

# Internal multiple attenuation by Kirchhoff extrapolation

Vincenzo Lipari<sup>1\*</sup>, Carlo Fortini<sup>1</sup>, Emmanuel Spadavecchia<sup>2</sup> and Nicola Bienati<sup>2</sup>

<sup>1</sup>*Politecnico di Milano, Dipartimento di Elettronica e Informazione, Piazza Leonardo da Vinci 32, 20133, Milano, and* <sup>2</sup>*Eni Exploration & Production Division, Italy*

Received August 2012, revision accepted November 2013

## ABSTRACT

Surface-related multiple elimination is the leading methodology for surface multiple removal. This data-driven approach can be extended to interbed multiple prediction at the expense of a huge increase of the computational burden. This cost makes model-driven methods still attractive, especially for the three dimensional case. In this paper we present a methodology that extends Kirchhoff wavefield extrapolation to interbed multiple prediction. In Kirchhoff wavefield extrapolation for surface multiple prediction a single round trip to an interpreted reflector is added to the recorded data. Here we show that interbed multiples generated between two interpreted reflectors can be predicted by applying the Kirchhoff wavefield extrapolation operator twice. In the first extrapolation step Kirchhoff wavefield extrapolation propagates the data backward in time to simulate a round trip to the shallower reflector. In the second extrapolation step Kirchhoff wavefield extrapolation propagates the data forward in time to simulate a round trip to the deeper reflector. In the Kirchhoff extrapolation kernel we use asymptotic Green's functions. The prediction of multiples via Kirchhoff wavefield extrapolation is possibly sped up by computing the required traveltimes via a shifted hyperbola approximation. The effectiveness of the method is demonstrated by results on both synthetic and field data sets.

**Key words:** Kirchhoff, Extrapolation, Attenuation.

## INTRODUCTION

Most current seismic imaging algorithms are based on the single scattering assumption. They, thus, require seismic data to contain primaries only. In order to satisfy this requirement we need to remove, or at least attenuate, free-surface and internal multiples from input data prior to the imaging step. Nowadays the suppression of surface-related multiples has become a routine in seismic data processing, mainly through the application of SRME-like techniques (Verschuur, Berkhout and Wapenaar 1992). On the other hand, the suppression of interbed multiples is still non-routine in most data processing flows. Free-surface multiples usually have stronger amplitudes than interbed multiples. However, strong impedance contrasts

at internal reflectors may create a set of interbed multiples that can obscure primary reflections from sedimentary reflectors. Moreover, the effect of interbed multiples on the interpretation is usually more troublesome than that due to surface-related multiples because the internal multiples have apparent velocities that are often close to those of primary reflections.

Different approaches to internal multiple prediction have been proposed in the past. Several works discuss the extension of SRME for internal multiples elimination (IME). In one of the earliest works on this topic (Berkhout 1982) interbed multiples are predicted by means of downward continuation to the multiple generator of both the source and receiver wavefield. Verschuur and Berkhout (1997) proposed the use of common focus point gathers. Jakubowicz (1998) Showed how to avoid downward continuation by decomposing the internal multiple into three wavefields by means of

\*E-mail: vincenzo.lipari@polimi.it

convolutions and cross-correlations. van Borselen (2002) presented an extension of this method for the removal of internal multiples that have crossed a pseudo-boundary between two strong internal reflectors. A similar strategy was proposed by Ikelle (2006) through the concept of virtual events. Weglein *et al.* (1997) and Matson *et al.* (1999) explained how to attenuate all internal multiples at once through a subseries of the inverse scattering series. This is the most general solution. It requires neither velocity models nor interpreted horizons but it is *even* more expensive compared to data-driven IME. We propose an alternative methodology based on wavefield extrapolation (Wiggins 1988; Pica *et al.* 2005; Pica and Delmas 2008). Spadavecchia *et al.* (2011) used a Kirchhoff wavefield extrapolation (KWE) operator for the prediction of surface multiples. The KWE operator predicts the multiples generated by an interpreted reflector by adding a round trip through the *subsurface* to the recorded data. Here we propose to apply the KWE operator twice for the prediction of the internal multiples generated between two interpreted reflectors (Lipari *et al.* 2011).

For the Kirchhoff extrapolation operator we use asymptotic Green’s functions. Moreover, in order to speed up the procedure, it is possible to calculate the traveltimes on-the-fly by means of a shifted hyperbola approximation, which is less accurate but also less expensive than ray tracing or eikonal solvers.

The main field of application of the proposed methodology is the one of full 3D data where the data-driven methodologies are too expensive or even troublesome. Kirchhoff wavefield extrapolation is also useful for fast-track multiple prediction.

This paper is organized as follows. We first describe the KWE operator and how it can be used for interbed multiple prediction. Secondly, we discuss some issues of the proposed algorithm. Thirdly, we show synthetic and field data examples. Finally, we discuss analogies and differences between the KWE algorithm and data-driven methods.

## KIRCHHOFF WAVEFIELD EXTRAPOLATION

Kirchhoff Wavefield Extrapolation allows to compute the wavefield at a generic point starting from the wavefield acquired over a measurement surface. Thus,  $\mathbf{d}$  being the data recorded over a set  $\{x_1, \dots, x_N\}$  of receivers at surface  $X$ , the wavefield  $\mathbf{r}$  propagated forward in time at a generic set  $\{y_1, \dots, y_m\}$  of virtual receivers at surface  $Y$  can be computed as  $\mathbf{r} = \mathbf{G}\mathbf{d}$ . Here the elements of  $\mathbf{d}$  and  $\mathbf{r}$  are the wavefields for one frequency at all receiver’s locations (*e.g.*,  $\mathbf{d} = (d(x_1, w), d(x_2, w), \dots, d(x_n, w))$ ) and  $\mathbf{G}$  is the forward KWE

Table 1 Notation for Kirchhoff extrapolation operators.

$\mathbf{G}$	Kirchhoff extrapolation operator forward in time and downward in space
$\mathbf{G}_{ab}$	Kirchhoff extrapolation operator forward in time and downward in space from reflector a to reflector b
$\mathbf{G}^T$	Kirchhoff extrapolation operator forward in time and upward in space
$\mathbf{G}^{-1}$	Kirchhoff extrapolation operator backward in time and upward in space
$\mathbf{G}^{-T}$	Kirchhoff extrapolation operator backward in time and downward in space
$\tilde{\mathbf{G}}$	Kirchhoff extrapolation operator $\mathbf{G}^T\mathbf{G}$

operator between  $X$  and  $Y$  (*i.e.*, a discretized version of the well-known Kirchhoff integral):

$$r(y_k, \omega) = \sum_i w_{ik} e^{j\omega\tau_{ik}} d(x_i, \omega). \quad (1)$$

Here  $\tau_{ik}$  is the traveltime from  $y_k$  to  $x_i$ , and  $w_{ik}$  is a weight.

The operator  $\mathbf{G}^T$ , applied to the wavefield measured at  $Y$ , computes the wavefield at  $X$  via forward extrapolation:

$$r(x_i, \omega) = \sum_k w_{ki} e^{j\omega\tau_{ki}} d(y_k, \omega). \quad (2)$$

In vector notation  $\mathbf{r} = \mathbf{G}^T\mathbf{d}$ .

The backward Kirchhoff extrapolation computes the data  $\mathbf{d}$  from the forward propagated data  $\mathbf{r}$ . Therefore, it corresponds to the inverse  $\mathbf{G}^{-1}$  of the forward KWE operator:

$$d(x_i, \omega) = \sum_k \tilde{w}_{ki} e^{-j\omega\tau_{ki}} r(y_k, \omega). \quad (3)$$

In vector notation  $\mathbf{d} = \mathbf{G}^{-1}\mathbf{r}$ .

The notation for the KWE operators is summarized in Table 1.

