GEOPHYSICS, VOL. 70, NO. 4 (JULY-AUGUST 2005); P. V97–V107, 19 FIGS.
10.1190/1.1997369

# Multiple attenuation in complex geology with a pattern-based approach

Antoine Guitton[1]

## ABSTRACT

Primaries (signal) and multiples (noise) often exhibit different kinematics and amplitudes (i.e., patterns) in time and space. Multidimensional prediction-error filters (PEFs) approximate these patterns to separate noise and signal in a least-squares sense. These filters are time-space variant to handle the nonstationarity of multioffset seismic data. PEFs for the primaries and multiples are estimated from pattern models. In an ideal case where accurate pattern models of both noise and signal exist, the pattern-based method recovers the primaries while preserving their amplitudes. In the more general case, the pattern model of the multiples is obtained by using the data as prediction operators. The pattern model of the primaries is obtained by convolving the noise PEFs with the input data. In this situation, 3D PEFs are preferred to separate (in prestack data) the multiples properly and to preserve the primaries. Comparisons of the proposed method with adaptive subtraction with an $\ell_2$ norm demonstrate that for a given multiple model, the pattern-based approach generally attenuates the multiples and recovers the primaries better. In addition, tests on a 2D line from the Gulf of Mexico demonstrate that the proposed technique copes fairly well with modeling inadequacies present in the multiple prediction.

## INTRODUCTION

In the presence of complex geology where multipathing, illumination gaps, and coherent noise are prevalent, the most advanced techniques need to be used for preprocessing and imaging. For multiple attenuation, Weglein (1999) shows that current technology may be divided into filtering methods, which exploit the periodicity and the separability (moveout discrepancies) of the multiples, and wavefield methods, which first predict and then subtract the multiples (Verschuur et al., 1992; Dragoset and MacKay, 1993 ; Weglein et al., 1997). Traditionally, filtering techniques are chosen because of their robustness and low cost. However, filtering techniques have some limitations when tackling multiples in complex media. For example, predictive deconvolution in the ray-parameter domain fails when the water bottom is not flat (Treitel et al., 1982). Furthermore, numerous authors (Matson et al., 1999; Bishop et al., 2001; Paffenholz et al., 2002) present cases where wavefield approaches such as surface-related multiple elimination (SRME) attenuate multiples much better than filtering techniques such as radon-based methods.

Wavefield techniques usually begin with a prediction step where surface-related multiples are modeled from the data with (Wiggins, 1988; Lokshtanov, 1999) or without any subsurface information. Then the multiples are subtracted from the data. Both aspects of the multiple attenuation procedure are important, but most of today's efforts are concentrated on prediction, not on subtraction. With SRME (Verschuur et al., 1992), two significant assumptions are usually made for subtraction. First, it is assumed that the signal has minimum energy, leading to the adaptive subtraction of the multiples with an $\ell_2$ norm. However, this assumption might not hold where primaries and multiples interfere (Spitz, 1999). For instance, Guitton and Verschuur (2004) show that when primaries are much stronger than multiples, the $\ell_1$ norm should be used instead. Second, it is assumed that the multiples are accurately modeled. This point relies on the acquisition or the interpolation/extrapolation of the data to provide the necessary traces for the prediction step. In practice, however, the data are never acquired densely enough and the interpolation schemes are never perfect, especially with sparse acquisition geometries. Consequently, the prediction is not as accurate as required. Therefore, other subtraction techniques are desirable when adaptive subtraction fails to recover the primaries and when the multiple model is not precise enough. As stated by Berkhout (personal communication, 2004), the subtraction step is the weakest component of SRME, and more work needs to be done in this direction.

Manuscript received by the Editor April 30, 2004; revised manuscript received October 20, 2004; published online July 7, 2005.
[1]Stanford University, Department of Geophysics, Stanford Exploration Project, Mitchell Building, Stanford, California 94305-2215. E-mail: antoine@sep.stanford.edu.
© 2005 Society of Exploration Geophysicists. All rights reserved.

A new class of multiple attenuation techniques has emerged recently to circumvent some of the limitations of adaptive subtraction and modeling. These techniques are called pattern based because they discriminate primaries from multiples according to their multidimensional spectra (Manin and Spitz, 1995; Guitton and Cambois, 1998; Brown and Clapp, 2000; Fomel, 2002). This paper presents one implementation of a pattern-based approach based on prediction-error filters (PEFs). By construction, PEFs approximate the inverse spectrum of the data from which they are estimated (Burg, 1975). Therefore, PEFs can serve as proxies for the patterns of both primaries and multiples and can be used for signal-noise separation (Guitton, 2002).

This paper begins with a description of nonstationary PEF estimation. The helical boundary conditions (Mersereau and Dudgeon, 1974; Claerbout, 1998) are used to estimate 2D and 3D filters. Both noise (multiples) and signal (primaries) PEFs are needed for the attenuation. Because the signal is usually unknown, a method that needs only the noise and the input data to derive a pattern model for the signal is presented. Having estimated the noise and signal PEFs, the multiple subtraction method (pattern-based) that separates primaries from multiples in a least-squares sense is then described.

To illustrate the efficiency of the pattern-based approach, multiples are attenuated for the Sigsbee2B synthetic data set. The results of multiple attenuation are analyzed after migration to assess the effects of the proposed technique on the primaries. This example illustrates that the pattern-based approach can lead to a very good elimination of the multiples if an accurate pattern model for both primaries and multiples is available. In addition, this data set shows that 3D PEFs preserve the primaries better than 2D filters. Then, adaptive and pattern-based subtractions are compared on a synthetic data example provided by BP. This example proves that when multiples and primaries are spatially uncorrelated (i.e., different patterns), PEFs attenuate multiples better than adaptive subtraction. Finally, multiple attenuation results on a Gulf of Mexico 2D line are shown. This last example illustrates that the pattern-based method is robust to model inadequacies.

## FILTER ESTIMATION

The key assumption of the proposed multiple attenuation technique is that primaries and multiples have different multidimensional spectra that PEFs can approximate (Claerbout, 1992; Spitz, 1999). Therefore, PEFs for primaries and multiples are needed prior to the subtraction step. This section describes how these filters are estimated.

The PEFs used in this paper are time-space domain nonstationary filters to cope with the variability of seismic data with time, offset, and shot position. Implementing nonstationary filters is not an easy task. A possible solution is to break up the data set into patches and then estimate a filter for each patch. However, reassembling these patches creates edge effects in the overlapping zones (Guitton and Verschuur, 2004). Alternatively, nonstationary convolution (i.e., each filter is in a column of a matrix) or combination (i.e., each filter is in a row of a matrix) can be used to estimate one filter per data point (Margrave, 1998) or, more realistically, per micropatch (Crawley, 2000). A micropatch is made of neighboring data points that share the same filter. With this technique, the filters can vary smoothly across the data set while leaving almost no edge effect.

A complete description of the nonstationary filters goes beyond the scope of this paper, which concentrates on the most important steps of the filter estimation procedure only. We call $\mathbf{Y}$ the nonstationary combination matrix (Margrave, 1998) with the data vector $\mathbf{y}$ from which we want to estimate the filters and $\mathbf{a}$ the unknown PEFs coefficients. One way to estimate the PEFs is to minimize the length of residual vectors $\mathbf{r_y}$ (Claerbout and Fomel, 2004):

$$\mathbf{0} \approx \mathbf{r_y} = \mathbf{Ya}. \tag{1}$$

Expression 1 is also called a fitting goal. By definition of the PEFs, the first coefficient of the unknown filters is always one (Figure 1). To take this requirement into account, equation 1 becomes

$$\mathbf{0} \approx \mathbf{r_y} = \mathbf{YKa} + \mathbf{y}, \tag{2}$$

where $\mathbf{K}$ is a masking operator that forces the first coefficient of the PEFs to be one. If one filter is used per data point, the matrix of unknown coefficients $\mathbf{a}$ can be enormous, making the problem very underdetermined. This difficulty can be overcome in two ways. First, the filter is kept constant inside a micropatch. Second, a smoothing operator $\mathbf{R}$ is introduced to penalize strong variations between neighboring filters. Both strategies are considered here.

