SEISMIC WAVELET ESTIMATION WITHOUT THE INVALID WHITENESS ASSUMPTION
Anton Ziolkowski
Department of Geology and Geophysics, University of Edinburgh, U.K.

ABSTRACT
Current seismic data processing practice processes the data to zero phase and requires an estimate of the causal wavelet in order to find the zero-phasing filter. Most wavelet estimation processes rely on some component of the earth response to be white. There is no evidence to support this assumption and it should be jettisoned. It follows that the value of predictive deconvolution in the processing sequence should be seriously questioned. Without the whiteness assumption, the wavelet estimation problem becomes hopeless. Source signature measurements provide a solution to this problem and, via careful signature deconvolution as a first step in processing, allow the wavelet to be made short and constant throughout the data. In tying the seismic data to well logs, it may be necessary to include the effects of attenuation. A comparison of the statistical and deterministic processes at a well favours the deterministic approach using source signature measurements, which needs only an estimate of Q, to balance the spectrum of the synthetic seismogram and make a good tie.

INTRODUCTION
In the last twenty years the seismic exploration industry has become very comfortable with the idea of estimating wavelets from seismic data by making assumptions. Very little seismic data is now acquired with the source signature measured specifically to help in the processing of the data. In the past it was recognised that the source signature can be quite complicated and that it needed to be measured and then deconvolved from the data (e.g. Mateker, 1971; Wood et al., 1978). The aim of seismic source development has been to produce sources whose signatures do not need to be measured. In the processing of much marine seismic data, for example, the source signature is assumed to be equal to a library signature, for which a signature deconvolution filter has already been computed. This filter is applied to all the data to make the signature into a minimum-phase wavelet. It is then common practice to assume that the earth filtering effect on the propagating wavelet is minimum-phase such that the wavelet in the data is still minimum-phase.

In practice, there are many factors that cause the source signature to vary from shot to shot and from line to line. In the case of air gun arrays, for example, the interaction between air-gun sub-arrays varies with the distance between them, and this can vary substantially due to wave action. The period of oscillation of an air gun bubble is very sensitive to the gun depth, and the depth at which the guns tow varies with currents and wave height. These variations occur, they are significant, and the signature cannot be identical with the library signature. How much it deviates from the library signature cannot be determined unless the signature is measured. In general these measurements are not normally made.

In the acquisition of Vibroseis data, there are several issues related to the mechanics of the vibrator, such as plate bending and harmonic distortion, that are not addressed by the vibrator control system. Therefore the ground force is not equal to the sweep used to correlate the data, and the force varies from vibrator to vibrator within a source array, and from place to place on the earth’s surface. It is possible to measure the true ground force, but this is not normally done.

Seismic wavelet estimation is always important at the point where we want to relate the seismic data to known geology. The interpretation of seismic data is most straightforward when the wavelet is zero phase; that is, when the wavelet is symmetrical with its central peak at the arrival time of the event. Since such non-causal wavelets cannot exist in practice, it is necessary to know what the true seismic wavelet is before a filter can be found to make the wavelet zero-phase. Normally a wavelet is estimated from the data and the zero-phasing filter removes the phase of this wavelet. This filter is then applied to the data to make the data ‘zero-phase’. In marine seismic data ‘processed to zero phase’, an indication of poor wavelet estimation is the pre-cursor on the sea-bottom reflection. If the wavelet had been estimated properly this pre-cursor would be very small. Sometimes it is enormous and this should worry anyone who wishes to see small reflectors just above or just below large ones. One can often see examples of such precursors at geophysical exhibitions in the booths of companies who sell marine seismic data from prospective areas around the world.

The seismic wavelet is required for inversion - to estimate acoustic impedances and reflection coefficients from the seismic data. The wavelet is also needed whenever the seismic data are to be tied to well logs. The zero-phase wavelet that is estimated to be in the data after zero-phasing can be used to compute a synthetic seismogram, by convolving it with the reflection coefficient sequence derived from the well logs. Often this procedure does not produce a very good match with the seismic data at the well. An example is provided towards the end of this paper. Adjustments can be made to the wavelet to improve the match. It is normally found that the adjustment required is different for every well. This raises the question: What is happening to the wavelet between the wells? How much of the lateral variation we see in the seismic data is caused by lateral variations in the wavelet? And how much is caused by lateral variations in the geology? If we cannot eliminate lateral variations in the wavelet, how can we attribute all lateral variations in the data to lateral variations in the geology?
SEISMIC WAVELET ESTIMATION WITHOUT THE INVALID WHITENESS ASSUMPTION

Continued from Page 18

How much can we trust the assumptions that underlie conventional seismic wavelet estimation? In this paper I examine the whiteness assumption that is at the heart of many wavelet estimation techniques, and argue that there is no evidence to support it. Once this assumption is jettisoned, the wavelet estimation problem becomes hopeless unless we measure the source signature. It also follows that we should have serious misgivings about the effects of predictive deconvolution on seismic data. I briefly describe existing methods for measuring source signatures and discuss signature deconvolution and how to maintain a short known wavelet through the processing sequence and obtain a good tie to well data. Finally I discuss the effect of the earth’s attenuation on the propagating wavelet and the adjustment that must be made to tie the well synthetic seismogram with the seismic data.

