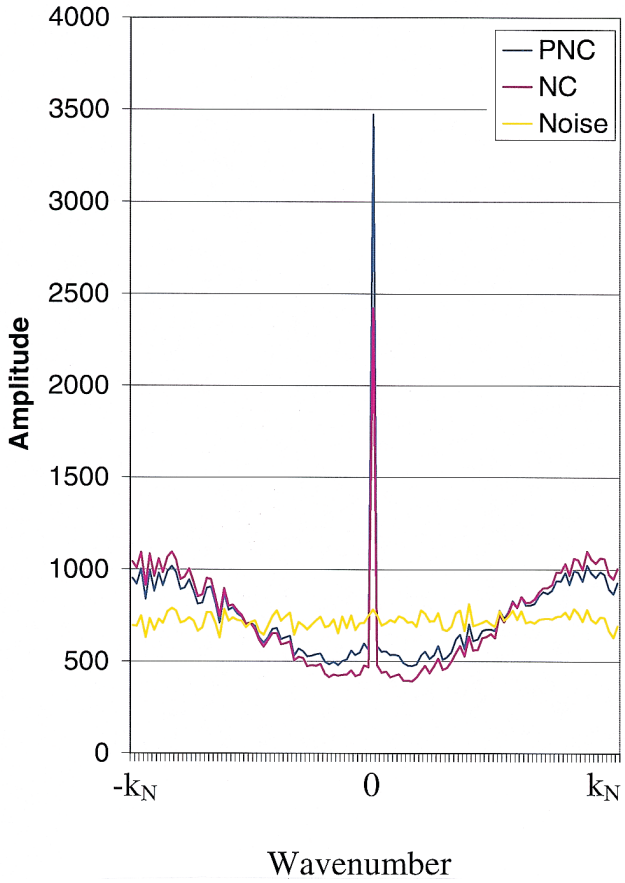

FIG. 19. The average wavenumber spectra for the difference sections for $L = 1$ runs. Averaging is done over temporal frequencies. The spike at $k = 0$ indicates that difference sections contain coherent energy pointing to signal damage.

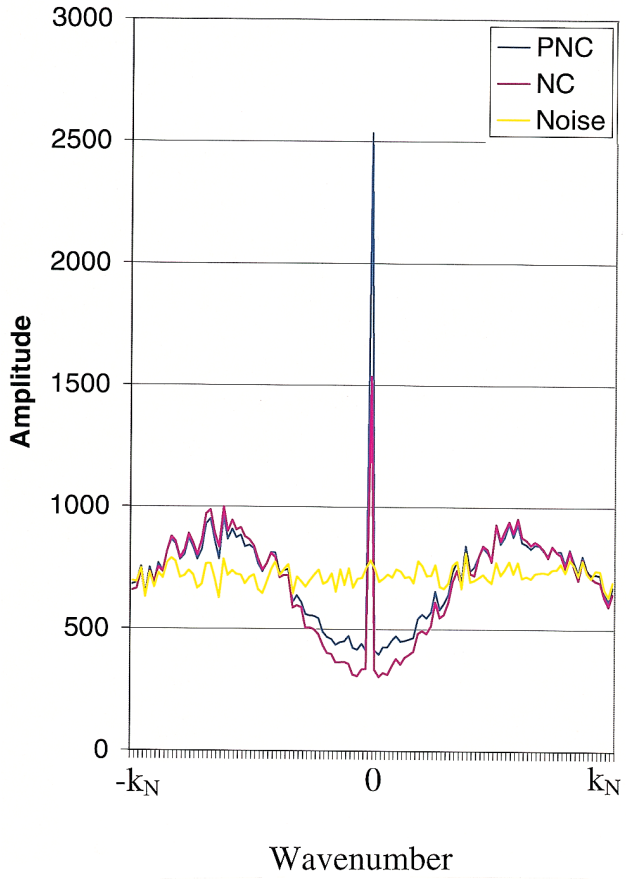

FIG. 20. The average wavenumber spectra for the difference section from $L = 2$ runs.