Introducing the Laplacian operator $\mathbf{R}$, equation 2 is augmented as follows:

$$\begin{aligned} \mathbf{0} &\approx \mathbf{r_y} = \mathbf{YKa} + \mathbf{y}, \\ \mathbf{0} &\approx \epsilon_1 \mathbf{r_a} = \epsilon_1 \mathbf{Ra}, \end{aligned} \tag{3}$$

where $\epsilon_1$ is a trade-off parameter between coefficient estimation and filter smoothing. In practice, $\epsilon_1$ is selected by trial and error. The Laplacian operator smoothes filter coefficients along two (i.e., time/offset) or three (i.e., time/offset/shot) axes, depending on the PEF's dimensions. A least-squares estimate of the PEF's coefficients leads to the following objective function:

$$f(\mathbf{a}) = \|\mathbf{r_y}\|^2 + \epsilon_1^2 \|\mathbf{r_a}\|^2, \tag{4}$$

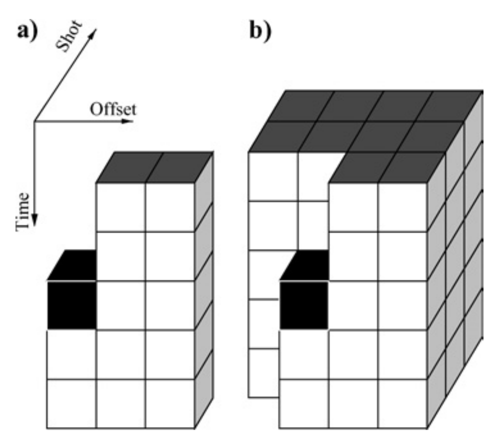


Figure 1. Prediction error filters in (a) two and (b) three dimensions. The first coefficient in black is always one.

where **a** is estimated iteratively with the conjugate gradient method.

The amplitudes of seismic data vary across offset, shot, and time. Large-amplitude variations can be troublesome with least-squares inversion because they tend to bias the final result (Claerbout, 1992) by ignoring or overfitting some areas in the data space. Therefore, it is important to make sure these amplitude variations do not affect processing. One solution is to apply a weight to the data prior to the inversion, such as automatic gain control (AGC) or a geometric spreading correction. Alternatively, a weight **W** can be incorporated inside the fitting goals in equation 3:

$$\mathbf{0} \approx \mathbf{r_y} = \mathbf{W}(\mathbf{YKa} + \mathbf{y}),$$
$$\mathbf{0} \approx \epsilon_1 \mathbf{r_a} = \epsilon_1 \mathbf{Ra}. \quad (5)$$

This weight is a diagonal operator and can be interpreted as a change of norm consistent with the data, similar to $\ell_1$ (Guitton and Verschuur, 2004). In this paper, the weight $W_{i,i}$ for one data point $i$ is obtained by computing

$$W_{i,i} = \left(\frac{y_{AGC_i}}{y_i}\right) \quad (6)$$

where $\mathbf{y}_{AGC}$ is the data vector after AGC, making sure that no division by zero occurs. This weight can also incorporate a mute zone where no data are present.

Thanks to the helical boundary conditions (Mersereau and Dudgeon, 1974; Claerbout, 1998), the PEFs can have any dimension. In this paper, both 2D and 3D filters (Figure 1) are used, but 3D filters lead to the best noise attenuation results. When 2D filters are used, the multiple attenuation is performed on one shot gather at a time. When 3D filters are used, the multiple attenuation is performed on one macrogather at a time. A macrogather is a cube made of adjacent shots with all offsets and time samples. There is an overlap of five shot gathers between successive macrogathers. When the multiple attenuation is finished, the macrogathers are reassembled to form the final result.

The next section describes how to choose the pattern model **Y** in equation 3 when PEFs for multiples and primaries are estimated.

## ESTIMATING SIGNAL AND NOISE PEFs

For multiple attenuation, nonstationary PEFs **N** and **S** for the multiples and the primaries, respectively, are required. Therefore, pattern models for the noise and the signal need to be constructed. For surface-related multiples, the multiple model may be provided by surface-related multiple prediction (SRMP) (Verschuur et al., 1992; van Dedem, 2002), which yields a correct prestack model of the multiples.

As for amplitude, an accurate surface-related multiple model can be derived if (1) the source wavelet is known, (2) the surface source and receiver coverage is large and dense enough, and (3) all terms of the series that model different orders of multiples are incorporated (Verschuur et al., 1992). In practice, a single convolution of the input data (i.e., one term of the series) is usually performed, giving a multiple model with erroneous relative amplitudes for high-order multiples. This single convolution can be interpreted as the first iteration of the recursive formulation of SRME (Berkhout and Verschuur, 1997). In addition, the limited size of the acquisition geometry can create inaccurate multiples if they bounce outside the recording array. Dragoset and Jericevic (1998) detail other possible flaws introduced in the prediction by defective acquisition parameters. Because PEFs estimate patterns, wrong relative amplitudes and kinematic errors can affect the multiple suppression results. However, as we see later, 3D filters seem to cope better with noise modeling inadequacies.

Signal PEFs are more difficult to estimate since the primaries are usually unknown. As a possible solution to this problem, Spitz (1999) estimates a signal PEF **S** by deconvolving a data PEF **D**, estimated from the data, by a noise PEF **N**. With this process, Spitz assumes that

$$\mathbf{D} = \mathbf{SN}. \quad (7)$$

I call equation 7 the Spitz approximation. Note that **D**, **N**, and **S** are matrices for the combinations with the nonstationary PEFs (Margrave, 1998). These matrices are very sparse and are never formed in practice (Claerbout and Fomel, 2004). Equation 7 can be retrieved by considering a simple 1D example using the $Z$-transform notations (Claerbout, 1976) for a data PEF $D_{1\mathrm{D}}(Z)$, a signal PEF $S_{1\mathrm{D}}(Z)$, and a noise PEF $N_{1\mathrm{D}}(Z)$. Extension to more dimensions is straightforward using the helical boundary conditions (Claerbout, 1998). Because PEFs have the inverse spectrum of the data from which they have been estimated (Burg, 1975), we have (omitting $Z$ for clarity)

$$\frac{1}{D_{1\mathrm{D}}^* D_{1\mathrm{D}}} = \frac{1}{S_{1\mathrm{D}}^* S_{1\mathrm{D}}} + \frac{1}{N_{1\mathrm{D}}^* N_{1\mathrm{D}}}, \quad (8)$$

where the asterisk denotes the complex conjugate. Equation 8 states that the spectrum of the data is equal to the spectrum of the noise plus the spectrum of the signal. Equation 8 can be written as follows:

$$D_{1\mathrm{D}}^* D_{1\mathrm{D}} = \frac{S_{1\mathrm{D}}^* N_{1\mathrm{D}}^* S_{1\mathrm{D}} N_{1\mathrm{D}}}{N_{1\mathrm{D}}^* N_{1\mathrm{D}} + S_{1\mathrm{D}}^* S_{1\mathrm{D}}}. \quad (9)$$

Because PEFs are important where they are small (i.e., where they attenuate seismic events), the denominator can be neglected:

$$D_{1\mathrm{D}}^* D_{1\mathrm{D}} \approx S_{1\mathrm{D}}^* N_{1\mathrm{D}}^* S_{1\mathrm{D}} N_{1\mathrm{D}}, \quad (10)$$

which leads to the Spitz approximation in equation 7. The PEFs **D** and **N** can be estimated easily because the data vector and a noise model are often available. However, estimating the signal PEFs requires a potentially unstable nonstationary deconvolution $\mathbf{S} = \mathbf{DN}^{-1}$ (Rickett, 2001) in equation 7. To avoid the deconvolution step, the noise PEFs are convolved with the data:

$$\mathbf{u} = \mathbf{Nd}, \quad (11)$$

where **u** is the result of the convolution and **d** is the input data vector (signal plus noise). Estimating the PEFs **U** for **u** gives by definition of the PEFs (Claerbout and Fomel, 2004)

$$\mathbf{0} \approx \mathbf{Uu}. \quad (12)$$

Then, from the Spitz approximation in equation 7, we have

$$\mathbf{0} \approx \mathbf{Uu} = \mathbf{UNd} = \mathbf{Dd} = \mathbf{SNd} \quad (13)$$

and $\mathbf{U} = \mathbf{S}$. Therefore, by convolving the data with the noise PEFs, signal PEFs consistent with the Spitz approximation can be computed. Again, an important assumption is that signal and noise are uncorrelated.

The PEFs for the primaries $\mathbf{S}$ and the multiples $\mathbf{N}$ are estimated directly from the data and the model of the multiples. These filters approximate the multidimensional spectra of the noise and signal. The next section describes the noise attenuation step.

## MULTIPLE ATTENUATION

When the noise and signal PEFs have been determined, the task of multiple attenuation follows. First, consider that the seismic data $\mathbf{d}$ are the sum of signal (primaries) and noise (multiples):

$$\mathbf{d} = \mathbf{s} + \mathbf{n}, \tag{14}$$

where $\mathbf{s}$ is the signal we want to preserve and $\mathbf{n}$ is the noise we wish to attenuate.