Surface-related multiples are predicted via KWE in two extrapolation steps (Spadavecchia *et al.* 2011). KWE first downward propagates the recorded data from the acquisition surface to the reflector where the multiple to be predicted made the last upward bounce. Secondly, the resulting output is upward propagated from the reflector to the acquisition surface. The cascade of the two datuming steps can be written as a single operator:

$$\mathbf{m} = \mathbf{G}^T\mathbf{G}\mathbf{d} = \tilde{\mathbf{G}}\mathbf{d}. \quad (4)$$

In our implementation (Spadavecchia *et al.* 2011) the KWE operator  $\tilde{\mathbf{G}}$  can be approximated by means of rays of

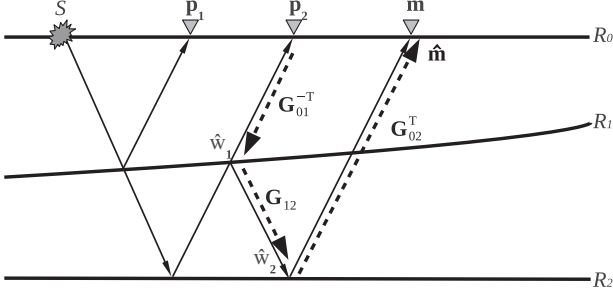


Figure 1 Sketch of the theoretical path for internal multiple reconstruction starting from the primary reflection at the deeper reflector.

minimal traveltime, allowing surface multiples prediction in a single datuming step:

$$m(x_k, \omega) = \sum_i \tilde{w}_{ik} e^{i\omega \tilde{\tau}_{ik}} d(x_i, \omega), \quad (5)$$

where  $\tilde{\tau}_{ik}$  is the Fermatian reflection traveltime (Spadavecchia *et al.* 2011).

## PREDICTION OPERATOR FOR INTERNAL MULTIPLES

Figure 1 shows the raypaths (solid arrows) of three events generated by reflections at horizons  $R_1$  and  $R_2$  and recorded at the measurement surface  $R_0$ . Events  $\mathbf{p}_1$  and  $\mathbf{p}_2$  are the primaries reflected at reflectors  $R_1$  and  $R_2$  respectively and  $\mathbf{m}$  is the interbed multiple generated between  $R_1$  and  $R_2$ . Accordingly, the recorded wavefield is:  $\mathbf{d} = \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{m} + \sum_i \mathbf{e}_i$ , where the events  $\mathbf{e}_i$  are all the recorded events not displayed in Fig. 1. Any KWE operator is linear consequently,  $\mathbf{Gd} = \mathbf{Gp}_1 + \mathbf{Gp}_2 + \mathbf{Gm} + \sum_i \mathbf{Ge}_i$ . Therefore, we start by isolating and analysing the effect on primaries  $\mathbf{p}_2$  of the cascade  $\mathbf{G}_{02}^T \mathbf{G}_{12} \mathbf{G}_{01}^{-T}$  (dashed arrows in the sketch of Fig. 1).

Hereinafter we will use the ‘hat’ notation to denote the outputs of the KWE operator, e.g.,  $\hat{\mathbf{m}}$  is an estimate of multiple  $\mathbf{m}$  obtained through KWE extrapolation.

The backward extrapolation operator  $\mathbf{G}_{01}^{-T}$  applied to  $\mathbf{p}_2$  returns the event  $\hat{\mathbf{w}}_1$ , which is the estimated upgoing primary reflections generated by reflector  $R_2$  and virtually recorded at  $R_1$ :

$$\hat{\mathbf{w}}_1 = \mathbf{G}_{01}^{-T} \mathbf{p}_2. \quad (6)$$

$\hat{\mathbf{w}}_1$  is then forward propagated to  $R_2$  by applying  $\mathbf{G}_{12}$ , resulting in the prediction of interbed multiple  $\hat{\mathbf{w}}_2$  as recorded at  $R_2$ .  $\hat{\mathbf{w}}_2$  is finally propagated forward in time and upward in space through  $\mathbf{G}_{02}^T$  in order to obtain the prediction of the interbed

multiple:

$$\hat{\mathbf{m}} = \mathbf{G}_{02}^T \mathbf{G}_{12} \mathbf{G}_{01}^{-T} \mathbf{p}_2. \quad (7)$$

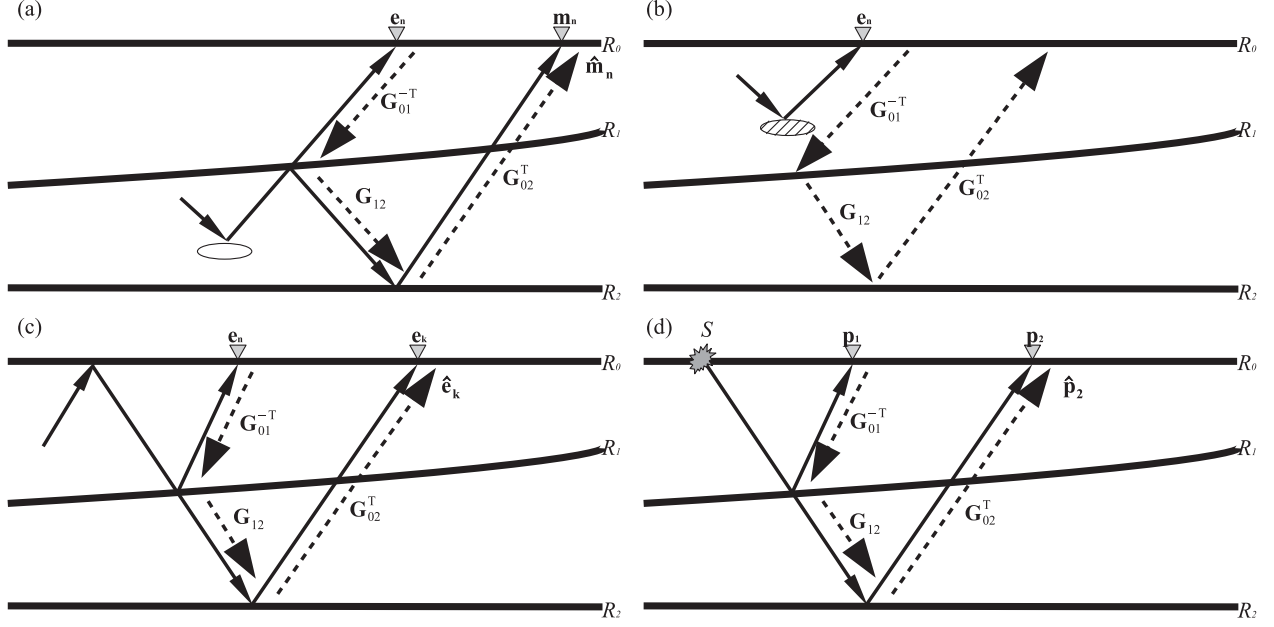
When applied to a specific primary,  $\mathbf{p}_2$ , the cascade  $\mathbf{G}_{02}^T \mathbf{G}_{12} \mathbf{G}_{01}^{-T}$  predicts the desired internal multiple. However, the input data contain other events besides  $\mathbf{p}_2$ . For the sake of completeness we illustrate all these events but in usual processing flows most of them are eliminated by previous processing (e.g., surface-related multiples). The four different cases sketched in Fig. 2 summarize the effect of the described procedure on all other recorded events:

1. Figure 2(a):  $\mathbf{e}_n$  is a generic event whose last upward bounce is deeper than reflector  $R_1$ . The cascade  $\mathbf{G}_{02}^T \mathbf{G}_{12} \mathbf{G}_{01}^{-T}$  transforms  $\mathbf{e}_n$  into  $\hat{\mathbf{m}}_n$ , i.e., the interbed multiple corresponding to  $\mathbf{e}_n$ . Primary  $\mathbf{p}_2$  belongs to this class of events.
2. Figure 2(b):  $\mathbf{e}_n$  is a generic event whose last upward bounce is shallower than reflector  $R_1$ . The cascade  $\mathbf{G}_{02}^T \mathbf{G}_{12} \mathbf{G}_{01}^{-T}$  transforms  $\mathbf{e}_n$  into a virtual event, which does not correspond to any of the actual recorded events.
3. Figure 2(c):  $\mathbf{e}_n$  is a multiple whose last upward bounce occurs on  $R_1$ . The cascade  $\mathbf{G}_{02}^T \mathbf{G}_{12} \mathbf{G}_{01}^{-T}$  transforms  $\mathbf{e}_n$  into a corresponding multiple whose last upward bounce occurs on  $R_2$ .
4. Figure 2(d):  $\mathbf{p}_1$  is the primary of  $R_1$ . The cascade  $\mathbf{G}_{02}^T \mathbf{G}_{12} \mathbf{G}_{01}^{-T}$  transforms this event into  $\hat{\mathbf{p}}_2$ , i.e., the primary of the deeper reflector  $R_2$ .

