ANALYSIS OF DATA-DRIVEN INTERNAL MULTIPLE PREDICTION

ADRIANA CITLALI RAMÍREZ

WesternGeco, 10001 Richmond Ave., Houston, TX 77042, U.S.A.
Present address: Statoil, Arkitekt Ebbels vey 10, N-7005 Rotvoll, Norway. ADPER@statoil.com

(Received July 25, 2012; revised version accepted February 10, 2013)

ABSTRACT

Ramírez, A.C., 2013. Analysis of data-driven internal multiple prediction. *Journal of Seismic Exploration*, 22: 105-128.

The internal multiple prediction (IMP) algorithm analyzed in this paper is almost entirely data-driven, requiring a convolution and a crosscorrelation of the input data and information about the main internal multiple generators. The generators or generating horizons are the reflectors where the internal multiples' energy was downward reflected. There are two common approaches to applying IMP:

1) The first is the layer-stripping approach in which internal multiples are predicted starting from the shallowest generator (top-down approach) and subtracted from the input data prior to attempting the prediction using the next horizon as generator. For each generator's prediction, there is a subtraction.

2) The second approach, referred to as the non-top-down approach, predicts the multiples using one horizon at a time, but does not remove the predicted multiples from the input data prior to running the IMP algorithm with the next horizon. The first approach is in agreement with the theory behind this algorithm. The second approach still provides value; however, the same internal multiple can be predicted more than once by different horizons. These predictions have different amplitude information and opposite polarity with respect to each other. Hence, it is not always easy to deal with these internal multiple models when attempting to subtract them from the input data. I provide an analysis of the prediction of internal multiples using IMP with the different approaches.

KEY WORDS: internal multiples, interbed multiples, adaptive subtraction, wave theory, monotonicity condition, seismic processing.

0963-0651/13/$5.00

INTRODUCTION

The purpose of this paper is to describe an analysis that helps in our understanding of the phase and polarity of the internal multiples' prediction using an almost entirely data-driven algorithm. The internal multiple prediction algorithm, IMP, predicts these events by using convolution and correlation operations of preprocessed input data. The input data are preprocessed using information about the reflectors in the subsurface; the need for subsurface information is what makes IMP not entirely data-driven. The IMP algorithm is very similar to the method proposed by the Delphi consortium (Verschuur and Berkhout, 1996; Berkhout and Verschuur, 1997; Verschuur et al., 1998) and further updated by Jakubowicz (1998). Some relevant references for the practical implementation of this method are Moore (2001), El-Emam et al. (2007), Baumstein (2008), and Terenghi et al. (2010). The algorithm used by IMP does not contain the obliquity factors required by the theory based on the wave equation (Araújo et al., 1994; Weglein et al., 1997). The wave-theoretical algorithm described by Araújo et al. (1994) and Weglein et al. (1997), based on the inverse scattering series, is entirely data-driven. It does not require any velocity or moveout discrimination, event picking, or any subsurface information. However, these advantages come at a price: the algorithm works in the frequency-wavenumber domain that requires proper sampling in both receiver and source domains and has a high computational cost. To simplify the requirements of the inverse scattering internal multiple algorithm and to improve efficiency, one could ignore the obliquity factors and use a high-frequency approximation to write the algorithm in the space-time (or space-frequency) domain to obtain an expression similar to the one presented by Jakubowicz (1998) that requires knowledge of the reflectors in the subsurface. This is the algorithm that I refer to as IMP. Thus, the IMP algorithm can be written in the space-time or space-frequency domain, which makes it more efficient but less data-driven and less accurate at the far offsets. IMP is constrained by a total traveltime monotonicity condition (ten Kroode, 2002), which makes it less effective than the inverse scattering algorithm that obeys a vertical time or pseudodepth (constant-velocity F-K migration depth) monotonicity condition, as explained by Nita and Weglein (2009). It is important to mention that a more complete representation of the algorithm in a high-frequency regime could be obtained without excluding the obliquity factors, as demonstrated for the layered media case by Jin et al. (2009).

Some of the assumptions and limitations of IMP are:

1) The most significant multiple-generating horizons were identified and their corresponding primaries were picked in the data.

2) The data are reasonably free of noise.

3) The predicted multiples do not have the correct amplitudes.

4) Adaptive subtraction will be used to match the predicted internal multiples (model) with the actual internal multiples in the data.

In principle, not only the most significant, but all generating horizons in a top-down approach must be identified and their corresponding primary reflections picked in the data. By all generating horizons, in a top-down approach I refer to the set of horizons starting at the shallowest (water-bottom horizon, in the marine case) and ending with the deepest generator of interest. In practice, this is almost never feasible and only the most significant generators are (or can be) properly identified. This more realistic scenario and its implications in the effectiveness on the prediction of internal multiples will be further analyzed and discussed throughout this paper.

The organization of this paper is as follows: The first section provides a general theoretical background. Once the algorithm under analysis is well established, a brief discussion of the practical application of IMP is provided. The practical application has two common approaches: a top-down or layer-stripping approach and a non-top-down approach (Ramírez et al., 2011). The field data application will, most generally, fall in the second approach. The non-top-down approach is not ideal; hence, the result from applying it must be thoroughly understood for a successful application. Therefore, the third section provides an analysis that explains exactly the general output of IMP with the non-top-down approach. These results are discussed in the final section.

METHOD

In general, the IMP algorithm will predict an estimate of the internal multiples for a target location and a target generator using the relation

$$D^{IM}(x_r|x_s;\omega) = -\int_{x_1} dx_1 \int_{x_2} dx_2\, D_1(x_s|x_1;\omega)\, D_2^*(x_1|x_2;\omega)\, D_3(x_2|x_r;\omega) \quad , \qquad (1)$$

where the first space argument, x_r, represents the receiver coordinate and the second argument, x_s, represents the source coordinate, ω is the temporal frequency, the superscript * represents complex conjugation, D_i are the preprocessed input data, and D^{IM} are the predicted internal multiples. The reason why I state that this equation will predict an estimate of the internal multiples for a specific generator is because the left hand side of this equation is a set of events associated to the internal multiples of a target generator with the following characteristics: 1) correct traveltime when compared to the true internal multiples in the input (raw) data, 2) incorrect phase and amplitude due to the fact that the events in D^{IM} contain the effect of three source signatures (the auto-crosscorrelation plus convolution of the raw data source signature), 2) the fact that the left hand side does not compensate for the accumulated radiation pattern (via the obliquity factors that appear in the inverse scattering theory), the left hand side of the equation, needs to be compensated by two obliquity factors, and 3) the amplitude information is an approximation to the one of the true internal multiples. The amplitude estimation has been thoroughly analyzed by Ramírez and Weglein (2005b). It is always less than the true amplitude of the internal multiple in the raw data by a factor of one set (up and down) of transmission coefficients corresponding to the generating reflector and the square of all pairs of transmission coefficients from reflectors above the generating one. In practice, IMP will also have differences due to obliquity and source signature factors which are not properly compensated for in the theory.

Eq. (1) is written in the time domain as a convolution (represented by o) and a crosscorrelation (represented by ⊗),

$$D^{IM}(x_r|x_s;t) = -\int_{x_1} dx_1 \int_{x_2} dx_2\, D_1(x_s|x_1;t) \otimes D_2(x_1|x_2;t) \circ D_3(x_2|x_r;t) \quad . \qquad (2)$$

The difference between the three datasets' input to the algorithm are:

• D_2 is the portion of the data that represents the generating horizon. In other words, it contains the primary event corresponding to the interface in the subsurface interpreted as an internal multiple generator. For a first-order internal multiple, the generating interface is the one where the downward reflection took place. For example, D_2 for the shallowest generator will only contain the information corresponding to the first primary in the data. This is achieved (in the time domain) by muting out all events except for the internal multiple generator.

• D_1 and D_3 have a top mute (in the time domain). It will mute all the information in the data with times shorter than or equal to the traveltime for the event selected as the generator in D_2.

