Stanford Exploration Project, Report 72, December 18, 1997, pages 1–200

Multiple suppression in the velocity domain

Carlos A. Cunha Filho and Jon F. Claerbout[1]

keywords: *velocity, deconvolution, multiples, synthetic data*

ABSTRACT

Multiple suppression in the velocity domain is often accomplished by a filtering process that removes the events inside a selected window in that domain. Alternately, the suppression goal can be formulated as an optimization problem in which we try to find a model that, when added to the data, minimizes the power of the output in some region of the velocity domain. Defining this region is of fundamental importance for the success of the method, that is, for attaining maximum suppression of multiples with minimum interference with the primaries. The predictive character of reverberation events in the velocity domain can be efficiently used to automatically design these multiple-related windows. Tests with synthetic data show the efficacy and robustness of the method, which is able to preserve primary events with stacking velocities inside the multiple velocity range.

INTRODUCTION

Thorson and Claerbout (1985) extensively studied the use of velocity stacks, of their pseudo-inverse and stochastic inverse, to solve the problem of missing data restoration, and to design a more robust stacking operator for data with missing traces. They stressed the importance of estimating the inverse covariance matrix of the model, and provided a theoretical basis for justifying the use of an iterative process to update the estimation. Harlan (1988) modeled primaries and multiples using two beams of velocity functions that were manually designed using the semblance spectrum as a guide. The modeled multiples were then convolved with a wavelet and subtracted from the original data. Some limitations of this approach include the absence of diffractions in the modeled data, the assumption that multiples and primaries are restricted to non-overlapping thin regions, and the need for human interaction in the design of the beams.

Although the velocity space is a good place to distinguish multiple from primary reflections, attempts to suppress energy associated with multiples in this domain

[1]**email:** not available

are usually spoiled not only by the cost of the method, but also by the presence of artifacts and significant changes in the amplitude and phase of primary events. Filtering a region in the velocity domain and transforming back to the space domain would be an expensive procedure if an exact (pseudo)inverse were to be used, because the velocity transform lacks any of the special structures that would allow for a less expensive exact inverse. Fortunately, only a few iterations are required to obtain a very good approximation for the least-squares inverse. Also, the filtering can be formulated as an optimization problem (power minimization) that can be solved by a fast-converging conjugate gradients method. In general, the method converges after a few iterations, which makes it affordable for routine processing.

The other undesirable aspects, artifacts and changes in the primaries, are caused by the fact that although primaries and multiples are concentrated, respectively, in the high- and low-velocity regions, the contribution from the near offset traces spreads through the whole velocity domain. Attempts to separate the velocity domain into two or three regions (filter design) involve two adverse factors: complete dependence on the designer's judgement, and the unavoidable truncation of both primary and multiple events. To overcome these limitations we use the predictability of the multiples in the velocity domain to automatically design a window function that separates primaries from multiples. The method uses a multichannel prediction error filter, which is also designed using conjugate gradients.

GENERAL FORMULATION OF THE METHOD

Multiple suppression in the velocity domain can be formulated as an optimization problem. The aim is to find some "model" in the space–time domain, which, when added to the original CMP gather, minimizes the power within a pre-specified window in the velocity–time domain. The precise mathematical formulation follows.

Let $\mathbf{g}$ represent the original CMP gather, $\mathbf{v}$ its velocity transform, and N the velocity transform operator

$$\mathbf{v} = \mathbf{N}\mathbf{g}. \tag{1}$$

Then, we want to find the unknown model $\hat{\mathbf{g}}$ that minimizes the objective function

$$\mathcal{O} = ||\mathbf{W}\mathbf{N}(\mathbf{g} + \hat{\mathbf{g}})||^2, \tag{2}$$

where $\mathbf{W}$ is a windowing operator that selects the region of the velocity spectrum in which the energy of the multiples is dominant. To find the solution $\hat{\mathbf{g}}$ that minimizes $\mathcal{O}$ we need to solve

$$\frac{d\mathcal{O}}{d\hat{\mathbf{g}}} = 0$$

leads to

$$\begin{aligned} -\mathbf{W}\mathbf{N}\mathbf{g} &= \mathbf{W}\mathbf{N}\hat{\mathbf{g}} \\ -\mathbf{W}\mathbf{v} &= \mathbf{W}\mathbf{N}\hat{\mathbf{g}} \\ \mathbf{d} &= \mathbf{A}\mathbf{m}, \end{aligned} \tag{3}$$

which is a system of linear equations to be solved for the unknown $\mathbf{m} = \hat{\mathbf{g}}$, where the data $\mathbf{d}$ is the windowed velocity transform of the original data. The system can be over- or under determined, depending on how the operators $\mathbf{W}$ and $\mathbf{N}$ are defined, but it will usually be large enough, since $\mathbf{m}$ is typically 10^5 points long. Of course $\hat{\mathbf{g}} = -\mathbf{g}$ is one of the solutions of equation 4. However the window operator $\mathbf{W}$ "transfer" the events that map outside the window (which are predominantly primaries) to the null space. The use of conjugate gradients (CG) seems to be a suitable choice for solving this problem especially because a few iterations would not only be enough but would also contribute to the stabilization of the process, when noise is present. We have chosen to use a very simple form for the velocity transform operator $\mathbf{N}$, the nearest neighbor normal moveout-stack. The advantages of choosing this operator are its computational speed and the simplicity of its transpose $\mathbf{N}'$, which is needed in the conjugate algorithm (Claerbout, 1991).