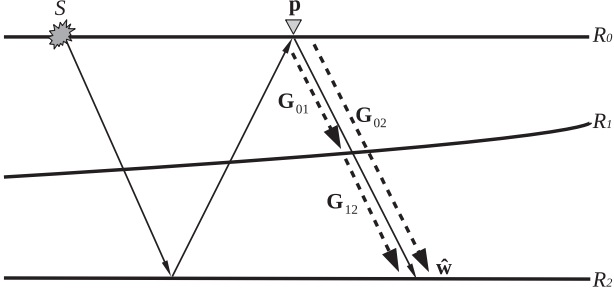
The goal of multiple attenuation is to remove the multiples while leaving the primaries untouched. This goal is achieved after prediction by adaptively subtracting the predicted multiples from the recorded data. Accordingly, case 4 must be specifically addressed as  $\hat{\mathbf{p}}_2$  is the prediction of a primary event and the adaptive subtraction filter would attenuate the corresponding primary  $\mathbf{p}_2$  in the multiple free data.

In order to avoid the presence of  $\hat{\mathbf{p}}_2$  on the predicted multiple gather it is sufficient to apply the KWE operator only to the portion of the data recorded after the reflection of the first interpreted reflector. This muting procedure will not affect the prediction of the later events because KWE is a linear operator. Alternatively, it is possible to mute out the event at an intermediate prediction step, as explained in the next section.

Although the events predicted in case 2 are undesired *too*, they do not have the same moveout as the primaries and a well-designed 2D adaptive filter should properly handle it (Wu and Dragoset 2011). Moreover, since before the KWE extrapolation we mute out the events recorded before the primary  $\mathbf{p}_1$ , this type of spurious event can only arise from



**Figure 2** Sketch of the effect of the internal multiple predictor operator on different events: (a) last upward reflection deeper than the shallower reflector; (b) last upward reflection above the shallower reflector; (c) multiple with last upward reflection on the shallower reflector; (d) primary of the shallower reflector.



**Figure 3** Sketch of the datuming between reflectors  $R_0$  and  $R_2$  as a single operator  $G_{02}$  and as the cascade of the operators  $G_{01}$  and  $G_{12}$ .

multiples. Therefore, they usually have lower amplitude than the surrounding events. This will ease the adaptive subtraction.

In case 3 the operator  $G_{02}^T G_{12} G_{01}^{-T}$  predicts peg-legs of the deeper interpreted reflector. These events are correctly removed by adaptive subtraction.

## COMPACT PREDICTION OPERATOR

Each one of  $G_0$ ,  $G_{02}$  and  $G_{12}$  requires the computation of the Green's function traveltimes between the points of the involved reflectors.

Figure 3 illustrates two different ways to obtain the prediction  $\hat{w}$  of the event  $w$  at  $R_2$  from the recorded event  $p$

at  $R_0$ . It is possible to directly propagate  $p$  to  $R_2$  obtaining  $\hat{w} = G_{02}p$ . Alternatively, it is possible to first compute the datum at  $R_1$  through  $G_{01}$ , then to apply the operator  $G_{12}$  to the resulting wavefield obtaining  $\hat{w} = G_{12}G_{01}p$ . Therefore, the operator  $G_{02}$  is equivalent to the cascade of  $G_{01}$  and  $G_{12}$ :

$$G_{02} = G_{12}G_{01}. \quad (8)$$

The last identity is exact only for ideal full-aperture operators. In practice this is not the case. However, this is not critical for the quality of the result, especially because in multiples an accurate amplitude estimation prediction is not essential. The summation surface must contain the involved raypaths or, for a more accurate identity, the involved Fresnel zones.

In order to avoid the computation of the traveltimes between  $R_1$  and  $R_2$  the operator  $G_{12}$  can be expressed in terms of the free surface-related Green's functions:

$$G_{12} = G_{02}(G_{01})^{-1}. \quad (9)$$

By doing this expansion the internal multiple prediction operator becomes the cascade of two extrapolation steps involving the fast KWE operator  $\tilde{G}$  described in equation (5), which is the *same* operator used for the prediction of surface multiples (Spadavecchia *et al.* 2011):

$$\hat{m} = (G_{02})^T G_{02} (G_{01}^T G_{01})^{-T} p_2 = \tilde{G}_{02} (\tilde{G}_{01})^{-T} p_2. \quad (10)$$

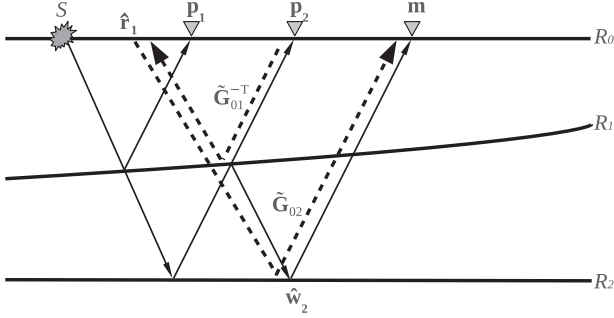


Figure 4 Sketch of the raypaths of the two operators for internal multiple prediction.

In principle it is possible to apply the prediction in a single step by precomputing the operator  $\tilde{\mathbf{G}}_{02}(\tilde{\mathbf{G}}_{01})^{-T}$ . However we decided to split the prediction in two steps in order to highlight the fact that the prediction of internal multiples is performed by a double application of the KWE operator used for surface multiple prediction. Therefore, it is possible to predict interbed multiples with the same code (with minimal modifications) used for the prediction of the surface multiples of interpreted reflectors. Moreover in our implementation we compute (or read) the traveltimes on-the-fly, which makes the implementation in two steps more efficient than the precomputation of the full prediction operator. Figure 4 shows the raypaths implied by this algorithm.

Figure 5 shows the outputs of the two extrapolation steps for a synthetic data set generated by finite differences on a 1D velocity model. There are only two layers and the corresponding reflectors are  $R_1$  and  $R_2$ . Since in the finite differences solution the upper boundary is not absorbing, there are surface multiples too but all the events in the input panel between the blue and red lines are interbed multiples.

The result shows that for this example the algorithm correctly predicts internal multiples.

As mentioned previously, it is necessary to avoid the prediction of the primaries  $\hat{\mathbf{p}}_2$ . One possibility is to extrapolate only a portion of the data, starting from the traveltime of the primary generated by reflector  $R_1$ . This strategy is followed in the example of Fig. 5. Another way is to apply a muting window around time  $t = 0$  and zero offset in the intermediate output. Indeed  $(\tilde{\mathbf{G}}_{01})^{-T}$  maps the primary  $\mathbf{p}_1$  into an impulse at the source. Figure 6 shows the intermediate output without (left) and with (right) windowing. When the window is not applied the primary  $\mathbf{p}_1$  is actually mapped into an impulse at the source position (highlighted by the square). The output after the second extrapolation step is shown in Fig. 7. With-

out muting (left)  $\mathbf{p}_1$  is mapped into  $\hat{\mathbf{p}}_2$ . The muting window (right) successfully rejects the unwanted prediction of  $\mathbf{p}_2$ .

