Stanford Exploration Project, Report 111, June 9, 2002, pages 157–170

Stanford Exploration Project, Report 111, June 9, 2002, pages 157–170

Adaptive subtraction of multiples with the ℓ^1-norm

Antoine Guitton and Eric Verschuur[1]

ABSTRACT

The estimation of shaping filters with the ℓ^1-norm as opposed to the ℓ^2-norm leads to a proper attenuation of multiples when significant amplitude discrepancies exist between multiples and primaries. The actual method implemented is the fairly standard iteratively re-weighted least-squares method which is an excellent approximation to ℓ^1. Synthetic and field data results illustrate the advantages of the ℓ^1-norm.

INTRODUCTION

A classical approach for attenuating multiples consists of building a multiple model (Verschuur et al., 1992) and adaptively subtracting this model from the multiple infested-data by estimating shaping filters (Dragoset, 1995; Liu et al., 2000; Rickett et al., 2001). The estimation of the shaping filters is usually done in a least-squares sense making these filters relatively easy to compute. In some cases, however, a least-squares criterion can lead to undesirable artifacts. This happens when, for example, the relatively strong primaries are surrounded by multiples, such that the filter tends to distort primary energy as well.

In this paper we show that the estimation of shaping filters with the ℓ^1-norm gives better results than with the ℓ^2-norm when multiples and primaries have noticeable amplitude differences. We first illustrate this with a simple 1D problem that highlights the limits of the least-squares approach. In a second synthetic example, we attenuate internal multiples and show that the ℓ^1-norm gives far better results than does ℓ^2. To finish, we utilize shaping filters on a multiple contaminated gather from a seismic survey showing that the ℓ^1-norm leads to a substantial attenuation of the multiples.

A SIMPLE 1D PROBLEM

In this section, we demonstrate on a 1D problem that the attenuation of multiples with least-squares adaptive filtering is not effective when amplitude differences exist between primaries and multiples. This simple example helps us to better understand the behavior of our adaptive scheme in more complicated cases.

[1]**email:** antoine@sep.stanford.edu, D.J.Verschuur@CTG.TUDelft.nl

Shaping filters and the ℓ^2-norm

h This section illustrates some limitations of the ℓ^2-norm for the estimation of shaping filters. In figure 1, we display a very simple 1D problem. On the top we have four events corresponding to one primary (on the left) and three multiples (on the right). Note that the primary has higher amplitude than the multiples. On the bottom we show a multiple model that exactly corresponds to the real multiples. Our goal is to estimate one shaping filter **f** that minimizes the objective function

$$e(\mathbf{f}) = \|\mathbf{d} - \mathbf{M}\mathbf{f}\|_2^2, \tag{1}$$

where **M** is the matrix representing the convolution with the time series for the multiple model (Figure 1b) and **d** the time series for the data (Figure 1a).

Now, if we estimate the filter **f** with enough degrees of freedom (enough coefficients) to minimize equation (1), we obtain for the signal Figure 2a, and for the noise Figure 2b. The estimated signal does not resemble the primary in Figure 1a. We show the corresponding shaping filter in Figure 3. This filter is not the single spike at $lag = 0$ that we desire. The problem stems from the least-squares criterion which yields an estimated signal that has, by definition, minimum energy. In this 1D case, the total energy in the estimated signal (Figure 2a) is $e = 2.4$, which is less than the total energy of the primary alone ($e = 4$). This is the fundamental problem if we use the ℓ^2-norm to estimate the shaping filter. In the next section, we show that the ℓ^1 norm should be used if amplitude differences exist between primaries and multiples.

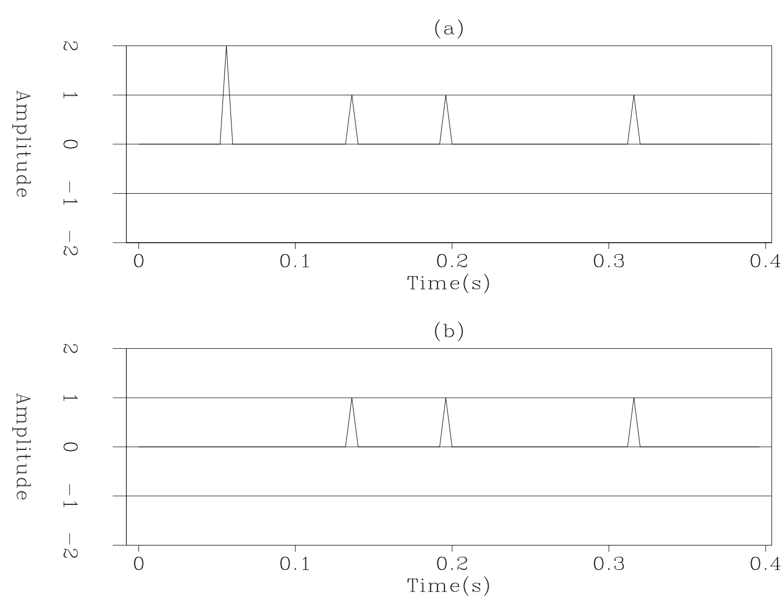


Figure 1: (a) The data with one primary on the left, and three multiples on the right. (b) The multiple model that we want to adaptively subtract from (a). antoine1-datmul [ER]