By definition of the signal and noise PEFs, the following relationships hold:

$$\begin{aligned} \mathbf{Nn} &\approx \mathbf{0}, \\ \mathbf{Ss} &\approx \mathbf{0}. \end{aligned} \tag{15}$$

Equations 14 and 15 can be combined to solve a constrained problem to separate signal from noise as follows:

$$\begin{aligned} \mathbf{0} &\approx \mathbf{r_n} = \mathbf{Nn}, \\ \mathbf{0} &\approx \epsilon\mathbf{r_s} = \epsilon\mathbf{Ss}, \\ \text{subject to} &\leftrightarrow \mathbf{d} = \mathbf{s} + \mathbf{n}. \end{aligned} \tag{16}$$

The scalar $\epsilon$ is related to the S/N and is usually estimated by trial and error. Replacing $\mathbf{n}$ by $\mathbf{s} - \mathbf{d}$ in equation 16 yields

$$\begin{aligned} \mathbf{0} &\approx \mathbf{r_n} = \mathbf{Ns} - \mathbf{Nd}, \\ \mathbf{0} &\approx \epsilon\mathbf{r_s} = \epsilon\mathbf{Ss}. \end{aligned} \tag{17}$$

Sometimes it is useful to add a masking operator for the noise and signal residuals $\mathbf{r_n}$ and $\mathbf{r_s}$ when performing the noise attenuation. For example, in parts of the data where no multiples are present, the signal should be preserved. In addition, a mute zone can be taken into account very easily. Calling $\mathbf{M}$ this masking operator, the fitting goals in equation 17 are weighted as follows:

$$\begin{aligned} \mathbf{0} &\approx \mathbf{r_n} = \mathbf{M}(\mathbf{Ns} - \mathbf{Nd}), \\ \mathbf{0} &\approx \epsilon\mathbf{r_s} = \epsilon\mathbf{MSs}. \end{aligned} \tag{18}$$

Solving for $\mathbf{s}$ in a least-squares sense leads to the objective function

$$f(\mathbf{s}) = \|\mathbf{r_n}\|^2 + \epsilon^2\|\mathbf{r_s}\|^2. \tag{19}$$

It is interesting to look at the least-squares inverse $\hat{\mathbf{s}}$ for $\mathbf{s}$:

$$\hat{\mathbf{s}} = (\mathbf{N}'\mathbf{MN} + \epsilon^2\mathbf{S}'\mathbf{MS})^{-1}\mathbf{N}'\mathbf{MNd}, \tag{20}$$

where prime stands for the adjoint. Because $\mathbf{M}$ is a diagonal matrix of zeroes and ones, we have $\mathbf{M}'\mathbf{M} = \mathbf{M}$. In practice, all computations are done in the time domain. In the Fourier domain, equation 20 demonstrates that the least-squares estimate of $\mathbf{s}$ is obtained by combining the spectra of both noise and signal. Abma (1995) shows that equation 20 is similar to Wiener filtering for random noise attenuation. Soubaras (1994) uses a very similar approach for random and coherent noise attenuation (Soubaras, 2001). Because the size of the data space can be quite large, $\mathbf{s}$ is estimated iteratively with the conjugate gradient method.

Therefore, multiple attenuation with prediction-error filters (in two or three dimensions) is a two-step process where (1) noise and signal PEFs are estimated and (2) signal and noise are separated according to their multidimensional spectra. The next section illustrates this technique with the Sigsbee2B data set.

## MULTIPLE ATTENUATION WITH THE SIGSBEE2B DATA SET

The Sigsbee2B data set was designed to generate strong surface-related multiples. Figure 2 shows the true stratigraphic interval velocity model for this data set. The data were created with a 2D acoustic finite-difference modeler with constant density. Two data sets were generated: one with a free surface in Figure 3a (FS) and one without a free surface in Figure 3b (NFS). Both receiver and source ghosts are included in the modeling with or without free surface (Technical University of Delft, 2002). We can directly subtract the two data sets to obtain a very accurate prestack model of the surface-related multiples — without the need for SRMP.

In complex geology, multiple attenuation results should be assessed after prestack migration; then, the effects of the multiple attenuation technique on the amplitudes of the primaries in angle-domain common-image gathers or on migrated images (Figure 4) can be inspected. For the Sigsbee2B data set, a split-step, double square-root (DSR) migration code with three reference velocities is used (Popovici, 1996). It is interesting to see in Figure 4a that the multiples are very weak after migration below the salt compared to the constant-offset sections in Figure 3a. In particular, the water-bottom multiple underneath salt seems to disappear. This is because the multiples are extremely distorted by the migration process in the vicinity of the complex salt structure. Inspecting the migration of the primaries only in Figure 4b, we find the multiples in Figure 4a mask many primaries in the deepest part of the model and need to be removed.

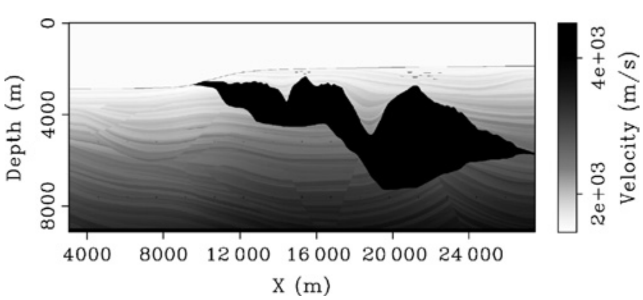


Figure 2. Stratigraphic interval velocity model of the Sigsbee2B data set.

Two important tests are carried out in this section. First, because the true primaries and multiples are known, the noise and signal PEFs can be estimated ideally without SRMP or the Spitz approximation and used for the separation. Second, in the more realistic case where only a pattern model of the multiples is known, noise attenuation results are shown with 2D or 3D filters. The next section demonstrates that when an accurate pattern model of the noise and signal is available, the signal can be recovered very well.

### Estimating biases

A bias is a processing footprint left by the multiple attenuation technique, e.g., edge effects from the nonstationary PEFs. In an ideal but unrealistic case, a pattern model for both the primaries and the multiples might be available. In this case, a bias is also any difference between the true primaries and the estimated primaries after attenuation of the multiples. In this section we learn that the bias is minimal with the pattern-based approach.

For the pattern model of the primaries, the answer, i.e., the data modeled without the free surface condition, is used. For the multiples, the difference between the FS (Figure 3a) and NFS (Figure 3b) data sets is used. Because the noise and signal PEFs are estimated from accurate pattern models, only 2D filters are estimated. Three-dimensional filters can help if the primaries and multiples are correlated in time and offset but are uncorrelated across shot position. With 2D filters, the attenuation is performed one shot gather at a time. Figure 5a displays the estimated primaries, and Figure 5b shows the difference with the true primaries (Figure 3b). The bias introduced by the attenuation method is very small; 3D filters probably would have given better results where the difference between Figures 5a and 3b is the strongest (e.g., near 20 km).

Looking now at the same estimated primaries after migration in Figure 6a, we see again that the attenuation gives a very good result with little bias. Some energy is visible in the difference plot in Figure 6b where no multiples are actually present, however. These artifacts have two origins. First, below 4000 m some primaries are affected by the multiple attenuation process, especially at far offset where primaries and multiples overlap. Second, above 4000 m the amplitude of the reflections for the sea floor and the top of salt are slightly different between the FS and NFS data sets. These differences are migrating at the reflector positions in Figure 6b but with a very small energy.