## PRACTICAL ISSUES

### Computation of the traveltimes

As shown in equation (1) the key elements of the operator  $\mathbf{G}$  are the Green's functions traveltimes  $\tau_{ik}$  and the weight function  $w_{ik}$ . The computation of Green's functions is not trivial. In order to limit the computational cost we decided to use the asymptotic approximation of the Green's function:  $g_{ij}(\omega) = A(\mathbf{x}_i, \mathbf{y}_j)e^{-j\omega\tau(\mathbf{x}_i, \mathbf{y}_j)}$ . Here  $\mathbf{x}_i$  and  $\mathbf{y}_j$  are the grid points over the selected surfaces and  $A(\mathbf{x}_i, \mathbf{y}_j)$  and  $\tau(\mathbf{x}_i, \mathbf{y}_j)$  are the geometrical spreading and the traveltime between  $\mathbf{x}_i$  and  $\mathbf{y}_j$  respectively. Accordingly we should use ray tracing as the computational engine. For further accelerating the internal multiple prediction we can use the shifted hyperbola approximation (Castle 1994) of the traveltimes that allows computing the operator  $\hat{\mathbf{G}}$  on the fly. The shifted hyperbola approximation handles also VTI anisotropy (Fomel and Grechka 2001) and provides a reasonable approximation of non-hyperbolic moveout at large offsets.

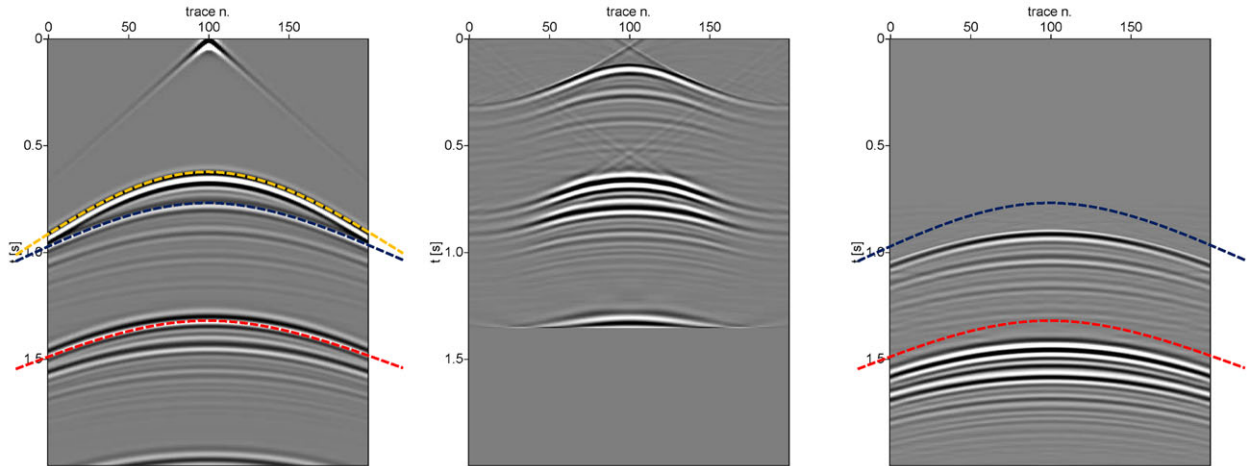
The errors in the velocity model, horizon picking and traveltimes computation can be expected to affect multiple prediction accuracy (more details can be found in Spadavecchia *et al.* 2011). Compared to the other sources of error the shifted hyperbola approximation has proved to be an acceptable approximation in many practical situations.

### Computation of weights

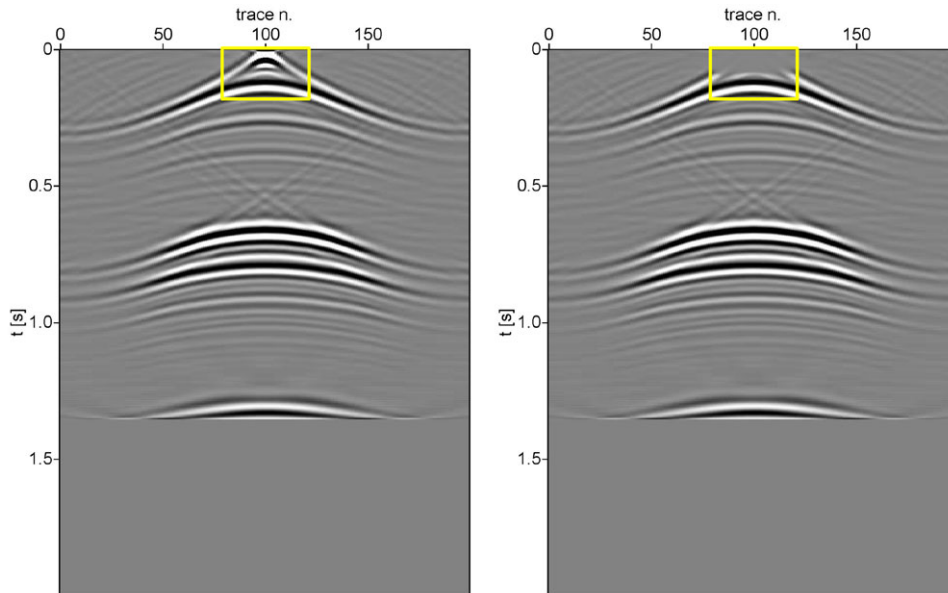
The weight function  $w_{ik}$  accounts for geometrical spreading, acquisition geometry, emergence angle and for the difference between the 2D and 3D operator. For the purpose of multiple prediction and attenuation, since we apply adaptive subtraction in order to obtain the multiple free data, an accurate computation of the weights  $w_{ik}$  is not required. We only use the weights with the goal of easing the adaptive filtering procedure. In practice we take into account only geometrical spreading in order to avoid the measured wavefield and the predicted multiples to be orders of different magnitude.

### Aliasing and regularity issues

With reference to the issues related to the acquisition geometry we distinguish between those related to sampling density and those related to regularity.



**Figure 5** Stages of the internal multiple prediction algorithm. The left panel shows the input common shot gather, superimposed with the traveltimes of the primaries (yellow and blue lines) and of the first-order surface multiple (red line). The central panel is the output of the backward extrapolation step. The right panel shows the predicted multiples.



**Figure 6** Intermediate step with complete data extrapolation (left) and the result of applying a window at the source position (right). The primary  $p_1$  is mapped into an impulse at the source position (highlighted by the square) in the left panel. The impulse is strongly attenuated in the panel at the right.

Like any other Kirchhoff operator anti-aliasing is an important component, particularly when dealing with 3D acquisitions, since the cross-line sampling is often more sparse than the in-line sampling (especially for marine acquisitions). In order to avoid aliasing noise, standard methods suppress some of the high-frequency components of the data by performing operator dip dependent low-pass filtering. An alternative method is to precompute multiple low-passed versions of the

input traces and *then* select *the* appropriate input data during summation (Gray 1992) according to the actual operator dip. In our implementation the storage requirements are reduced through linear interpolation among a limited set of precomputed low-passed versions of the input data.

The approximation of the Kirchhoff integral by a summation implicitly assumes a regular distribution of the integrand samples. If we are dealing with an irregular sampling of the



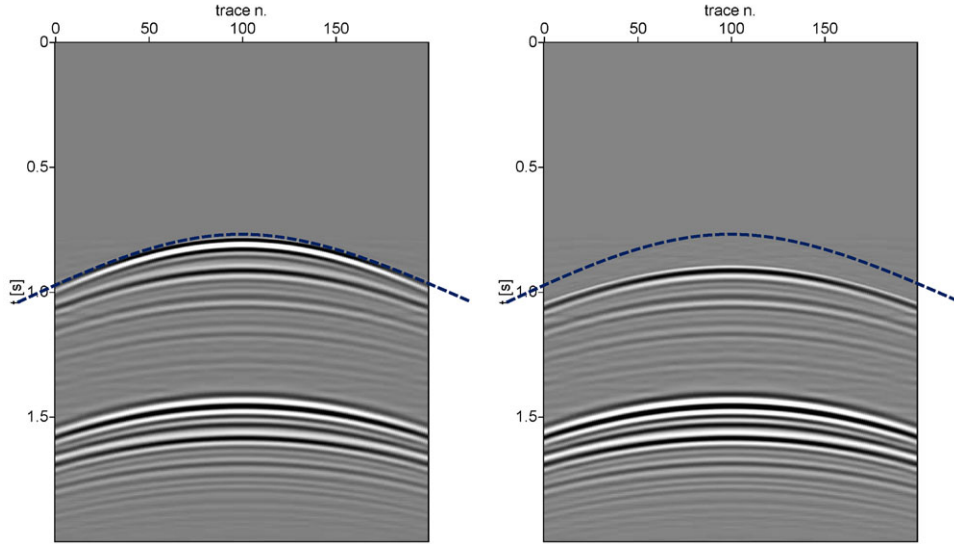


Figure 7 Final prediction result with unmuted data extrapolation (left) and after extrapolating the windowed data panel (right). The dotted lines correspond to the traveltime of the primary  $p_2$ .