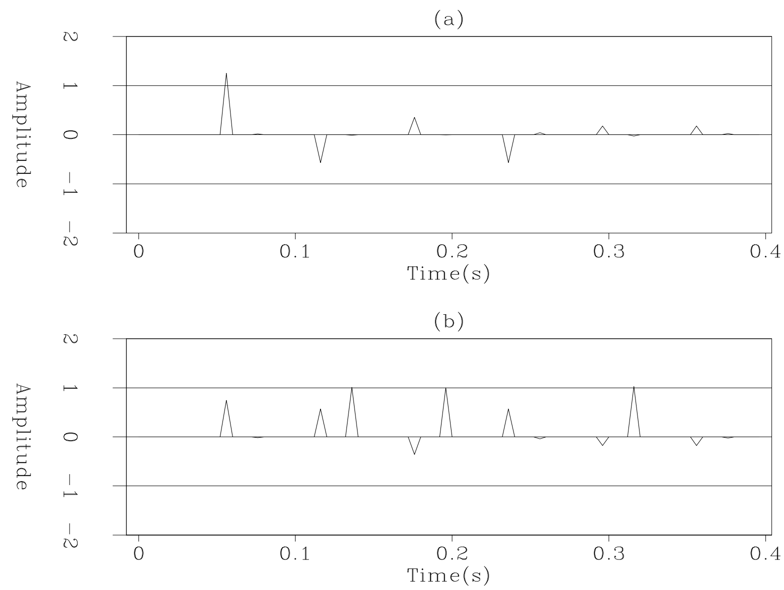
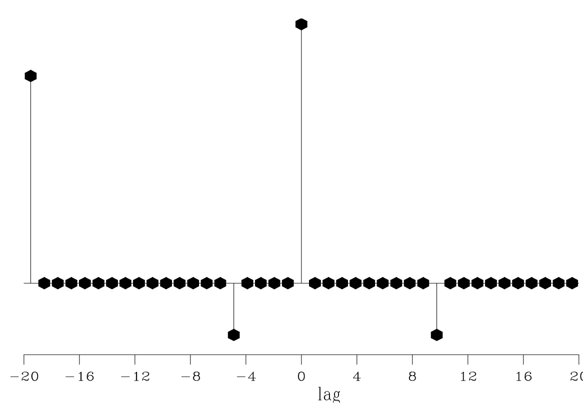


Figure 2: (a) The signal estimated with the ℓ^2-norm. (b) The noise estimated with the ℓ^2-norm. antoine1-1Dl2 [ER]

Figure 3: Shaping filter estimated for the 1D problem with the ℓ^2-norm. This filter is not a single spike at $lag = 0$. antoine1-filterl2 [ER]



Shaping filters and the ℓ^1-norm

We prove that the ℓ^1-norm solves the problem highlighted in the preceding section. Now our goal is to estimate one shaping filter **f** that minimizes the objective function

$$e(\mathbf{f}) = |\mathbf{d} - \mathbf{Mf}|_1. \tag{2}$$

To achieve this, the shaping filter is estimated iteratively using a nonlinear conjugate gradient solver (NLCG) as described in Claerbout and Fomel (2001). The objective function we actually minimize is

$$e(\mathbf{f}) = \|\mathbf{W}(\mathbf{d} - \mathbf{Mf})\|_2^2, \tag{3}$$

with

$$\mathbf{W} = \mathbf{diag}\left(\frac{1}{(1 + r_i^2/\epsilon^2)^{1/4}}\right), \tag{4}$$

where r_i is the residual for one component of the data space, and ϵ a constant we choose a priori. Equation (3) is minimized with the standard iteratively re-weighted least-squares approach (Nichols (1994); Bube and Langan (1997); Guitton (2000)) The objective function in equation (3) amounts to the ℓ^1 measure when r_i/ϵ is large and amounts to the ℓ^2 measure when $r_i/\epsilon << 1$ with a smooth transition between the two.

In Figure 4, we display the result of the adaptive subtraction when the ℓ^1-norm is utilized to estimate the shaping filter [equation (3) with a small ϵ]. The estimated signal in Figure 4a is perfect, and so is the estimated noise. It is easy to check that the energy in Figure 4a ($e = 2$) is less than the energy in Figure 2a ($e = 3.2$) if we use the ℓ^1 norm. Figure 5 shows the shaping filter associated with the ℓ^1-norm. This filter is a spike at $lag = 0$. This simple 1D example demonstrates that the ℓ^1 should be utilized each time significant amplitude differences exist between multiples and primaries. In the next section, we show another synthetic example where internal multiples are attenuated.

ATTENUATION OF INTERNAL MULTIPLES

In this section we illustrate the efficiency of the ℓ^1-norm when internal multiples are attenuated in 2D.

The synthetic data

Figure 6a shows a synthetic shot gather for a 1D medium. This gather is corrupted with internal multiples only. In Figure 6b, we display the internal multiple model obtained using the CFP approach (Berkhout and Verschuur, 1999). This internal multiple model is perfect and could be directly subtracted from the data in Figure 6a. Note that the amplitude of the

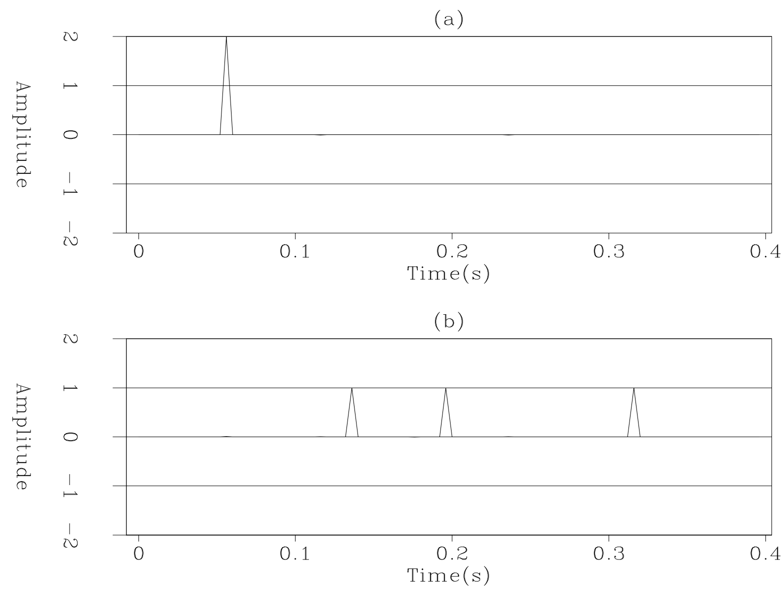


Figure 4: (a) The signal estimated with the ℓ^1-norm. (b) The noise estimated with the ℓ^1-norm. antoine1-1Dl1 [ER]