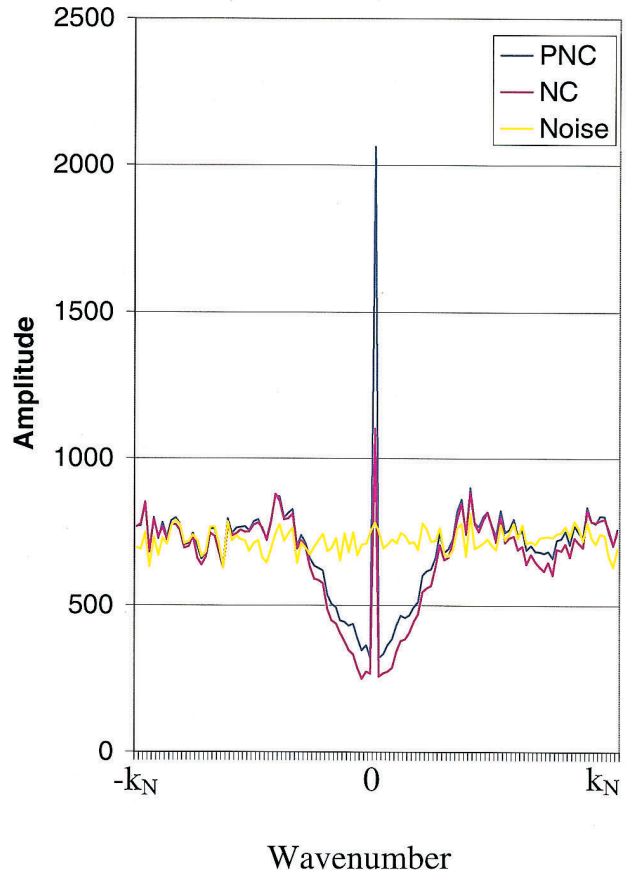

FIG. 21. The average wavenumber spectra for the difference section from $L = 3$ runs.

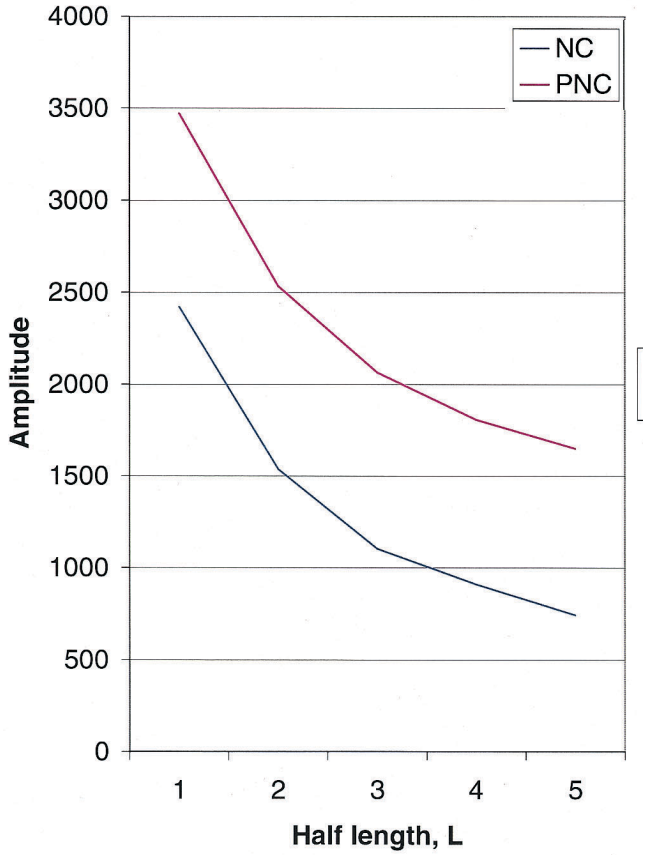

FIG. 22. Signal degradation plot. A higher amplitude means more damage. This plot is obtained by plotting the values at $k = 0$ from Figures 19, 20, and 21.