Testing the method on a synthetic dataset

To illustrate the method and compare the results from the different schemes described here, we have generated a simple synthetic dataset using an elastic modeling program. Figure **??**-a shows the reverberation-free data obtained with the elimination of the surface (air-water) boundary. The pure P wave reflections are the events at approximately 0.4, 1.6, and 3.1 seconds. Some converted PSPP and PSSP reflections are also visible at the far offsets between 2.0 and 3.0 seconds, and at about 4.5 seconds. The more realistic dataset (with the surface boundary present), which has been used as input for the several multiples suppression schemes, is shown in Figure **??**-b. Figure **??**-a shows the semblance spectrum of the data overlaid with the manually designed window in which reverberations are dominant. The result of applying the basic scheme described by equation (4) to the data in **??**-b is shown in Figure **??**-b.

WEIGHTED LEAST-SQUARES

Linear scheme

As indicated in the examples, the basic scheme of equation (4) suppress part of the multiples but also interferes with the nearby primaries. Furthermore, the converted waves are suppressed in conjunction with the multiples. The deterioration of the primaries is caused by the fact that some of energy associated with the primaries extends over the windowed region and the operator, as defined in equation (4), acts uniformly over all the window. A possible solution for that is to use a weighting function, designed to privilege the more energetic regions inside the window, which are more likely related to multiples. Equation (4) is replaced in this case by

$$\begin{aligned} -\mathbf{WJNg} &= \mathbf{WJN\hat{g}} \qquad (4) \\ -\mathbf{WJv} &= \mathbf{WJN\hat{g}}, \end{aligned}$$