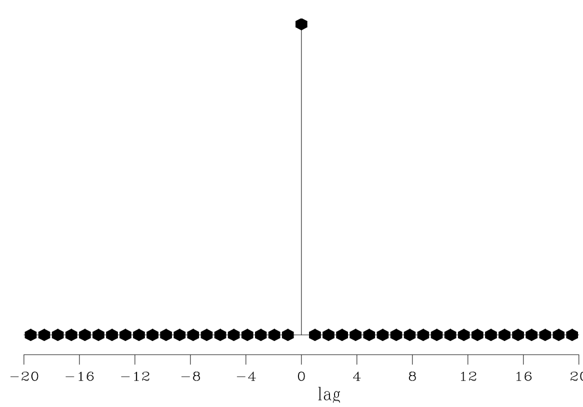


Figure 5: Shaping filter estimated for the 1D problem with the ℓ^1-norm. This filter is a single spike at $lag = 0$. antoine1-filterl1 [ER]

internal multiples is significantly less than the amplitude of the primaries, making the ℓ^2-norm unsuitable for estimating the shaping filters. Figure 7 displays the histograms of both the data and the internal multiples. The narrow peak of the noise indicates that the ℓ^1-norm should be used.

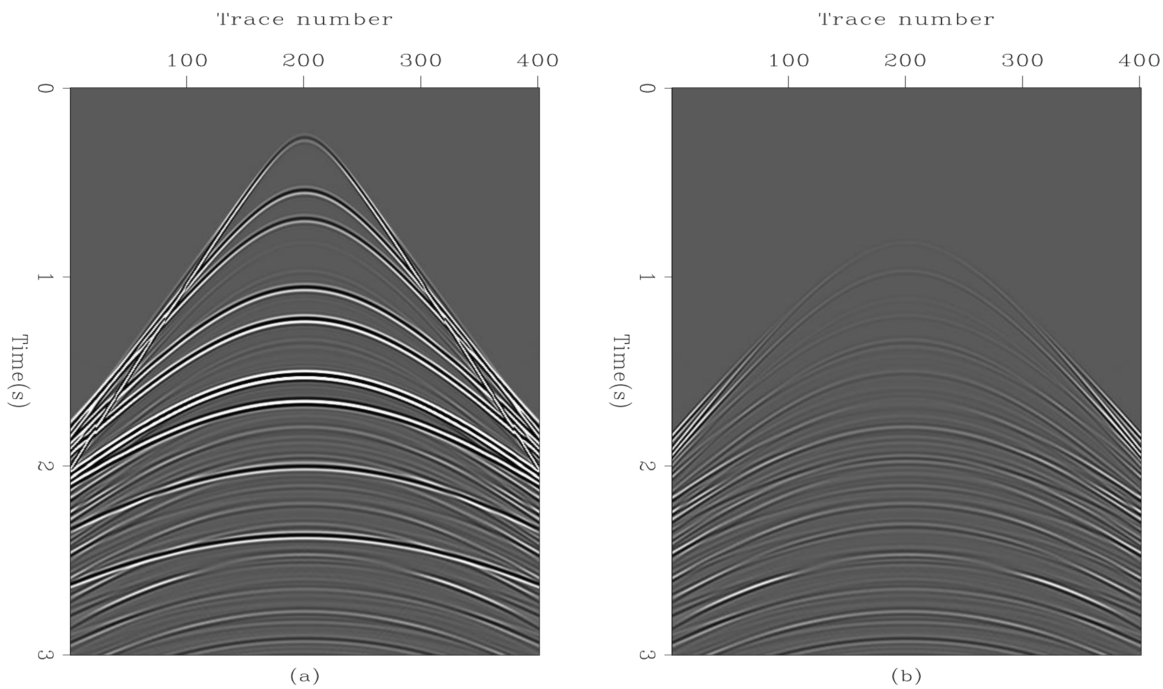


Figure 6: (a) A synthetic shot gather infested with internal multiples. (b) The internal multiples model obtained using the CFP technology. This model matches the internal multiples in (a). antoine1-inter [ER]

Adaptive filtering with non-stationary helical filters

To handle the inherent non-stationarity of seismic data, we estimate a filter bank of non-stationary filters using helical boundary conditions (Claerbout, 1998). This approach has been successfully utilized by Rickett et al. (2001) to attenuate surface-related multiples. We use a NLCG solver with the ℓ^1-norm and a standard conjugate gradient solver with the ℓ^2-norm. The filter coefficients vary smoothly across the output space thanks to a preconditioning of the problem (Rickett et al., 2001). In the following results, the non-stationary filters are 1D. We estimate the same number of coefficients per filter with the ℓ^2- and ℓ^1-norm.