In eq. (2), D_1 and D_3 contain a set of recorded events each with a traveltime longer than the traveltime of the generator (event in D_2). Let the events in D_1 have traveltimes represented by the subscript i, events in D_3 will be then represented with a subscript k, and the primary in D_2 will be represented with the subscript j. The interpretation of eq. (2) is the prediction of internal multiples having traveltimes equal to $(t_i - t_j + t_k)$, corresponding to a combination of longer-shorter-longer events in total traveltime. The traveltimes of the events in D_1 and D_3 are added and the traveltime from the generator, D_2, is subtracted. The three events combined in the algorithm are referred to as subevents of the predicted internal multiple. In general, the predicted traveltimes correspond to the actual traveltimes of true internal multiples in the data.

The IMP algorithm has been tested and validated in terms of its ability to predict the traveltime of the internal multiples in the data (see El-Emam et al., 2007; Hembd et al., 2010, and references therein); it was also extended and validated for marine surveys acquired with poor crossline sampling and for cross-spread land surveys by Terenghi et al. (2010). This paper is an extension that focuses on understanding the phase and polarity of the predicted internal multiples.

IMP IN PRACTICE

There are two common approaches to apply IMP to predict internal multiples:

- The first is the layer-stripping approach in which internal multiples are predicted starting from the shallowest generator (top-down approach) and subtracted from the input data prior to attempting the prediction using the next horizon as generator. For each generator's prediction there is a subtraction. No generators are to be skipped in this approach.

- The second approach, the non-top-down approach, predicts the multiples using one horizon at a time but does not remove the predicted multiples from the input data prior to running IMP with the next horizon.

The first approach is in agreement with the theory behind this algorithm. The second approach is not ideal but still provides value. In the non-top-down approach, the same internal multiple can be predicted more than once by different horizons. These predictions have different amplitude information and opposite polarity. As recognized by Baumstein (2008), the second approach is the most likely to be used in the field data case because we are either not always able to identify each generating horizon (primary in D_2), or to run the algorithm in a top-down approach due to the computational cost involved. Baumstein (2008) analyzed the second approach using synthetic 2D data with two generating primaries (three primaries in total). In the analysis, Baumstein (2008) recognized that making predictions for the second generating horizon without removing the internal multiples generated at the shallowest horizon produces a multiple model for the second horizon dominated by multiples from the first horizon with wrong polarity and amplitude. In other words, the prediction contains internal multiples corresponding to the first and second horizon, where the first set has wrong polarity and the second set has the correct polarity. The prediction has both incorrect amplitudes and mixed polarities. This fact has severe implications in the practical application of this algorithm. The main implication is related to the adaptive subtraction step, making it difficult to accomplish with conventional global search and energy minimization (least-square-based) techniques. In the following, I will provide an analysis of the prediction of internal multiples using IMP with the non-top-down approach. I will show that, in the case studied by Baumstein (2008), the prediction using the second generator does not contain all internal multiples due to the first and second generator. It contains a partial set of internal multiples due to the first generator and all multiples (within the limitations and assumptions of the algorithm) for the second horizon. Furthermore, I will extend and generalize this analysis for any number of horizons and provide a specific set of conditions that are sufficient and necessary to predict an artifact with IMP.

In the next section, a mathematical analysis of the predictions of internal multiples using the non-top-down approach and any number (N) of generating horizons is given. The deeper generators ($3 \leq n \leq N$) can generate artifacts (Ramírez et al., 2011). A set of conditions that the data must satisfy to create these artifacts is also provided.

ANALYSIS

An analysis is useful to understand the inner workings of an algorithm. The amplitude predicted by IMP is not correct; it is an attenuated amplitude further scaled by the effect of two obliquity factors (containing information related to angle of incidence/reflection and geometrical spreading). Ignoring the missing obliquity factors, the predicted amplitude is, in general, smaller than the amplitude of the true internal multiple (Weglein and Matson, 1998; Ramírez and Weglein, 2005a; Lira et al., 2010; Zhang and Shaw, 2010). Furthermore, if the source signature in the data is not deconvolved or compensated for in the prediction process, the prediction contains the effect of three wavelets.

For an analysis of the amplitude prediction the reader is referred to Ramírez and Weglein (2005a) and Ramírez Pérez (2007) where the internal multiple algorithm (using the inverse scattering algorithm and normal-incidence, planewave data that do not require obliquity factor compensation) is analyzed in terms of its effectiveness to predict information corresponding to the amplitude of the true internal multiples in the data. The result of this analysis is an analytic expression of the amplitude prediction of first-order internal multiples that is generalized for a model with any number of interfaces or internal multiple generators. The analysis was done in the context of the inverse scattering series internal multiple prediction algorithm, but it is also useful to understand the amplitude prediction in IMP. Ramírez and Weglein's analysis is an extension of an earlier work by Weglein and Matson (1998) where a model with a single generating horizon was considered. Zhang and Shaw (2010) extended the analysis considering higher-order internal multiples in the single-generator model used by Weglein and Matson (1998).

Ramírez and Weglein's (2005) earlier analysis focused on internal multiples predicted using only primary subevents. The analysis by Zhang and Shaw (2010) includes internal multiples as subevents; as mentioned before, it only considers a model with a single generator that is always considered to be a primary event. Here, I extend both works by considering both primaries and internal multiples as subevents as well as a model with any given number of generators. The analysis is performed in 1D normal-incidence planewave data; and, hence, it is valid for both the inverse scattering series algorithm and the IMP algorithm. In practice and in multi-D the analysis provided here remains valid given the assumptions in each algorithm. In other words, the amplitude, phase, and traveltime prediction remains valid for the inverse scattering series algorithm and it is only constrained by the pseudo-depth or vertical time monotonicity condition and the effect of the wavelet (if not properly dealt with). The result is constrained by the total traveltime monotonicity condition for the IMP algorithm and the wavelet effect (if not compensated for); IMP will also suffer the effects of the missing obliquity factors affecting the amplitude and traveltime predictions analyzed here.

Consider planewave data, without the effect of a wavelet, containing primaries and internal multiples. Using Einstein's summation convention, these data can be expressed as

$$D(t) = R'_i\delta(t - t_i) - R'_{ijk}\delta[t - (t_i + t_k + - t_j)]$$

$$+ R'_{ijklm}\delta[t - (t_i + t_k + t_m - t_j - t_l)] + \dots \ , \qquad (3)$$

where R' represents the amplitude of recorded events (it includes the effect of reflection and transmission effects), the subscripts i, j, k, ... represent reflection points in the subsurface, and t_i is the two-way traveltime from source to interface i and back to the receiver position; sources and receivers are assumed to be located at zero depth. The first term on the right side represents primaries, the second term represents first-order internal multiples, the third term represents second-order internal multiples, and the higher-order terms represent higher-order internal multiples. The negative sign used to describe first-order internal multiples is there because each of these events contains a single downward reflection with a negative reflection coefficient (considering positive those reflection coefficients representing an upward reflection). For the same reason, the second-order internal multiples have a positive sign - the result of the multiplication of two downward (negative) reflection coefficients.

The notation in eq. (3) can be illustrated using the example raypaths in Fig. 1, where:

- The first term in the equation is illustrated in Figs. 1a and 1b. The primary in Fig. 1a can be mathematically described as

$$(\text{Primary})_1 \;=\; R_1\delta(t - t_1) \;, \tag{4}$$

thus, for this primary $R_1' = R_1$. The primary in Fig. 1b is mathematically described as

$$(\text{Primary})_2 \;=\; T_{01}R_2T_{10}\delta(t - t_2) \;, \tag{5}$$

thus, $R_2' = T_{01}T_{10}R_2$, where T_{ij} is the transmission coefficient for a wave traveling from medium i to j.

- The second term in eq. (3) corresponds to Fig. 1c. The first-order internal multiple in the figure can be mathematically described as

$$(\text{Internal multiple})_{212} = T_{01}R_2(-R_1)R_2T_{10}\delta[t - (t_2 - t_1 + t_2)]$$

$$= -T_{01}T_{10}R_2^2R_1\delta[t - (2t_2 - t_1)] \;, \tag{6}$$

thus, $-R_{212}' = -T_{01}T_{10}R_2^2R_1$.