HOW MANY ELIGIBLE WAVELETS ARE THERE?

The purpose of this section is to estimate the number of different eligible seismic wavelets we can find in a given seismogram. A seismogram is the convolution of the seismic source time function with the impulse response of the earth and with the impulse response of the receiver and recording system, plus noise. For the remainder of this paper we assume the impulse response of the receiver and recording system is known, and we ignore it in the rest of the analysis. Because of the presence of multiples, the impulse response of the earth is infinitely long. However, after some time, the signal amplitude falls to a level where it can no longer be detected, and recording is discontinued. Let the recording be digitised and the resulting seismogram be \( x = x_0, x_1, \ldots, x_m \), which we know.

Ignoring the truncation of the complete response (which is an issue, as shown in the next section), we may define as

\[ x = s \ast g + n, \tag{1} \]

in which is the source time function, is the impulse response of the earth, is uncorrelated noise, and the asterisk * denotes convolution. Equation (1) is one equation with three unknown functions. We would like to find the unknown wavelet. For the time being we neglect the noise, reducing equation (1) to

\[ x = s \ast g. \tag{2} \]

We also assume we know the number of samples \( n+1 \) of \( s \); that is, \( s = s_1, s_2, \ldots, s_m \). Since we have a perfect convolution, it follows that . Using the well-known Z-transform on equation (2), we have

\[ X(Z) = S(Z)G(Z), \tag{3} \]

in which

\[ X(Z) = x_0 + x_1Z + \cdots + x_mZ^m = \sum_{k=0}^{m} x_kZ^k, \tag{4} \]

each of these Z-transforms is a polynomial in \( Z \), which may be factorised into its roots:

\[ X(Z) = x_0 \prod_{k=1}^{m} (Z - Z_k) = s_0 \prod_{k=n+1}^{m} (Z - Z_k), \tag{5} \]

in which \( \prod \) denotes the product.

The Z-transform of the wavelet \( s \) has \( n \) roots, and the Z-transform of the seismogram \( x \) has \( m \) roots, where \( m > n \). The problem has now been reduced to finding the \( n \) roots of the wavelet, from the \( m \) roots of the seismogram. Without any more information, we have no means of knowing which combination of \( n \) roots forms the wavelet. The number of different ways we can select \( n \) roots from \( m \) roots is

\[ \frac{m!}{n!(m-n)!}. \]

In a typical seismic trace the number of samples \( m \) is about 3,000 and the number of samples \( n \) of the wavelet is, say, 100. The largest factorial I can compute on my pocket calculator is only \( 69! = 1.71 \times 10^{98} \). The number of possible wavelets we are trying to calculate here is about \( 10^{100} \), within a few orders of magnitude. That is, there are about \( 10^{100} \) different eligible wavelets, each with 100 samples, that we can, in principle, estimate from a seismogram of 3,000 samples. Note that we have got this far only after neglecting the noise, assuming a perfect convolution, and assuming the length of the wavelet.

BLIND PRE-STACK WAVELET ESTIMATION FROM COMMON SHOT GATHERS

Now consider a shot gather. Every seismogram in the shot gather is generated by the same source in the same place. The source time function is the same for all traces in the shot gather, whereas the impulse response of the earth is different for every receiver position. Can we use this information to extract the wavelet from the data? This problem has been discussed by Ziolkowski and Slob (1991) using the Z-transform approach.

The roots of the Z-transform of the source time function are the same from trace to trace, while, with luck, the roots of the earth impulse response are different from trace to trace. If the roots of the Z-transform of each trace are plotted in the complex plane and summed, the roots of the source time function are repeated, while the roots of the different impulse responses are scattered over the complex plane and are not repeated. The roots of the source time function can be identified as those that are repeated. Note that this approach is trying to enhance characteristics of the wavelet relative to the characteristic of the earth impulse response.

We (Ziolkowski and Slob) showed that this idea worked with synthetic seismograms, only if (1) there was no truncation of the seismogram, and (2) there was no noise. These two conditions kill the idea for all practical purposes. We concluded that “any distortion of the data amplitudes by the addition of noise, windowing, or
SEISMIC WAVELET ESTIMATION WITHOUT THE INVALID WHITENESS ASSUMPTION

Continued from Page 19

truncation of the data makes it impossible to extract the wavelet by identifying multiple zeros of the Z-transform of traces in a shot gather.” We further concluded that “it is in general impossible to identify and extract the source signature from real seismic data using no statistical assumptions about the impulse response of the earth.”

This result may not be the last word on wavelet estimation using the roots of the Z-transform, but it is certainly not good news for this method.

