Adaptive multiple subtraction with wavelet-based complex unary Wiener filters

Sergi Ventosa*, Sylvain Leroy†, Irène Huard†, Antonio Pica†, Hérald Rabeson*, Patrice Ricarte* and Laurent Duval*

* *IFP Energies nouvelles,*

1-4 avenue de Bois-Préau, 92852 Rueil-Malmaison, France

† *CGGVeritas*

27 Avenue Carnot

91341 Massy, France

(August 22, 2011)

Running head: **CWT unary adaptive multiple filtering**

ABSTRACT

Multiple attenuation is a crucial task in seismic data processing because multiples usually cover primaries from fundamental reflectors. Predictive multiple suppression methods remove these multiples by building an adapted model, aiming at being subtracted from the original signal. However, before the subtraction is applied, a matching filter is required to minimize amplitude differences and misalignments between actual multiples and their prediction, and thus to minimize multiples in the input dataset after the subtraction. In this work we focus on the subtraction element. We propose an adaptive multiple removal technique in a 1-D complex wavelet frame combined with a non-stationary adaptation performed via single-sample (unary) Wiener filters, consistently estimated on overlapping windows in the transformed domain. This approach greatly simplifies the matching filter estimation

and, despite its simplicity, compares promisingly with standard adaptive 2-D methods, both in terms of results and retained speed and efficiency.

INTRODUCTION AND BACKGROUND

Multiples correspond to unwanted coherent events related to more than one reflection of wave fields onto given surfaces. Their attenuation (Verschuur and Berkhout, 1992; Matson and Dragoset, 2005) represents one of the greatest challenges in past and present seismic processing. Several types are usually recognized: ghosts, short- and long-period multiples. Multiples may be classified according to different criteria: number of downward bounces, or order; delay; multiple generator. Other classifications exist (Verschuur, 2006, p. 7 sq.). In practice, they are also distinguished by two main mitigating approaches (Verschuur, 2006, p. 16):

1. based on a discrimination due to different velocities and reflections suffered by primaries and multiples. Representations which map data to new domains with minimum overlap between primaries and multiples (based on dip, location, velocity or curvature). They can be subsequently either reverted, provided a (near) perfect inverse exists, or subtracted — with limited distortion inversion — from original data, after transformed domain primary selection or multiple mute;

2. based on periodicity and predictability, in which multiple location is predicted (either directly or based on external models) and subtracted.

First approach, termed multiple filtering, encompasses stacking combined with Normal Move Out, homomorphic filtering (Buttkus, 1975), $f-k$ in Wu and Wang (2011) (or Yilmaz, 2001, p. 877 sq.), $\tau-p$ representations (Nuzzo and Quarta, 2004) or alternative — parabolic, hyperbolic — breeds of the Radon transform (Hampson, 1986; Beylkin and Vassiliou, 1998; Herrmann et al., 2000; Trad et al., 2003; Nowak and Imhof, 2006), with pros & cons discussion in Kabir and Verschuur (2007). Second approach includes prediction filters and

variations thereof (Taner, 1980; Spitz, 1999; Abma et al., 2005; Spitz et al., 2009), also termed identification, updating, shaping, sequential or matching filters in electrical engineering literature (Ristow and Kosbahn, 1979) or predictive deconvolution (Taner et al., 1995). Recently, modeling based techniques, such as surface related multiple estimation, have demonstrated excellent performance, allowing data-driven multiple removal (Verschuur and Berkhout, 1992; Weglein et al., 1997; Lin et al., 2004) or model-driven, wave-equation based, multiple removal (Pica et al., 2005; Weisser et al., 2006). These methods, based on multiple models, consist in first predicting, second subtracting, multiple events from original seismic data.

Widespread adaptive subtractions are designed under minimum least-squares criteria (Verschuur and Berkhout, 1997; Dragoset et al., 2010). Wiener-type filters (Robinson and Treitel, 1967) reveal optimal in minimizing the energy of the difference between the data and the adapted model. To reduce primary distortion or over-adaptation in practice, those matching filters are usually applied in two steps. Long filters compensate global amplitude, waveform and time-shift differences, followed by short filters correcting time variant discrepancies. Still, primaries and multiples, generated from the same source, are not fully uncorrelated, or orthogonal, which poses major challenges in the design of practically optimal matching filters. This observation leads to more hybrid methods, mixing both aforementioned approaches. Multichannel adaptation is a straightforward option (Wang, 2003). Adaptive subtraction may be assisted by a suitable representation of the data (Taner (1980) with radial suppression, Berkhout and Verschuur (2006) with focal transform). More robust objective matching criteria such as based on the L_1 norm (Guitton and Verschuur, 2004; Pang et al., 2009) or independence-based assumption subtraction (Kaplan and Innanen, 2008) are efficient alternatives as well. Recent works focus on multiscale or wavelet-like ap-

proaches (Jacques et al., 2011), due to their ability to exploit slight differences in primaries and multiple spectra and to promote sparsity. Pokrovskaia and Wombell (2004); Ahmed (2007) use standard discrete wavelet transforms. Herrmann and Verschuur (2004); Donno et al. (2010); Neelamani et al. (2010) follow a twist towards curvelets, sometimes regarded as related to Radon transforms (de Hoop et al., 2009).

In many cases, the type and length of the matching filter, whatever the domain is, yield additional parameter choices and demand computational resources. We here aim at subduing this issue by decomposing a complicated wide-band problem into a set of more tractable narrow-band problems. Namely, through a wavelet transform, the filter length aspect is shrunk down to a plurality of single-sample filter estimations. In the proposed approach, model-based multiple removal is solely based on 1-D traces, and hence does not compete directly with higher dimensional (Dragoset et al., 2010), regularized (Guitton, 2004; Fomel, 2009) or robust approaches (Herrmann and Verschuur, 2005). Its advantages lie in low complexity, speed and possible parallelization. More precisely, comparing with existing literature, we retain the following three concepts:

- Wiener-like or adapted matching filter (Lines, 1996) in a non-stationary fashion (Fomel, 2009),
- complex trace processing (Monk, 1993; Wang, 2003; Huo and Wang, 2009),
- transformed domain based on continuous wavelet frames (Sinha et al., 2005).