From these results it appears that the quality of the multiple attenuation depends essentially on the filters. If accurate

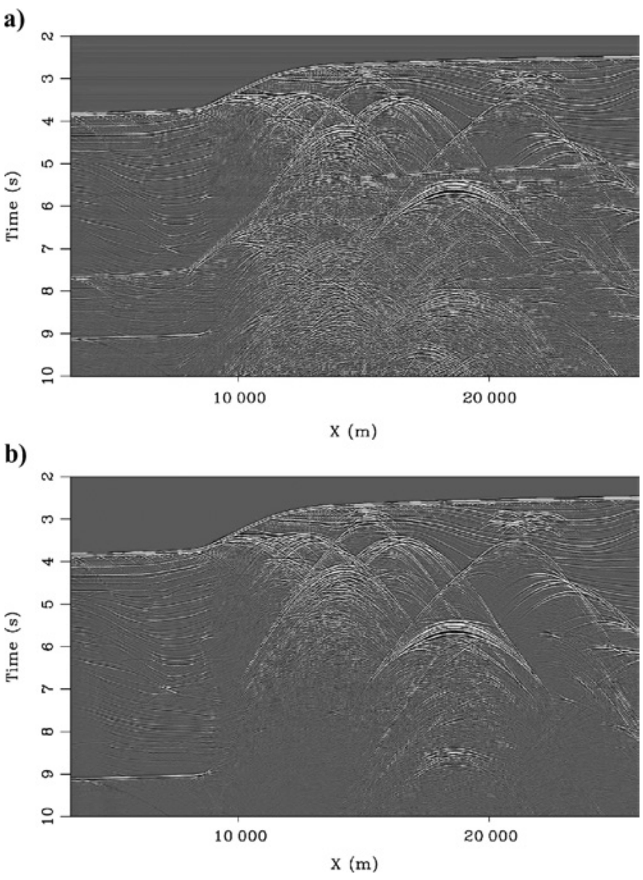


Figure 3. Two constant-offset sections ($h = 342$ m) of the Sigsbee2B data set (a) with and (b) without free-surface condition. The multiples are very strong below 5 s. The weak horizontal striping in (a) comes from a source effect only present with the free-surface condition modeling.

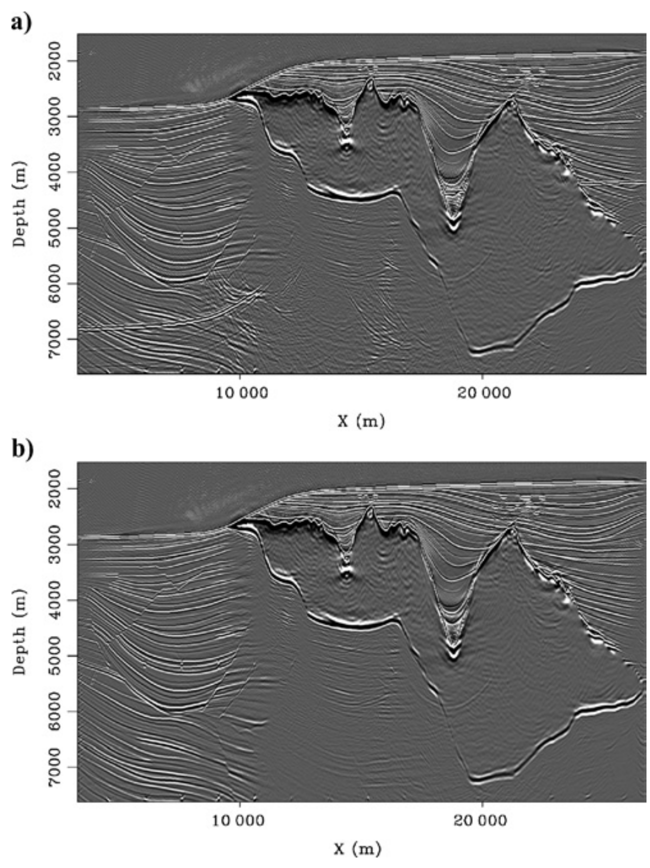


Figure 4. Migrated images for the data (a) with and (b) without free-surface condition. Comparing with Figure 3, the multiples appear much weaker below the salt after migration. However, some reflectors near 10 000 m are hidden in (a).

pattern models of the primaries and multiples are available, the primaries are recoverable while preserving their amplitude. Therefore, we should always try to find the best pattern models for the signal and the noise. In practice, a very accurate pattern model of the multiples can often be estimated with the autoconvolutional process of the Delft approach (Verschuur et al., 1992). For the primaries, the next section shows that the Spitz approximation gives a very good pattern model if 3D filters are used for multiple removal.

### Testing the Spitz approximation

Now we assume that only a pattern model of the multiples is known. The Spitz approximation in equation 7 shows how the PEFs for the signal can be estimated. The primaries are recovered with 2D and 3D filters. Figure 7 displays two constant-offset sections after multiple attenuation with 2D and 3D PEFs. Three-dimensional PEFs give by far the best results and attenuate multiples very well. In particular, comparing with Figure 3b, the 2D PEFs leave more multiple energy below 6 s, between 10 000 and 15 000 m.

After migration, we see again in Figure 8 that the 3D PEFs attenuate the multiples more effectively. The circles in Figure 8 indicate areas where the 3D filters are the most competent. A close-up in Figure 9 illustrates in more detail how the two results with 2D or 3D filters differ below the salt. Events are more continuous and are preserved better with 3D filters. Comparing with the true reflectors in Figure 9a, important primaries (shown as 1 in Figure 9a) are attenuated with both 2D and 3D filters. Notice that event 3 in Figure 9a and c looks very similar to a multiple (shown as 2) left with the 2D filters in Figure 9a.

These important observations could not have been made before migration in the prestack domain because the primaries are much weaker than the surface-related multiples below the salt. This indicates that for complex geology, the quality of a multiple removal technique should be assessed in the image space. Realizing that some primaries are attenuated in Figure 9 should motivate us in devising improved strategies for building more accurate noise and signal pattern models.

The fact that 3D PEFs attenuate the multiples better than 2D PEFs is not surprising. With higher dimensions, primaries and multiples are less likely to be correlated. Therefore, the noise and signal PEFs are less prone to annihilate similar data components. This is particularly important with the Spitz approximation, which implicitly assumes that primaries and multiples are uncorrelated.

The next section compares the pattern-based approach with adaptive subtraction on a synthetic data set provided by BP.

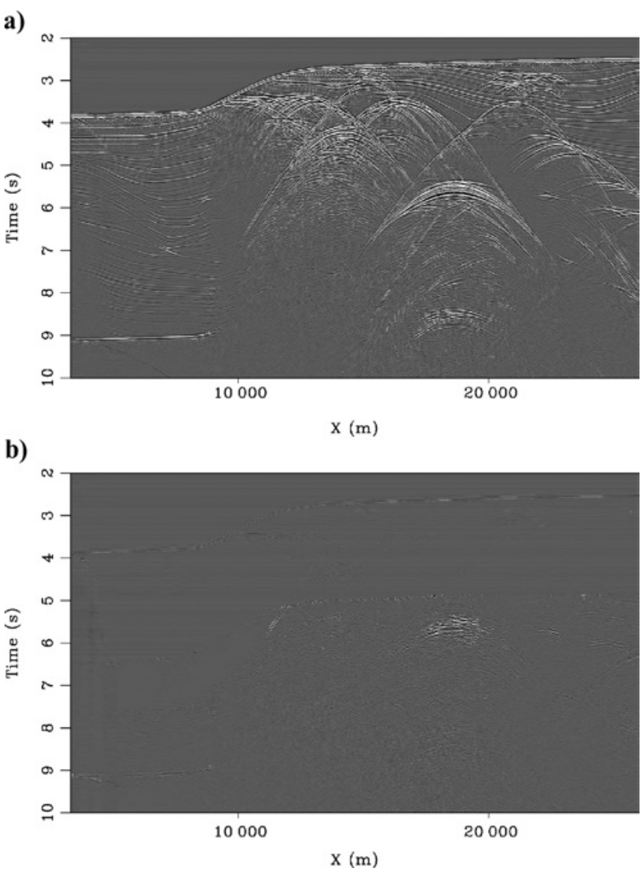


Figure 5. Two constant offset panels at $h = 342$ m for (a) the estimated primaries and (b) the difference with the true primaries. The true primaries and multiples are used to estimate the PEFs.

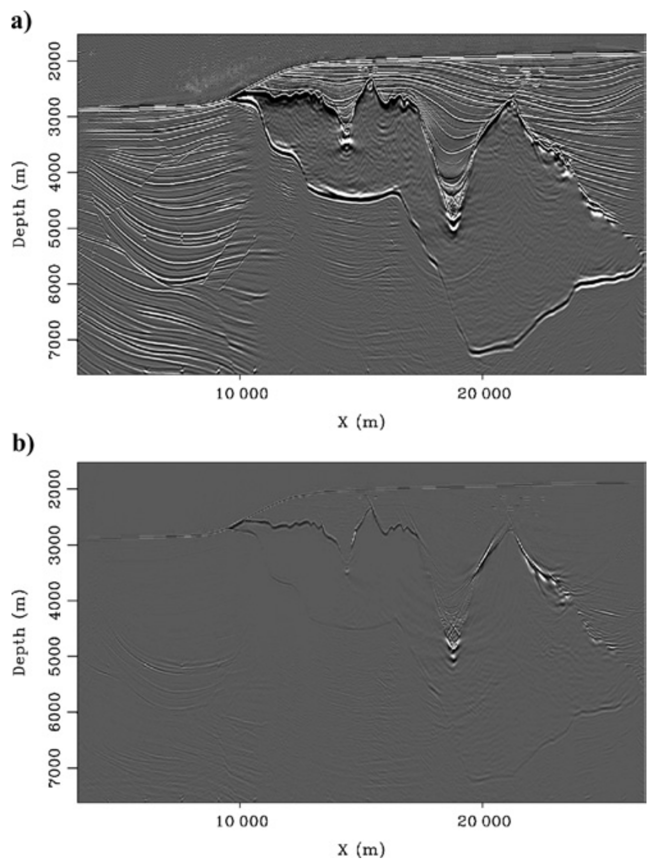


Figure 6. (a) Migration result after multiple attenuation when the true primaries and multiples are used to estimate the PEFs. (b) Difference between (a) and Figure 4b. The primaries are well recovered.