THE WHITENESS ASSUMPTION FOR THE EARTH RESPONSE MAKING THE GEOLOGY GO AWAY

Enders Robinson (1957) was the first to pose and tackle the problem of deconvolution using statistical assumptions. His work was part of a large research programme at Massachusetts Institute of Technology (MIT), begun in 1942, in the application of prediction techniques to non-stationary time series in which all the calculations had been carried out using desk calculators (Wadsworth et al. 1953). The assumptions Robinson (1957) made about the earth response were: (1) “from knowledge of the arrival time of one wavelet we cannot predict the arrival time of another wavelet”; (2) “from knowledge of the strength of one wavelet we cannot predict the strength of another wavelet”; (3) “the seismic trace is an automatic volume control (AVC) recording so that the seismic trace has a constant standard deviation (or variance) with time”. These assumptions were required to enable the power spectrum of the seismogram to be equated to the power spectrum of the wavelet.

In this section we discuss why this step is so important in reducing the number of eligible wavelets. In the next section we explain why this step is invalid.

The discrete Fourier transform of the seismogram may be obtained from equation (4) by substituting $Z = e^{j2\pi f}$, where $\Delta f$ is the sample interval:

$$X(\omega) = \sum_{n} x(n)e^{-j2\pi fn},$$

(6)

The Fourier transform corresponding to equation (3) is then

$$X(\omega) = S(\omega)G(\omega).$$

(7)

This complex multiplication in the frequency domain can be written as amplitude and phase components:

$$|X(\omega)|e^{j\theta(\omega)} = |S(\omega)|e^{j\theta_1(\omega)}|G(\omega)|e^{j\theta_2(\omega)}.$$  

(8)

where $|X(\omega)|=|S(\omega)||G(\omega)|$ is the amplitude spectrum and $\theta(\omega) = \theta_1(\omega) + \theta_2(\omega)$ is the phase spectrum. $\phi_\omega$ is the power spectrum of the seismogram is defined as

$$\phi_\omega(\omega) = X(\omega)X* (\omega),$$

(9)

in which the asterisk superscript denotes complex-conjugate: $X(\omega)\cdot X(\omega) = X(\omega)X^*(\omega)$. Therefore,

$$\Phi_\omega(\omega) = |X(\omega)|^2 = |S(\omega)|^2|G(\omega)|^2.$$  

(10)

If the impulse response of the earth $g$ is white, random and stationary, as originally assumed by Robinson, then the amplitude spectrum of the earth response $G(\omega)$ is flat: it has a value, independent of frequency, equal to the standard deviation of the time series $g$. Equation (10) then becomes

$$\Phi_\omega(\omega) = \sigma_g^2|S(\omega)|^2.$$  

(11)

That is, the power spectrum of the seismogram is the same as the power spectrum of the wavelet. Note that the phase spectrum is lost.

To see how this reduces the number of possible wavelets, we write equation (11) in terms of Z-transforms:

$$\Phi_\omega(Z) = \sigma^2(s_1 + s_2 Z + ... + s_n Z^n)(s_1^2 + s_2^2 Z^2 + ... + s_n^2 Z^n),$$

(12)

which may be factorised into $2n$ roots:

$$\Phi_\omega(Z) = \sigma^2(s_1 + s_2 Z + ... + \prod (Z - Z_j)(Z^{-1} - Z_j).$$

(13)

The roots occur in pairs: $Z_j$ and $1/Z_j$. We do not know which of these roots to choose. This is a consequence of losing the phase spectrum. To form the $n+1$-length wavelet $s$ we need to choose one root from each of these $n$ pairs of roots. There are $2^n$ possible ways we can do this. For $n = 100$, $2^n$ is about $10^{30}$.

That is, the whiteness assumption has reduced the number of eligible wavelets from about $10^{30}$ to about $10^{32}$. Robinson then made the famous assumption that the wavelet is minimum-phase, which forced the choice that $Z_j > 1$, thereby reducing the number of eligible wavelets to 1. (Since $n$ roots of the Z-transform of the autocorrelation function lie outside the unit circle, and $n$ roots lie inside, there is only one way to choose a wavelet with all its $n$ roots outside the unit circle.)

Before we discuss the assumptions Enders Robinson made nearly fifty years ago, it should be said that many other authors since that time have tried to replace his assumptions by other assumptions, in each case with the objective of reducing the number of possible wavelets from 100 to 1. And in each case the resulting estimated wavelet is different, as it should be. In nearly all these methods the assumption of whiteness is critical. Three readily accessible books of collected papers on this subject have been published by the Society of Exploration Geophysicists: Webster (1978), which has two volumes; Robinson and Osman (1996); and Osman and Robinson (1996).
WHY THE WHITENESS ASSUMPTION IS INVALID

Robinson (1957) was the first to use the whiteness assumption for the response of the earth. Other authors who wish to estimate wavelets from seismic data have found that they need this assumption for the earth response in one form or other. Walden and Hosken (1985) give the following definition of a white reflectivity sequence: “The reflectivity sequence consists of a sequence of mutually independent, identically distributed random variables. The corresponding theoretical autocorrelation function is zero except at lag zero (where it takes the value unity) and the power spectrum is flat to Nyquist frequency”.

Is the whiteness assumption valid?