The adopted complex Morlet wavelet frame emulates limited-scale complex derivatives (Monk, 1993). The adaptation in the wavelet domain is performed on each complex scale or voice, see Figure 6, with a "unary" complex filter, *e.g.*, a one-coefficient matching operator (Pes-

quet and Leporini, 1997), accounting for localized phase and amplitude variations between data and model. Due to the fine time-frequency localization afforded by the continuous wavelet, we turn to non-stationary with sliding overlapping windows operating along each voice. The redundancy allowed by the wavelet frame allows a near shift invariance which better manages the time variability in model misalignment errors. The closest work in spirit is Neelamani et al. (2010), while trying to deflect from the relatively high memory footprint of the curvelet (either real or complex).

METHODOLOGY

A classical 1-D trace observation model is:

$$d[n] = p[n] + m[n] + w[n]\,, \tag{1}$$

where $d[n]$, $p[n]$, $m[n]$ and $w[n]$ denote the recorded data, primary events, multiples and background noise, respectively, at discrete time index n.

We perform a time-scale decomposition of each data $d[n]$ and model $x[n]$ traces with discrete wavelet frame approximations to continuous wavelets. We refer to Jorgensen and Song (2009) and references therein for comparison between discrete and continuous wavelet transforms.

The associated discrete family of functions is defined as a sampling of a mother wavelet $\psi(t)$, t denoting the continuous time variable:

$$\psi^{v}_{r,j}[n] = \frac{1}{\sqrt{2^{j+v/V}}}\psi\left(\frac{nT - r2^{j}b_0}{2^{j+v/V}}\right)\,, \tag{2}$$

with T the sampling rate and V the number of voices per octave. Indices $r, j \in \mathbb{Z}$ and $v \in [0, \ldots, V-1]$ denote, respectively, discretized time, scale, and octave. Finally, b_0 stands

for the sampling period at scale zero. Its redundancy, approximately of $2b_0V$, may be adapted to the expected computational efficiency.

The time-scale representation of a given trace $d[n]$ is given by the inner product:

$$Wd_{j,v}[r] = \left\langle d[n], \psi^v_{r,j}[n] \right\rangle = \sum_n d[n]\overline{\psi^v_{r,j}[n]}\,. \tag{3}$$

We choose $\psi(t)$ as the complex Morlet wavelet since it yields a simple interpretation of amplitude and phase delay in the transformed domain. Additionally, the straightforward group-delay estimation that this wavelet allows together with the redundancy provided by the frame, helps avoiding reconstruction artifacts observed with non-redundant discrete wavelet processing (Gilloire and Vetterli, 1992).

The mother Morlet wavelet is approximately analytic and writes:

$$\psi(t) = \pi^{-1/4} e^{-j\omega_0 t} e^{-t^2/2}\,, \tag{4}$$

where ω_0 is the central frequency of the modulated Gaussian. A standard central frequency $\omega_0^s = \pi\sqrt{2/\ln(2)}$ halves side-lobe amplitude with respect to the main lobes.

The discriminative power of the wavelet frame helps reformulate the design of a long matching filter into a combined global and local complex optimum unary filter, which minimizes the error between multiple events and their model, as described in the following section. Once the multiple model is matched and subsequently subtracted from the original data in the wavelet domain, each trace is synthesized back in time domain with an inverse Morlet wavelet transform.

Amplitude and phase estimation

Following a minimum energy approach, the optimum unary filter in a given wavelet voice is the complex value a_j^v that, multiplied by the decomposed multiple model $\mathbf{x} = Wx_j^v[r]$, minimizes the mean square error $\xi(a)$ with the decomposed input dataset $\mathbf{d} = Wd_j^v[r]$:

$$a_{\text{opt}} = \arg\min_a \xi(a) = \arg\min_a \|\mathbf{d} - a\mathbf{x}\|^2 \tag{5}$$

or equivalently,

$$a_{\text{opt}} = \arg\min_a \sum_r \left|Wd_j^v[r] - a_j^v Wx_j^v[r]\right|^2 , \tag{6}$$

where sequences $Wd_j^v[r]$ and $Wx_j^v[r]$ are complex and the $\sum$ symbol denotes a locally weighted sum along odd L consecutive samples around index r.

Optimal a_{opt} yields orthogonality between filtered signal $\mathbf{y} = \mathbf{d} - a_{\text{opt}}\mathbf{x}$ and adapted model $a\mathbf{x}$, which writes in inner product form as:

$$\langle \mathbf{d} - a_{\text{opt}}\mathbf{x}, a\mathbf{x} \rangle = 0 \quad \forall a \neq 0 \,. \tag{7}$$

By developing and applying linearity on the first argument and conjugate linearity on the second one, we obtain:

$$\overline{a} \left(\langle \mathbf{d}, \mathbf{x} \rangle - a_{\text{opt}} \langle \mathbf{x}, \mathbf{x} \rangle \right) = 0 \quad \forall a \neq 0 \tag{8}$$

and therefore

$$a_{\text{opt}} = \frac{\langle \mathbf{d}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} = \frac{\sum_r Wd_j^v[r]\overline{Wx_j^v[r]}}{\sum_r \left|Wx_j^v[r]\right|^2} , \tag{9}$$

i.e., the unary Wiener filter solution for complex signals. This algorithm may be applied either locally or globally, depending on problem constraints.

Note that careful proceeding is important: instead of performing a direct measurement of the fractional delay, or equivalently of the group delay, we attempt to estimate the phase

delay at each scale or frequency component. When the signal-to-noise (S/N) ratio is high, all meaningful phase components are sufficiently well estimated. Alternatively, when the S/N ratio decreases, moderate phase errors made at key scale components may induce huge errors on group delay. As a consequence, a heedful processing of a_{opt} at each scale is performed, to provide a low dip varying phase for meaningful components.