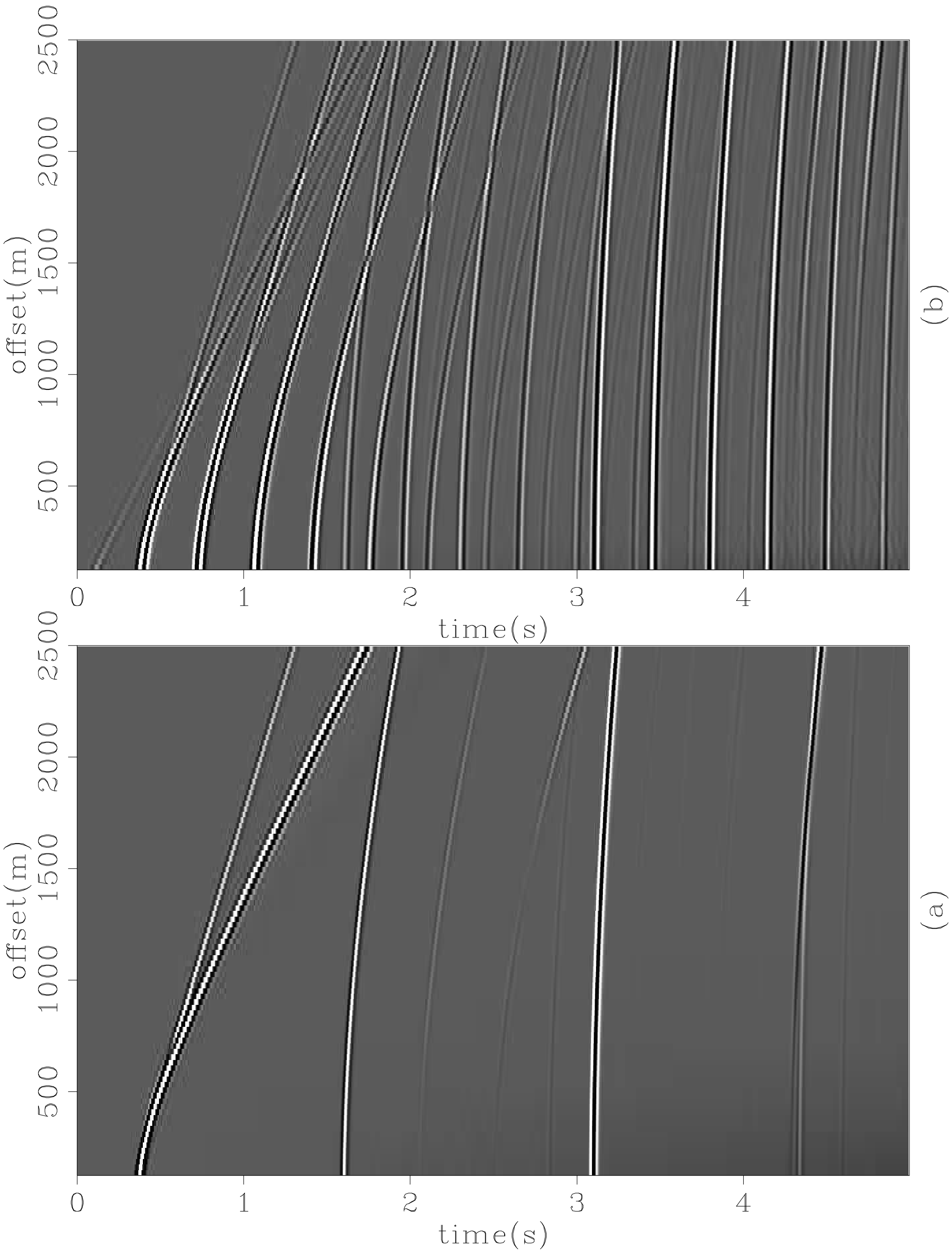

Figure 1: *(a)* Synthetic data, free of surface-related reverberation. *(b)* Synthetic data with surface-related reverberations. carlos1-syntdata [ER]

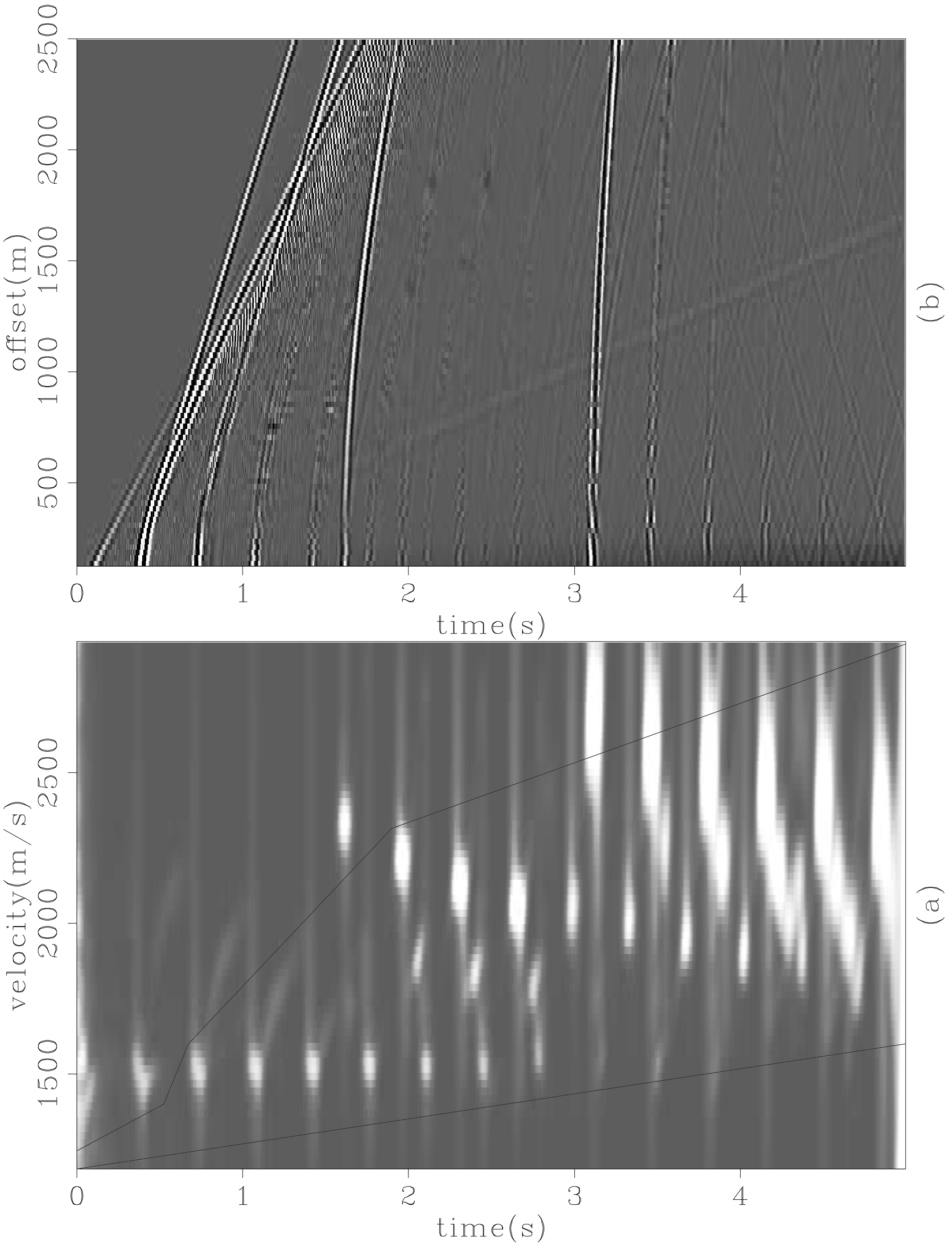

Figure 2: *(a)* Semblance panel of the data of Figure **??**-b, overlaid with the manually designed window function. *(b)* Result of applying the scheme defined by equation (4) to the same data. carlos1-semrm0 [ER]

where $\mathbf{J}$ is a diagonal weighting operator chosen to represent an approximation for the noise (multiples) covariance matrix. We have chosen the operator $\mathbf{J}$ to be the smoothed semblance velocity spectrum of the original data (Figure **??**-a).