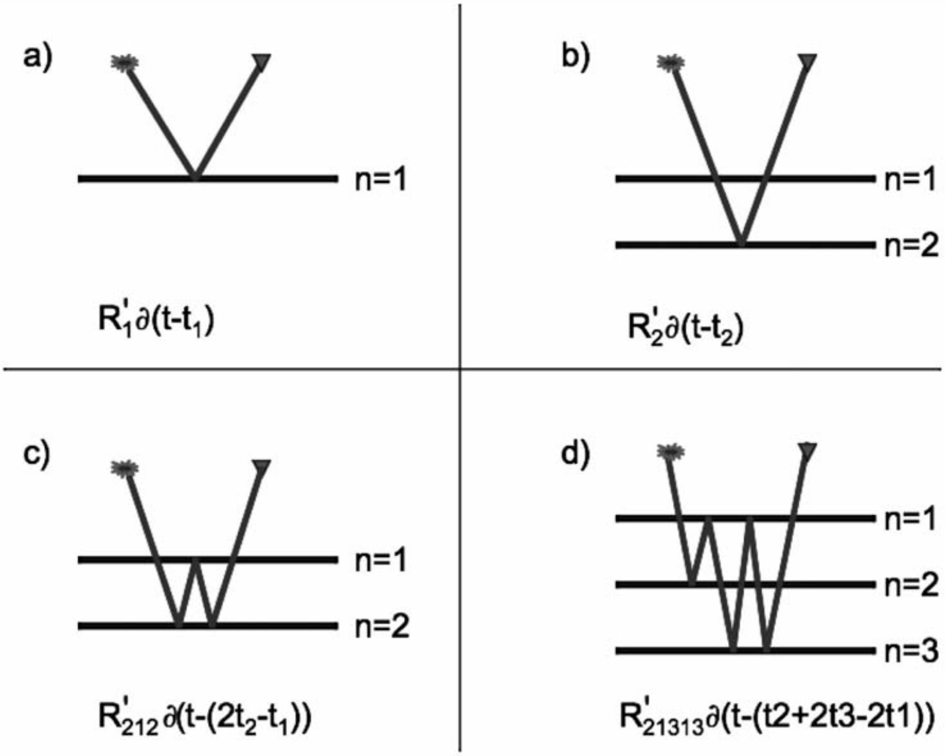

Fig. 1. Raypaths illustrating the different types of seismic events described in eq. (3): a) and b) represent primary events, c) and d) represent first-and second-order internal multiples, respectively.

- The third term is represented by Fig. 1d. The second-order internal multiple in the figure can be mathematically described as

$$(\text{Internal multiple})_{21313}$$

$$= T_{01}R_2(-R_1)T_{12}R_3T_{21}(-R_1)T_{12}R_3T_{21}T_{10}\delta[t - (t_2-t_1+t_3-t_1+t_3)]$$

$$= -T_{01}T_{10}(T_{12}T_{21})^2R_1^2R_2R_3^2\delta[t - (t_2+2t_3-2t_1)] \quad , \quad (7)$$

thus, $R'_{21313} = T_{01}T_{10}(T_{12}T_{21})^2R_1^2R_2R_3^2$.

Thus, data in eq. (3) can be separated into primaries, and different orders of internal multiples. The reflection points for first-order internal multiples satisfy the conditions $j < i$ & $j < k$. This condition leads to the well-known 'w' diagram representing internal multiples in data space. The 'w' diagram was used in the original derivation of the inverse scattering internal multiple prediction equation (Weglein et al., 1997). The reflection points for second-order internal multiples obey the relations $j < i$ & $j < k$ and $l < k$ & $l < m$. The primaries, in the context of this discussion, are defined as singly upward reflected events and can include headwaves. The headwaves as subevents of data-driven internal multiple prediction via ISS was first studied by Nita and Weglein (2004), however, only the case with two reflectors was part of this study - thus, no artifacts, nor higher order effects were analyzed. The internal multiples are defined as multiply reflected events with at least one downward reflection within the subsurface. Observe that this classification of events is only complete for data generated from a 1D layered medium without a free surface, where each layer is considered to be of constant velocity. When working with multidimensional data, more types of events would be generated (like prismatic waves or composite events, diving waves, and others). The interested reader is referred to Weglein and Dragoset (2005) for a thorough classification of events generated in multidimensional models or predictions.

A predicted first-order internal multiple using eq. (2) will have the general form

$$PIM_1 = R'_iR'_jR'_k\delta[t - (t_i + t_k - t_j)] \quad ; \quad (8)$$

analogously, the second-order predicted internal multiple is

$$PIM_2 = R'_iR'_jR'_kR'_lR'_m\delta[t - (t_i + t_k + t_m - t_j - t_l)] \quad . \quad (9)$$

Observe that the subscripts in amplitudes and traveltimes can either refer to primaries or internal multiples' traveltimes and amplitudes. In other words, not only primaries contribute to the prediction of internal multiples; internal multiples themselves can be combined to predict other internal multiples.

Internal multiple prediction for the first horizon

Following the general workflow of IMP, the first step is to select the generating horizons. The process starts with the shallowest of the generators by picking the corresponding primary and applying the proper mutes to the data to obtain two datasets: D_2 with the generating primary isolated and D_1 containing all the events with longer total traveltimes than the primary in D_2. The third data, D_3, in eq. (2) is the same as D_1. Once the data are prepared for the shallowest horizon, the related internal multiples are predicted.

Assuming a layered medium with N layers where the interfaces between layers are numbered with positive integers starting with the shallowest location, the shallowest generator is represented by the primary $R_1'\delta(t - t_1)$. Using eq. (2) with

$$\begin{aligned} D_1 &= R_i'\delta(t - t_i) - R_{ijk}'\delta[t - (t_i + t_k - t_j)] \\ &\quad + R_{ijklm}'\delta[t - (t_i + t_k + t_m - t_j - t_l)] + \dots , \quad \forall i \neq 1, \\ D_2 &= R_1'\delta(t - t_1) \ , \\ D_3 &= D_1 \ , \end{aligned} \tag{10}$$

the prediction gives multiples of the forms in equations 8 and 9 for first and second order, respectively,

$$PIM_1 = R_i'R_1'R_k'\delta[t - (t_i + t_k - t_j)] \ , \tag{11}$$

$$\begin{aligned} PIM_2 &= R_i'R_1'R_k'R_l'R_m'\delta[t - (t_i + t_k + t_m - t_1 - t_l)] \\ &\quad + R_i'R_j'R_k'R_1'R_m'\delta[t - (t_i + t_k + t_m - t_j - t_1)] \ . \end{aligned} \tag{12}$$

Theoretically, IMP must be applied one horizon at a time and the predicted internal multiples must be removed from the data before selecting the next horizon and predicting the next set of internal multiples.

Internal multiple prediction for the second horizon

In the following, I will only consider the traveltime of the events.

In general and under the assumption that the event selected and isolated as generator in D_2 is a primary, we can write:

$$t_{klm} = t_k - t_l + t_m \tag{13}$$

as the formula for the traveltime of the first-order internal multiples predicted using primaries only;

$$t_{ijklm} = t_{ijk} - t_l + t_m = (t_i - t_j + t_k) - t_l + t_m \quad (14)$$

would be the traveltime formula for an internal multiple predicted using two primaries and a first-order internal multiple as subevents;

$$t_{ijklmno} = t_{ijk} - t_l + t_{mno} = (t_i - t_j + t_k) - t_l + (t_m - t_n + t_o) \quad (15)$$

would be the traveltime formula for an internal multiple predicted combining one primary and two first-order internal multiples, and combinations using higher-order internal multiples. The assumption, of course, is that the algorithm is selecting subevents that meet the condition longer-shorter-longer in traveltime, and that this condition transfers to a lower-higher-lower condition in actual depth (ten Kroode, 2002; Ramírez and Weglein, 2005a; Malcolm and de Hoop, 2005; Nita and Weglein, 2007). This assumption, or total traveltime monotonicity condition, is equivalent (only for 1D layered media) to the pseudodepth monotonicity condition, governing the inverse scattering algorithm. If this condition is not met, the result is the prediction of artifacts. In the problem considered here, the monotonicity condition is always satisfied for internal multiples predicted with primary subevents, eq. (13), as demonstrated by Ramírez and Weglein (2005a), but it is not always satisfied when internal multiples are used as subevents (Ramírez et al., 2011). This is further explained at the end of this section.