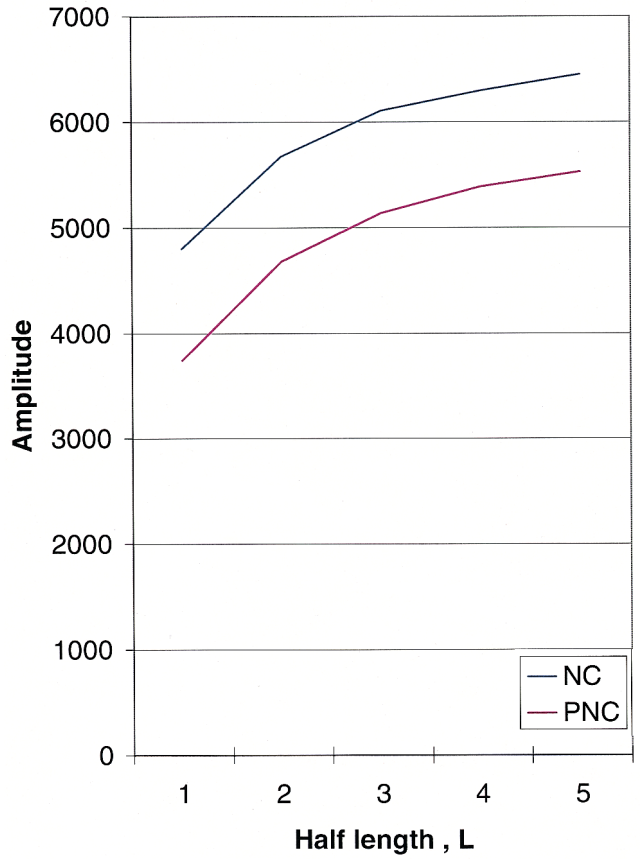

FIG. 23. Signal preservation plot. Higher amplitudes mean better signal preservation. This plot is obtained by plotting the value of average wavenumber spectra at $k = 0$ from the signal outputs of the filters.

ratio, NC f-xy filters preserve faults in both the in-line and cross-line directions. Such filters are less sensitive to the size of the gate than the PNC operators. An NC f-x prediction filter can handle lateral amplitude variations along dip directions much better than a PNC f-x prediction filter can.

The comparison of the signal and noise spectra of the PNC and NC f-x prediction filters on noisy single-event synthetics shows that for the same total filter length NC f-x prediction filters preserve more of the signal and reject less noise than PNC filters. The spectral shape of the rejected noise is more strongly dependent on the length of the filter than on the type of filter. Spectral comparisons of the noise outputs of PNC and NC filters do not show significant differences in rejected noise character. Both are nonwhite.

ACKNOWLEDGMENTS

My thanks go to my former colleagues at Halliburton Geophysical Services: Vasudaven Sudhakar; Federico Martin, who tested the software; Bob Frisbee; Clive Gerrard, who coded the production version; Dave Monk; and Bob Stephens, who offered helpful ideas during the development of the method. Mike Reed developed Western's PNC version of the prediction filter and incorporated the NC filter efficiently into the processing system of Western Geophysical. I thank Ray Abma, Ruben Martinez, Sven Treitel, and my colleagues at Western Geophysical—Subashis Mallick, Diana Pattberg, Bill Dragoset, Laurent Meister, and Luis Canales—for reviewing the manuscript. Finally, I am thankful for the reviewers of GEOPHYSICS—Robert Soubaras, Warren S. Ross, and the anonymous reviewers—for pointing out areas that needed strengthening. I am grateful to Western Geophysical for allowing me to publish this work.

REFERENCES

Abma, R., and Claerbout, J., 1995, Lateral prediction for noise attenuation by t-x and f-x techniques: Geophysics, **60**, 1887–1896.

Berkhout, A. J., 1977, Least-squares inverse filtering and wavelet deconvolution: Geophysics, **42**, 1369–1383.

Bunch, A. W., and White, R. E., 1985, Least squares filters without transient errors: Geophys. Prosp., **33**, 657–673.

Burg, J. P., 1975, Maximum entropy spectral analysis: Ph.D. dissertation, Stanford Univ.

Canales, L. L., 1984, Random noise reduction: 54th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 525–527.

Chase, M. K., 1992, Random noise reduction by FXY prediction filtering: Expl. Geophys., **23**, 51–56.

Claerbout, J., 1991, Earth sounding analysis, processing vs. inversion: Stanford Univ. course notes.

Galbraith, M., 1991, Random noise attenuation by f-x prediction: 61st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1428–1431.

Gülünay, N., 1986, F-x deconvolution and complex Wiener prediction filter: 56th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 279–281.