Nonlinear scheme

The first few iterations of the CG algorithm lower the energy in the regions where the semblance of the original data was high, but the remaining problem is that the subsequent iterations still focus to much attention in the same places because the weighting function stays the same. To overcome this restriction we need a variable weighting operator that is reevaluated after a given number of steps of the CG solver, resulting in the reformulation of equation (4) as an iterative linearization of a nonlinear problem. Each external step i of the CG solver will work on the residuals $\mathbf{d^i}$ defined as

$$\Delta \mathbf{d^i} = -\mathbf{W}\mathbf{J^{i-1}}(\mathbf{v} - \mathbf{N}\,\hat{\mathbf{g}}^{\mathbf{i-1}}), \tag{5}$$

where

$$\mathbf{J^{i-1}} = \text{semblance}\{\mathbf{g^{i-1}} = \mathbf{g} + \hat{\mathbf{g}}^{\mathbf{i-1}}\}.$$

Figure **??** show the results of the linear and nonlinear semblance weighting schemes. Although with less artifacts in the shallow reflections, and with less deterioration in the primaries (notably the near offset of the primary at 1.6 seconds) relative to the non-weighted algorithm, they fail to eliminate the set of reverberations in the shallow part of the gather. Furthermore, at the near offset traces the multiples are not suppressed, as is usually the case in velocity-domain-based methods.

USING PREDICTION TO DESIGN THE WINDOW FUNCTION

The partial failure of using the windowing operator to isolate the multiples is caused mainly by the presence of reverberating energy outside the window range. This failure becomes critical for low-order reverberations, which are located close to the primaries and are therefore near the boundary of the window. A possible solution would be the use of an offset-dependent weighting function in the transform, which would favor the contribution of more velocity-sensitive large offsets. As a result, the velocity stack would have a better resolution, with less spreading of energy outside the central velocity. However this procedure would also imply a decreased resolution in the near-offset traces where the multiples are already hard to remove. To overcome this limitation, we need a way to *separate primaries and multiples in the whole range covered by the velocity domain.* A look at the semblance spectrum (Figure **??**) shows that the reverberation trends are aligned semi-vertically in a reasonably predictable way. It is possible, then, to design a short operator that predicts the reverberation of order $n+1$ from the reverberation of order n.