I have been unable to find any evidence to support the whiteness assumption. It is not valid for the normal-incidence reflection coefficient series determined from logs; it is not valid for the transmission response of a sequence of elastic layers; and it is not valid for the normal-incidence reflection response of a sequence of layers. Further, it is not valid for the non-normal reflection response of a sequence of layers, except in the case where the angle of incidence is so great that only total internal reflections are seen: this is the complicated zone on reflection records that is normally muted out. Thus, for all practical purposes in seismic reflection, the whiteness assumption is invalid. This section gives evidence to support this statement.

Walden and Hosken (1985) investigated the spectral properties of primary reflection coefficients from a wide variety of rock sequences around the world. Each sequence showed a ‘pseudo-white’ spectrum above a corner frequency, below which its power spectrum fell off as \( f^{-\beta} \), where \( \beta \) is between 0.5 and 1.5. The corner frequency was different for each log, and typically about 100 Hz or greater. The sampling interval was chosen to be 1 ms, so the Nyquist frequency was 500 Hz. The ‘pseudo-white’ section of the spectral band was less than two octaves and often not greater than one octave. Relative to these broad trends each log exhibited rapid spectral fluctuations of, typically, ±5 dB. Walden and Hosken went to considerable trouble to try to model the main trends using ARMA models and found that it was necessary to have knowledge of the sedimentary environments in order to choose the right parameters for these models.

Since the work of Walden and Hosken, no work has been published to suggest that primary normal-incidence reflection coefficient sequences are white. They are not even approximately white. Even if the main low-cut trend is known and removed, there are still large rapid spectral variations over the whole bandwidth of the sequence. The autocorrelation of the reflection coefficient series exhibits significant peaks and troughs.

The reflection response contains multiples as well as primary reflections. So the reflection response does not have the same spectral properties as the primary reflection sequence. The multiples are totally dependent on the primaries. In their famous paper ‘Reflections on amplitudes’, O’Doherty and Anstey (1971) showed that it is internal multiples that carry seismic energy down to deep reflectors and back to the surface. Without this mechanism very deep reflections would be undetectable. In other words, it is only because of these internal multiples that the seismic reflection process works at all. It is a consequence of this mechanism that the transmission response of a layered sequence is low-pass and the reflection response loses high frequency energy progressively with time. That is, the generation of internal multiples acts as a progressive high-cut filter on the reflection response. Even if the primary reflection coefficient series has a pseudo-white spectrum above about 100 Hz, the high-cut filtering effect of the internal multiples ensures that there is no pseudo-flat zone in the frequency spectrum of the reflection response. Since this high-cut filtering effect increases progressively down the seismogram, we very rarely see seismic reflection energy above 50 Hz at typical target depths of 2 - 4 s two-way time.

The relation between the primary reflection coefficients and the reflection response has been studied in great detail (e.g. Ewing et al., 1957; Kennett, 1983). Calculation of the response given the layer parameters - the forward problem - is now well understood and there are various commercial packages available to do this. Calculation of the primary reflection coefficient sequence from the responses - the inverse problem - is difficult. It is known that, even in the absence of noise, it is numerically unstable to perform the inversion as the direct inverse of the forward problem (e.g. Ziolkowski et al., 1989). Therefore we must consider the reflection response of the earth as some non-white sequence containing both primary reflections and internal multiples.

Are there any conditions in which the reflection response can be regarded as ‘white’? Treitel and Robinson (1966) modelled the earth as a sequence of horizontal elastic layers sandwiched between two half spaces. They gave each layer the same one-way travel time and studied the normal incidence reflection and transmission responses. They found that the reflection response to a wave incident from the upper half space is all-pass, or white, only if there is a perfect reflector somewhere in the sequence. At the perfect reflector all the downgoing energy is reflected back to the upper half space. Since none of the intervening layers may absorb energy, all the energy of the input signal is returned to the upper half space and the amplitude spectrum of the reflection response is the same as the amplitude spectrum of the input signal. A perfect reflector is an interface at which the ratio of the acoustic impedances on the two sides is zero or infinity. The reflection coefficient at the interface is -1 or +1. In the real earth there are no perfect reflectors in the subsurface. So a white normal-incidence reflection response is actually impossible.

Fokkema and Ziolkowski (1987) used a model similar to that of Treitel and Robinson, but allowed each layer to have arbitrary

Continued from Page 20

SEISMIC WAVELET ESTIMATION WITHOUT THE INVALID WHITENESS ASSUMPTION

Continued from Page 20

WHY THE WHITENESS ASSUMPTION IS INVALID

Robinson (1957) was the first to use the whiteness assumption for the response of the earth. Other authors who wish to estimate wavelets from seismic data have found that they need this assumption for the earth response in one form or other. Walden and Hosken (1985) give the following definition of a white reflectivity sequence: “The reflectivity sequence consists of a sequence of mutually independent, identically distributed random variables. The corresponding theoretical autocorrelation function is zero except at lag zero (where it takes the value unity) and the power spectrum is flat to Nyquist frequency”.

Is the whiteness assumption valid?

I have been unable to find any evidence to support the whiteness assumption. It is not valid for the normal-incidence reflection coefficient series determined from logs; it is not valid for the transmission response of a sequence of elastic layers; and it is not valid for the normal-incidence reflection response of a sequence of layers. Further, it is not valid for the non-normal reflection response of a sequence of layers, except in the case where the angle of incidence is so great that only total internal reflections are seen: this is the complicated zone on reflection records that is normally muted out. Thus, for all practical purposes in seismic reflection, the whiteness assumption is invalid. This section gives evidence to support this statement.