Integer delay estimation

When the group delay difference between model and multiple sequences is less than the half a signal period at all scales, it can be rectified with the previously mentioned approach, as the cross-correlation between both sequences is close to the zero delay. However, when the group delay difference crosses this limit, and the cross-correlation peak departs from the zero delay sample, the estimation of a_{opt} done in equation 9 is no longer valid. As a consequence, a redesign of the filter length is necessary.

Two options can be followed to add robustness against high fractional delays while keeping with the unary filter approach: design filters with a high varying phase along scales, or add a gross group delay correction first and then apply the previous unary filter for fine group delay correction. In contrast with the Fourier transform, wavelets' fast decay strongly limits the maximum group delay correction on the first option. Hence, we focus exclusively on the second one. However, high varying phase filters enable a relatively simple scale-wise unary filter design that could reinforce matching in noisy scenarios, in applications where the group delay range specifications are small enough to fit their limitations.

We thus introduce a delay term into the above unary filter to yield a higher delay robustness, while keeping a minimum L_2 criterion. This amounts to finding optimum a and

delay l that minimize:

$$\xi(a,l) = \sum_r \left| W d_j^v[r] - a_j^v W x_j^v[r-l] \right|^2 = \left\| \mathbf{d} - a\mathbf{x}_l \right\|^2 , \tag{10}$$

where $\mathbf{x}_l$ denotes a delayed $\mathbf{x}$ of l samples. The optimum a at a given delay, as shown above, is:

$$a_{\text{opt}}[l] = \frac{\sum_r W d_j^v[r] \overline{W x_j^v[r-l]}}{\sum_r \left| W x_j^v[r-l] \right|^2} = \frac{\langle \mathbf{d}, \mathbf{x}_l \rangle}{\left\| \mathbf{x}_l \right\|^2} . \tag{11}$$

Several criteria, well adapted to the nature of the seismic signals, can be chosen in the selection of the optimum delay. Keeping with minimum mean square error criterion, the optimum integer delay is selected among all sub-optimal $a_{\text{opt}}[l]$ as the one minimizing energy of error term $\xi(a,l)$:

$$l_{\text{opt}} = \arg\min_l \xi(a_{\text{opt}}[l]) . \tag{12}$$

Alternatively, due to the importance of shape over amplitude, we can define the optimum delay as the one that maximizes the normalized cross-correlation between the adapted model and the data, commonly called coherence (Neidell and Taner, 1971; Taner et al., 1979; Schimmel and Paulssen, 1997),

$$l_{\text{opt}} = \arg\max_l \; C[l] \tag{13}$$

being

$$C[l] = \text{Re} \left[\frac{\sum_r W d_j^v[r] \overline{\left(a_{\text{opt}}[l] W x_j^v[r-l] \right)}}{\sqrt{\sum_r \left| W d_j^v[r] \right|^2} \sqrt{\sum_r \left| a_{\text{opt}}[l] W x_j^v[r-l] \right|^2}} \right] \tag{14}$$

or equivalently,

$$l_{\text{opt}} = \arg\max_l \; \text{Re} \left[\frac{\langle \mathbf{d}, a_{\text{opt}}[l]\mathbf{x}_l \rangle}{\|\mathbf{d}\| \, \|a_{\text{opt}}[l]\mathbf{x}_l\|} \right] , \tag{15}$$

where we focus on the real part of the normalized cross-correlation.

Note that the use of unary filters enables two important simplifications of the previous equations. In first place, applying conjugate linearity in the second argument of the inner

product and positive scalability of the norm,

$$l_{\text{opt}} = \arg\max_{l} \text{ Re} \left[\frac{\overline{a_{\text{opt}}[l]}}{|a_{\text{opt}}[l]|} \frac{\langle \mathbf{d}, \mathbf{x}_l \rangle}{\|\mathbf{d}\| \, \|\mathbf{x}_l\|} \right] \tag{16}$$

we notice that the normalized cross-correlation between the adapted model and the data depends only on the phase of the optimum unary filter and the normalized cross-correlation between the original model and the data.

But even more important for us, when the weighed sums used in the optimum unary filter estimation, equation 11, and in the normalized cross-correlation estimator, equation 14, are equal, the normalized cross-correlation of the corrected sequence is equivalent to the cosine of the angle between $\mathbf{d}$ and $\mathbf{x}_l$. This can be shown noticing that $\overline{a_{\text{opt}}[l]} / |a_{\text{opt}}[l]|$ can be written depending on $\mathbf{d}$ and $\mathbf{x}_l$ from its definition on equation 11,

$$\frac{\overline{a_{\text{opt}}[l]}}{|a_{\text{opt}}[l]|} = \frac{\overline{\langle \mathbf{d}, \mathbf{x}_l \rangle}}{|\langle \mathbf{d}, \mathbf{x}_l \rangle|} \, . \tag{17}$$

As a consequence, the normalized cross-correlation is totally independent of the optimum unary filter in this particular case,

$$l_{\text{opt}} = \arg\max_{l} \frac{|\langle \mathbf{d}, \mathbf{x}_l \rangle|}{\|\mathbf{d}\| \, \|\mathbf{x}_l\|} \tag{18}$$

and the coherence value measured is always real and positive.

This feature enables an important reduction on the computational cost when the normalized cross-correlation estimator is used, as it enables the estimation of the optimum integer delay, without having to estimate and apply first the optimum unary filter for each candidate.

EXAMPLE AND DISCUSSION

[Figure 1 about here.]

[Figure 2 about here.]

[Figure 3 about here.]

[Figure 4 about here.]

[Figure 5 about here.]

[Figure 6 about here.]