Gülünay, N., Sudhakar, V., Gerrard, C., and Monk, D., 1993, Prediction filtering for 3-D poststack data: 63rd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1183–1186.

Harris, P. E., and White, R. E., 1997, Improving the performance of f-x prediction signal-to-noise ratios: Geophys. Prosp., **45**, 269–302.

Hornbostel, S., 1991, Spatial prediction filtering in the t-x and f-x domains: Geophysics, **56**, 2019–2026.

Huang, T.-S., 1981, Two-dimensional digital signal processing I—Linear filters: in Springer-Verlag.

IEEE, 1986, Selected papers in multidimensional digital signal processing: Multidimensional signal processing committee of IEEE ASSP Society, IEEE Press.

Jain, A. K., 1981, Advances in mathematical models for image processing: Proc. IEEE, **69**, 502–528.

——— 1989, Fundamentals of digital image processing: Prentice-Hall, Inc.

Kay, S. M., and Marple, S. L., 1981, Spectrum analysis—A modern perspective: Proc. IEEE, **69**, No. 11, 1380–1419.

Makhoul, J., 1975, Linear prediction—A tutorial review: IEEE Transactions, **63**, 561–580.

Marple, S. L., 1987, Digital spectral analysis with applications: Prentice-Hall, Inc.

Nuttal, A. H., 1976, Spectral analysis of a univariate process with bad data points, via maximum entropy and linear predictive techniques: Naval Underwater Systems Center technical report 5303.

Robinson, E. A., 1967, Predictive decomposition of time series with application to seismic exploration: Geophysics, **32**, 418–484.

Soubaras, R., 1990, Deconvolution with prespecified final wavelet length: 60th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1665–1669.

——— 1994, Signal preserving random noise attenuation by the f-x projection: 64th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1576–1579.

Spitz, S., 1991, Seismic trace interpolation in the f-x domain: Geophysics, **56**, 785–794.

Spitz, S., and Deschizeaux, B., 1994, Random noise attenuation in data volumes—FXY predictive filtering versus a signal space technique: 56th Mtg., Eur. Assoc. Expl. Geophys., Expanded Abstracts, paper H033.

Treitel, S., 1974, The complex Wiener filter: Geophysics, **39**, 169–173.

Tufts, D. W., and Kumaresan, R., 1982, Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood: IEEE Transactions, **70**, No. 9, 975–989.

Ulrych, T. J., and Clayton, R. W., 1976, Time series modeling and maximum entropy: Phys. Earth Plan. Int., **12**, 188–200.

Van Valkenburg, M. E., Ed., 1986, Selected papers in mutidimensional digital signal processing: IEEE Press.

Wang, W., and West, G., 1991, F-X filters with dip rejection: 61st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1436–1438.

Xu, L., and Azimi-Sadjadi, R., 1993, Two-dimensional modeling of image random field using artificial neural networks: Internat. Conf. Acoustics, Speech, and Signal Processing, Proceedings, IEEE, **I**, 581–584.

APPENDIX A

THE NC FILTER AND SPATIAL AMPLITUDE VARIATIONS

This appendix demonstrates that an exponential amplitude variation along the space direction is predictable without error by an NC f-x filter but not by a PNC filter. Consider a complex series represented as $(1, a, a^2, \ldots, a^{N-1})$ where $a = e^{\alpha} z_0$. When $z_0 = e^{j\omega\Delta t}$, this represents the f-x response of a dipping event with exponential amplitude variation. Here the dip is Δt seconds per trace, and ω is the angular frequency. The first three lags of the autocorrelation series are

$$r_0 = S_{N-1}(\alpha)$$
$$r_1 = e^{\alpha} e^{j\omega\Delta t} S_{N-2}(\alpha)$$
$$r_2 = e^{2\alpha} e^{j2\omega\Delta t} S_{N-3}(\alpha),$$

where

$$S_N(\alpha) = 1 + e^{2\alpha} + e^{4\alpha} + \cdots + e^{2N\alpha}.$$

For $L = 1$, the three-point PNC f-x filter obtained from the Yule–Walker approach is

$$0.5e^{\alpha} \frac{S_{N-2}(\alpha)}{S_{N-1}(\alpha)} (z_0^*, 0, z_0).$$

The dip component is buried in the complex number z_0. This filter, when applied to data, produces a constant prediction error (it is the same at every sample except the ones at the edges of the data):

$$1 - \left(\frac{1 + e^{2\alpha}}{2} \right) \frac{S_{N-2}(\alpha)}{S_{N-1}(\alpha)}.$$