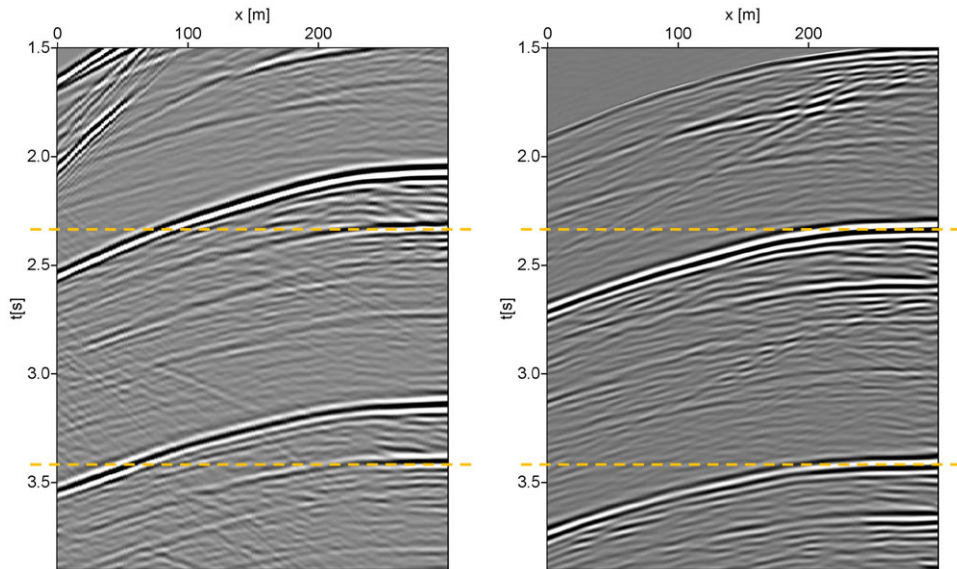


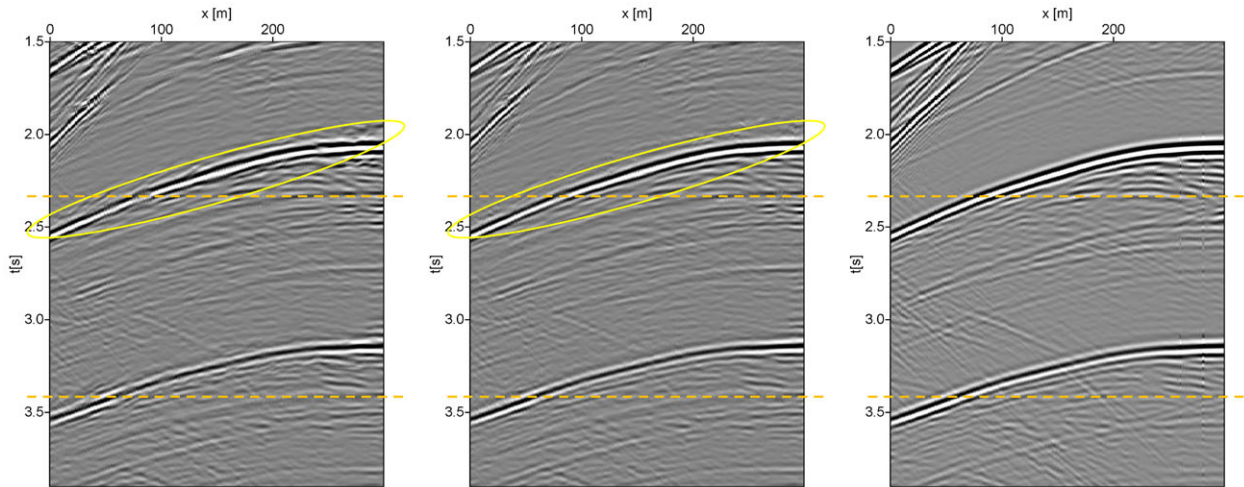
Figure 8 Detail of an input data shot gather (left) and internal multiples prediction (right) for synthetic data set 1.

integral (i.e., an irregular geometry of the receivers) we will introduce an error in the approximation that could cause, in some cases, summation artefacts. The irregularity of the geometry of the receivers can be taken into account through the weight function  $w_{ik}$ .

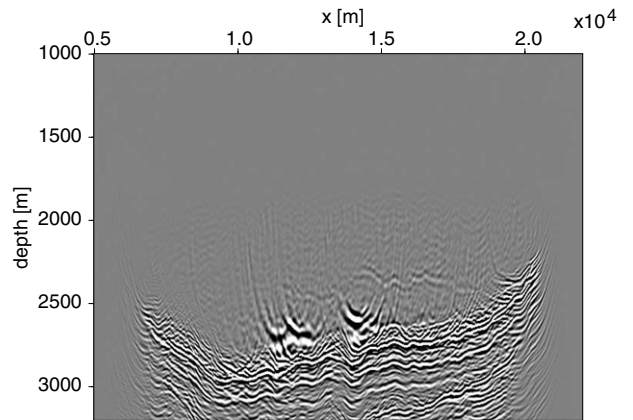
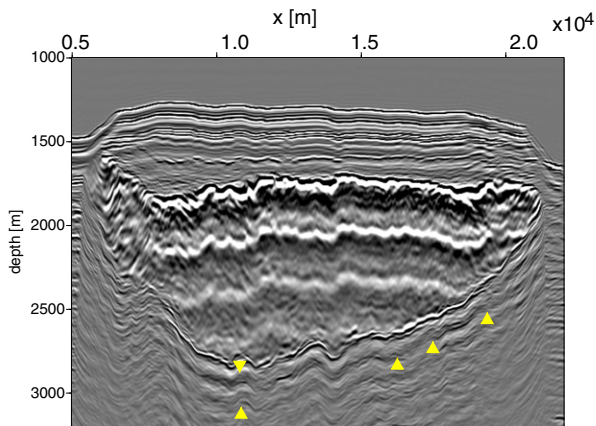
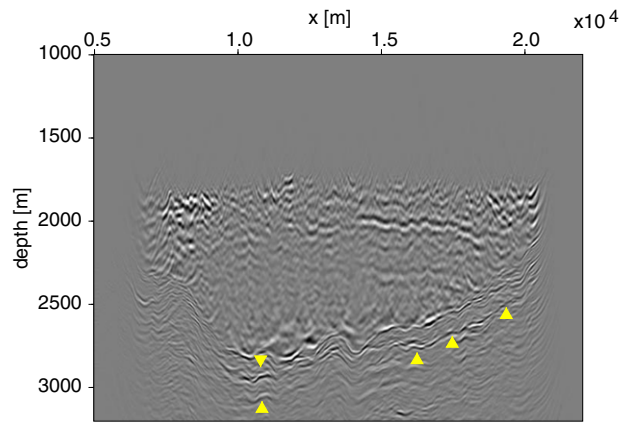
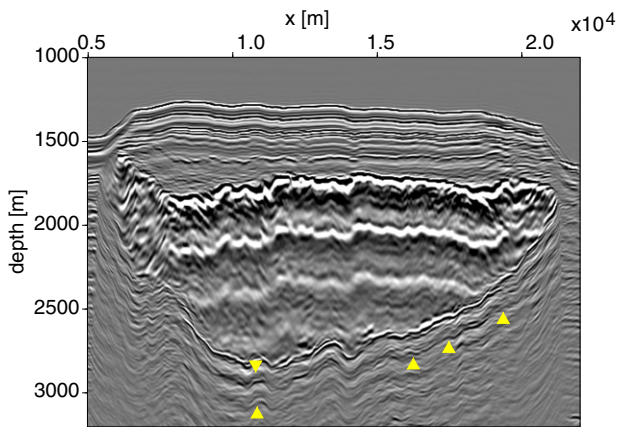
It is also possible to obtain a cleaner result by prior regularization but in this case it would be just a regularization over the receivers, which is much simpler than the multidimensional regularization needed by data-driven IME. In-

deed KWE just needs a regular receiver distribution for each shot gather in order to have a regular tessellation of the integration surface, while data-driven IME needs a regular distribution of sources and receivers and one shot location for each receiver.