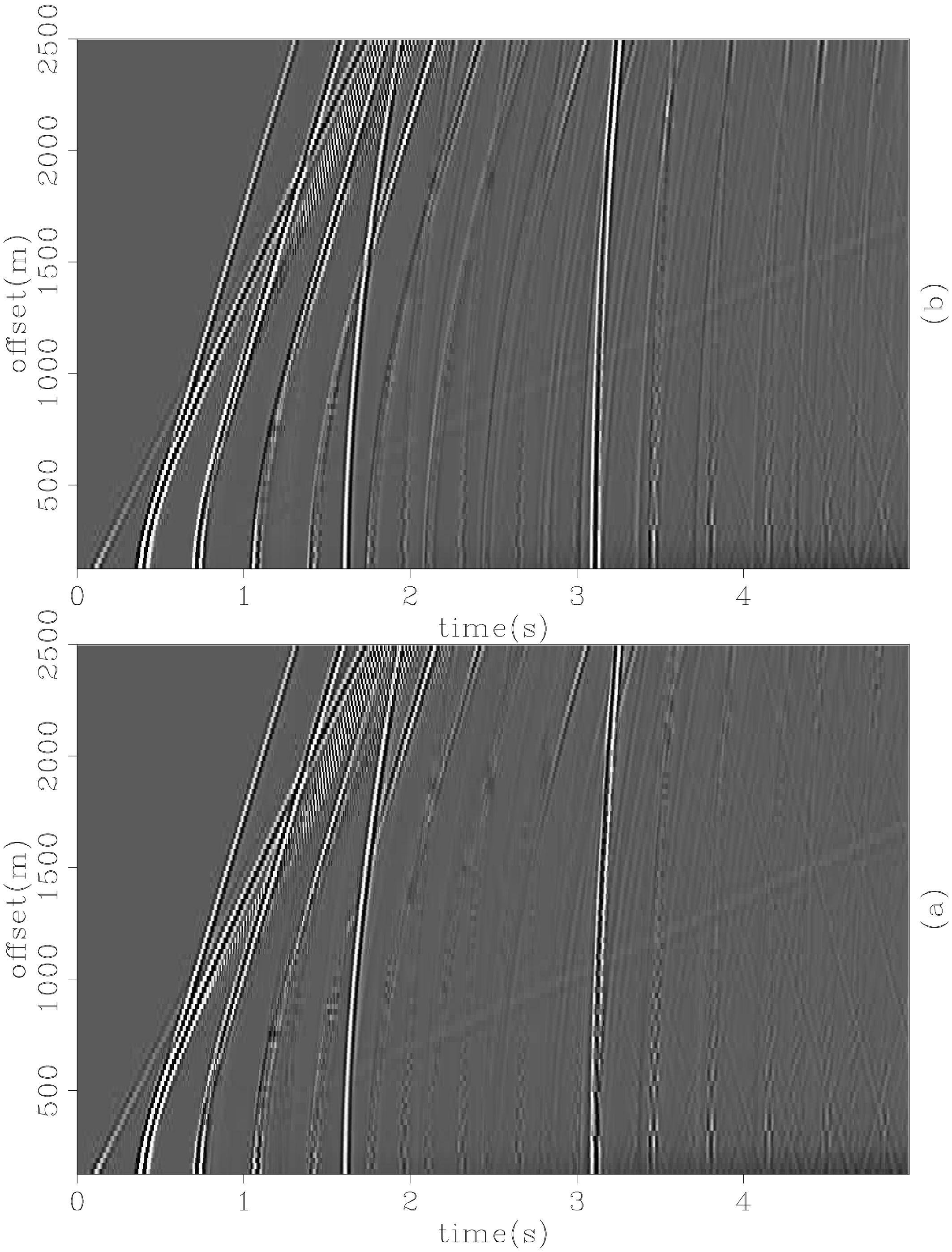

Figure 3: Results of *(a)* the multiple suppression using the linear pre-weighted scheme described by equation (4) and *(b)* multiple suppression using the nonlinear preweighted scheme described by equation (5). In both cases the synthetic dataset of Figure **??** is used as input. carlos1-multw01 [ER]

The prediction error has the semblance of the primaries almost unaffected and the semblance of the reverberations reduced, while the prediction itself is predominantly formed by the semblance of the reverberations. As a result, the prediction can be used as the weighting operator in the **CG** regression. However, instead of using the predicted semblance as a weighting function, we use it to define a window function according to

$$\begin{aligned} \mathbf{W_p} &= 1 \qquad \text{if} \quad \mathbf{J_p} \geq J_{cut} \\ \mathbf{W_p} &= 0 \qquad \text{if} \quad \mathbf{J_p} < J_{cut}, \end{aligned}$$

where $\mathbf{J_p}$ is the predicted semblance, and J_{cut} defines the cutting value for the window. The actual window function also has a tapering zone to make the transition smoother.

Designing the prediction operator

Figure **??** shows the form of the desired prediction-error operator in the velocity domain. The prediction lags in the velocity and time directions and the filter size are controlled by the particular characteristics of the dataset, following the same basic designing rules as with the standard 1-D prediction error filter. The coefficients a_{ij} are defined by a power minimization criterion (Claerbout, 1991). Therefore, if $\mathbf{S}$ is the semblance of the original data and $\mathbf{F}$ is the prediction filter, the coefficients of the filter are obtained by minimizing the L_2 error of equation (6), as follows:

$$\mathbf{F} \star \mathbf{S} = 0. \tag{6}$$

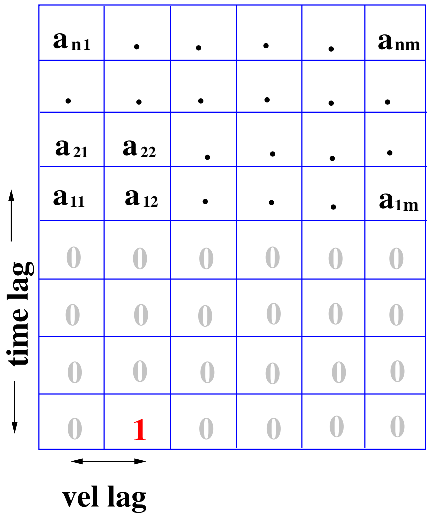

Figure 4: The prediction-error filter for suppressing the regions of the velocity semblance spectrum that are associated with primary reflections. carlos1-predoper [NR]

Once the filter coefficients are defined, the weighting operator $\mathbf{J_p}$ is obtained by

$$\mathbf{J_p} = \mathbf{S} - \mathbf{F} \star \mathbf{S}. \tag{7}$$

Because of the large size of the problem to be solved in equation (6) and since the semblance is a smooth function of time, it is convenient to subsample S in time