Adaptive subtraction results

Figure 8a shows the estimated primaries when the ℓ^2-norm is used to compute the shaping filters. Figure 8b displays the estimated internal multiples. As expected, because of the amplitude differences between the signal (primaries) and the noise (multiples), the adaptive sub-

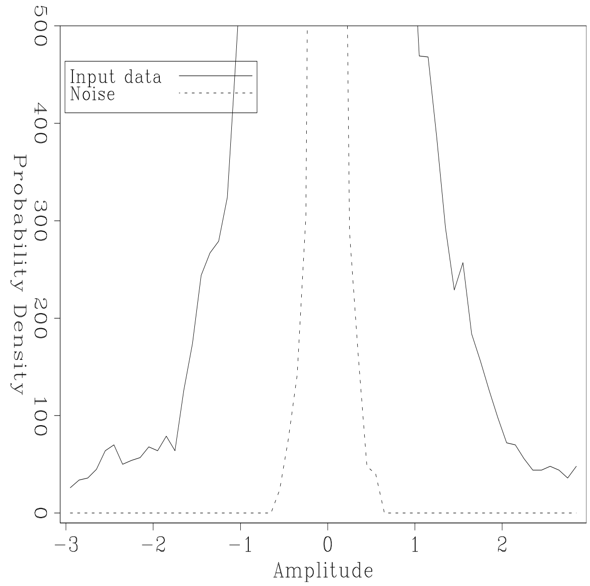


Figure 7: Histograms of the input data (Figure 6a) and of the noise (Figure 6b). The density function of the noise is much narrower than for the data. The ℓ^1-norm should be used to estimate the signal. antoine1-histodata [ER]

traction fails and we retrieve the behavior explained in the preceding section with the 1D example. Now, in Figure 9, we see the beneficial effects of the ℓ^1-norm. Figure 9a shows the estimated primaries and Figure 9b the estimated multiples. The noise subtracted almost perfectly matches the internal multiple model in Figure 6b, as anticipated.

POSTSTACK LAND DATA MULTIPLE REMOVAL EXAMPLE

In this section we attenuate in the poststack domain surface-related multiples with shaping filters that we estimate with the ℓ^2- and ℓ^1-norm. These filters are non-stationary. Figure 10a shows the multiple-infested data. Figure 10b displays the multiple model computed with the Delft modeling approach (Kelamis et al., 1999). Note that for this gather, the amplitude differences between the primaries and the multiples are not very strong. Our goal is to illustrate the use of the ℓ^1-norm in a more general case when surface-related multiples are present in the data. We specifically focus on the event at 1.6s in Figure 10. This event is a primary that we want to preserve during the subtraction.

Figures 12 displays the estimated signal when the non-stationary shaping filters are computed with the ℓ^2 and ℓ^1-norm. The amplitude of the primary at 1.6s is well preserved with the ℓ^1-norm in Figure 11a. However, the amplitude of this primary is attenuated with the ℓ^2-norm as displayed in Figure 11b. Figure 10 shows a comparison between the subtracted noise with the ℓ^1 (Figure 10a) and the ℓ^2-norm (Figure 10b). We conclude that the ℓ^2-norm tends to subtract too much energy.

This last example proves that the estimation of shaping filters can always be done with the ℓ^1-norm. The good thing about our inversion scheme and the objective function in equation (3) is that only one parameter (ϵ) controls the $\ell^1 - \ell^2$ behavior. Thus we can decide to switch from one norm to another very easily. In Figure 13, I show a histogram of the input data and

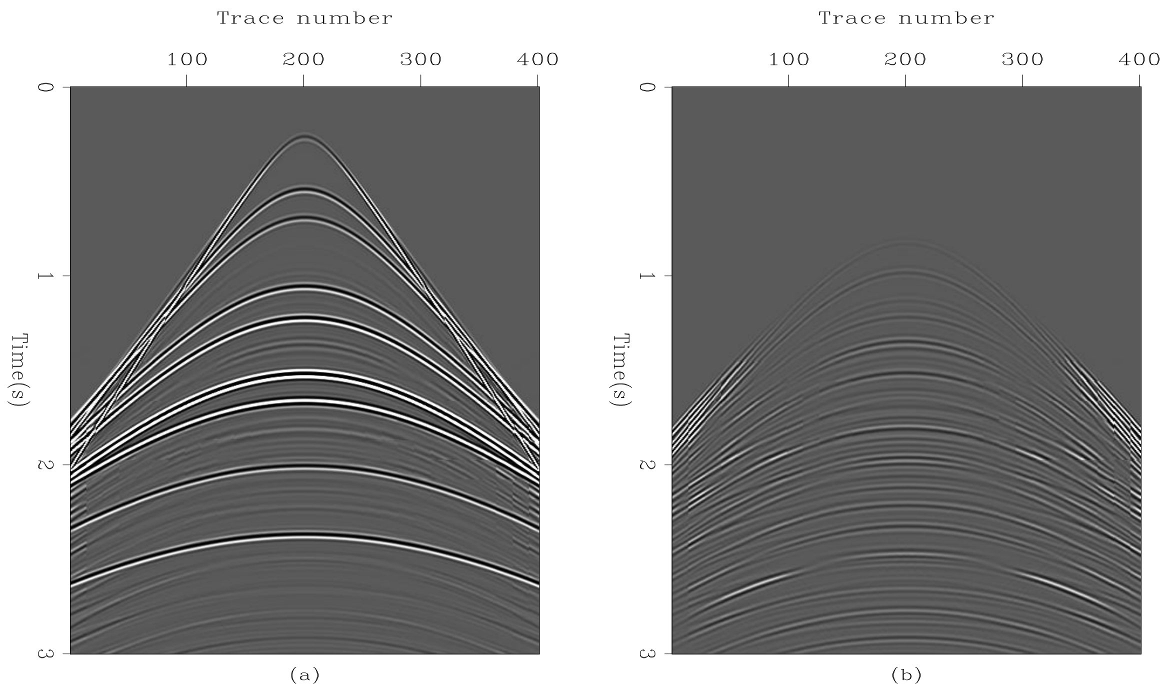


Figure 8: (a) The estimated primaries with the ℓ^2-norm. (b) The estimated internal multiples with the ℓ^2-norm. Ideally, (b) should look like Figure 6b, but it does not. antoine1-interl2 [ER]