## ADAPTIVE VERSUS PATTERN-BASED SUBTRACTION

The goal of this section is to compare the pattern-based method with the more conventional adaptive subtraction on a synthetic data set provided by BP. These two methods have very different properties. On the one hand, adaptive subtraction assumes implicitly that the signal has minimum energy. In addition, the separation is very fast because the matching filters and the primaries are usually estimated simultaneously with Wiener-Levinson methods. On the other hand, the proposed approach assumes that the signal and noise have different patterns. Unfortunately, high-dimension filters for primaries and multiples need to be estimated first before starting the separation, which adds to the total cost of the method. Therefore, the pattern-based approach can be much slower than adaptive subtraction.

Computing considerations aside, the two methods can perform very differently according to the geologic setting. Here, both are tested in a salt environment with a synthetic data set. This data set was primarily designed to conduct blind tests for velocity estimation methods. Consequently, no structural information is known. The adaptive subtraction technique used in this section is based on the estimation of 2D, time-space-varying matching filters (Rickett et al., 2001; Guitton and Verschuur, 2004). The filters are computed for one shot gather at a time. With the pattern-based approach, 3D filters are used for the multiple attenuation. Ideally, 3D filters should also be used with adaptive subtraction. However, matching filters are generally not estimated that way.

The multiple model is computed with SRMP. The synthetic model has an offset spacing of 12.5 m and a shot separation of 50 m. To make the multiple prediction work, the offset axis is sampled down to 50 m. This choice is clearly not ideal to obtain an accurate model of the multiples, and the missing shots should be interpolated. In practice with field data, however, it is never possible to obtain all of the necessary traces, either during the acquisition or with interpolation. Thus, the multiple model is never perfect and this synthetic model, with this choice of offset spacing, illustrates this problem. Figure 10 shows one constant-offset section from −15 000 m to +15 000 m with primaries and multiples. This area is particularly interesting because of the presence of diffraction hyperbolas. Because no velocity model or sedimentary section is available, a possible interpretation of these diffractions is the presence of salt bodies with a rugose top (similar to what we see with the Sigsbee2B data set). The multiple model is shown in Figure 11 for the same offset. The label DT points to diffraction tails where the model is not properly matching the multiples in the data. Besides these few imperfections in the model, the model looks very faithful to the actual multiples.

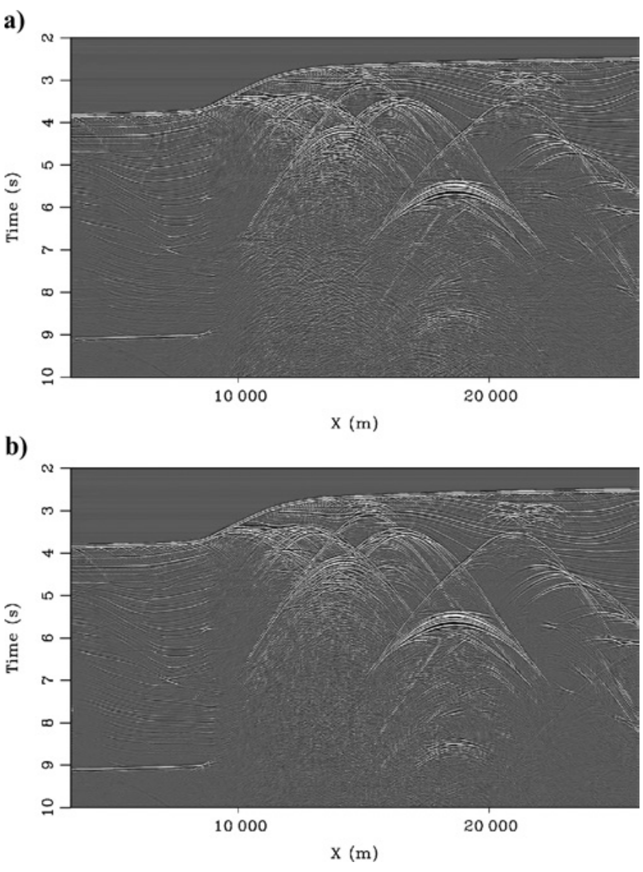


Figure 7. Two constant offset sections ($h = 342$ m) after multiple attenuation with the Spitz approximation using (a) 2D and (b) 3D filters.

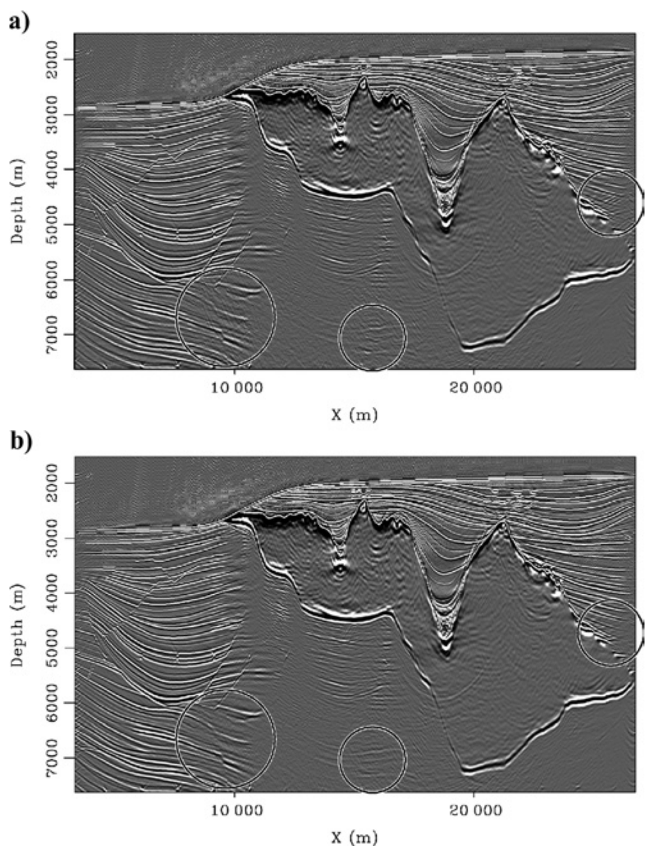


Figure 8. Two migration results of the estimated primaries with (a) 2D and (b) 3D filters. The circles show areas where multiples are better attenuated with 3D filters than with 2D filters.

Because no velocity analysis was conducted with this data set, no stacks are presented. Alternatively, close-ups of constant-offset sections are shown to illustrate strengths and weaknesses of the two different approaches. Figure 12 shows a comparison between the input data, the multiple model, the estimated primaries with the adaptive subtraction, and the estimated primaries with the pattern-based technique. Parameters for the adaptive subtraction such as filter and micropatch sizes were determined by trial and error to give the most satisfying multiple attenuation results. The offset is 700 m. As shown by the arrows, the pattern-based method generally performs better. The same conclusions hold in Figure 13. Note the presence of aliasing artifacts in Figure 13b as a result of the coarse sampling of the offset axis for the multiple prediction (van Dedem, 2002). These artifacts do not seem to affect the estimated signal for either method.

Sometimes, it can be rather difficult to see if multiples are removed by looking at 2D planes, as exemplified in Figure 14. The multiple model in Figure 14b indicates that most of the events seen in Figure 14a are multiples. The multiple attenuation result with adaptive subtraction in Figure 14c shows one event at 2 that seems to be a primary. However, by looking at the shot gathers (not shown here), it appears that this event is a multiple that the pattern-based approach was able to attenuate.