Walden and Hosken (1985) investigated the spectral properties of primary reflection coefficients from a wide variety of rock sequences around the world. Each sequence showed a ‘pseudo-white’ spectrum above a corner frequency, below which its power spectrum fell off as \( f^{-\beta} \), where \( \beta \) is between 0.5 and 1.5. The corner frequency was different for each log, and typically about 100 Hz or greater. The sampling interval was chosen to be 1 ms, so the Nyquist frequency was 500 Hz. The ‘pseudo-white’ section of the spectral band was less than two octaves and often not greater than one octave. Relative to these broad trends each log exhibited rapid spectral fluctuations of, typically, ±5 dB. Walden and Hosken went to considerable trouble to try to model the main trends using ARMA models and found that it was necessary to have knowledge of the sedimentary environments in order to choose the right parameters for these models.

Since the work of Walden and Hosken, no work has been published to suggest that primary normal-incidence reflection coefficient sequences are white. They are not even approximately white. Even if the main low-cut trend is known and removed, there are still large rapid spectral variations over the whole bandwidth of the sequence. The autocorrelation of the reflection coefficient series exhibits significant peaks and troughs.

The reflection response contains multiples as well as primary reflections. So the reflection response does not have the same spectral properties as the primary reflection sequence. The multiples are totally dependent on the primaries. In their famous paper ‘Reflections on amplitudes’, O’Doherty and Anstey (1971) showed that it is internal multiples that carry seismic energy down to deep reflectors and back to the surface. Without this mechanism very deep reflections would be undetectable. In other words, it is only because of these internal multiples that the seismic reflection process works at all. It is a consequence of this mechanism that the transmission response of a layered sequence is low-pass and the reflection response loses high frequency energy progressively with time. That is, the generation of internal multiples acts as a progressive high-cut filter on the reflection response. Even if the primary reflection coefficient series has a pseudo-white spectrum above about 100 Hz, the high-cut filtering effect of the internal multiples ensures that there is no pseudo-flat zone in the frequency spectrum of the reflection response. Since this high-cut filtering effect increases progressively down the seismogram, we very rarely see seismic reflection energy above 50 Hz at typical target depths of 2 - 4 s two-way time.

The relation between the primary reflection coefficients and the reflection response has been studied in great detail (e.g. Ewing et al., 1957; Kennett, 1983). Calculation of the response given the layer parameters - the forward problem - is now well understood and there are various commercial packages available to do this. Calculation of the primary reflection coefficient sequence from the responses - the inverse problem - is difficult. It is known that, even in the absence of noise, it is numerically unstable to perform the inversion as the direct inverse of the forward problem (e.g. Ziolkowski et al., 1989). Therefore we must consider the reflection response of the earth as some non-white sequence containing both primary reflections and internal multiples.

Are there any conditions in which the reflection response can be regarded as ‘white’? Treitel and Robinson (1966) modelled the earth as a sequence of horizontal elastic layers sandwiched between two half spaces. They gave each layer the same one-way travel time and studied the normal incidence reflection and transmission responses. They found that the reflection response to a wave incident from the upper half space is all-pass, or white, only if there is a perfect reflector somewhere in the sequence. At the perfect reflector all the downgoing energy is reflected back to the upper half space. Since none of the intervening layers may absorb energy, all the energy of the input signal is returned to the upper half space and the amplitude spectrum of the reflection response is the same as the amplitude spectrum of the input signal. A perfect reflector is an interface at which the ratio of the acoustic impedances on the two sides is zero or infinity. The reflection coefficient at the interface is -1 or +1. In the real earth there are no perfect reflectors in the subsurface. So a white normal-incidence reflection response is actually impossible.

Fokkema and Ziolkowski (1987) used a model similar to that of Treitel and Robinson, but allowed each layer to have arbitrary
thickness. We (Fokkema and Ziolkowski) studied the reflection response at non-normal incidence and found that the reflection response can be white, but only if the critical angle at the deepest interface is exceeded. Since velocities generally increase with depth, this critical angle exists, and post-critical reflections are seen at large offsets in real data. However, since these post-critical reflections arrive together with head waves, and are difficult to separate from them, they are very difficult to analyse, and are normally removed from the seismic data at a very early stage in the processing sequence. It is on the data from which post-critical reflections and refractions have been removed that wavelet estimation is performed. This is exactly the part of the seismic data in which the whiteness assumption is invalid.

In summary, the reflection response of the earth is non-white. There is a low-cut aspect to the response, observed all over the world, and attributable to the primary reflection coefficients themselves. There is a high-cut effect caused by internal multiples. Within these two main trends there are many peaks and troughs in the spectrum that depend on the properties of the layered earth. At pre-critical incidence wave theory shows that it is impossible for the reflection response to be white.