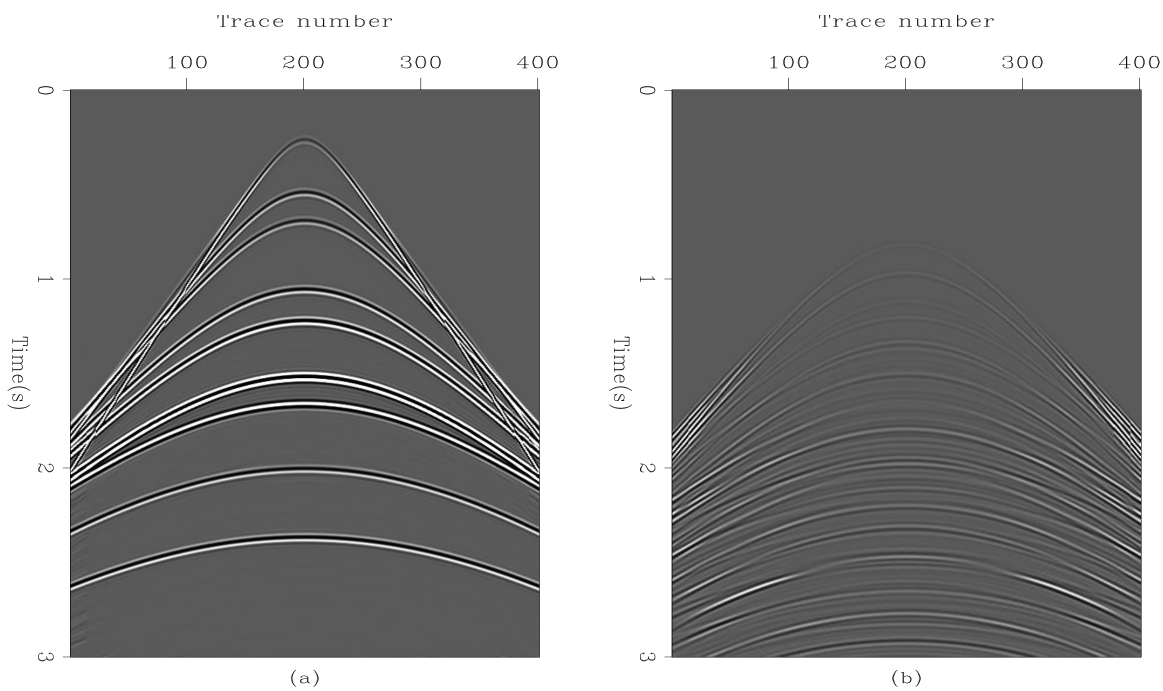


Figure 9: (a) The estimated primaries with the ℓ^1-norm. (b) The estimated internal multiples with the ℓ^1-norm. Beside some edge-effects, (b) resembles closely Figure 6b. The adaptive subtraction worked very well. antoine1-interl1 [ER]

of the estimated noise with the ℓ^1 and ℓ^2-norms. The theory predicts that the distribution of the ℓ^2 result should be gaussian and that distribution of the ℓ^1 result should be exponential. Figure 13 corroborates this.

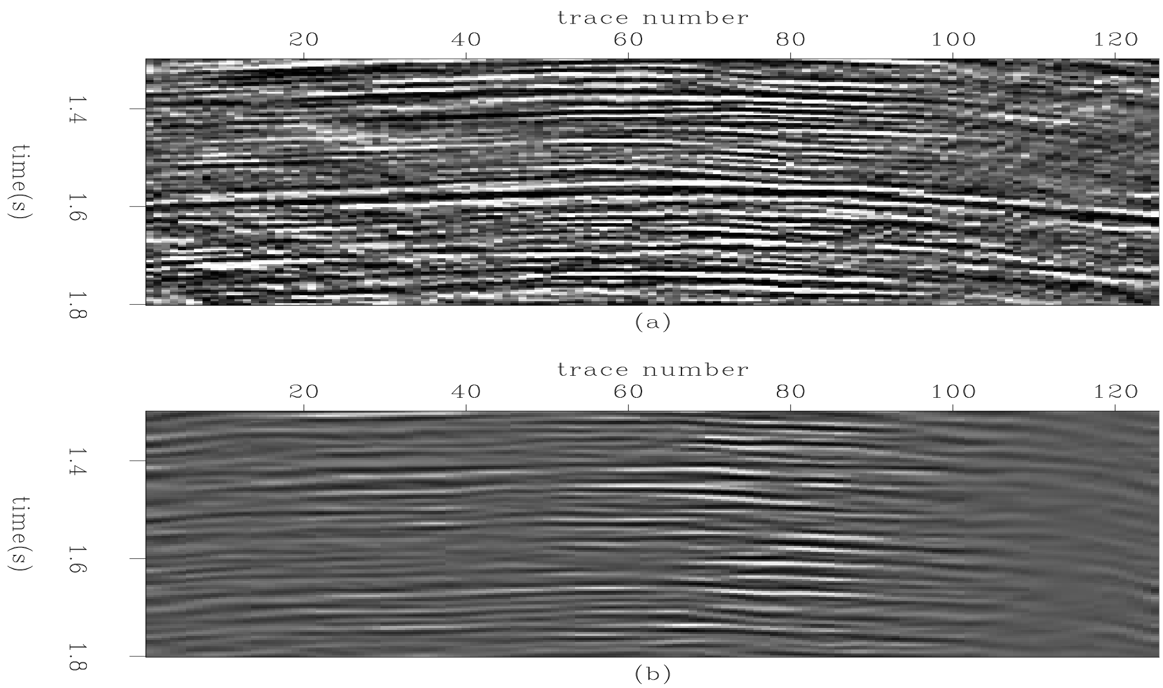


Figure 10: (a) Stack infested with multiples. (b) The multiple model computed with the Delft modeling approach. The subtraction is done poststack. antoine1-win3 [ER]