For the second horizon, one sets time $t_l = t_2$. Eq. (13) is the prediction of first-order internal multiples with downward reflection at horizon 2, H2. Eq. (14) will predict first- and second-order internal multiples. Eq. (15) will predict different orders of internal multiples. The first, in traveltime, internal multiple event that can be selected as a subevent in the algorithm's prediction for H2 is the internal multiple t_{212}. A simple confirmation of this statement is the fact that the subevent IM^1_{212} has a total traveltime longer than the primary of the second reflector, P_2 , which satisfies the monotonicity condition as long as the traveltime of the third subevent, t_m, is longer than that of P_2. This particular combination results in

$$t_{2122m} = (t_2 - t_1 + t_2) - t_2 + t_m \quad , \quad \forall m > 2. \quad (16)$$

With the help of simple algebra, one obtains

$$t_{2122m} = t_{21m} = t_2 - t_1 + t_m \quad , \quad \forall m > 2, \quad (17)$$

the traveltime of internal multiples with downward reflection at the first interface and upward reflection at interfaces 2 and m. Observe that operating with the primary for the second horizon as generator predicts multiples generated at the first (shallowest) horizon because the internal multiple subevent, IM^1_{212}, contains an instance of $+t_2$.

In general, if there is an extra instance of $+t_2$ in eqs. (14) or (15), the prediction is formed of internal multiples generated at the first horizon. Selecting $t_k = t_2$ forces the selection of $t_j = t_1$, and allows us to generalize the result just obtained to

$$t_{i122m} = (t_i - t_1 + t_2) - t_2 + t_m = t_i - t_1 + t_m = t_{i1m} \quad , \tag{18}$$

$$t_{i122mno} = (t_i - t_1 + t_2) - t_2 + (t_m - t_n + t_o)$$

$$= t_i - t_1 + t_{mno} = t_{i1mno} \quad , \tag{19}$$

where i > 1 for both equations, m > 2 for eq. (18), and $t_{mno} > t_2$ in the last equation. In fact, the latter condition is always true in a 1D earth, because any internal multiple traveltime t_{mno} is longer than the traveltime of the second shallowest horizon. Thus, eqs. (18) and (19) demonstrate that these combinations always result in the traveltime prediction of internal multiples associated with the first horizon.

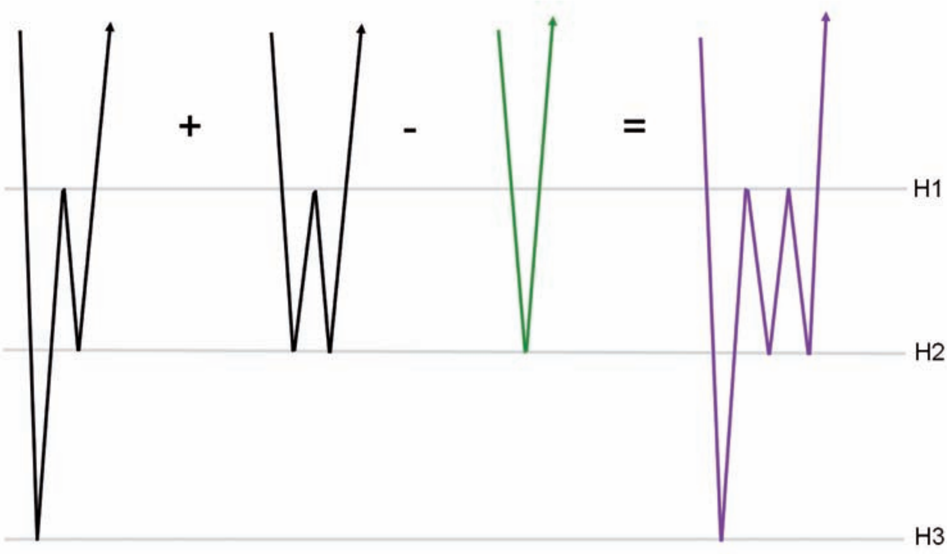

Fig. 2. The three diagrams on the left represent the subevents combined to predict the internal multiple on the right. In this example, the two internal multiples (black) with traveltimes t_{312} and t_{212} are convolved together and crosscorrelated with the generating primary (green) with traveltime t_2. The combination results in the second-order internal multiple (purple) with traveltime t_{31212}.

One can conclude that operating with the primary P_2 as generator in IMP, results in the prediction of all internal multiples generated at horizon H2, all first-order internal multiples generated at H1 except for IM^1_{212}, and all higher-order internal multiples generated with at least one downward reflection at H1. Fig. 2 illustrates the prediction of an internal multiple using the result in eq. (19).

The condition of subevents having an instance of $+t_2$ in their traveltime is a sufficient but not necessary condition to predict the traveltime of a true internal multiple with the primary generator P_2 using IMP (Fig. 3). In fact, specifically for P_2 acting as generator, all events will have the opportunity to act as subevents in the prediction except for the primary with time t_1 and the predicted events will have the true internal multiple traveltimes (no artifacts in 1D or multi-D as long as the monotonicity condition is satisfied). Hence,

$$t_{ijk2m} = (t_i - t_j + t_k) - t_2 + t_m \quad , \tag{20}$$

where i can have any value, $j > i$, $k > i$, and $m > 2$, and

$$t_{ijk2mno} = (t_i - t_j + t_k) - t_2 + (t_m - t_n + t_o) \quad , \tag{21}$$

where i, j, and k satisfy the same condition as in eq. (20) and $t_{mno} > t_2$ corresponds to traveltimes of true internal multiples.

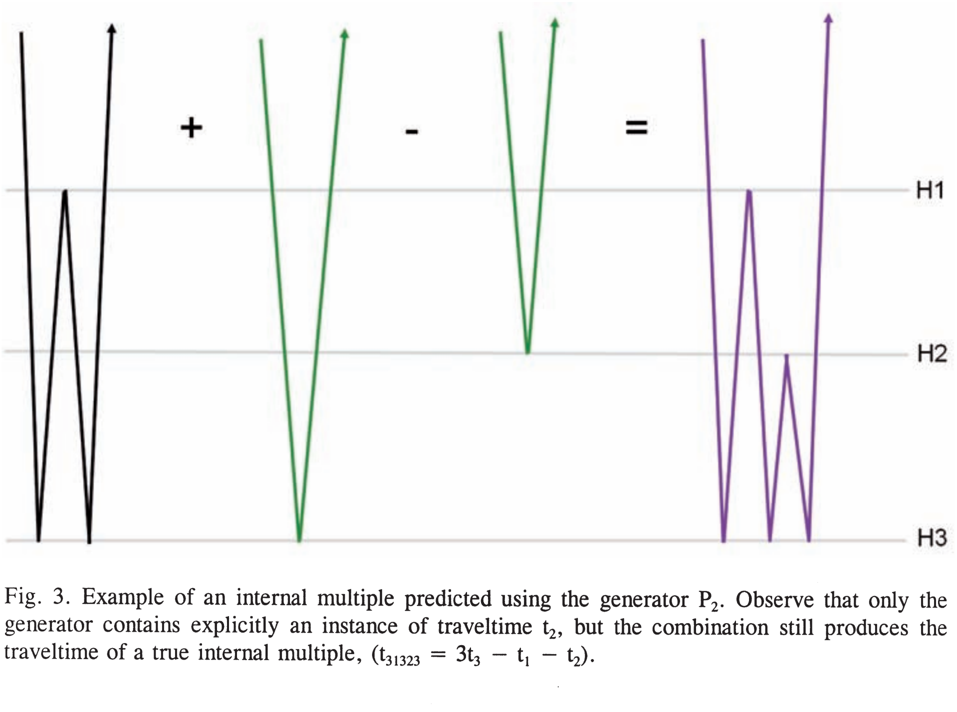

Fig. 3. Example of an internal multiple predicted using the generator P_2. Observe that only the generator contains explicitly an instance of traveltime t_2, but the combination still produces the traveltime of a true internal multiple, $(t_{31323} = 3t_3 - t_1 - t_2)$.