CONSEQUENCES FOR WAVELET ESTIMATION AND PREDICTIVE DECONVOLUTION

The case for jettisoning the whiteness assumption is not new, but it does not yet seem to have entered the consciousness of many practising geophysicists. The consequences for seismic data processing are profound: we are right back to our eligible wavelets, and we must question the value of predictive deconvolution.

Current data processing practice appears to ignore this. In a typical fairly recent example, Foster et al. (1997) employ Robinson’s (1957) theory to estimate the wavelet. They first equate the amplitude spectrum of the wavelet to the amplitude spectrum of the data in a given window; that is, they assume the earth response is white in this window. They then assume the wavelet is minimum-phase. Minimum-phase wavelets are estimated in this way over a number of windows in the data and are averaged to form a single estimated wavelet for the whole data set. The phase of this average wavelet is then subtracted from all the data to create a ‘zero-phase’ section. Before this wavelet estimation and zero-phasing step, they use gap deconvolution before and after stack.

Given that the whiteness assumption is not even remotely correct, what are the consequences for wavelet estimation and predictive deconvolution?

First, the amplitude spectrum of the wavelet is not equal to the amplitude spectrum of the data. Any wavelet estimation based on the whiteness assumption will give the wrong answer. Whatever use of this assumption is made in processing is likely to lead to misinterpretations. We can check this using well data, and an example is given later on.

Continued on Page 23
Second, since the whiteness assumption is invalid, so is gap deconvolution. This can easily be demonstrated, as follows. The gap deconvolution filter is the prediction-error filter of Peacock and Treitel (1969). It aims to predict and subtract undesirable components of the seismogram. These are determined from the autocorrelation of the seismogram which can be defined as

\[
\phi_s(\tau) = \sum_i x_i x_{i-\tau}
\]

where \(\tau\) is known as the lag. The autocorrelation may have significant values at lags in the region \(\alpha < \tau < (\alpha + n)\), which are interpreted as caused by undesirable effects such as multiples or bubble pulses (in marine data). The least-squares prediction filter \((p_0,p_1,\ldots,p_n)\) that predicts these undesirable effects is calculated from the following equations

\[
\begin{bmatrix}
\phi_s(0), \phi_s(1), \ldots, \phi_s(n) \\
\phi_s(1), \phi_s(0), \ldots, \phi_s(n-1) \\
\vdots \\
\phi_s(n), \phi_s(n-1), \ldots, \phi_s(0)
\end{bmatrix} \begin{bmatrix}
p_0 \\
p_1 \\
\vdots \\
p_n
\end{bmatrix} = \begin{bmatrix}
\phi_s(\alpha) \\
\phi_s(\alpha + 1) \\
\vdots \\
\phi_s(\alpha + n)
\end{bmatrix}.
\]

The prediction-error filter is then formed as \((1, \alpha-1\) zeros, \((p_0,p_1,\ldots,p_n)\). Convolution of this filter with the seismogram yields a sequence \(y\), say, in which the predictable part has been removed. The sequence \(y\) has an autocorrelation function which is essentially zero in the range \(\alpha < \tau < (\alpha + n)\). We can see from equation (15) that \(p_i = 0\), for all \(i\), if the right-hand side coefficients are all zero: in other words there is nothing to predict if the autocorrelation function is zero in this range. In predictive deconvolution, the "predictable part" is by definition whichever component of the seismogram contributes to non-zero values of the autocorrelation function within the chosen range of lags. Removing this component ensures that the autocorrelation is zero in the desired range.

The problem with predictive deconvolution is this: it throws out the baby with the bath water. Since the sequence of primary reflection coefficients is non-white, its autocorrelation function is in general non-zero. So prediction-error filtering removes primary reflections as well as multiples. In any case, as we know from the study of wave propagation in layered media, most multiples are not convolutional and cannot, therefore, be removed with a convolutional filter.

The value of predictive deconvolution should therefore be seriously reconsidered.

**PRE-STACK WAVELET ESTIMATION BY SOURCE SIGNATURE MEASUREMENT**

The scope for imaginative assumptions that will help us to find possible solutions to equation (1) is virtually unlimited. However, if we are serious about constraining the range of possible solutions to what is physically possible, we must try to measure the source signature.

In my paper 'Why don’t we measure seismic signatures?' (Ziolkowski, 1991) I outlined published and tested methods to measure seismic signatures for dynamite and vibroseis on land, and for marine air guns. At that time there was no agreement on the need to measure any source signature, except for the land vibrator, and I pointed out that we still did not have it right for the vibrator. In the decade that has elapsed since then, progress has been made on source signature measurements, both for marine and land seismic operations. However, little use is being made of these developments, and one must assume that there is virtually no consciousness among those who place seismic data acquisition contracts of the need to make source signature measurements.

In marine seismic exploration, one contractor has gone to considerable trouble to implement, test and refine the 'notional source method' of Ziolkowski et al. (1982) and Parkes et al. (1984). This method relies on the principle that the acoustic wavefield of an array of interacting oscillating air gun bubbles can be described as a superposition of spherical waves radiating from the rising air bubbles. Each spherical wave propagates at the speed of sound and has amplitude that decays inversely with distance. This description is valid at distance greater than about 1 m from the bubble centres. At closer distances, nonlinear terms become important. The contractor has been offering this source signature estimation to the oil industry since 1995, although there has not been great use of this offer. The patent for this particular invention expired in May 2001 - after twenty years. The method may now be implemented by anyone without incurring patent licence fees.