PRESTACK LAND DATA MULTIPLE REMOVAL EXAMPLE

The above methodology has also been tested on a shot record from a land data survey. The preprocessing and multiple prediction is described by Kelamis et al. (1999). I display in Figure 14a the selected shot record. Figure 15a shows the predicted multiples. Note that these multiples are the ones that are directly generated by the shot record based convolutions (Berkhout and Verschuur, 1997) and no adaptation has been applied yet. Both ℓ^2 and ℓ^1 adaptive subtraction has been carried out for this data. The resulting records are displayed in Figures 14b and c, respectively. The removed multiples are shown in Figures 15b and c.

Figures 15b and 15c demonstrate that the multiples are better attenuated with the ℓ^1-norm for long offsets. Although the truth cannot be revealed from these results, it appears that the ℓ^1 results are more reliable.

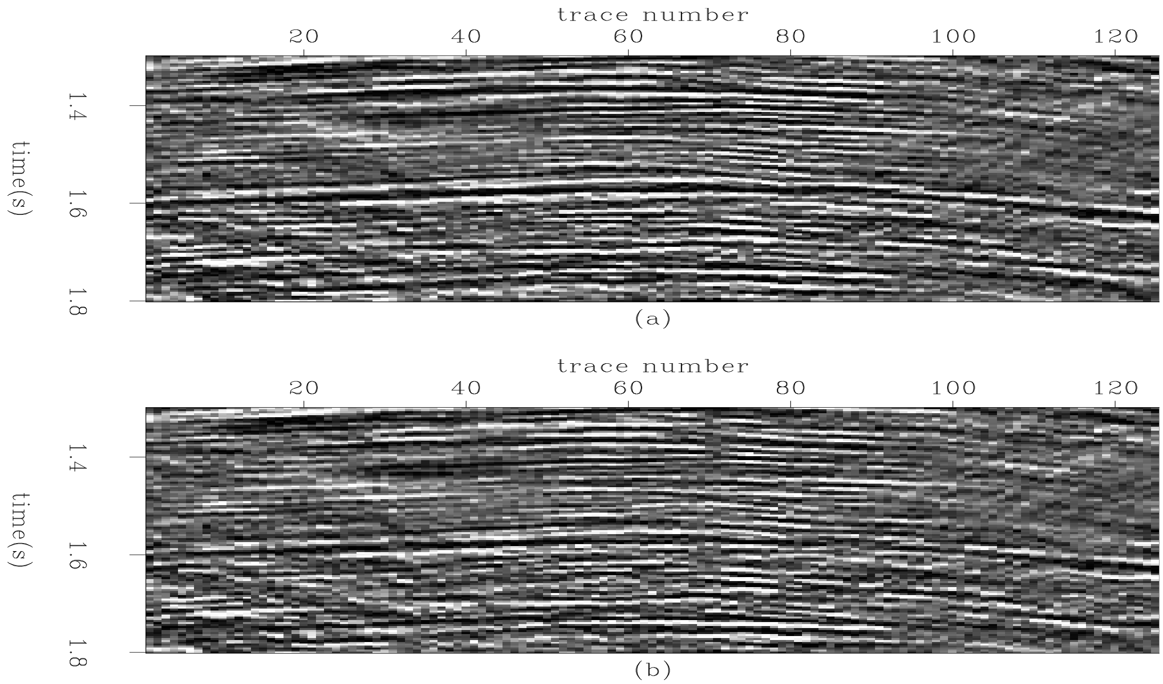


Figure 11: (a) The estimated primaries with ℓ^1-norm adaptive subtraction. (b) The estimated primaries with ℓ^2-norm subtraction. The primary at 1.6s is very attenuated with the ℓ^2-norm. The ℓ^1 technique preserves its amplitude better. antoine1-win [ER,M]