For the next horizons t_l (l = 3,4,...), eqs. (14) and (15) will predict internal multiples for the selected generating horizon, Hl, a set of internal multiples associated with any horizon in the overburden of Hl, and artifacts. Fig. 4 illustrates the prediction of an artifact.

If one considers eq. (14), and selects an internal multiple subevent satisfying:

(i) $t_{ijk} > t_l$, where i, j,and k are integers smaller or equal to $l - 1$, then we predict an artifact. One example would be the combination of the internal multiple t_{212} with the primaries t_3 and t_m (m > 3),

$$t_{2123m} = (t_2 - t_1 + t_2) - t_3 + t_m = t_{212} + (t_m - t_3) \quad , \tag{22}$$

which creates an artifact. Note that the condition to predict this artifact is $t_{212} > t_3$. This condition is not always satisfied. For example, if $t_{212} < t_3$, the top mute in D_1 and D_3 would mute the event t_{212} and the artifact in eq. (22) would not be predicted. However, one can always find a subevent combination that creates an artifact like this; for example, t_{21212} could have a longer traveltime than t_3 (the artifact will be very weak in amplitude, though). If one considers eq. (15) with $l > 2$, then having an internal multiple subevent satisfying relationship (i) is not a sufficient condition to predict an artifact. The artifact will be predicted when

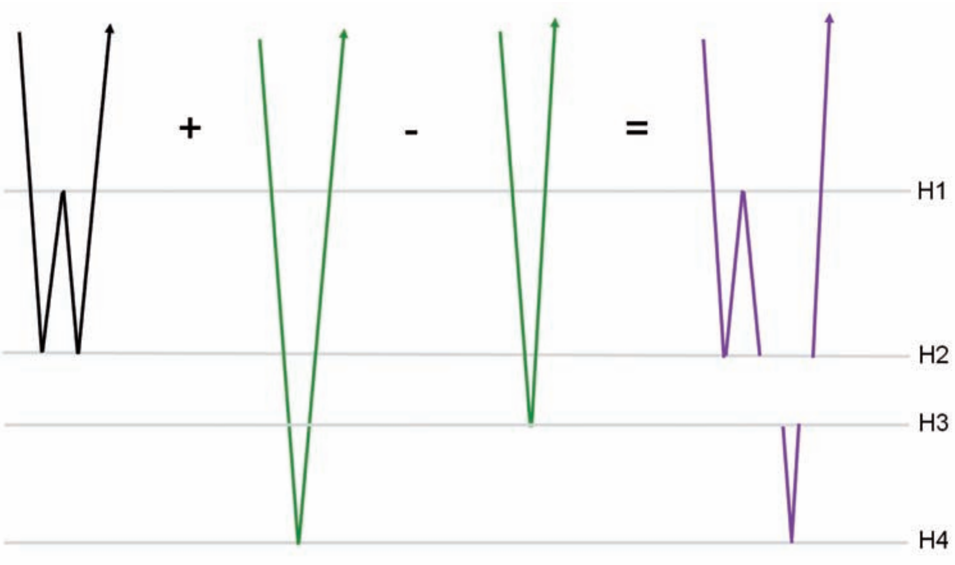

Fig. 4. The artifact on the right is predicted with the subevent combination described in eq. (22) with m = 4, only if the condition $t_{212} > t_3$ is satisfied.

both internal multiple subevents satisfy relation (i), or when one internal multiple satisfies relationship (i) and the second one satisfies:

(ii) $t_{mno} > t_l$, where n and o are integers smaller than or equal to $l - 1$ and m is an integer larger than l.

Observe that true internal multiples will be predicted if one internal multiple satisfies relationship (i) and the second one satisfies:

(iii) $t_{mno} > t_l$, where n is an integer smaller than or equal to $l - 1$, and m and o are integers bigger than or equal to l. An example of such an internal multiple is:

$$t_{2123424} = (t_2 - t_1 + t_2) - t_3 + (t_4 - t_2 + t_4)$$
$$= t_2 - t_1 + t_4 - t_3 + t_4 = t_{21434} \quad . \tag{23}$$

Polarity

If we assume that the data contain primaries (let them be positive) and internal multiples (negative for first-order, positive for second-order, and so forth) then, for horizon 1:

(j) Eq. (13) will predict first-order internal multiples with positive polarity. A diagrammatic example is displayed in Fig. 5.

(jj) Eq. (14) will predict second-order internal multiples with negative polarity. A diagrammatic example is displayed in Fig. 6.

(jjj) Eq. (15) will predict third-order internal multiples with positive polarity.

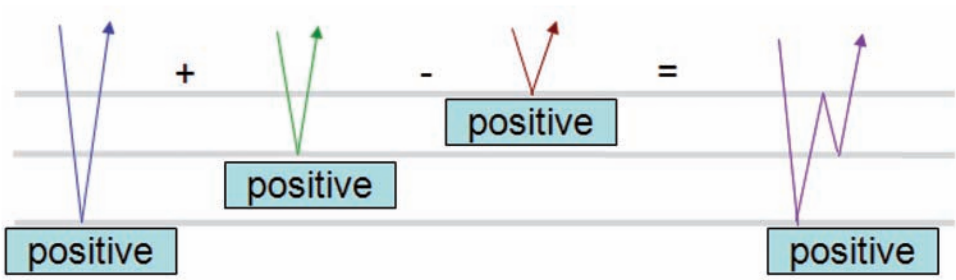

Fig. 5. The three diagrams on the left represent the primaries combined to predict the internal multiple on the right. In this example, the red primary (first horizon) is used as the predicting generator for the internal multiple on the right.

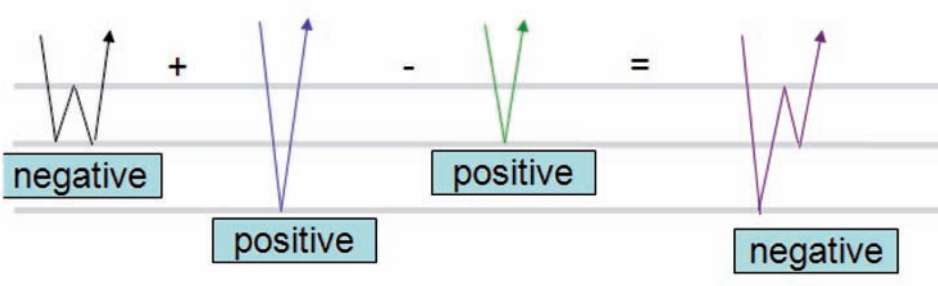

Fig. 6. The three diagrams on the left represent the subevents (two primaries and one internal multiple) combined to predict the internal multiple on the right. In this example, the green primary (second horizon) is used as the predicting generator for the internal multiple on the right. Observe that the predicted internal multiple has the same traveltime as the internal multiple predicted with horizon 1 in Fig. 5. The polarity of the internal multiple predicted with horizon 2 is the opposite of the polarity of the prediction using horizon 1.

When we run the algorithm for horizon 2, items (j)-(jjj) remain valid for internal multiples generated at the second horizon; however, for internal multiples related to the first horizon:

(jjb) Eq. (14) will predict first-order internal multiples with negative polarity because instead of three primary subevents, two primaries and one internal multiple from the data are combined.

(jjjb) Eq. (15) will predict second-order internal multiples with positive polarity because the prediction uses two internal multiples and one primary from the data.

Remember that IMP multiplies the prediction by a factor of -1 [see eqs. (1) and (2)]. Hence, the internal multiples corresponding to a given horizon, H_n, predicted using the primary representing H_n will have the correct polarity.