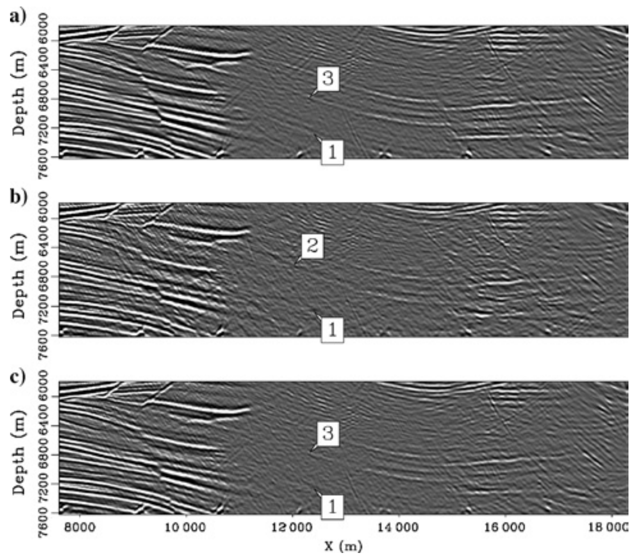


Figure 9. Close-up of Figure 8 showing two migrated images when (b) 2D and (c) 3D filters are used. The true primaries are shown in (a). Arrow 1 points to primaries that are attenuated with the pattern-based approach. Arrow 3 is a primary that the 3D PEFs recover. Arrow 2 points to a multiple that is not attenuated with the 2D PEFs but which resembles 3 quite closely. The 3D filters remove this event.

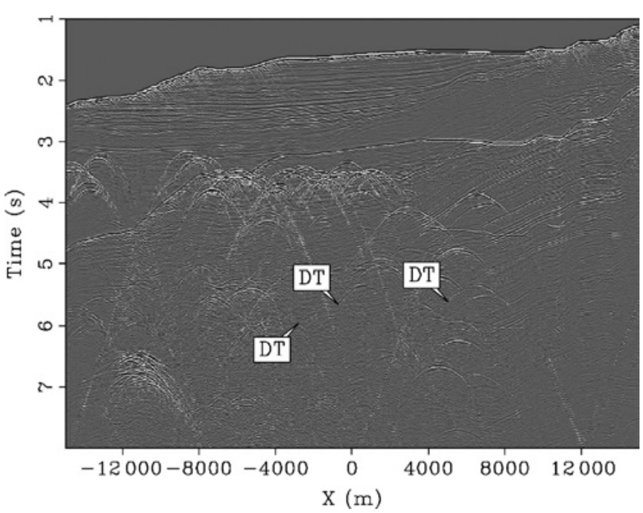


Figure 10. Constant-offset section ($h = 500$ m) of the BP synthetic data set with multiples. Tails of diffracted multiples are denoted by DT.

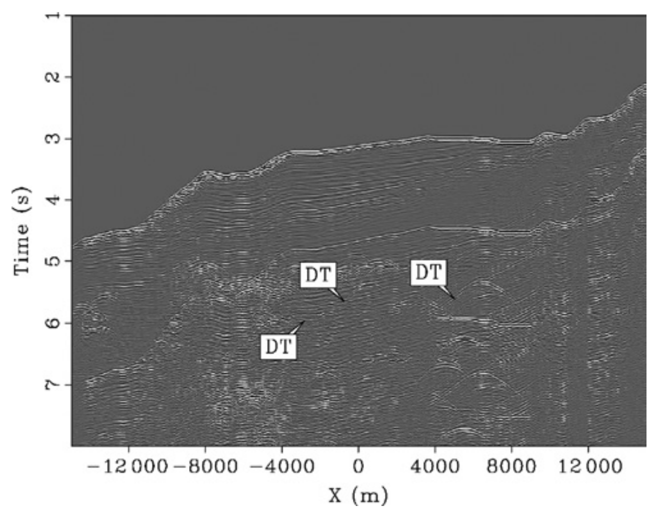


Figure 11. Constant-offset section ($h = 550$ m) of the estimated multiples for the BP synthetic data set. The multiples are accurately modeled except for the diffracted multiples, shown as DT, for which the limited range of offsets and number of shots hamper any attempt at modeling the diffraction tails.

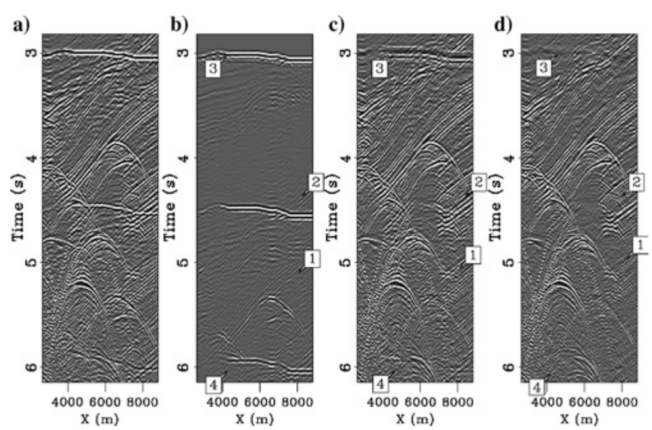


Figure 12. Constant-offset sections ($h = 700$ m) for (a) the input data, (b) the multiple model, (c) the estimated primaries with adaptive subtraction, and (d) the estimated primaries with the pattern-based approach. Arrows point to locations where the pattern-based approach attenuates multiples significantly better than the adaptive subtraction.

One shortcoming of the pattern-recognition technique is that it relies on the Spitz approximation to provide a signal pattern model. By construction, the signal and noise filters will span different components of the data space. Therefore, the estimated primaries and multiples are uncorrelated. This fact proves that with the Spitz approximation, higher dimension filters are preferred because primaries and multiples have fewer chances to look similar.

Figure 15 shows an example where primaries are damaged by the pattern-based method. For instance, in Figure 15a we see at 2 a primary that is attenuated by the PEFs (Figure 15d) but is well preserved by the adaptive subtraction (Figure 15c). Here the primaries and multiples (Figure 15b) exhibit similar patterns, and the signal may have minimum energy. Using the Spitz approximation, event 2 is identified as noise and is removed. For event 3, it is quite difficult to say if multiples are removed in Figure 15d or if primaries are preserved in Figure 15c. Looking at the corresponding shot gathers did not help to make a decision; migration would probably help answer this question. Otherwise, event 4 is preserved with adaptive subtraction, and 1 and 5 are well recovered with the pattern-based approach.

This synthetic example indicates that the pattern-based approach tends to attenuate the multiples more accurately than adaptive subtraction when the multiples are not correlated with the primaries. This illustrates that higher-dimension filters should be preferred to better discriminate between noise and signal. The next section shows how the pattern-based approach performs on a field data set from the Gulf of Mexico.

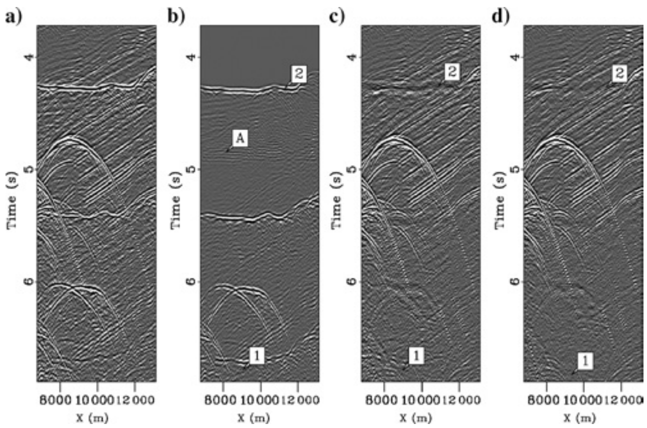


Figure 13. Constant-offset sections ($h$ = 4550 m) for (a) the input data, (b) the multiple model, (c) the estimated primaries with adaptive subtraction, and (d) the estimated primaries with the pattern-based approach. Arrow A points to aliasing effects as a result of the offset sampling of the shot gathers. The pattern-based approach attenuates the multiples better than the adaptive subtraction in 1 and 2.

## GULF OF MEXICO EXAMPLE

The pattern-based approach is tested on a 2D line from the Mississippi Canyon. This data set has been used extensively to benchmark multiple attenuation techniques (The Leading Edge, 1999). A stacked section of this data set is shown in Figure 16. Strong surface-related multiples are visible below 3 s. The tabular salt near the water bottom generates peg-legs in the data. The shot and receiver spacing is 26.6 m, and the first hydrophone is 100 m away from the source. The missing short-offset traces were interpolated on common-midpoint (CMP) gathers with a radon-based technique (Kabir and Verschuur, 1995) before multiple prediction. The actual separation is performed with the original traces only, without the near-offset data.

