# Criteria for optimum estimated pulse inverse filtering

Ivan J. P. de Vasconcelos[1], Liliana A. Diogo[1], Renato L. Prado[1].
[1]Applied Geophysics Laboratory, University of Sao Paulo, Brazil.



## Abstract

This work revolves around the determination of critical criteria for optimum performance of the estimated pulse inverse filtering algorithm introduced by Porsani and Ursin (1998). Their algorithm assumes that for a given signal with, according to Z transform representation, $\alpha$ zeros (out of a total of N zeros) inside the unit circle, one may estimate a suite of inverse filters using the same autocorrelation function (ACF) for each value of $\alpha$; finally the suite of filters (and its corresponding value of $\alpha$) which best represents the source pulse should maximize the Lp norm of the filtered trace. Tests in synthetic datasets have showed that the algorithm is extremely sensible to the choice of certain parameters, such as the numbers of coefficients of the ACF and the time window in which the objective function for finding $\alpha$ is calculated. The observation of algorithm's behavior on synthetic data enabled us to establish quantitative criteria for the selection of optimum parameters based on the nomalized Cross-Correlation (nCC) between the input and the deconvolved trace. The criteria were then applied to deconvolve real data, showing good preliminary results as to enhance vertical resolution and to eliminate non-geological features.

## Introduction

The deconvolution is an absolutely necessary step in seismic processing, since its function is to remove the presence of the source wavelet from the dataset, recovering the best representation of the subsurface reflection coefficient series. Such process is essential if we are to interpret, and to, most importantly, quantitatively characterize subsurface geological information.

The most widely used technique for deconvolution is the Wiener spiking deconvolution (Robinson&Treitel,1980; Berkhout, 1977), which assumes that the media has the statistical properties of random white noise and the source pulse must be of minimum phase. Because of these assumptions, the inverse filtering operator can be calculated using only the ACF of the pulse. Such technique, however, yields very poor results in the presence of a mixed phase pulse. In cases where the input pulse is mixed phase but known, one can adapt the Wiener filter into a pulse-shaping filter (Robinson&Treitel, 1980; Yilmaz, 1987), phase-shifting the operator with a delay in the desired output.

In most cases, however, we cannot apply pulse-shaping filters to data embedded with mixed phase pulses since the input pulse is not known, therefore bringing up the necessity for alternative deconvolution procedures.

## Estimated pulse inverse filtering

In an attempt to fulfill such need, several techniques have been developed so far, and amongst them a mixed-phase deconvolution algorithm by Porsani and Ursin (1998). This algorithm assumes that the mixed phase pulse, $p_t$ , with N zeros Z-transform, is the result of the convolution of a minimum-phase component, $a_t$ (with N-$\alpha$ zeros outside the unit circle) with a maximum-phase component (Eisner&Hampson, 1990), $b_t$ (with $\alpha$ zeros outside the unit circle); and it must not have any zeros on the unit circle. Thus, rendering:

$$p_t = a_t \times \overline{b}_t \times \delta_{t-\alpha} \quad (1)$$

Written with the Z-transform:

$$P(Z) = Z^{\alpha} A(Z) B\left(Z^{-1}\right) \quad (2)$$

Note that the maximum phase component is anticausal, with $\alpha$ samples in negative time. The minimum phase pulse with the same amplitude spectrum of P(Z)

$$P'(Z) = A(Z)B(Z) \quad (3)$$

So that we may write P(Z) in terms of P'(Z)

$$P(Z) = P'(Z)\frac{Z^{\alpha} B\left(Z^{-1}\right)}{B(Z)} \quad (4)$$

The inverse filter H(Z) of P(Z) can then be expressed as

$$H(Z) = \frac{1}{A(Z)}\frac{1}{Z^{\alpha} B\left(Z^{-1}\right)} = H'(Z)\frac{B(Z)}{Z^{\alpha} B\left(Z^{-1}\right)} \quad (5)$$

Being H'(Z) the inverse filter of P'(Z). From this we can see that by convolving the minimum-delay inverse filter with an anticausal all-pass filter we obtain the mixed-delay inverse filter.