DISCUSSION

Let's examine the non-top-down approach a bit further. Once we generate the internal multiple predictions (or models) for a set of n horizons using IMP with an unchanged input dataset, we have to deal with n models. Assuming that all generating horizons (up to a level of interest) were identified and used, then

1. The first model, H1, contains events corresponding to all orders of internal multiples associated with the first horizon. The predictions have the correct

polarity (correct means that it corresponds to the polarity of the true internal multiples in the data). No artifacts and no internal multiples from other horizons are present.

2. The second model, H2, contains events corresponding to all orders of internal multiples generated by the second horizon with the correct polarity, and a set of internal multiples with incorrect polarity (opposite to the true polarity) corresponding to traveltimes of most of the multiples predicted in (1). No artifacts are present.

3. The third horizon, H3, will contain the internal multiples related to H3, plus a set of events corresponding to internal multiples from H2 and H1 with opposite polarity. It will also have the possibility of containing artifacts (although they are less common than the repeated predictions of internal multiples).

4. This can be generalized to an arbitrary number of horizons.

If one decides to ignore the fact that each prediction (except for H1) contains events with traveltimes corresponding to internal multiples generated at earlier horizons, then we can imagine running a set of sequential adaptive subtractions. The multiples from H1 will be attenuated by adaptive subtracting the H1 model from the input data, but these same multiples (the residual left after the first subtraction) will be enhanced when we attempt to subtract the multiples from the second horizon. The problem with this path is that in each model (except for the one corresponding to H1), there is a combination of events with different polarities, some correct and some incorrect. This must be dealt with during the adaptive subtraction process. For example, one can imagine that the traveltime, amplitude, and polarity analysis provided here (complemented by the traveltime and amplitude analyses done by Weglein and Matson (1998), Ramírez and Weglein (2005a), Ramírez Pérez (2007), and Zhang and Shaw (2010)) would be helpful in improving the removal of internal multiples from the data. For example, when Baumstein (2008) analyzed the prediction for the second horizon (as generator) in a non-top-down approach, he proposed to use a curvelet-based adaptive subtraction method that is able to adjust the polarity of events individually. An example of this adaptive subtraction method is given by Neelamani et al. (2008).

To illustrate the theory and analysis in this manuscript, I used 3D finite difference, 2-parameter acoustic synthetic data generated with the 2.5D velocity model in Fig. 7 and the density model in Fig. 8. The internal multiples were predicted using a 3D implementation of IMP, equation 2 in a non-top-down approach. The zero-offset gather of the input data is shown in Fig. 9, where the green arrow point at the direct wave (not used in this algorithm) and the red arrows point at the three primaries; the rest of the events are internal multiples

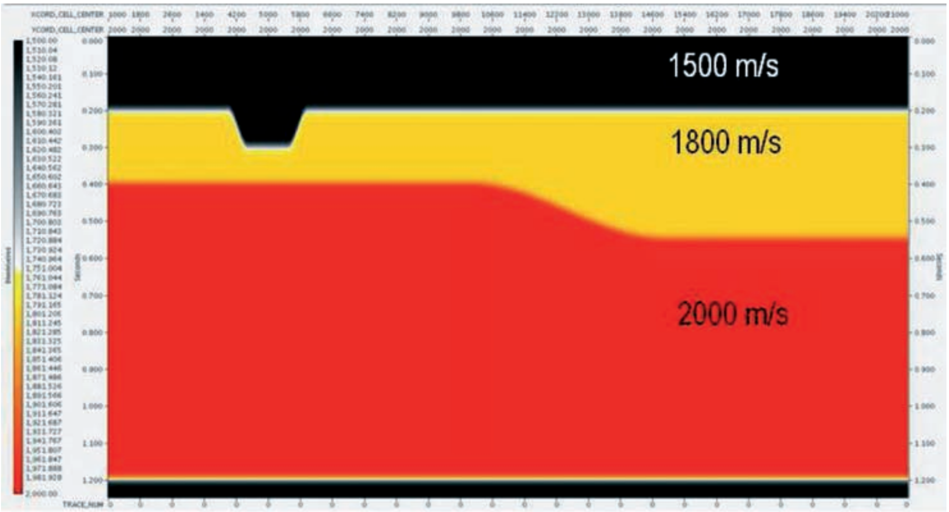

Fig. 7. Velocity model: 2D slice of a 2.5D model.

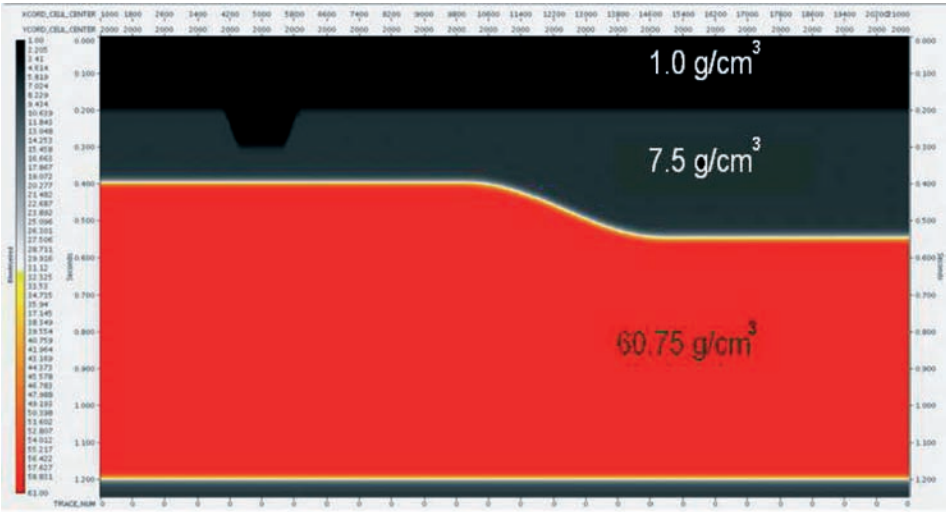

Fig. 8. Density model: 2D slice of a 2.5D model.

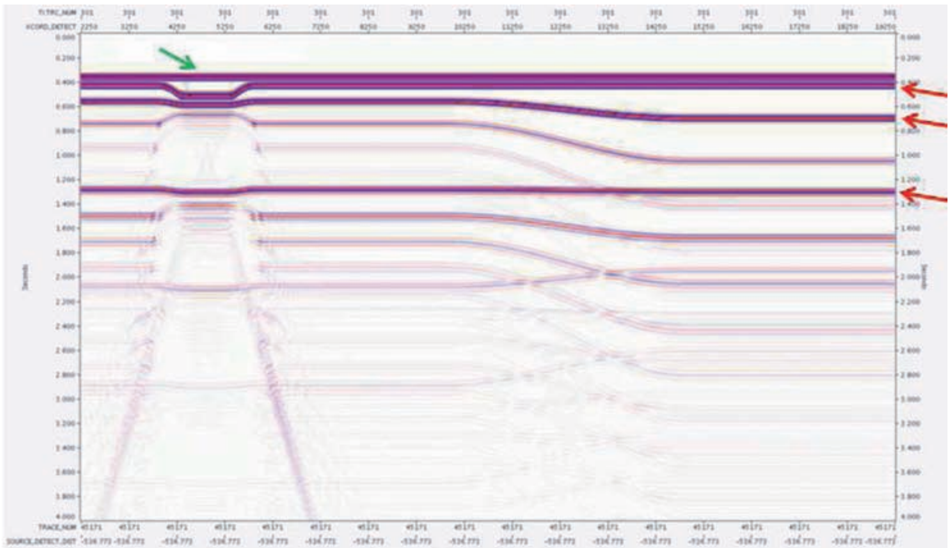

Fig. 9. Zero-offset gather of the input data: The green arrow points at the direct wave and the red arrows point at the primaries.

as the data were generated without the effect of the free surface. Because there are only three horizons in the model, there are two generating horizons corresponding to the two shallowest horizons. Thus, the first two primaries in Fig. 9 are the ones that will be used as generating primaries in the algorithm.