The example shown in Figures 1-a and 3-a is taken from a 3-D real marine dataset, in common shot and receiver gather, respectively. The corresponding model, in Figure 1-b and 3-b, is computed with a 3-D convolutional algorithm, which achieves good precision on the full offset range. The data is decomposed with a Morlet wavelet with $\omega_0 = 1.2\omega_0^s$, $j \in [1, 4]$, $v \in [0, 3]$ and $b_0 = 1$. The optimum unary filter estimation for each sample uses a rectangular window of 0.636 s. Additionally, to provide robustness against delays higher than half of the signal period, we have chosen the maximum normalized cross-correlation criterion with a range of ± 12 ms.

The main objective is to uncover primaries masked by strong multiple events using a multiple model, see Figures 1 and 3. We compare complex 1-D unary complex filters in time-scale domain and a standard 2-D adaptive filter in time-space domain. The latter (Lin et al., 2004) is performed with a two-pass approach, starting with a long filter to take into account mainly the seismic wavelet squaring introduced into the model during the convolution process, followed by a shorter filter to handle local amplitude and phase variations of multiple events. As Figures 2, 4 and 5 illustrate, our 1-D adaption based on scale-wise unary filters yield a quality equivalent to a standard 2-D approach, despite not

using additional information conveyed by neighboring traces. Being single-pass, wavelet-based unary filters achieve slightly better noise reduction while a standard 2-D approach exhibits slightly superior multiple attenuation.

Figure 6 depicts intermediate results on a portion around the first multiple arrival of the trace located at the intersection of the common receiver and shot plane section shown in Figures 1 and 3. The left panels show the recorded signal and the multiple model in time, together with their modulus in the time-scale domain; while on the right ones the filtered model and the filtered recorded signal. We appreciate how the scale-wise unary filter successfully reduces the differences between multiple model and actual multiple events, to attain sufficient multiple event attenuation to uncover main primary events.

The continuous wavelet frame provides an inherent scale-adaptive windowing that, when combined with the complex unary filter, offers a potentially seamless matching procedure in 1-D. Additional constraints may be added in scale and along the space dimensions to improve the fractional delay estimation and reinforce the lateral coherence of the method.

In scenarios where the multiples are stronger than primaries, the distortion on the primaries can be heavily reduced by applying filters based on minimum mean absolute error criteria, instead of mean square error ones. The high level of sparsity that this approach provides (Pang et al., 2009; Guitton and Verschuur, 2004) could enable a better separation of primaries and multiples events and could lead to better solutions on this cross-correlation limited scenario (in opposition to a noise limited one). In contrast, when the multiples are weak, the mean absolute error criteria should be applied with great care to avoid primaries overadaption.

CONCLUSION

We propose a model-based multiple subtraction whose non-stationary adaptation aptitude results from combining complex Morlet wavelet frame with unary complex Wiener filters. We therefore balance the standard matching filter estimation complexity by a wavelet transform which allows for simple non-stationary model adaptation. Unary filters with a low varying phase along scale can adapt fractional delays up to half the signal period at a given frequency. To add robustness against higher delays, we integrate an integer delay term in the kernel of the unary filters, which can be limited and estimated globally or locally depending on the application. Although redundant and 1-D, its results on a real dataset compare promisingly with standard 2-D least-square methods. Even if it does not directly contend more involved multiple removal methods, either regularized or in higher dimension spaces, its computational efficiency allows for reduced memory footprint and higher code parallelization. Further developments include combination with higher dimensional multiscale transforms, aimed at better apparent velocity separation and residual noise attenuation. The presented multiple removal technique gears toward updating the portfolio of hybrid demultiple algorithms.

ACKNOWLEDGEMENTS

The authors thank Statoil for allowing them to show the Norwegian Sea results. They also acknowledge IFP Energies nouvelles and CGGVeritas (with a special mention to Philippe Herrmann) for the authorization to present this work.

REFERENCES

Abma, R., N. Kabir, K. H. Matson, S. Michell, and S. A. Shaw and B. McLain, 2005, Comparisons of adaptive subtraction methods for multiple attenuation: The Leading Edge, **24**, 277–280.

Ahmed, I., 2007, 2D wavelet transform-domain adaptive subtraction for enhancing 3D SRME: 77th Annual International Meeting, SEG, Expanded Abstracts, 2490–2494.

Berkhout, A. J., and D. J. Verschuur, 2006, Focal transformation, an imaging concept for signal restoration and noise removal: Geophysics, **71**, A55–A59.

Beylkin, G., and A. Vassiliou, 1998, Fast Radon transform for multiple attenuation: 68th Annual International Meeting, SEG, Expanded Abstracts, 1351–1352.

Buttkus, B., 1975, Homomorphic filtering — theory and practice: Geophysical Prospecting, **23**, 712–748.

de Hoop, M. V., H. Smith, G. Uhlmann, and R. D. van der Hilst, 2009, Seismic imaging with the generalized Radon transform: a curvelet transform perspective: Inverse Problems, **25**, 025005 (21pp).

Donno, D., H. Chauris, and M. Noble, 2010, Curvelet-based multiple prediction: Geophysics, **75**, WB255–WB263.

Dragoset, B., E. Verschuur, I. Moore, and R. Bisley, 2010, A perspective on 3D surface-related multiple elimination: Geophysics, **75**, 75A245–75A261.

Fomel, S., 2009, Adaptive multiple subtraction using regularized nonstationary regression: Geophysics, **74**, V25–V33.

Gilloire, A., and M. Vetterli, 1992, Adaptive filtering in sub-bands with critical sampling: analysis, experiments, and application to acoustic echo cancellation: IEEE Transactions on Signal Processing, **40**, 1862–1875.

Guitton, A., 2004, A pattern-based approach for multiple removal applied to a 3D Gulf of Mexico data set: Geophysical Prospecting, **54**, 135–152.

Guitton, A., and D. J. Verschuur, 2004, Adaptive subtraction of multiples using the l_1-norm: Geophysical Prospecting, **52**, 27–38.