Several competing methods have been proposed and I outline four of them here. Landrø and Sollie (1992) used a forward-modelling scheme, following the method introduced by Ziolkowski (1970), in conjunction with a ministreamer towed beneath the air gun array. They used an iterative solution of the forward-modelling equations to solve for adjustable parameters in the equation of motion of each air gun bubble in order to match the ministreamer data in a least-squares sense. The array signature may then be produced by running the forward modelling scheme again using the best-fitting parameters.

Amundsen (1993) used ministreamer measurements to solve for the notional source signatures by a linear least-squares method. He neglected the relative motions of the ministreamer and rising air bubbles to obtain a time-invariant problem amenable to a least-squares approach.

These two new ministreamer methods were compared with the notional source method in an experiment described by Laws et al. (1998). In the meantime Ziolkowski and Johnston (1997) showed how errors in the notional source method could be estimated for each shot using additional near-field hydrophones.

In a third new method Hobbs and Jakubowicz (2000) have res-
SEISMIC WAVELET ESTIMATION WITHOUT THE INVALID WHITENESS ASSUMPTION

Continued from Page 23

The wavelet that returns from the target point, then, the task of wavelet estimation is hopeless. However, we have seen that this is invalid. At this point, the wavelet is unknown after predictive deconvolution and forces the use of wavelet estimation techniques to find out what the wavelet might be.

In the wavelet estimation step, as we have seen, it is assumed that the wavelet is minimum phase. This assumption would be helpful if the amplitude spectrum of the wavelet can be estimated from the data. Normally this estimate requires the whiteness assumption. However, we have seen that this is invalid. At this point, the task of wavelet estimation is hopeless.

THE EFFECT OF ATTENUATION AND THE PROBLEM OF WELL TIES

One argument that is sometimes put forward for not measuring the source signature is that the wavelet that returns from the target looks nothing like the source signature after it has been filtered by the earth. Earth filtering has been studied in considerable detail. See, for example, Toksöz and Johnston (1981) for a superb collection of papers on this subject. Kjartansson (1979) is especially relevant to this discussion. Kjartansson found that, although Q is weakly dependent on frequency, there is no indication that non-constant Q models provide a better description of attenuation than the constant Q theory does.
Each layer has its own wave velocity and $Q$. In principle this effect can be modelled. The overall effect on the propagating wavelet is to broaden the wavelet along the wave path. Thus the wavelet gets longer, as if it had been convolved with a low-pass causal filter. (Something that has always bothered me about the wavelets that are estimated from seismic data is that they are invariably shorter than the original source signatures. This is the case, for example, in the paper of Foster et al. (1997)).

One can check the processing scheme using wells, and I use an example that has already been published (Ziolkowski, 2000). The University of Edinburgh collaborated with Enterprise Oil and Geco-Prakla to use source signature measurements in a small part of a 3-D data set obtained in the North Sea in 1995. Near-field pressure measurements were made according to the notional source method. At this point the data could be processed two ways: the conventional way, resulting in a ‘zero-phase’ section, and the more deterministic way, using the method proposed above. A new deviated well was drilled and logged.

Using the ‘zero-phase’ processed section near the well, we constructed a zero-phase wavelet from the data over a window 1800-2400 ms and convolved it with the normal-incidence reflection coefficient series computed from the well logs. Extensive corrections for the well deviation were made. The result of this exercise is shown in Figure 1. The “tie” is poor. This is normally what happens with this approach, as mentioned above.

Figure 1: Statistical well tie: seismic data processed to “zero phase” without using source signature measurements; zero-phase wavelet extracted from the data and convolved with normal-incidence reflection coefficients to form a synthetic seismogram. (Reproduced by permission of Enterprise Oil plc.)
However, because we had the near field measurements, we were able to estimate the far field wavelet. Signature deconvolution was applied as a first step in processing to convert this wavelet to the shortest possible zero-phase wavelet within the bandwidth. As in the conventional approach to the processing of the data, predictive deconvolution was applied before stack (but not after stack) to attenuate water layer multiples. In this case, though, the gap was chosen to be larger than the length of the wavelet, to ensure that no lateral variations were introduced into the wavelet by the processing. Subsequent processing was the same as for the conventional route, except that there was no proprietary conversion to zero phase.

The result of convolving the known short zero-phase wavelet with the reflection coefficient series is shown in Figure 2. Again, the "tie" is not good, but we notice that the bandwidth of the synthetic seismogram is much broader than the bandwidth of the data. This part of the North Sea has high intrinsic attenuation: the peak frequency of the data at 2000 ms is only 25 Hz. We therefore filtered the synthetic seismogram by applying a minimum-phase filter to the synthetic seismogram with an amplitude spectrum of the form $\exp\left(-\alpha \omega^\alpha\right)$ in which $\omega$ is the angular frequency and $\alpha$ is a constant found by trial and error. This corresponds to a constant $Q$ model for attenuation. Using a value of $Q$ equal to 37.5, the "tie" is as shown in Figure 3, and is good. Higher up in the section the tie is of course not so good. Because the tie is so good in Figure 3, we can state that the wavelet in the data at this depth in the section is our zero-phase wavelet convolved with the minimum-phase $Q$ filter. We have achieved this result deterministically, but with the choice of a single number $\alpha$ (or $Q$) to cause the spectrum of our synthetic seismogram to more nearly match the spectrum of the data.