The multiple attenuation starts with the estimation of 3D PEFs from SRMP with one convolution and the Spitz

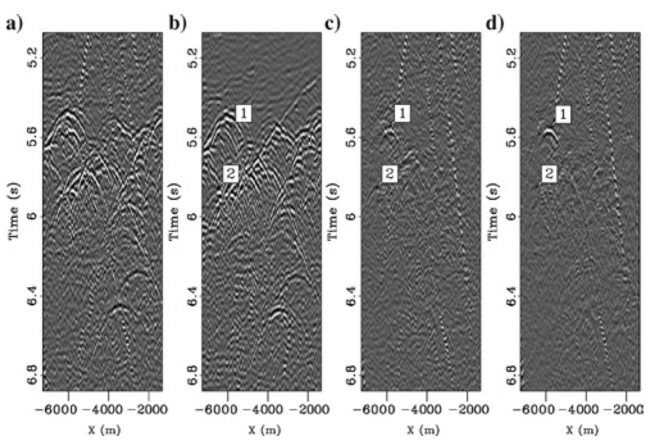


Figure 14. Constant-offset sections ($h$ = 3300 m) for (a) the input data, (b) the multiple model, (c) the estimated primaries with adaptive subtraction, and (d) the estimated primaries with the pattern-based approach. The 1 points to a primary that the pattern-recognition preserves very well. The 2 points to an event that is attenuated with the pattern-based approach but not with the adaptive subtraction in (c). Though not shown here, a close inspection of the corresponding shot gathers suggests that 2 is actually a multiple.

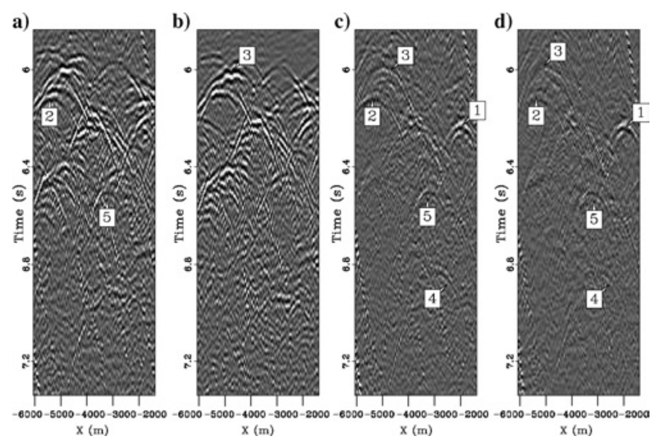


Figure 15. Constant-offset sections ($h$ = 5050 m) for (a) the input data, (b) the multiple model, (c) the estimated primaries with adaptive subtraction, and (d) the estimated primaries with the pattern-based approach. The 1 and 5 show events better preserved with the pattern-based method; 2 and 1 are better recovered with the adaptive subtraction; and 4 seems to point to a primary that the adaptive subtraction is able to save. Because the area is contaminated with strong multiples, it is difficult to know without a stratigraphic model if 3 is a primary.

approximation. Then multiples and primaries are separated according to their multidimensional spectra. Figure 17 displays the multiple attenuation result for one shot gather outside the salt boundaries ($x = 4500$ m). The multiple model in Figure 17c looks fairly accurate. The estimated primaries (Figure 17b) and multiples (Figure 17d) show that the subtraction is working at short and far offsets.

A second shot gather is shown in Figure 18 from inside the salt boundaries ($x = 12\,000$ m). Below the salt, diffracted multiples with bounces outside the acquisition grid affect the prediction (Kabir, 2003). However, the estimated primaries (Figure 18b) and multiples (Figure 18d) are again separated well. Note that second-order multiples below 5 s are also attenuated. At short offset, between 0 and 1000 m, one primary above the first water-bottom multiple has been removed in Figure 18d. This is because this event has a pattern similar to the water-bottom multiple.

Stacked sections of the input data, estimated multiples, and primaries are displayed in Figure 19. These stacks are located below the salt, as shown in Figure 19a. The estimated primaries in Figure 19b are still contaminated with diffractions (shown as M). The multiples are attenuated very well otherwise. Comparing the stack of the multiple model (Figure 19c) with the stack of the extracted multiples (Figure 19d) indicates that no prominent primaries have been attenuated.

The Gulf of Mexico example demonstrates that the pattern-based approach is an effective tool for multiple attenuation in complex geology. Although the multiple model obtained with SRMP presented some flaws (e.g., diffracted multiples), the proposed approach is able to attenuate the multiples while preserving the primaries.

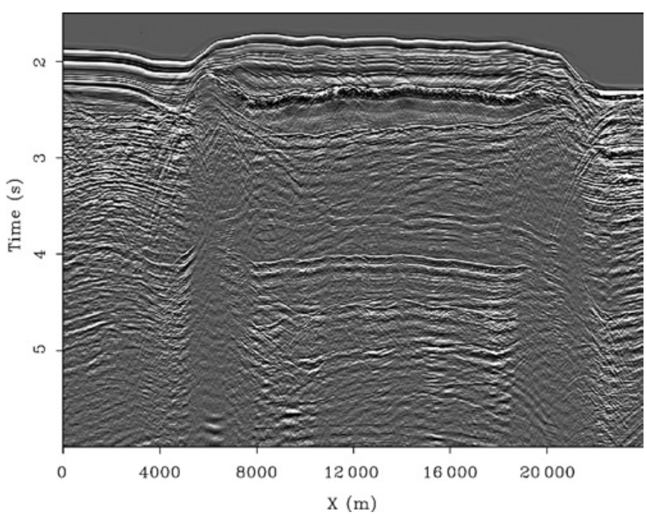


Figure 16. Stacked section of a 2D line from the Gulf of Mexico. Surface-related multiples appear below 3 s.

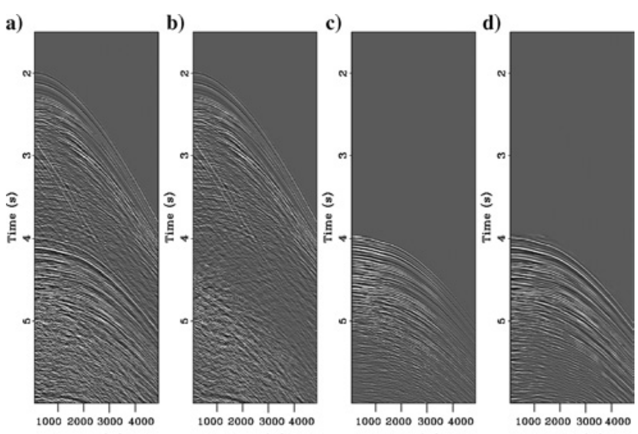


Figure 17. Shot gathers outside the salt boundaries at 4500 m for (a) the input data, (b) the estimated primaries, (c) the multiple model with SRMP, and (d) the estimated multiples.

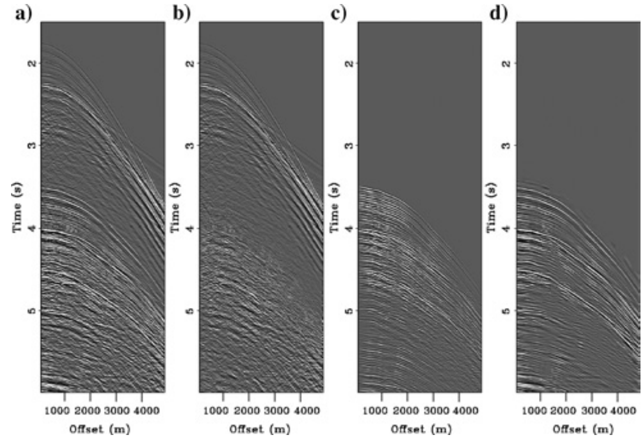


Figure 18. Shot gathers inside the salt boundaries at 12 000 m for (a) the input data, (b) the estimated primaries, (c) the multiple model with SRMP, and (d) the estimated multiples.

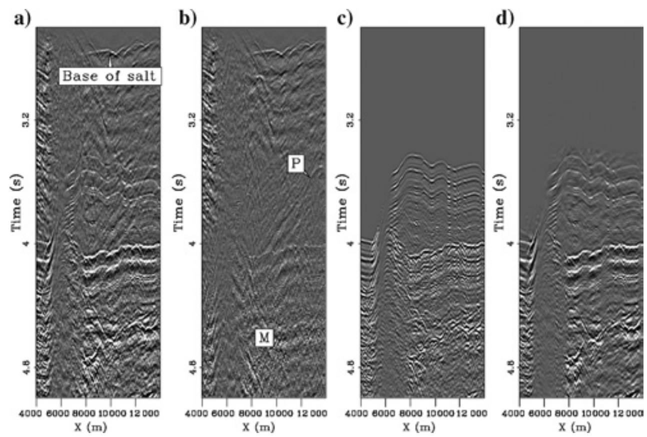


Figure 19. A comparison of stacked sections for (a) the input data, (b) the estimated primaries, (c) the multiple model with SRMP, and (d) the estimated multiples. Some primaries (P) in (b) are recovered while some multiples (M), not properly modeled by SRMP, are remaining.