In order to obtain a practical inverse filtering operator, one must estimate the ACF of the pulse (e.g. by using the ACF of a trace or an average ACF of a CMP gather), and with it on the diagonal of the nonsymmetrical Toepliz coefficient matrix, solve the Extended Yule-Walker system to estimate the inverse of the minimum-delay component, C(Z). The estimation of the minimum-phase pulse P'(Z) is performed by solving the Yule-Walker (YW) set of equations twice. The maximum-delay component is hence obtained as the result of the convolution between the minimum-delay pulse and the minimum-delay component inverse filter. The final estimated pulse

inverse filter may be expressed as in equation (5) or, in terms of the estimated information:

$$H(Z) = H'(Z)\frac{P'(Z)C(Z)}{Z^{\alpha}P'(Z^{-1})C(Z^{-1})} \quad (6)$$

From this relation and from and the formulation of the EYW and YW equations sets (Porsani&Ursin, 1998), we can see that each value of $\alpha$ allows us to estimate a single inverse operator, so to choice of the best value of $\alpha$ is done by normalizing the *Lp* norm (Porsani&Ursin, 1998) of the deconvolved trace. When the maximum value of this criterion is reached for a given *p* (empirically, must be larger than 2, used equal to 5) it corresponds to the best value for $\alpha$.

**Tests on Synthetic Data**

Before applying the mixed-phase deconvolution to real datasets, controlled experiments were put up to evaluate the algorithm's behavior in face of input parameter changes. The first step was to build a reflectivity series, as seen in figure 2. The series was then convolved with several signals (Gabor, Ricker, First and Third order derivatives of the Gaussian signal) to simulate seismic data. Tests involved checking the algorithm's capacity to estimate and filter different source pulses, filtering source pulses of the same form but with different peak frequencies, and finally evaluating the effectiveness of the filtering procedure for a given series with different signal to noise ratios (S/R).

With the wish of further interpreting the differences caused on the output of the algorithm by varying parameters, we made use of the Autocorrelation (AC) and normalized Cross-correlation (CC), which respectively are strictly statistically defined (Robinson&Treitel, 1980) as:

$$\phi_{\tau} = \lim_{T\to\infty}\frac{1}{2T+1}\sum_{t=-T}^{T} x_{t+\tau}x_t \quad (9)$$

$$\phi_{12}(\tau) = \lim_{T\to\infty}\frac{1}{2T+1}\sum_{t=-T}^{T} x_{1,t+\tau}\,x_{2,t} \quad (10)$$

Being $\phi_{11}$ the autocorrelation, and $\phi_{12}$, the cross-correlation. But for actual calculations, correlations are considered in the deterministic form. Nevertheless, in order to compare different correlation plots, we must have normalized Cross-Correlations (nCC) and normalized Autocorrelations (nAC). Letting each series, in terms of time average, have zero mean and unit variance and by applying the Schwarz inequality, we have the desired normalization, with correlation values that are less or equal to 1. In practice, that is to say that input data must be normalized with respect to their Root Mean Square (RMS) values. Instead of normalizing each dataset, we can directly incorporate the normalization into correlation computations by

$$\phi'_{xx}(\tau) = \frac{\phi_{xx}(\tau)}{\phi_{xx}(0)} \quad (13)$$

$$\phi'_{12}(\tau) = E\{X'_1(t+\tau)X'_2(t)\} = \frac{1}{\sqrt{\phi_{11}(0)\,\phi_{22}(0)}}E\{X_1(t+\tau)X_2(t)\} = \frac{\phi_{12}(\tau)}{[\phi_{11}(0)\phi_{22}(0)]^{1/2}} \quad (14)$$

Where E denotes Expected Value and $x_1$' represents the series $x_1$ normalized by its RMS.

After completing the tests, it has been generally observed that the algorithm performed well in estimating pulses and deconvolving data, as previously showed by Porsani and Ursin (1998). However, getting adequate results depended critically on the choice of certain input parameters. Two parameters were of greater importance: the number of coefficients of the ACF ($nR_{xx}$), and the time window in which the objective function to estimate the optimum value of $\alpha$ (the inversion was implemented using the genetic algorithm).

The parameter that has the most critical effect on the pulse estimation is the $nR_{xx}$ parameter. Differently from what we would expect from conventional Wiener filtering, the systematic increase of $nR_{xx}$ does not yield respectively better results, instead, the best results are only to be achieved under optimum values. The departure from these values can quickly produce results that by far fail to represent true reflectivity information. Figure 2 presents the reflectivity series convolved with a 36Hz peak frequency Gabor signal, it also depicts its mixed-phase deconvolution with $nR_{xx}$ values of 47, 51 and 55, plus their minimum-phase counterpart with $nR_{xx}$=51. Undoubtedly, the mixed-phase deconvolution with $nR_{xx}$=51 is the best representative of the reflectivity series, whereas the others not only have poorly recovered reflectivity information, but also appear to have an increased S/R when compared to the optimum result. This behavior is obviously analogous with concern to pulse estimation, since it was the seed for inverse filtering operator calculations; figure 3 shows the input pulse, and the pulses estimated with $nR_{xx}$ value of 51. We clearly see that the best pulse estimation is the one used to calculate the best inverse filtering operator of figure 2, that of $nR_{xx}$=51.