CONCLUSIONS

Current practice in the seismic exploration industry relies on the estimation of wavelets from seismic data to tackle a number of problems that arise whenever the seismic data must be related to known geology. In the absence of constraints the number of 100-length wavelets one can estimate from a seismogram of 3,000 samples is about $10^{30}$. With the whiteness assumption, that is so critical to most wavelet estimation techniques, this can be reduced to about 1. There is no evidence to support the whiteness assumption. In fact, the evidence is all to the contrary. The reflection response of the earth is non-white. There is a low-cut aspect to the response, observed all over the world, and attributable to the primary reflection coefficients themselves. There is a high-cut effect caused by internal multiples. Within these two main trends there are many peaks and troughs in the spectrum that depend on the properties of the layered earth. At pre-critical incidence wave theory shows that it is impossible for the reflection response to be white.

A consequence of the invalidity of the whiteness assumption is that predictive deconvolution must remove primary reflections as well as multiples.

Rejection of the whiteness assumption renders the estimation of wavelets by statistical methods hopeless. However, source signature measurements can be made for most seismic sources. Using these measurements and a careful approach to signature deconvolution, it is possible to ensure that the wavelet is short, constant, and known at every stage in processing. This has the potential to give much more trustworthy interpretations of the data than are possible with current processing methods. This may be checked using wells. In making the well tie, it may be necessary to allow for attenuation, but this involves the search for a single real number - the effective $Q$ - that is required to equalise the spectra of the synthetic seismogram and the real seismic data at the well.
ACKNOWLEDGMENTS

I am very grateful to the Canadian Society of Exploration Geophysicists for inviting me to write this article for this special issue of the ‘Recorder’. I thank Floris Strijbos of Shell Expro and my colleague Bruce Hobbs for interesting discussions.

REFERENCES

Kennett, B.L.N., 1983, Seismic wave propagation in stratified media: Cambridge University Press.
Kjartansson, E., 1979, Constant Q-wave propagation and attenuation: Journal of Geophysical Research, 84, 4737-4748.
Mateker, E.J., 1971, Big benefits seen with Maxipulse system use: Oil and Gas Journal, 69, 116-120.

Continued on Page 28

SEISMIC WAVELET ESTIMATION WITHOUT THE INVALID WHITENESS ASSUMPTION

Continued from Page 26

LYNX INFORMATION SYSTEMS LTD.

REPEAT FROM MAY ISSUE

PAGE 31
SEISMIC WAVELET ESTIMATION WITHOUT THE INVALID WHITENESS ASSUMPTION

Continued from Page 27


Toke’s, M.N. and Johnston, D.H., 1981, Seismic wave attenuation: Geophysics Reprints Series 2, SEG.


Webster, G.M., Deconvolution: Geophysics Reprints Series 1, SEG.


Continued from Page 2

EDITOR’S MESSAGE

Jon Downton and Larry Lines in their paper ‘AVO feasibility and reliability analysis in the presence of random noise’ show that in the presence of noise, constraints are needed to make the reflectivity attribute estimates for linearized AVO inversion more certain. These could potentially introduce theoretical error into the problems. However, the correctness of different approximations could be viewed in terms of how geologically plausible the used constraints are.

In ‘AVO and the general inverse theory’, Guillaume Cambois raises some intriguing questions about how reliable conventional AVO extraction is, and offers an alternative which perhaps avoids some of the potential pitfalls. He does this using general inverse theory. Cambois presents unsettling evidence showing a statistical correlation between intercept and gradient, and the possibility of systematic non-random extractions of intercept and gradient from pure noise.

Subhashis Mallick in his paper ‘Prestack waveform inversion using a genetic algorithm - the present and the future’ discusses the simple assumption that is made in AVO analysis - that there is no contamination of P wave amplitudes by other modes. This assumption is usually not valid. Mallick proposes that full waveform prestack inversion needs to be used, which gives an estimate of the optimum earth model at a given CMP location.

Mrinal Sen in his paper ‘Prestack waveform inversion on plane wave seismograms’ focuses on the challenging issues of waveform inversion. In his methodology Sen shows that τ−p appears to be the most appropriate domain for carrying out prestack waveform inversion, in that the inherent forward modeling is very fast and it generalizes easily to anisotropy and multicomponent analysis.

These 11 articles represent some of the outstanding technical work being done in our field. We trust this issue will provide you, our readers, stimulating reading, inspire you in your own technical work, and perhaps challenge you to submit an article for a future issue of the RECORDER. Keep in mind that you still have the September RECORDER to look forward to, the second part of the “Special Issue”, and it will contain 12 more excellent articles on a variety of topics.