This error is large when the exponential variation is large (regardless of the sign of alpha). For example, for $\alpha = 1$ (or $\alpha = -1$) and $N = 30$ there is a 43% prediction error in all samples of the output (excluding the two samples at the edges, where it is worse). On the other hand the NC f-x prediction filter can be shown to be

$$(p_{-1}, 0, p_1) = \frac{1}{e^{\alpha} + e^{-\alpha}} (z_0^*, 0, z_0)$$

since equation (5) gives

$$p_1 = \frac{r_1 r_0 - r_2 r_1^*}{r_0^2 - |r_2|^2} = e^{j\theta} e^{\alpha} S_{N-2} \frac{S_{N-1} - e^{2\alpha} S_{N-3}}{S_{N-1}^2 - e^{4\alpha} S_{N-3}^2}$$
$$= e^{j\theta} e^{\alpha} S_{N-2} \frac{1}{S_{N-1} + e^{2\alpha} S_{N-3}}$$

and the identity

$$S_{N-1} + e^{2\alpha} S_{N-3} = e^{\alpha}(e^{-\alpha} + e^{\alpha}) S_{N-2}$$

can be used. This NC prediction filter is independent of N and produces no prediction error except at one sample at each end of the window. This is because P_1 aligns with a^{k-1} and P_{-1} aligns with a^{k+1} during convolution, producing a^k as the output. Also, when $\alpha = 0$, this filter reduces to the filter given in equation (7).

APPENDIX B

THE NC *F-XY* PREDICTION FILTER FOR A SINGLE DIPPING EVENT

In this appendix I give a simple numerical example for the NC f-xy prediction filter. I have two purposes for doing so: (1) to provide the reader a concrete example to check an implementation of the method and (2) to illustrate that gate effects are negligible with this method.

I choose a very small space gate, six lines, each with six CDPs. For simplicity I choose a single dipping event with constant amplitude and look at a single frequency slice: 25 Hz. I assume that dips are 4 ms/CDP and 8 ms/line (36° of phase per trace in-line and 72° per trace cross-line). I design a 3×3 NC filter ($L_x = 1$, $L_y = 1$) . The left side of the normal equations becomes a 9×9 block matrix. The 3×3 NC prediction filter solution is

$$P_{\text{amp}} = \frac{1}{4} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

and

$$P_{\text{phase}} = \begin{bmatrix} -72 & 36 & 144 \\ 72 & 0 & -72 \\ -144 & -36 & 72 \end{bmatrix}.$$

Convolution of this filter with the input produces zero prediction error in a 6×6 output area (except at each corner), despite the fact that the input data length is so short. This robustness to small data gates occurs because (1) there is no noise, (2) NC prediction filter design is used, and (3) biased autocorrelations are used.

In general, if the filter length is $(2L_x + 1, 2L_y + 1)$, then the first and last L_x samples of the output in the x-direction and the first and last L_y samples in the y-direction have prediction errors. For this reason, these output points should be omitted from the output (except when they are at the edges of the frequency slice) and another space window for which these points are not at the edges should be used to process these samples properly. This leads to the processing of data in overlapping space windows and the corresponding

amplitude recovery if samples are predicted more than once.

If I design a 5×5 NC f-xy filter ($L_x = 2$, $L_y = 2$) for the same input, I obtain a 5×5 prediction error filter that is almost identical to the filter in the 3×3 zone and almost zero at other lags, thereby showing the stability of the solution. That is, the filter is center loaded. This is similar to an FXDECON filter being front loaded (see Abma and Claerbout, 1995). With NC design technique, we do not need to know how many events are present in the data, and we can safely use more operator points than we actually need. However, noise leaks into extra coefficients, so it is a good idea not to overspecify the filter length (Bunch and White, 1985).

The 3×3 NC prediction filter for flat events is real at all frequencies and is equal to

$$P = \frac{1}{4}\begin{bmatrix} -1 & 2 & -1 \\ 2 & 0 & 2 \\ -1 & 2 & -1 \end{bmatrix}.$$