Hampson, D., 1986, Inverse velocity stacking for multiple elimination: 66th Annual International Meeting, SEG, Expanded Abstracts, 422–424.

Herrmann, F. J., and D. J. Verschuur, 2004, Curvelet imaging and processing: adaptive multiple elimination: CSEG National Convention.

———, 2005, Robust curvelet-domain primary-multiple separation with sparseness constraints: 67th Annual Meeting, EAGE, Expanded Abstracts, P226.

Herrmann, P., T. Mojesky, T. Magesan, and P. Hugonnet, 2000, De-aliased, high-resolution Radon transforms: 70th Annual International Meeting, SEG, Expanded Abstracts, 1953–1956.

Huo, S., and Y. Wang, 2009, Improving adaptive subtraction in seismic multiple attenuation: Geophysics, **74**, 59–67.

Jacques, L., L. Duval, C. Chaux, and G. Peyré, 2011, A panorama on multiscale geometric representations, intertwining spatial, directional and frequency selectivity: Signal Processing, **91**, 2699–2730.

Jorgensen, P. E. T., and M.-S. Song, 2009, Comparison of discrete and continuous wavelet transforms: Springer, **1-10**.

Kabir, N., and D. J. Verschuur, 2007, Expert answers: Does parabolic radon transform multiple removal hurt amplitudes for avo analysis?: CSEG Recorder, 10–14.

Kaplan, S. T., and K. A. Innanen, 2008, Adaptive separation of free-surface multiples through independent component analysis: Geophysics, **73**, V29–V36.

Lin, D., J. Young, Y. Huang, M. Hartmann, 2004, 3D SRME application in the Gulf of Mexico: 74th Annual International Meeting, SEG, Expanded Abstracts, 1257–1260.

Lines, L., 1996, Suppression of short-period multiples — deconvolution or model-based inversion?: Canadian Journal of Exploration Geophysics, **32**, 63–72.

Matson, K., and B. Dragoset, 2005, An introduction to this special section — multiple attenuation: The Leading Edge, **24**, 252.

Monk, D. J., 1993, Wave-equation multiple suppression using constrained gross-equalization: Geophysical Prospecting, **41**, 725–736.

Neelamani, R. N., A. Baumstein, and W. S. Ross, 2010, Adaptive subtraction using complex-valued curvelet transforms: Geophysics, **75**, V51–V60.

Neidell, N. S., and M. T. Taner, 1971, Semblance and other coherency measures for multichannel data: Geophysics, **36**, 482–497.

Nowak, E. J., and M. G. Imhof, 2006, Amplitude preservation of Radon-based multiple-removal filters: Geophysics, **71**, V123–V126.

Nuzzo, L., and T. Quarta, 2004, Improvement in GPR coherent noise attenuation using $\tau - p$ and wavelet transforms: Geophysics, **69**, 789–802.

Pang, T., W. Lu, and Y. Ma, 2009, Adaptive multiple subtraction using a constrained l_1-norm method with lateral continuity: Journal of Applied Geophysics, **6**, 241–247.

Pesquet, J.-C., and D. Leporini, 1997, A new wavelet estimator for image denoising: IEE Sixth International Conference on Image Processing and Its Applications, 249–253.

Pica, A., G. Poulain, B. David, M. Magesan, S. Baldock, T. Weisser, P. Hugonnet, and P. Herrmann, 2005, 3D surface-related multiple modeling: The Leading Edge, **24**, 292–296.

Pokrovskaia, T., and R. Wombell, 2004, Attenuation of residual multiples and coherent noise in the wavelet transform domain: 66th Annual Meeting, EAGE, Expanded Abstracts,

P175

Ristow, D., and B. Kosbahn, 1979, Time-varying prediction filtering by means of updating: Geophysical Prospecting, **27**, 40–61.

Robinson, E. A., and S. Treitel, 1967, Principles of digital Wiener filtering: Geophysical Prospecting, **15**, 311–332.

Schimmel, M., and H. Paulssen, 1997, Noise reduction and detection of weak, coherent signals through phase-weighted stacks: Geophysical Journal International, **130**, 495–505.

Sinha, S., P. S. Routh, P. D. Anno, and J. P. Castagna, 2005, Spectral decomposition of seismic data with continuous-wavelet transform: Geophysics, **70**, P19–P25.

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

Spitz, S., G. Hampson, and A. Pica, 2009, Simultaneous source separation using wave field modeling and PEF adaptive subtraction: EAGE Marine Seismic Workshop, M11.

Taner, M. T., 1980, Long period sea-floor multiples and their suppression: Geophysical Prospecting, **28**, 30–48.

Taner, M. T., F. Koehler, and R. E. Sheriff, 1979, Complex seismic trace analysis: Geophysics, **44**, 1041–1063.

Taner, M. T., R. F. O'Doherty, and F. Koehler, 1995, Long period multiple suppression by predictive deconvolution in the $x - t$ domain: Geophysical Prospecting, **43**, 433–468.

Trad, D., T. Ulrych, and M. Sacchi, 2003, Latest views of the sparse Radon transform: Geophysics, **68**, 386–399.

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

Verschuur, D. J., 2006, Seismic multiple removal techniques: past, present and future: EAGE

Publications.

Verschuur, D. J., and A. J. Berkhout, 1997, Estimation of multiple scattering by iterative inversion, part II: Practical aspects and examples: Geophysics, **62**, 1596–1611.

Wang, Y., 2003, Multiple subtraction using an expanded multichannel matching filter: Geophysics, **68**, 346–354.

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.

Weisser, T., A. L. Pica, P. Herrmann, and R. Taylor, 2006, Wave equation multiple modelling: acquisition independent 3D SRME: First Break, **24**, 75–79.

Wu, M., and S. Wang, 2011, A case study of $f-k$ demultiple on 2D offshore seismic data: The Leading Edge, **30**, 446–450.

Yilmaz, Ö., 2001, Seismic data analysis: processing, inversion, and interpretation of seismic data: SEG.