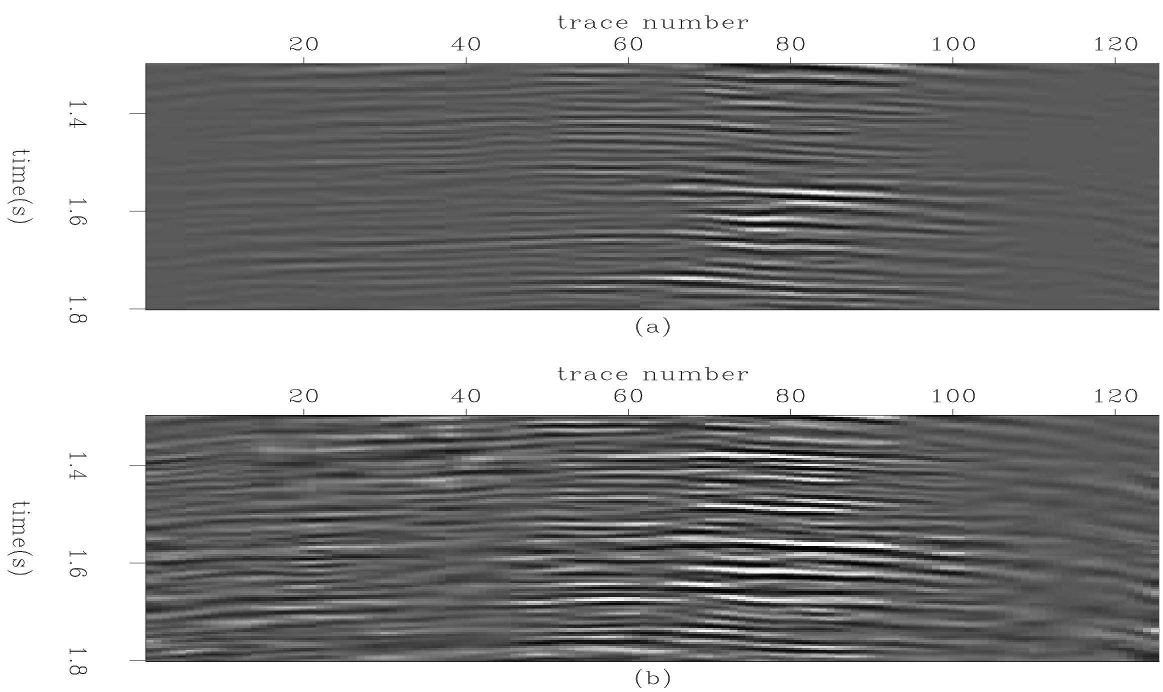


Figure 12: (a) The estimated multiples with the ℓ^1-norm subtraction. (b) The estimated multiples with the ℓ^2-norm subtraction. The ℓ^2-norm tends to over-fit some multiples that creates some leaking of primaries in the estimated noise. antoine1-win2 [ER,M]

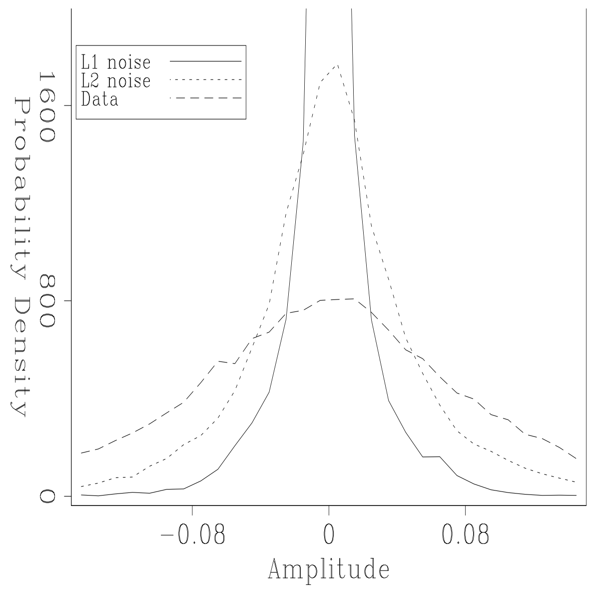


Figure 13: Histograms of the input data and of the estimated noise with the ℓ^1- and ℓ^2-norms. As predicted by the theory, the density function with the ℓ^1-norm is much narrower than with the ℓ^2-norm. antoine1-hist7636 [ER]

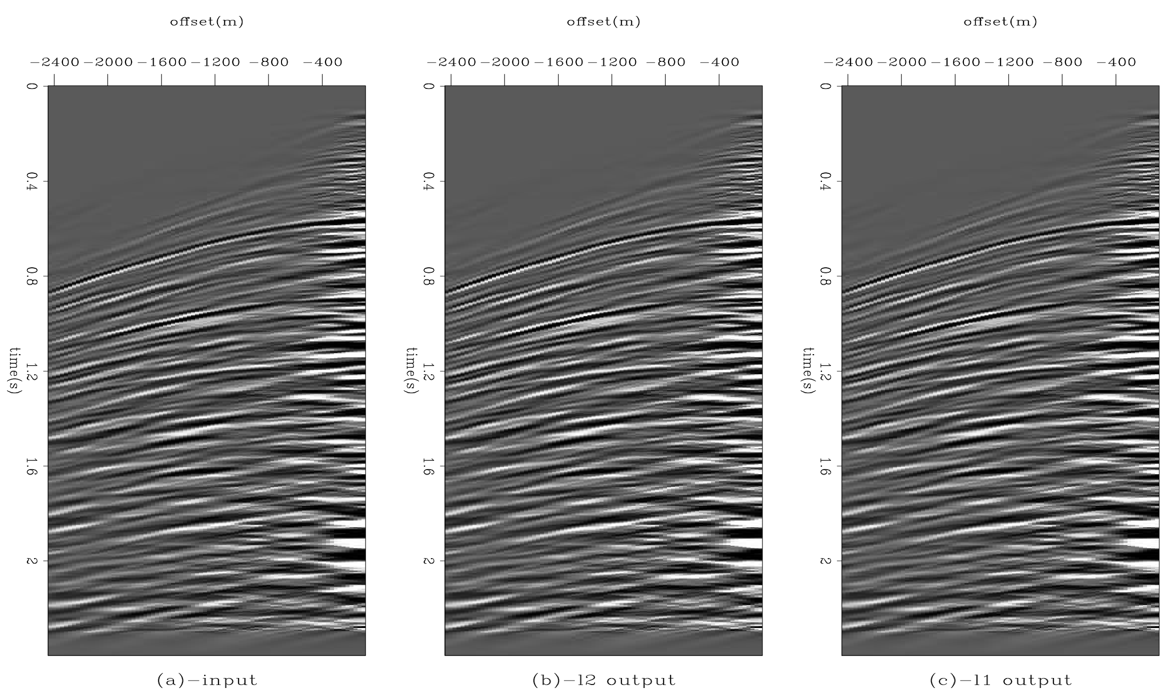


Figure 14: Example of adaptive multiple subtraction for land data. a) One selected shot record from a land survey. b) Estimated signal after ℓ^2 adaptive subtraction. c) Estimated signal after ℓ^1 adaptive subtraction. antoine1-comp.s.4 [ER,M]