Fig. 10 shows the zero-offset gather of the predicted internal multiples using the first primary as generator. The corresponding location of the first primary in the zero-offset gather is displayed with the dark green line, the light green line illustrates the location of the second primary. All internal multiples generated with the shallowest primary as generator, have the correct kinematics and the correct polarity. The polarity of the true internal multiples in the data are represented by the overlayed dotted and dashed lines. The dotted lines are polarities for internal multiples of different orders generated at the shallowest horizon, while the dashed lines correspond to the polarities of the two most prominent internal multiples generated at the second horizon. The lines are overlayed with the predicted events to demonstrate that the polarity in the predicted events is correct. The polarities were picked at the peak amplitude of each event, the negative sign of the polarity are represented with the red lines, and the positive polarity with blue ones. As expected from the theory and analysis, all predicted internal multiples correspond to different orders of internal multiples generated at the shallowest horizon (there are no events under the dashed lines). No artifacts are present, the kinematics and polarity are correct. Only the amplitudes and the wavelet have to be corrected in the adaptive subtraction step (not shown here).

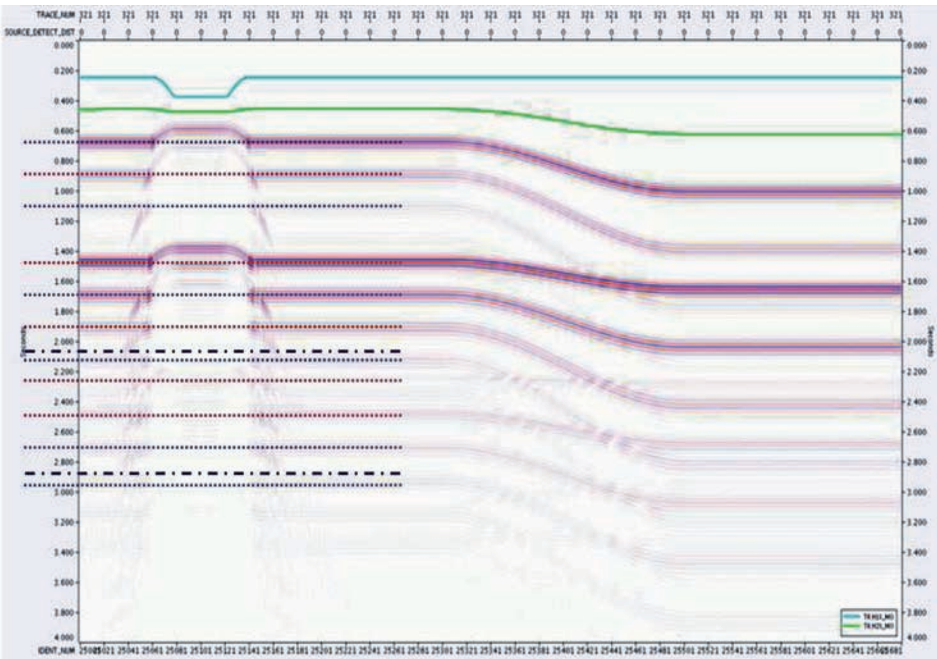

Fig. 10. Zero-offset gather of the predicted internal multiples for the first horizon: The dark green and the light green lines illustrate the first and second primary location in the zero-offset gather. The dotted and dashed lines are polarity indicators. They represent the polarity at the peak amplitudes of the true internal multiples. A negative polarity is displayed in red and a positive polarity in blue.

Fig. 11 show the zero-offset gather of the predicted internal multiples using the second primary as generator. The dotted/dashed lines in blue and red are the same as the ones in Fig. 10, indicating polarity and generating horizon of the true internal multiples in the data. As expected from the analysis of the non-top-down approach, when this approach is followed, the prediction for the second generator generates all internal multiples for that horizon with correct kinematics and polarity, and also generates all internal multiples for the shallowest horizon (except for the internal multiple with traveltime t_{212} illustrated by the shallowest blue dotted line) with correct kinematics and incorrect polarity. Fig. 12 shows a zoom of the zero-offset gathers for the generated internal multiples. The leftmost side of the figure shows the predicted internal multiples using the shallowest horizon, and the rightmost side shows the multiples predicted with the second generator. It is easy to see that the repeated multiples have opposite polarity; those correspond to true multiples generated at the shallowest horizon. In this figure, only one multiple generated by the second horizon is visible; it starts on the right at 1.9 seconds. This multiple has the correct polarity, as it corresponds to a true multiple generated at the second horizon and it was predicted using the second primary as a generator.

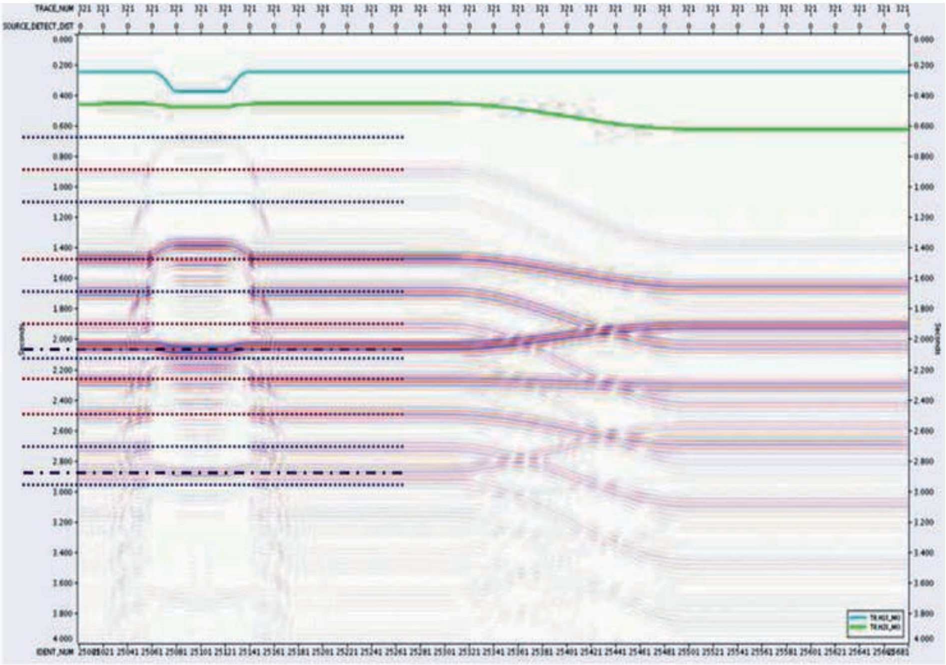

Fig. 11. Zero-offset gather of the predicted internal multiples for the second horizon: The dark green and the light green lines illustrate the first and second primaries. The dotted and dashed lines in red and blue are negative and positive polarity indicators. They represent the polarity at the peak amplitudes of the true internal multiples. This prediction is a non-top down approach, and hence, the internal multiples generated at the second horizon are predicted with correct polarity, while the predicted internal multiples corresponding to the first generator have opposite polarity.

CONCLUSIONS

The practical application of the almost entirely data-driven internal multiple prediction algorithm, IMP, can be expected to happen in a non-top-down approach or in a combination of a partial top-down-approach with a non-top-down. One reason for this is related to the sparsity of real data. Most of the computation time in the prediction of internal multiples is related to input data selection and interpolation done at each run of the IMP algorithm.

There are several reasons not to follow the layer-stripping method; among others, the high cost of the algorithm is a factor as well as the fact that not all generating horizons can always be identified when working with complex real data. Also, residual multiples in a layer-stripping approach would have a non-top-down-type of approach contribution. Hence, it is important to understand the output of IMP when it is applied in non-ideal conditions.