LIST OF FIGURES

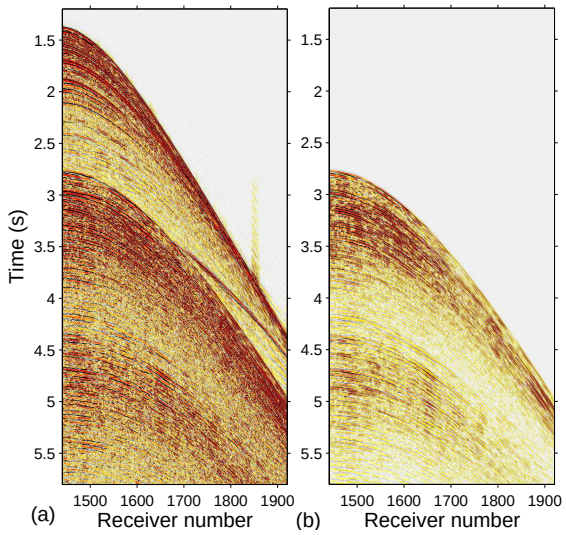

Figure 1: Portion of the common shot plane 1901. (a) Data recorded. (b) Multiple model.

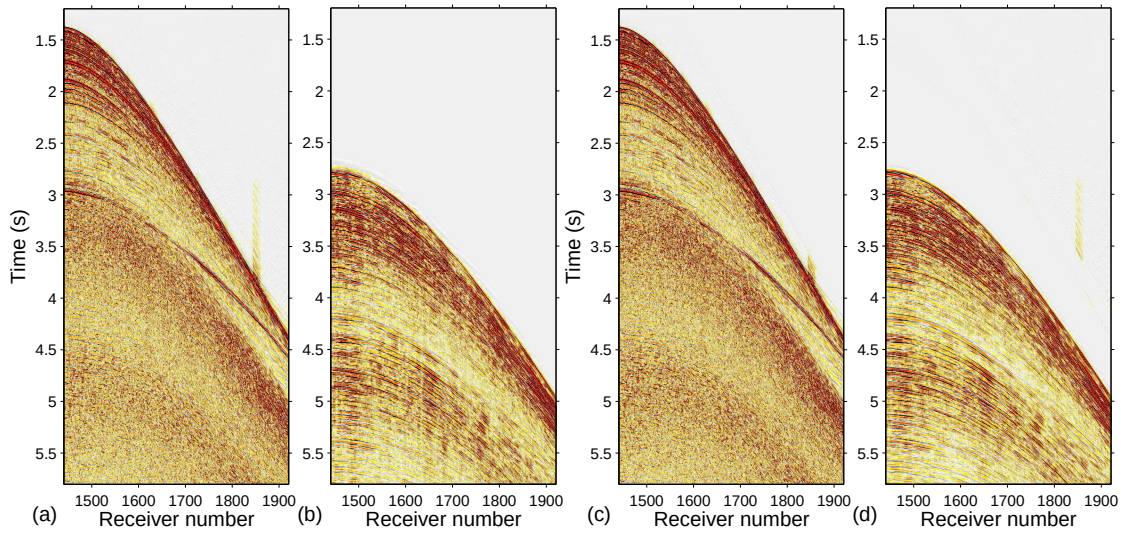

Figure 2: Subtraction results on the common shot plane 1901. (a) Filtered data and (b) adapted multiples with the 1-D unary complex filters in the time-scale domain. (c) Filtered data and (d) adapted multiples with a standard 2-D adaptive filter in time-space domain.