In multiple attenuation, however, summation artefacts have certainly a milder impact than in Kirchhoff migration since the final result is obtained after adaptive subtraction and we expect multidimensional adaptive subtraction to be



**Figure 9** Subtraction results for a shot gather of synthetic data set 1: L2 adaptive subtraction result (left) and hybrid norm adaptive subtraction result center. The yellow ovals highlight a strong primary event. The data set modelled without internal multiples is shown at the right.



**Figure 10** PSDM before (top) and after (bottom) internal multiples prediction and subtraction. The yellow arrows point to events related to internal multiples.

**Figure 11** Difference between the images of Fig. 10 (top). The yellow arrows point to events related to internal multiples. PSDM of the predicted multiples (bottom)



able to deal with these artefacts (Dragoset 1995). Thus, since the method we propose is especially addressed to fast-track projects, the KWE operator can be applied directly to the actual acquisition geometry without any kind of regularization. Moreover, in the 2D case, we did not apply any kind of anti-aliasing filtering, while still obtaining satisfactory results.

## EXAMPLES

### 2D synthetic data

We tested our internal multiple prediction algorithm on synthetic data generated by applying 2D acoustic modelling to a section of the velocity model taken from a 3D PSDM project. The velocity is rather smooth and reflectors were introduced as discontinuities in the density model. By applying Kirchhoff extrapolation we were able to predict internal multiples generated by the reverberation between two reflectors. After the prediction step we adaptively subtracted the multiples from the input data in order to obtain the estimated primaries. The adaptive subtraction was performed in the shot gathers domain. Fig. 8 shows a detail of an input data shot gather and the prediction of internal multiple events; Fig. 9 shows the result of the 2D adaptive subtraction using two different shaping filters: adaptive subtraction using the L2 norm (left) and adaptive subtraction using a hybrid norm centre. In the case of interbed events it is better to perform the adaptive subtraction in a L1 or an hybrid norm instead of using the least-squares method in order to better preserve primaries. Indeed the most common situation with interbed multiples is characterized by strong events (primaries) and very close weak multiples. In this case the estimated L2 filter mainly subtracts energy from the primaries rather than eliminating the estimated multiples. Accordingly, it is better to perform adaptive filtering using a hybrid L2/L1 norm, which provides the benefits of the L1 norm while preventing the filter to become unstable (Costagliola, Mazzucchelli and Bienati 2011; Guitton and Verschuur 2004). For comparison the right panel in Fig. 9 shows the same portion of the data generated without internal multiples. These data have been obtained as the summation of two distinct datasets, each one generated via finite difference modelling on a velocity model with only one of the two discontinuities.

### 2D field data

For a field data test we used the data set from the Mississippi Canyon proposed for the Multiple Attenuation Workshop at

the 1997 SEG meeting. As multiple generator horizons we picked the water bottom and the top of salt. After applying the Kirchhoff extrapolation process we subtracted the predicted multiples from the original data by applying the hybrid-norm adaptive subtraction on the shot gathers before prestack depth migration (PSDM). Figure 10 shows the comparison between the PSDM before and after multiple subtraction (from top to bottom). It is apparent that the multiples highlighted by the arrows have been almost completely removed after prediction and subtraction. Likewise there are other multiples that have been strongly attenuated. The prediction is not perfect (mainly due to inaccuracies in the velocity model between the water bottom and top of salt) and not everywhere adaptive subtraction is able to recover the inaccuracies in the prediction.

In the image at the top of Fig. 11 the difference between the previous images is shown. Note that in some locations the salt bottom has been deteriorated. In the image at the bottom of Fig. 11 the PSDM of the predicted multiples is shown.

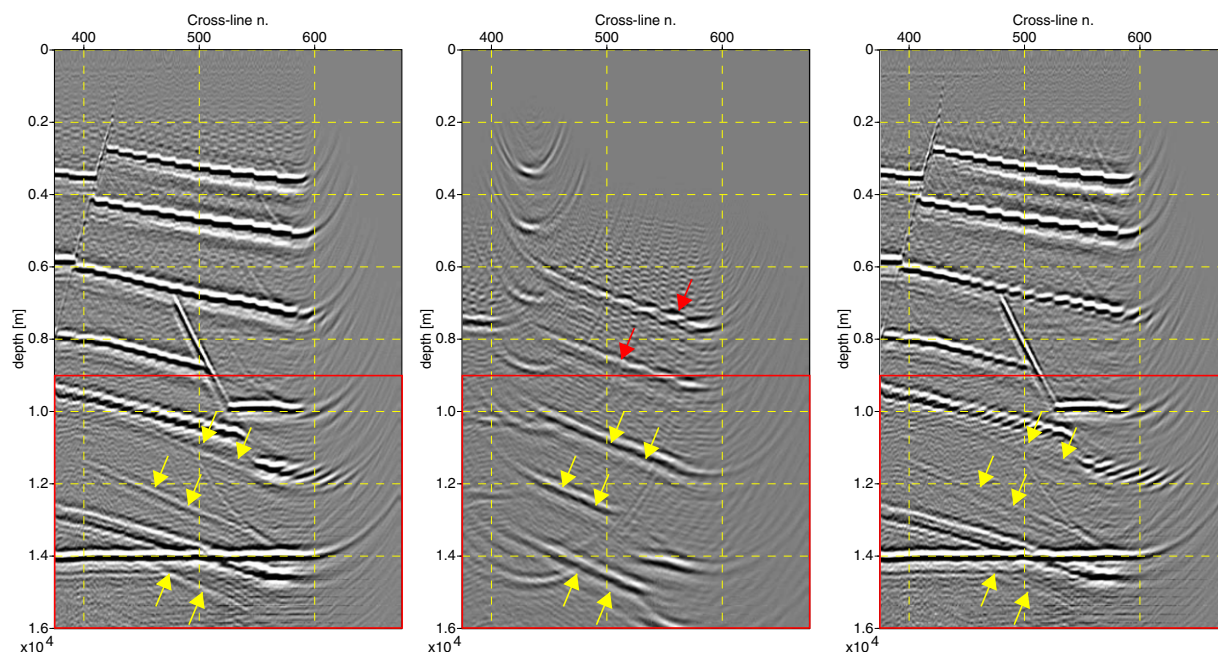
Finally, it is to be noticed that we did not apply any preprocessing to the data. Therefore, the image could contain further internal multiples not related to the reflectors that we picked for our prediction, or effects from the airgun bubble.