The analysis in this paper provides general formulas that can explain the output of IMP. Stated succinctly, *if the layer-stripping approach is followed, and if for each prediction, there is a perfect subtraction, then IMP will always predict internal multiples with correct polarities*. Some artifacts will still be predicted. For the non-top-down approach, the story is different: IMP will predict true internal multiples associated with the horizon used as a generator with correct polarity, internal multiples associated with all horizons in the overburden of the generator, with opposite polarity, and some artifacts. Due to

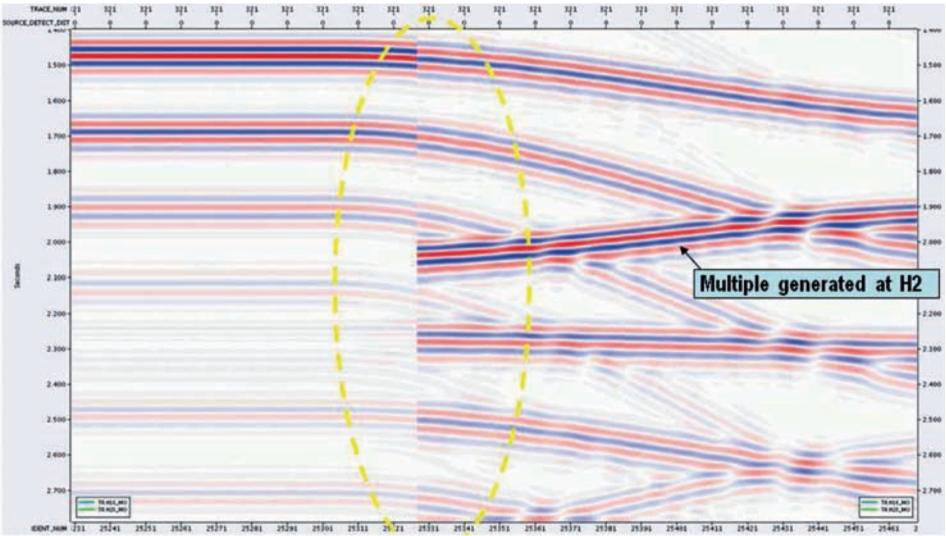

Fig. 12. Zoom of the zero offset gathers in Fig. 10 (leftmost) and in Fig. 11. The predicted multiples associated with the shallowest generator have opposite polarities. The polarities on the left are correct. The event that appears on the right with no counterpart on the left is associated with the second reflector and has the correct polarity.

the fact that the prediction of artifacts requires internal multiple subevents, these artifacts are weak in amplitude. Furthermore, there are specific conditions for these artifacts to be predicted, which makes them a lot less common than the prediction of events with the traveltime of a true internal multiple.

ACKNOWLEDGEMENTS

I thank WesternGeco for permission to publish this work. Ian Moore is thanked for useful technical discussions. I am also grateful for the support and manuscript revision of Konstantin Osypov and Einar Otnes. The suggestions by Art Weglein and by Kris Innanen really helped to improve the quality of the manuscript and the clarity of its intended message. Bogdan Nita is thanked for useful discussions in 2004 that triggered the analysis presented in this paper.

REFERENCES

Araújo, F.V., Weglein, A.B. Carvalho, P.M. and Stolt, R.H., 1994. Inverse scattering series for multiple attenuation: An example with surface and internal multiples. Expanded Abstr., 64th Ann. Internat. SEG Mtg., Los Angeles, 13: 1039-1041.

Baumstein, A., 2008. An upside-down approach to efficient surface-related and interbed multiple prediction. Expanded Abstr., 78th Ann. Internat. SEG Mtg., Las Vegas, 27: 2466-2470.

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

El-Emam, A., Moore, I. and Shabrawi, A., 2007. Interbed multiple prediction and attenuation: Case history from Kuwait. The Leading Edge, 26: 41-45.

Hembd, J., Griffiths, M., Ting, C.-O. and Chazalnoel, N., 2010. Application of interbed multiple attenuation in the Santos Basin, Brazil. Expanded Abstr., 80th Ann. Internat. SEG Mtg., Denver, 29, 3451-3455.

Jakubowicz, H., 1998. Wave-equation prediction and removal of interbed multiples. Expanded Abstr., 68th Ann. Internat. SEG Mtg., New Orleans: 1527-1530.

Jin, D., Yang, F., Zeng, J., Liu, Y. and Chang, X., 2009. A simplified method for 1.5D interbed multiples prediction based on inverse scattering series. Expanded Abstr., 79th Ann. Internat. SEG Mtg., Houston, 28: 3064-3067.

Lira, J.E.M., Innanen, K.A., Weglein, A.B. and Ramírez, A.C., 2010. Correction of primary amplitudes for plane-wave transmission loss through an acoustic or absorptive overburden with the inverse scattering series internal multiple attenuation algorithm: an initial study and 1D numerical examples. J. Seismic Explor., 19: 103-120.

Malcolm, A.E. and de Hoop, M.V., 2005. A method for inverse scattering based on the generalized Bremmer coupling series. Inverse Probl., 21: 1137-1167.

Moore, I., 2001. Practical implementation of interbed multiple attenuation: Explor. Geophys., 32: 80-88.

Neelamani, R., Baumstein, A. and Ross, W.S., 2008. Adaptive subtraction using complex curvelet transform. Extended Abstr., 70th EAGE Conf., Rome.

Nita, B.G. and Weglein, A.B., 2004. Imaging with tau = 0 versus t = 0: implications for the inverse scattering internal multiple attenuation algorithm. Expanded Abstr., 74th Ann. Internat. SEG Mtg., Denver, 23: 1289-1292.

Nita, B.G. and Weglein, A.B., 2007. Inverse-scattering internal multiple-attenuation algorithm: An analysis of the pseudodepth and time-monotonicity requirements. Expanded Abstr., 77th Ann. Internat. SEG Mtg., San Antonio, 26: 2461-2465.

Nita, B.G. and Weglein, A.B., 2009, Pseudo-depth/intercept-time monotonicity requirements in the inverse scattering algorithm for predicting internal multiple reflections. Commun. Computat. Phys., 5: 163-182.

Ramírez, A.C. and Weglein, A.B., 2005a. An inverse scattering internal multiple elimination method: Beyond attenuation, a new algorithm and initial tests. Expanded Abstr, 75th Ann. Internat. SEG Mtg., Houston, 24: 2115-2118.

Ramírez, A.C. and Weglein, A.B., 2005b. Progressing the analysis of the phase and amplitude prediction properties of the inverse scattering internal multiple attenuation algorithm. J. Seismic Explor., 13: 283-301.

Ramírez Pérez, A.C., 2007. I: Inverse Scattering Subseries for Removal of Internal Multiples and Depth Imaging Primaries; II: Green's Theorem as the Foundation of Interferometry and Guiding New Practical Methods and Applications. Ph.D. thesis, University of Houston, Houston, TX.

Ramírez-Pérez, A.C., Teague, A.G, Walz, M.A. and Wu, Z., 2011. Attenuating internal multiples from seismic data. U.S. Patent Application 592-25695-US, filed 2011.

ten Kroode, F., 2002. Prediction of internal multiples. Wave Motion, 35: 315-338.

Terenghi, P., Dragoset, W.H. and Moore, I., 2010. Interbed multiple prediction: US Patent Application Publication, N 20100074052 A1.

Verschuur, D.J. and Berkhout, A.J., 1996. Removal of interbed multiples. Extended Abstr., 58th EAGE Conf., Amsterdam.

Verschuur, D.J., Berkhout, A.J., Matson, H., Weglein, A.B. and Young, C.Y., 1998. Comparing the interface and point scatterer methods for attenuating internal multiples: A study with synthetic data, Part I. Expanded Abstr., 68th Ann. Internat. SEG Mtg., New Orleans, 17: 1519-1522.

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

Weglein, A.B. and Matson, K., 1998. Inverse scattering internal multiple attenuation: an analytic example and subevent interpretation. Proc. SPIE, Mathemat. Meth. Geophys. Imaging V. Siamak Hassanzadeh (Ed.), SPIE Homepage, 3453: 108-117.

Weglein, A.B. and Dragoset, W.H., 2005. Multiple Attenuation. SEG, Tulsa, OK.

Zhang, H. and Shaw, S., 2010. 1D analytical analysis of higher order internal multiples predicted via the inverse scattering series based algorithm. Expanded Abstr., 80th Ann. Internat. SEG Mtg., Denver, 29: 3493-3498.