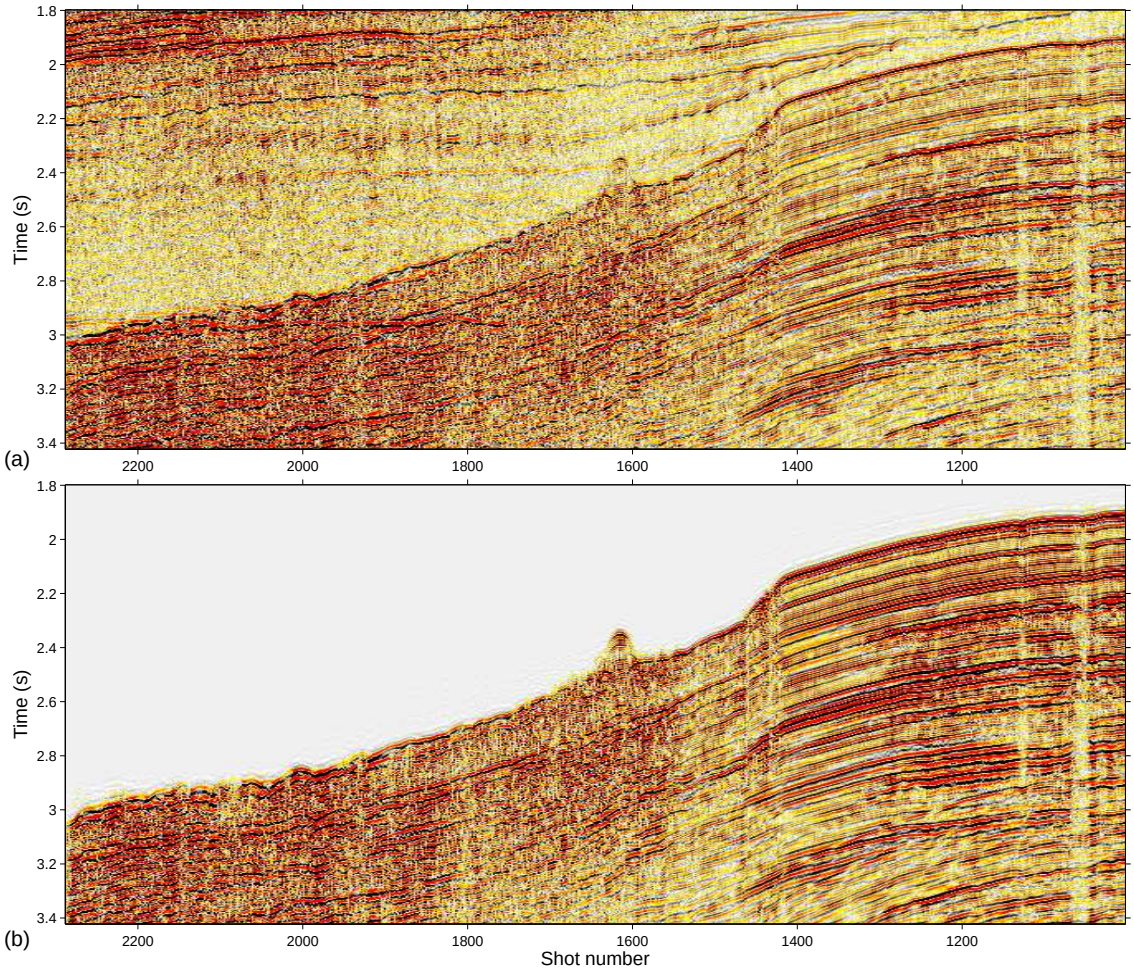

Figure 3: Portion of a near receiver plane. (a) Recorded data. (b) Multiple model.

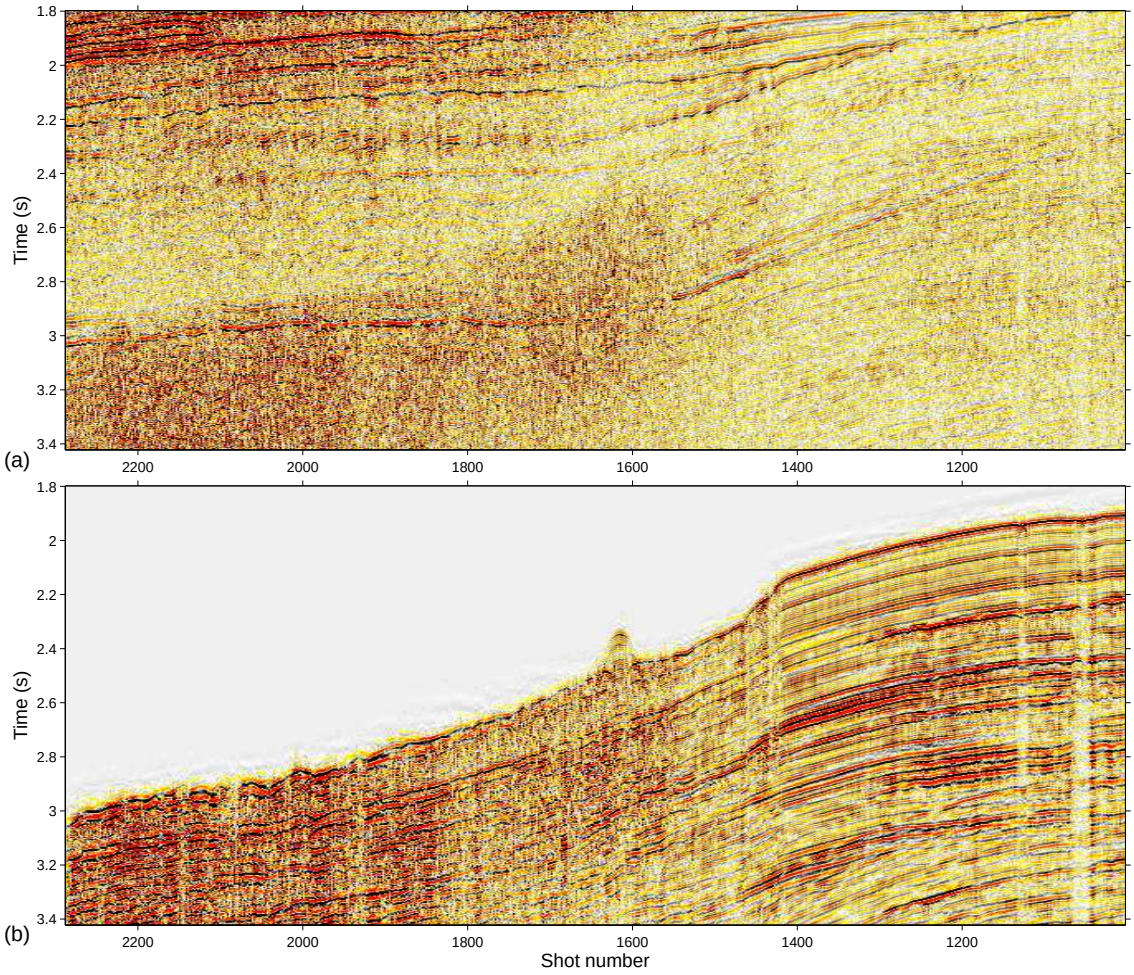

Figure 4: Subtraction results with the 1-D unary complex filters in the time-scale domain on a near receiver plane. (a) Filtered data. (b) Adapted multiples.