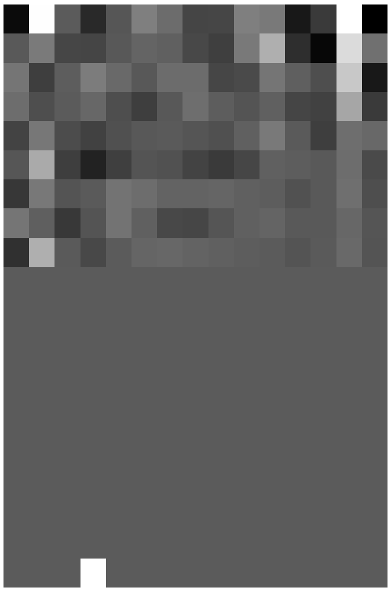

Figure 5: The estimated prediction-error filter for the synthetic data of Figure **??**-b. carlos1-realoper [ER]

before solving equation (6) and applying equation (7). The resulting operator $\mathbf{J_p}$ is then interpolated to the original time sampling interval. The prediction-error filter estimated for the synthetic data of Figure **??**-b is shown in Figure **??**.

Figure **??** compares the original semblance $\mathbf{S}$ with the result of applying the prediction operator $\mathbf{J_p}$ to it for the synthetic dataset of Figure **??**. It is impressive that the filter was able to correctly predict the different multiples patterns present in the data. None of the primaries are visible on the output. Figure **??**-a shows the window function $\mathbf{W_p}$ constructed from $\mathbf{J_p}$, and Figure **??**-b shows the semblance spectrum of the synthetic data after 15 iterations of the optimization process using $\mathbf{W_p}$ as the window function. In **??**-b the events associated with reverberations are barely visible. Most importantly, the parts of the spectrum associated with the converted waves were preserved, as show the events at 2.05, 2.5, and 4.3 seconds.

A comparison of the output of the optimization process with the "ideal" multiple free data (Figure **??**) shows that not only have the multiples been substantially attenuated, but also the amplitude and phase of the primaries were correctly preserved, including the weak, low-velocity converted waves.

CONCLUSIONS

The velocity domain is a suitable place to separate reverberations associated with the water layer. Using windows to separate primaries and multiples fails to suppress low-order reverberations and degrades primary reflections. Better results can be achieved with the use of a prediction-error operator to construct a primaries-free window function. Some advantages of this method are the substantial attenuation of multiples even at the near-offset traces, preservation of the amplitude and phase of