### 3D synthetic data

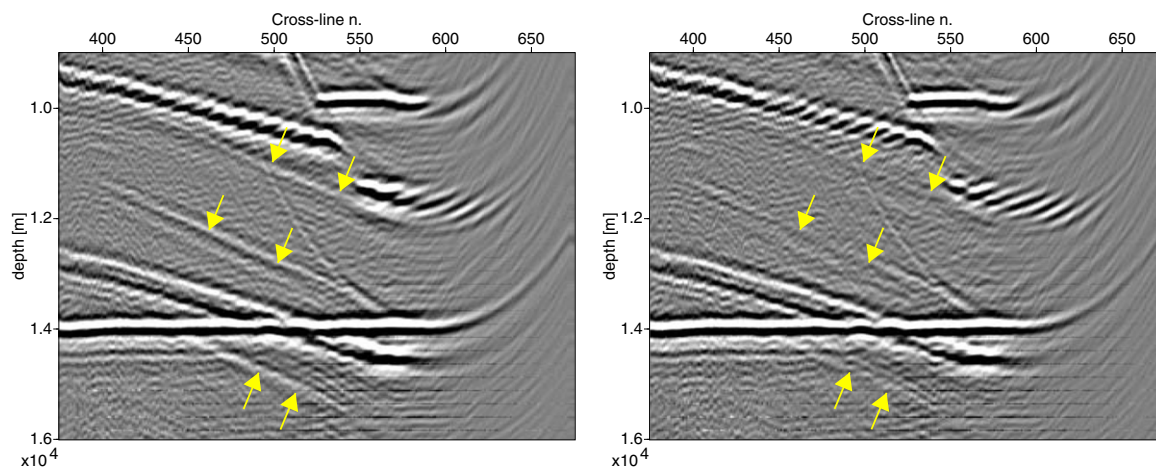
For a test on 3D synthetic data we used the EAGE/SEG Salt3D C3NA data set. Since we had no clearly identified internal multiple generators the prediction procedure was tentatively applied by selecting several different couples of reflectors. Here we show only the result from the couple of reflectors that seem to introduce the strongest internal multiples. After applying the Kirchhoff extrapolation we subtracted the predicted multiples from the original data by using a hybrid-norm adaptive subtraction applied before PSDM.

In the example of Figs. 12 and 14 we picked the sea-bed and the next shallowest reflector. The two figures show two different PSDM sections (in-line 100 and cross-line 500 respectively), before (left) and after (right) multiple subtraction whilst the central images show the PSDM of the predicted multiples. The yellow arrows highlight some of the multiples that have been successfully attenuated. The predicted multiples do not show the same lateral extent as the original multiples because of the limited receiver aperture.

In Figs. 13 and 15 we show a close-up of the areas highlighted by the red boxes in Figs 12 and 14.



**Figure 12** Migrated in-line sections of the Salt3D C3NA data set, before (left) and after internal multiple subtraction (right). The central panel shows the PSDM of the predicted multiples. The yellow arrows highlight some of the subtracted internal multiples. The red arrows highlight some internal multiples superimposed to primaries.



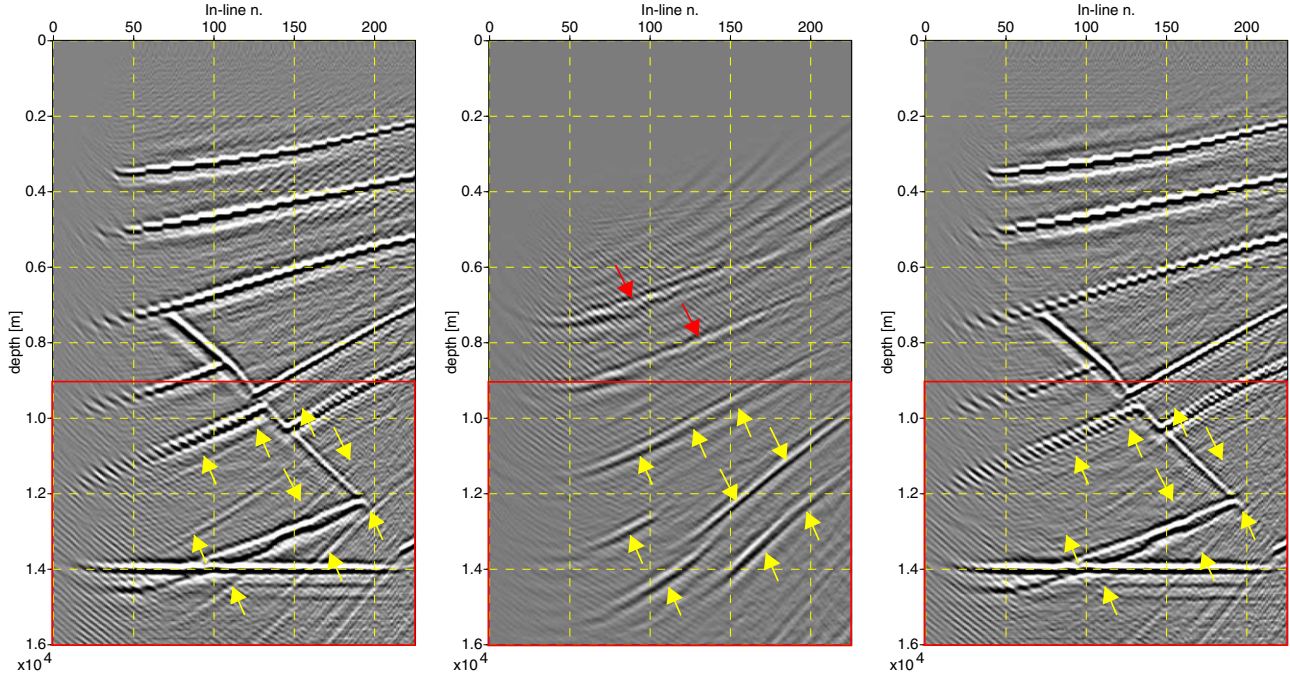
**Figure 13** Zoom of the migrated sections of Fig. 12 before (left) and after internal multiple subtraction (right).

This case was very challenging for adaptive subtraction because many of the predicted multiples were often crossing primaries (e.g., the events pointed out by red arrows).

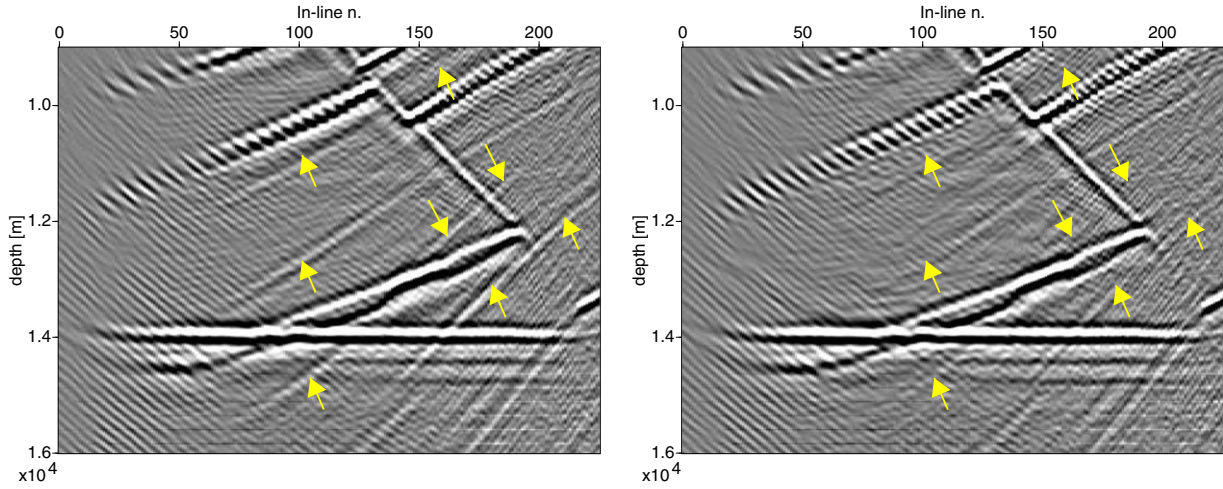
## DISCUSSION

The decomposition of multiple prediction into two wavefield extrapolations (backward and forward in time) allows to make a simple analogy between our algorithm and some of the

data-driven algorithms proposed in the past. In Jakubowicz (1998) it was shown that interbed multiples can be reconstructed by cross-correlating the reflected wavefield generated below the shallowest multiple generator with the primary of this reflector, followed by a convolution of this result with the wavefield reflected below the multiple generator. While in Jakubowicz (1998) the three wavefields that are combined to predict interbed multiples are separated by windowing the data, in our method the first component is the measured



**Figure 14** Migrated cross-line sections of the Salt3D C3NA data set, before (left) and after internal multiple subtraction (right). The central panel shows the PSDM of the predicted multiples. The yellow arrows highlight some of the subtracted internal multiples. The red arrows highlight some internal multiples superimposed to primaries.