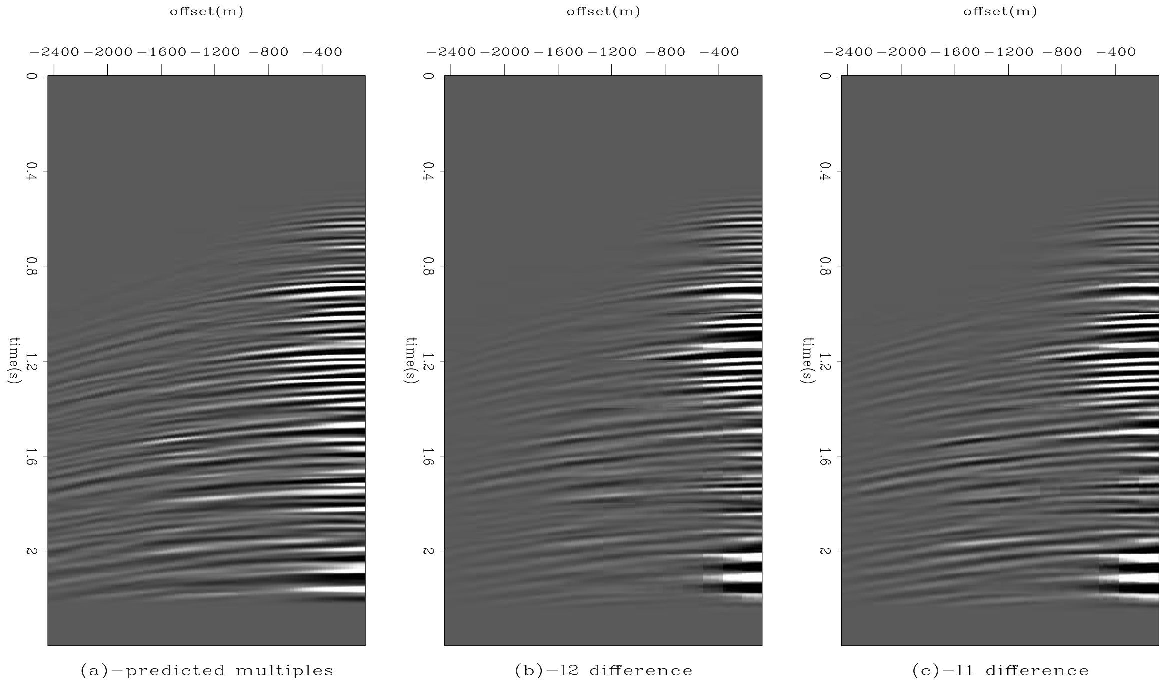


Figure 15: a) Predicted multiples for Figure 14a. b) Removed multiples with ℓ^2 adaptive subtraction. c) Removed multiples with ℓ^1 adaptive subtraction. Far offset multiples are better attenuated with the ℓ^1-norm between 1.2 and 1.8 seconds. antoine1-comp.n.4 [ER,M]

CONCLUSION

Significant amplitude differences between signal and noise make the ℓ^2-norm an unsuitable choice to estimate shaping filters. We showed that the ℓ^1-norm should always be considered in these circumstances. In addition, applications to real data with surface-related multiples tend to prove that the ℓ^1-norm should also be considered in the simplest cases.

ACKNOWLEDGMENTS

We would like to thank Saudi Aramco for providing the land data and the members of the SMAART 2 JV for their financial support.

REFERENCES

Berkhout, A. J., and Verschuur, D. J., 1997, Estimation of multiple scattering by iterative inversion, Part I: Theoretical considerations: Geophysics, **62**, no. 05, 1586–1595.

Berkhout, A., and Verschuur, D., 1999, Removal of internal multiples: 69th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 1334–1337.

Bube, K. P., and Langan, R. T., 1997, Hybrid ℓ^1/ℓ^2 minimization with applications to tomography: Geophysics, **62**, no. 04, 1183–1195.

Claerbout, J., and Fomel, S., 2001, Geophysical Estimation by Example: Class notes, http://sepwww.stanford.edu/sep/prof/index.html.

Claerbout, J., 1998, Multidimensional recursive filters via a helix: Geophysics, **63**, no. 05, 1532–1541.

Dragoset, B., 1995, Geophysical applications of adaptive-noise cancellation: 65th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 1389–1392.

Guitton, A., 2000, Huber solver versus IRLS algorithm for quasi L1 inversion: SEP-Report, **103**, 255–271.

Kelamis, P., Verschuur, D., Berkhout, A., and Erickson, K., 1999, Velocity-independent datuming of seismic data: 69th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 441–444.

Liu, F., Sen, M. K., and Stoffa, P. L., 2000, Dip selective 2-D multiple attenuation in the plane-wave domain: Geophysics, **65**, no. 01, 264–274.

Nichols, D., 1994, Velocity-stack inversion using $\mathbf{L_p}$ norms: SEP-Report, **82**, 1–16.

Rickett, J., Guitton, A., and Gratwick, D., 2001, Adaptive Multiple Subtraction with Non-Stationary Helical Shaping Filters: 63rd Mtg., Eur. Assn. Geosci. Eng., Expanded Abstracts, Session: P167.

Verschuur, D. J., Berkhout, A. J., and Wapenaar, C. P. A., 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, no. 09, 1166–1177.