## DISCUSSION AND CONCLUSIONS

Multiple attenuation can be cast as a problem where events are separated according to their patterns (e.g., multidimensional spectra). A pattern is made up of both the kinematic and amplitude information that PEFs can approximate. The pattern-based technique is a two-step procedure. First, nonstationary time-space domain PEFs are estimated for both primaries and multiples. These filters are estimated from a pattern model of the multiples usually computed with SRMP and a pattern model of the primaries computed with the Spitz approximation. The Spitz approximation assumes that the noise and signal are uncorrelated. Second, multiples are separated from the primaries in a least-squares sense according to their multidimensional spectra.

As illustrated with the Sigsbee2B data set, this approach has the potential to separate primaries and multiples very well as long as an accurate pattern model exists for PEF estimation. When no pattern model of the primaries exist, the Spitz approximation, which convolves the data with the noise PEFs, leads to a very good attenuation of the multiples if high-dimensional filters are used (i.e., three versus two dimensions). Indeed, primaries and multiples are less likely to look similar. Comparing adaptive and pattern-based subtraction indicates that the latter almost always removes the multiples better, except in areas where the primaries and multiples are correlated. An important property of the pattern-based approach is that it seems to cope well with modeling inadequacies.

Multiple attenuation can be viewed as a two-step process where multiples are first predicted and then subtracted. Both steps are important, but most of the efforts are usually concentrated on the prediction step and not the subtraction step. Since in practice it remains impossible to get a perfect multiple model because of the limitations of the acquisition geometry and interpolation/extrapolation techniques, new subtraction methods are needed. The pattern-based method presented in this paper is a successful tool for removing coherent energy in seismic data. This technique offers a viable alternative to adaptive subtraction by being less sensitive to errors in the multiple model. In addition, compared to adaptive subtraction, the pattern-based technique alleviates the strong assumption that primaries have minimum energy.

## ACKNOWLEDGMENTS

The author thanks SMAART-JV and BP for providing the two synthetic data sets used in this study, as well as WesternGeco for the Gulf of Mexico example. The author also thanks Paul Sava for providing the migration code and the sponsors of the Stanford Exploration Project for their financial support. The author is also grateful to the reviewers and Yonghe Sun for their constructive remarks on the manuscript.

## REFERENCES

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

Berkhout, A. J., and D. J. Verschuur, 1997, Estimation of multiple scattering by iterative inversion, part I: Theoretical considerations: Geophysics, **62**, 1586–1595.

Bishop, K., P. Keliher, J. Paffenholz, D. Stoughton, S. Michell, R. Ergas, and M. Hadidi, 2001, Investigation of vendor demultiple technology for complex subsalt geology: 71st Annual International Meeting, SEG, Expanded Abstracts, 1273–1276.

Brown, M., and R. Clapp, 2000, T-x domain, pattern-based groundroll removal: 70th Annual International Meeting, SEG, Expanded Abstracts, 2103–2106.

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

Claerbout, J., 1976, Fundamentals of geophysical data processing: Blackwell Scientific Publications.

———, 1992, Earth sounding analysis, processing versus inversion: Blackwell Scientific Publications.

———, 1998, Multidimensional recursive filters via a helix: Geophysics, **63**, 1532–1541.

Claerbout, J., and S. Fomel, 2004, Image estimation by example: Geophysical soundings image construction: ⟨http://sep.stanford.edu/sep/prof/gee/toc_html/⟩.

Crawley, S., 2000, Seismic trace interpolation with nonstationary prediction-error filters: Ph.D. dissertation, Stanford University.

Dragoset, B., and S. MacKay, 1993, Surface multiple attenuation and subsalt imaging: 63rd Annual International Meeting, SEG, Expanded Abstracts, 1099–1102.

Dragoset, W. H., and Z. Jericevic, 1998, Some remarks on surface multiple attenuation: Geophysics, **63**, 772–789.

Fomel, S., 2002, Applications of plane-wave destruction filters: Geophysics, **67**, 1946–1960.

Guitton, A., 2002, Coherent noise attenuation using inverse problems and prediction-error filters: First Break, **20**, 161–167.

Guitton, A., and G. Cambois, 1998, Prestack elimination of complex multiples: A Gulf of Mexico subsalt example: 68th Annual International Meeting, SEG, Expanded Abstracts, 1329–1332.

Guitton, A., and D. Verschuur, 2004, Adaptive subtraction of multiples using the $L_1$-norm: Geophysical Prospecting, **52**, 27–38.

Kabir, M., 2003, Weighted subtraction for diffracted multiple attenuation: 73rd Annual International Meeting, SEG, Expanded Abstracts, 1941–1944.

Kabir, M. M. N., and D. J. Verschuur, 1995, Restoration of missing offsets by parabolic radon transform: Geophysical Prospecting, **43**, 347–368.

Leading Edge, The, 1999. Special edition on multiple attenuation techniques, **18**.

Lokshtanov, D., 1999, Multiple suppression by data-consistent deconvolution: The Leading Edge, **18**, 115–119.

Manin, M., and S. Spitz, 1995, 3D attenuation of targeted multiples with a pattern recognition technique: 57th Annual International Meeting, European Association of Geoscientists and Engineers, Extended Abstracts, B046.

Margrave, G. F., 1998, Theory of nonstationary linear filtering in the Fourier domain with application to time-variant filtering: Geophysics, **63**, 244–259.

Matson, K. H., D. Paschal, and A. B. Weglein, 1999, A comparison of three multiple-attenuation methods applied to a hard water-bottom data set: The Leading Edge, **18**, 120–126.

Mersereau, R. M., and D. E. Dudgeon, 1974, The representation of two-dimensional sequences as one-dimensional sequences: IEEE Transactions on Acoustic, Speech and Signal Processing, **22**, 320–325.

Paffenholz, J., B. McLain, J. Zaske, and P. Keliher, 2002, Subsalt multiple attenuation and imaging: Observations from the Sigsbee2b synthetic data set: 72nd Annual International Meeting, SEG, Expanded Abstracts, 2122–2125.

Popovici, A. M., 1996, Prestack migration by split-step DSR: Geophysics, **61**, 1412–1416.

Rickett, J., 2001, Spectral factorization of wavefields and wave operators: Ph.D. dissertation, Stanford University.

Rickett, J., A. Guitton, and D. Gratwick, 2001, Adaptive multiple subtraction with non-stationary helical shaping filters: 63rd Annual International Meeting, European Association of Geoscientists and Engineers, Extended Abstracts, P167.

Soubaras, R., 1994, Signal-preserving random noise attenuation by the *f-x* projection: 64th Annual International Meeting, SEG, Expanded Abstracts, 1576–1579.

———, 2001, Dispersive noise attenuation for converted wave data: 71st Annual International Meeting, SEG, Expanded Abstracts, 802–805.

Spitz, S., 1999, Pattern recognition, spatial predictability, and subtraction of multiple events: The Leading Edge, **18**, 55–58.

Technical University of Delft, 2002. Publicly released 'SMAART' data sets, available to the geophysical industry. Updated 17 October 2002, accessed 28 February 2005 ⟨http://www.delphi.tudelft.nl/SMAART/S2Breadme.htm⟩.

Treitel, S., P. R. Gutowski, and D. E. Wagner, 1982, Plane-wave decomposition of seismograms: Geophysics, **47**, 1375–1401.

van Dedem, E. J., 2002, 3D surface-related multiple prediction: Ph.D. dissertation, Delft University of Technology.

Verschuur, D. J., A. J. Berkhout, and C. P. A. Wapenaar, 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, 1166–1177.

Weglein, A. B., 1999, Multiple attenuation: An overview of recent advances and the road ahead (1999): The Leading Edge, **18**, 40–44.

Weglein, A. B., F. A. Gasparotto, P. M. Carvalho, and R. H. Stolt, 1997, An inverse scattering series method for attenuating multiples in seismic reflection data: Geophysics, **62**, 1975–1989.

Wiggins, J. W., 1988, Attenuation of complex water-bottom multiples by wave equation-based prediction and subtraction: Geophysics, **53**, 1527–1539.