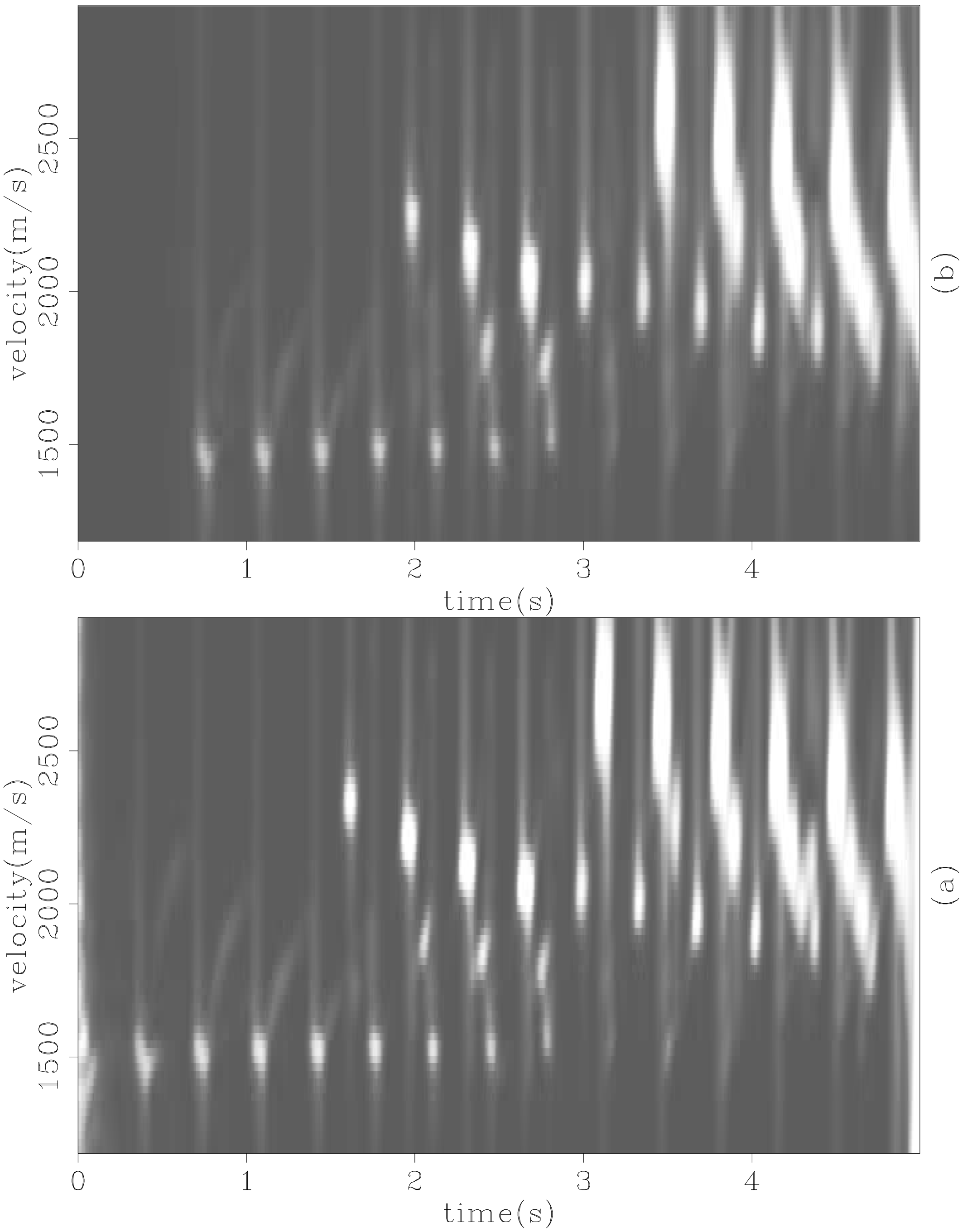

Figure 6: *(a)* The smoothed semblance spectrum of the data of Figure **??** and *(b)* the output of the prediction operator applied to it. carlos1-sempred [ER]

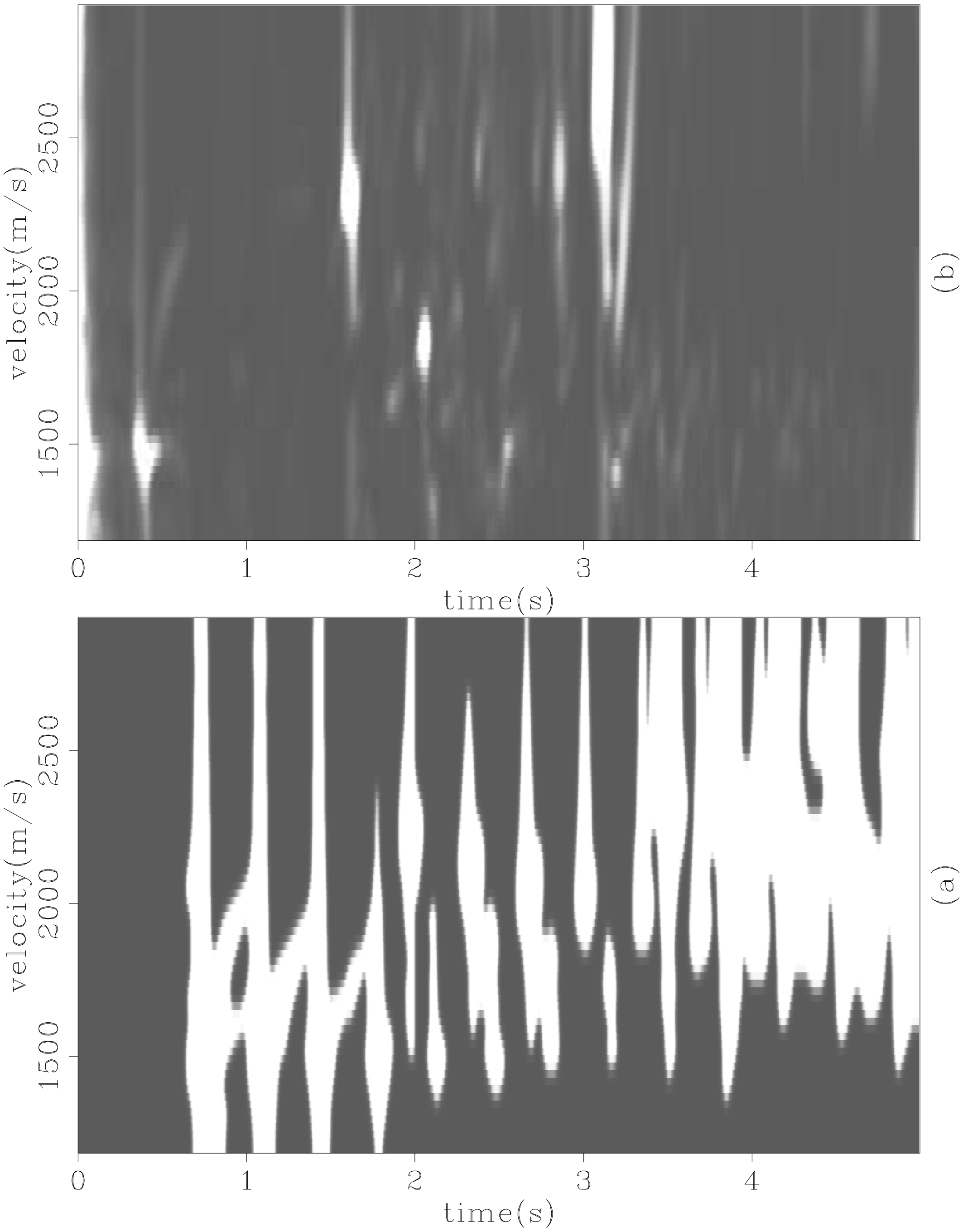

Figure 7: *(a)*The window function constructed from the predicted semblance spectrum shown in Figure **??**-b. *(b)* The semblance spectrum of the output of the multiples suppression scheme, using (a) as the window function. carlos1-semwind [ER]