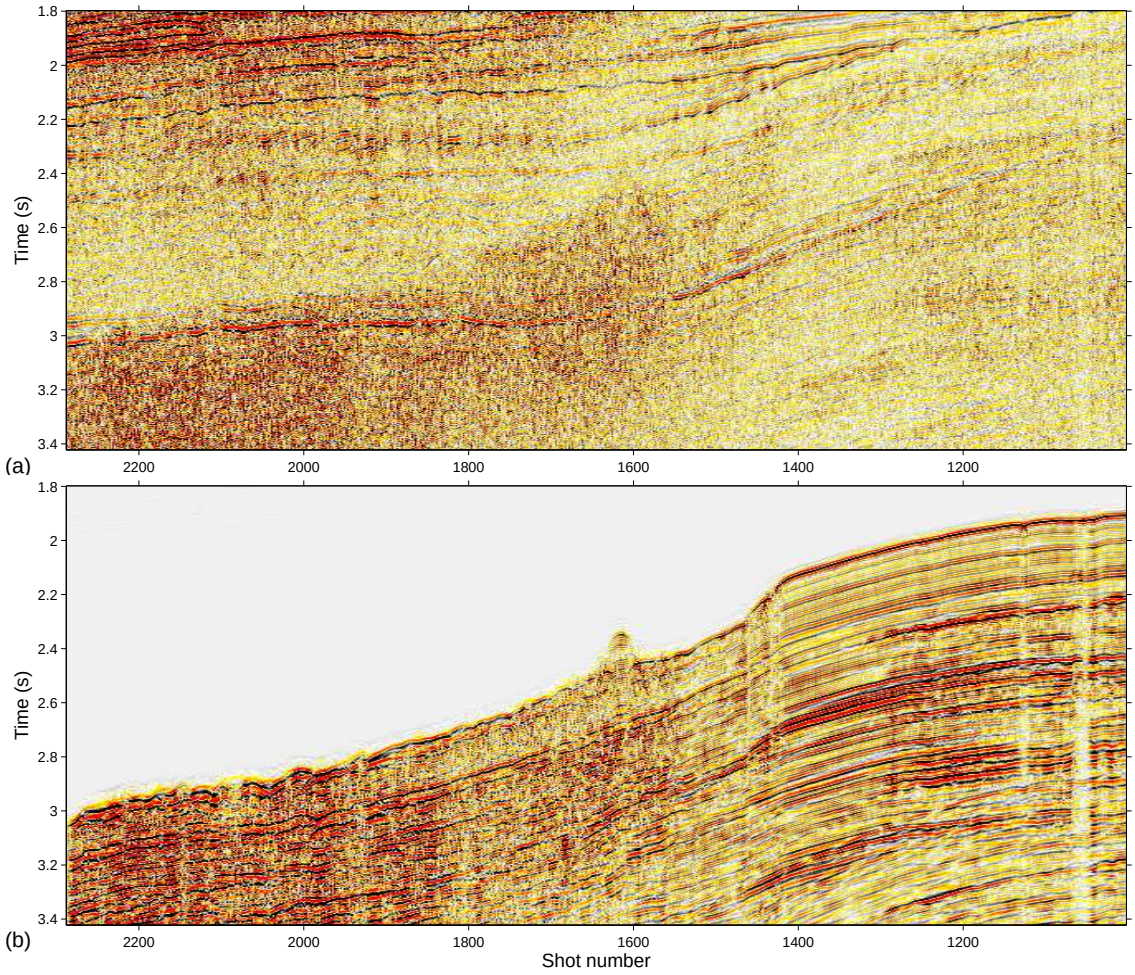

Figure 5: Subtraction results a standard 2-D adaptive filter in time-space domain on a near receiver plane. (a) Filtered data. (b) Adapted multiples.

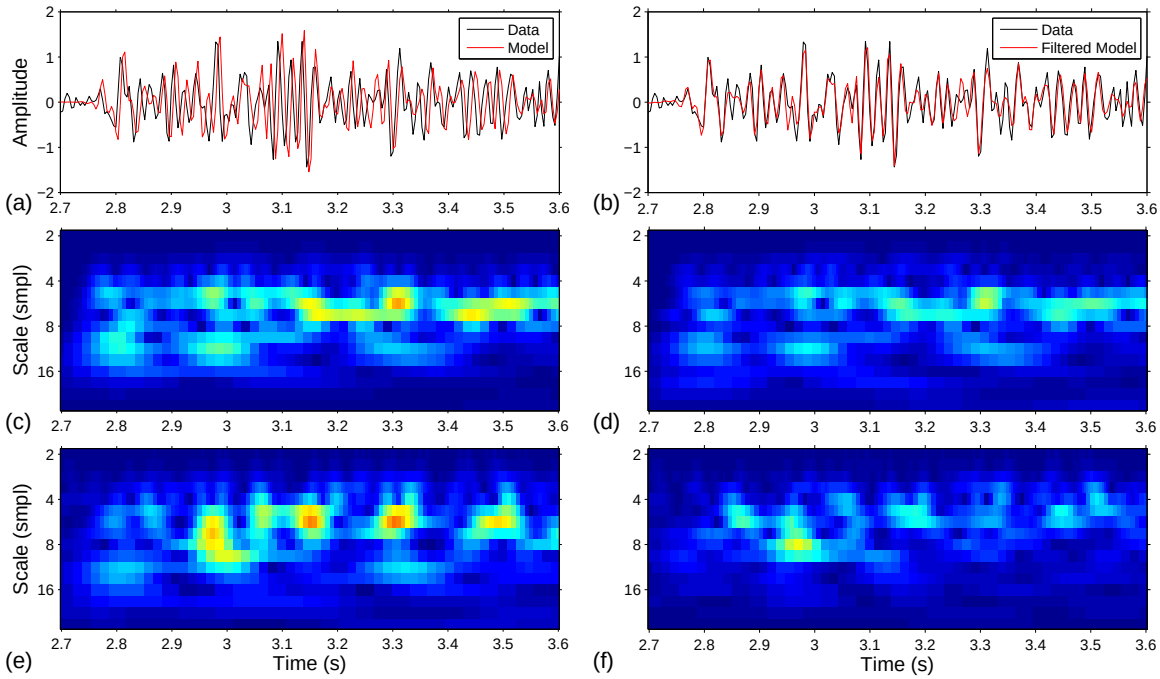

Figure 6: Subtraction algorithm details at the intersection between the planes shown on Figures 1 and 3. (a) Recorded data with respect to the multiple model in time and (b) with respect to adapted model. Modules in the time-scale domain of (c) the multiple model, (d) the adapted multiples, (e) the recorded data, and (f) the filtered data. The signals in (a) and (b) are attenuated by a factor of 1000 except the multiple model that is 2×10^7. The sampling rate is 250 samples/s.