The existence of an optimum value for $nR_{xx}$ is confirmed for all of the synthetic traces, but different data called for different optimum values of $nR_{xx}$, and we were, until this moment, unable to define a predictable pattern for this parameter, thus its determination is being entirely based on trial and error. Like any other process that involves finding optimum values, we set out to determine criteria that would serve the purpose of choosing such values. Hence, based on the convolutional theory and on the fact that the best deconvolution should represent the Reflection Series, one would expect that the nCC between the deconvolved trace and the series approximates the nAC of the Series, with its maximum close to zero lag, and with the property of being symmetrical. The only deconvolution to follow such rule (figure 4) would be the one with $nR_{xx}$=51. Still, that does not solve our problem, since with real datasets the Reflection Series is exactly what we want to recover, we must rely on a comparison based only on input and output of the deconvolution process. So, turning to figure 1 right

bellow, a properly triggered mixed-phase source pulse would have a peak at a non-zero time $t_f$ that when convolved with a Reflectitity Series spike at a given $t_s$ results in a data pulse with its maximum energy at $t_s+t_f$. A near-perfect deconvolutional procedure over the convolved signal should remove most of the source pulse signature and place the output peak close to the Reflectivity spike at a time $t_s+t_d$, with $t_d$ anywhere between zero and $t_f$. That has a direct consequence on the nCC of the input data with the deconvolved output as moving filter: since the output is "behind" the input on time, the maximum value of the nCC must have a positive lag corresponding to $t_f - t_d$ .

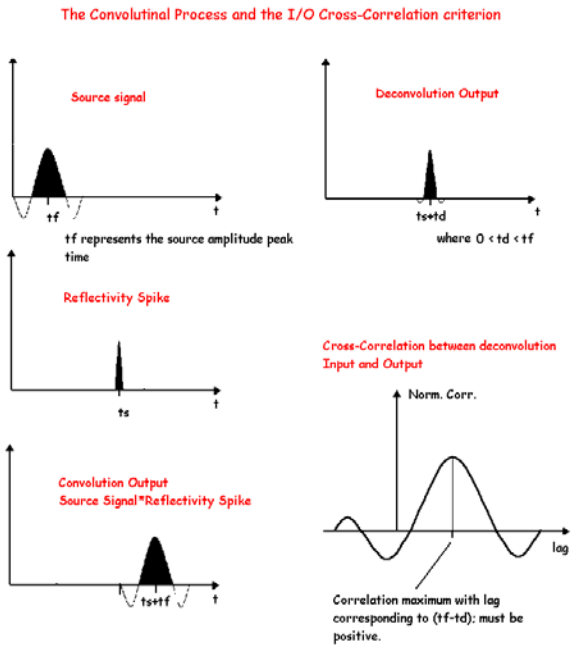


*Figure1- Illustration of the cross-correlation criterion.*

Indeed, such characteristic relationship between Input/Output (I/O), is observable in figure 4, where only the plot related to $nR_{xx}$=51 follows the exact same pattern. If the normalized Cross-Correlation were to be considered as an objective function for inverting for nRxx values, then perhaps not only should it seek the maximum correlation value, but its corresponding lag value should perhaps be taken into consideration, as a weighting factor.

**Real Data example**

To test the performance of the estimated pulse inverse filtering algorithm under the light of the above-mentioned criteria, we introduced it into the processing of a shallow seismic reflection survey over sediments of the Itaquaquecetuba formation, in the vicinities of Itaquaquecetuba, Sao Paulo state. The source was comprised of a 10 lb sledgehammer, being its signature visibly of mixed-phase. The data consisted of up to 12 fold conventional CMP gathers, and pre-processing involved band-pass frequency filtering and fk filtering to remove Ground Roll features (which are always quite a hurdle in shallow seismic data) and spatial aliasing. Refraction events were simply muted from common shot gathers. Velocity analysis was carried out through CVS panels, as coherence measures are often compromised in the face of shallow seismic data nature. In figure 5 we can see the final stacked product of the processing without any deconvolution, the target reflector is displayed in red, representing the base of a sand paleochannel.