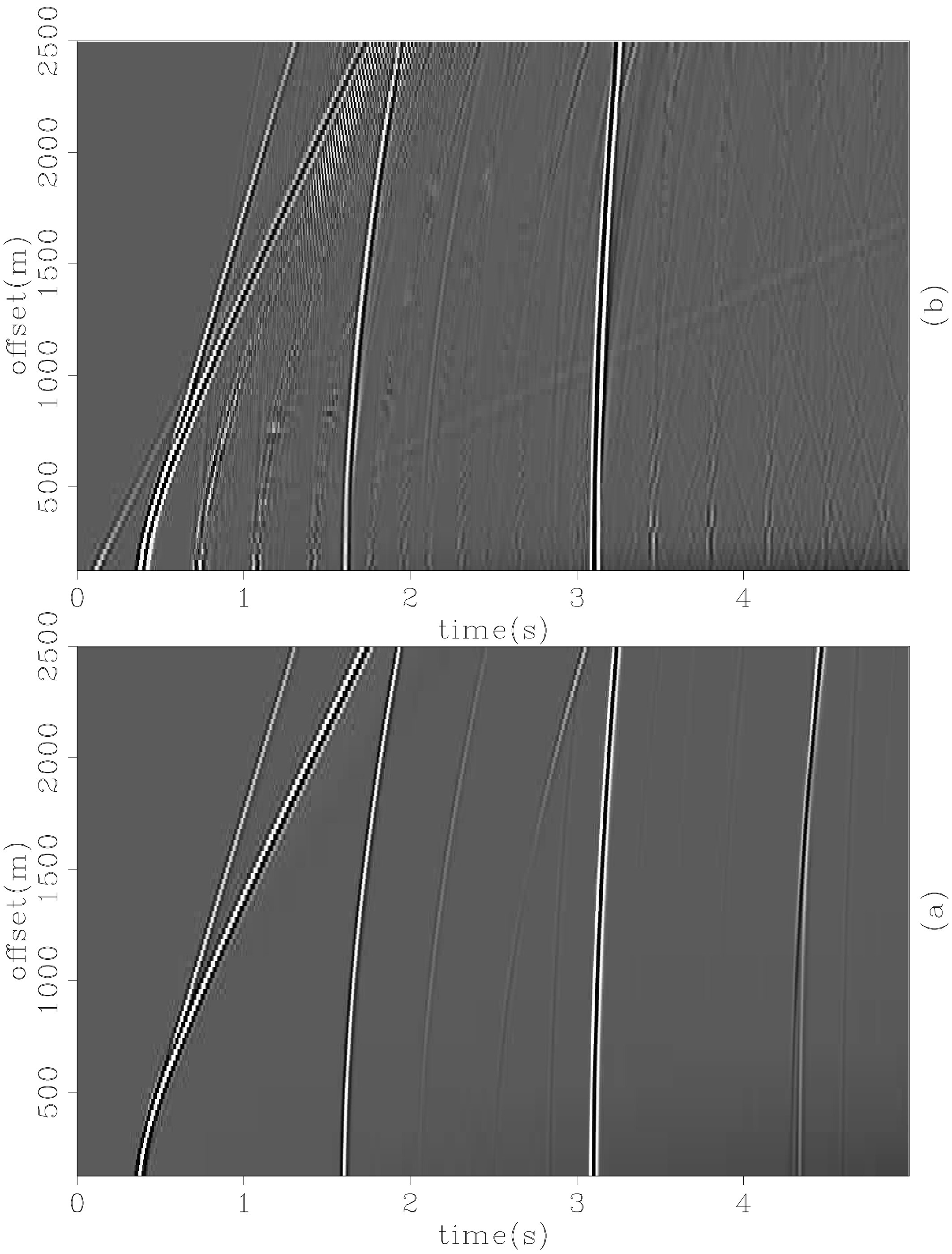

Figure 8: *(a)* Synthetic data, free of surface related reverberation (same as Figure **??**-a). *(b)* The output of the multiples suppression scheme applied to the synthetic data of Figure **??**-b, when the prediction-designed function of Figure **??**-a is used as the window operator. carlos1-predrm [ER]

the primaries, preservation of weak primary events with stacking velocities similar to the multiples, and the complete automatization of the process.

REFERENCES

Claerbout, J.F., 1991, Earth Soundings Analysis: Processing versus Inversion: Blackwell Scientific Publications.

Harlan, W.S., 1988, A model for phase balancing and velocity filtering: SEP–**57**, 313–346.

Thorson, J.R., and Claerbout, J.F., 1985, Velocity-stack and slant-stack stochastic inversion: Geophysics, **50**, 2727–2741.