**Figure 15** Zoom of the migrated sections of Fig. 14 before (left) and after internal multiple subtraction (right).

wavefield and then the numerical extrapolations provide the other wavefields needed for the prediction. The backward extrapolation is analogous to the correlation with the primary of the multiple generator while the forward extrapolation step is analogous to the convolution in Jakubowicz's method.

Indeed it is possible to give the following interpretation to equation (5):

$$\hat{r}(\mathbf{y}_k, \omega) = \sum_i \hat{w}_{ik} e^{j\omega \hat{\tau}_{ik}} d(x_i, \omega) = \sum_i U(\mathbf{x}_i, \mathbf{y}_k, \omega) d(\mathbf{x}_i, \omega), \quad (11)$$



where  $U(\mathbf{x}_i, \mathbf{y}_k, \omega) = \hat{w}_{ik} e^{j\omega \hat{t}_{ik}}$  is an asymptotic approximation of the reflected wavefield between  $\mathbf{x}_i$  and  $\mathbf{y}_k$ . In this way the forward KWE multiple predictor is a sum of convolutions, similarly to SRME but with the difference that we are convolving synthetic and real traces taken only from the analysed shot. Consequently the backward operator performs the following sum of correlations:

$$\hat{d}(\mathbf{x}_i, \omega) = \sum_k \hat{w}_{ik} e^{-j\omega \hat{t}_{ik}} r(\mathbf{y}_k, \omega) = \sum_k U^*(\mathbf{x}_i, \mathbf{y}_k, \omega) r(\mathbf{y}_k, \omega). \quad (12)$$

The windowing procedure required in Jakubowicz's method in order to isolate the primary event related to the shallower interpreted reflector corresponds to the horizon picking needed for wavefield extrapolation.

The main difference between data-driven IME techniques and our methodology is that KWE-based multiple prediction makes no assumptions on the regularity of the acquisition geometry. Furthermore, the prediction is done on each common shot gather independently, allowing a straightforward parallel implementation of the algorithm. Since the KWE operator for surface-related multiples is used twice, the computational cost is twice that of KWE-based surface multiple prediction. Moreover, since we are not convolving or correlating input traces among themselves, the KWE predicted multiples have the same wavelet as the data whereas in the data-driven IME the wavelet is actually cubed. Another strength of the proposed algorithm is that the prediction is performed without borrowing input from the neighbouring shots. This is an advantage since the cost of I/O and data movement is going to become more and more significant in modern computing hardware. There are, of course, also drawbacks: KWE-based prediction requires also the picking of a second primary generator, moreover it predicts only the interbed multiples generated by strict specular reflection between the two picked reflectors while in data-driven IME all the multiples generated between the multiple generator and the deeper reflectors are predicted in a single step.

Given all these facts the KWE method is not intended to substitute data-driven methods but it can be helpful in some specific situations. The main area of application of the KWE methodology is that of full 3D data. There are several implementations of 2D internal multiple prediction methodologies that provide a more accurate and complete result. Therefore, in the 2D case the field of application of the proposed method is that of fast-track projects when there is not enough time for more accurate techniques.

On the other hand the 3D case is probably the more appropriate area of application for the KWE methods since

the other methodologies are expensive or even troublesome because of the large amount of missing data in 3D data sets.

## CONCLUSIONS

We developed a model-driven algorithm for interbed multiple prediction based on Kirchhoff wavefield extrapolation. Dependency on a subsurface model is both the main advantage and disadvantage of the method. In particular, model dependency provides the information needed to predict internal multiples of a given shot without borrowing input traces from neighbouring shots. This can be an advantage for implementation on modern computing hardware where the cost of data movement is going to become more and more significant. On the other hand the algorithm requires the identification and the interpretation of the involved reflectors, which is not a trivial task and may imply a trial-and-error approach. Moreover, only a restricted class of interbed multiples can be predicted. The computational kernel exploits two shortcuts in order to speed up performance. First, in the asymptotic Green's function computation we use the shifted hyperbola to compute traveltimes. Second, we apply a stationary phase to approximate the sandwich of cascaded datuming steps into a single datuming step, obtaining a further performance improvement. We tested the algorithm successfully on both synthetic and field data. The results show that the Green's function computation by means of a shifted hyperbola is reliable over a reasonably large offset range. Nonetheless, the most difficult task when dealing with field data remains the identification of the internal multiple generating reflectors.

## ACKNOWLEDGEMENTS

The authors wish to thank Eni E&P management for permission to publish this work and WesternGeco for releasing the Mississippi Canyon data.

## REFERENCES

- Berkhout A. 1982. *Seismic migration: imaging of acoustic energy by wave field extrapolation*. Elsevier Scientific Publishing Company.
- Castle R.J. 1994. A theory of normal moveout. *Geophysics* 59, 983–999.
- Costagliola S. Mazzucchelli P. and Bienati N. 2011. Hybrid Norm Adaptive Subtraction for Multiple Removal. 73rd EAGE Conference & Exhibition, Expanded Abstracts.

- Dragoset B. 1995. Geophysical applications of adaptive noise cancellation. 65th SEG Meeting, *Expanded Abstracts* 14, 1389–1392.
- Fomel S. and Grechka V. 2001. Nonhyperbolic reflection moveout of P-waves: An overview and comparison of reasons. Technical Report CWP-372.
- Gray S.H. 1992. Frequency-selective design of the kirchhoff migration operator. *Geophysical Prospecting* 40, 565–572.
- Guitton A. and Verschuur D.J. 2004. Adaptive subtraction of multiples using the L1-norm. *Geophysical Prospecting* 52, 27–38.
- Ikelle L. 2006. A construct of internal multiples from surface data only: the concept of virtual seismic events. *Geophysical Journal International* 164, 383–393.
- Jakubowicz H. 1998. Wave equation prediction and removal of interbed multiples. 60th EAEG Conference & Exhibition, *Expanded Abstracts* 1527–1530.
- Li Y., Zhang Y. and Claerbout J. 2010. Geophysical applications of a novel and robust L1 solver. 80th SEG Meeting, *Expanded Abstracts*, 29, 3519–3523.
- Lipari V., Fortini C., Spadavecchia E. and Bienati N. 2011. Internal multiple attenuation by Kirchhoff extrapolation. 73th EAGE Conference & Exhibition, Expanded Abstracts.
- Matson K., Corrigan D., Weglein A., Young C. and Carvalho P. 1999. Inverse scattering internal multiple attenuation: Results from complex synthetic and field data examples. 69th SEG Meeting, Expanded Abstracts 18, 1060–1063.
- Pica A. and Delmas L. 2008. Wave equation based internal multiple modeling in 3D. 78th SEG Meeting, Expanded Abstracts 27, 2476–2480.
- Pica A., Poulain G., David B., Magesan M., Baldock S., Weisser T. et al. 2005. 3D surface related multiple modeling, principles and results. 75th SEG Meeting, Expanded Abstracts 24, 2080–2083.
- Spadavecchia E., Lipari V., Bienati N., Drufuca G. and Cazzola L. 2011. Water bottom multiple attenuation by Kirchhoff extrapolation. 73th EAGE Conference & Exhibition, Expanded Abstracts.
- van Borselen R. 2002. Fast track, Data driven interbed multiple removal Application to the North Sea. 64th EAGE Conference & Exhibition, Expanded Abstracts.
- Verschuur D.J. and Berkhout A.J. 1997. Estimation of Multiple scattering by Iterative inversion, part II: Practical Aspects and Examples *Geophysics* 1596–1611.
- Verschuur D.J., Berkhout A.J. and Wapenaar C.P.A. 1992. Adaptive surface related multiple elimination *Geophysics* 57, 1166–1177.
- Weglein A.B., Gasparotto F.A., Carvalho P.M. and Stolt R.H. 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.
- Wu Z.J. and Dragoset B. 2011. Robust internal multiple prediction algorithm. 81st SEG Meeting, Expanded Abstracts.