Deconvolution before stacking was applied to CMP gathers. To choose the optimum parameters, far-offset traces from CMP number 48 were analyzed with the criteria developed with synthetic data tests. Figure 5 shows the trace with a 24.0 m offset, its mixed-phase deconvolutions with $nR_{xx}$ values of 295 and 161, and minimum-phase deconvolution, $nR_{xx}$=295 on the time window for objective function calculation of 0.05 to 0.07 seconds. Figure 7 shows the nCC's for each $nR_{xx}$ value, and figure 6 shows the respective estimated pulses. We interpreted $nR_{xx}$=295 as being the best deconvolution as it does improve vertical resolution and because the estimated pulse seems to better represent the pulse in the data, even though it shows a lower correlation value in the maximum peak when compared to the $nR_{xx}$=161 result, which in turn presents a sensible smaller displacement in time of the series and shows little difference from the minimum-phase deconvolution. Upon examination of the deconvolution of the entire CMP 48, in figure 8, with $nR_{xx}$=295, we can see that mixed-phase deconvolution has given some improvement to vertical resolution and it has also helped to better separate the events. Finally, figure 9 shows the final stacked data after deconvolution has been applied to the CMP gathers with $nR_{xx}$=295. Firstly, the results obtained with CMP 48 were not the same as for the other gathers with the same parameters, suggesting that optimum values of $nR_{xx}$ should be sought independently for each gather, again calling for an automatic manner to perform the choice. Nevertheless, mixed-phase deconvolved data yielded a better section in terms of vertical resolution and mainly keeping only interpretable features without false, numerically created reflectors.

When trying to deconvolve already stacked sections, processing faced some difficulty in finding parameters that would approximate the nCC-based criteria. The best value that could be found was of $nR_{xx}$=555 and the time window for objective function computation were from 0.076 to 0.096 seconds. As observed in figures 9 and 7, results are not appropriate as an interpretational asset, since the deconvolution failed to characterize the reflector and the nCC reflects the inadequate pulse estimation.

**Summary and Conclusions**

Synthetic data testing has showed that the process of estimated pulse inverse filtering is particularly sensible to choice of the number of coefficients of the Autocorrelation Function (ACF), referred to as $nR_{xx}$. Such parameter usually has optimum values regarding each dataset, so we came up with selection criteria based on the normalized Cross-Correlation between input and output of the deconvolution. The criteria matched well of the optimum results obtained over synthetic data.

The application of the criteria to the real data has shown good results over the CMP 48, and its respective far-

offset traces, analyzed under the established criteria. The choice of the optimum value was not only based on the correlation but also on the visual interpretation of the outputs. Furthermore, this data also suggest that if employed as an inversion's objective function, the nCC must have its maximum value weighted by its corresponding lag. We were, however, unable to repeat the results in all CMP's using the same optimum parameters of CMP 48. Even so, mixed-phase deconvolved data yielded a slightly superior image than its minimum-phase counterpart. The fact that the synthetic data results were better than those obtained with real data with the nCC criteria might suggest that the estimation of the optimum value of $nR_{xx}$ has to be carried out separately to each gather, and that one must make sure that the convolutional model is valid, e.g. deconvolving Common Offset gathers to avoid the presence of the effect of high frequency attenuation with distance. Over stacked sections results were inadequate probably due to the lack of waveform preservation after stack, which dramatically effects deconvolution.

**References**

**Berkhout, A. J**., 1977, Least squares inverse filtering and wavelet deconvolution: Geophysics, 42, 1369-1383.

**Eisner, E., Hampson, G.,** 1990, Decomposition into minimum and maximum phase components: Geophysics, 55, 897-901

**Porsani, M. J., Ursin, B.,** 1998, Mixed-phase deconvolution: Geophysics, 63, 637-647.

**Robinson, E. A., Treitel, S.,** 1980, Geophysical signal analysis: Prentice Hall.

**Yilmaz, O., 1987**, Seismic Data Processing: Soc. Expl. Geoph.

**Acknowledgements**

We would like to express our gratitude towards Prof. Milton J. Porsani, Federal University of Bahia (UFBa), who has kindly shared the algorithm, data, and some valuable discussions. We are also thankful to Rodrigo Nunes, who took part in the acquisition of field data.

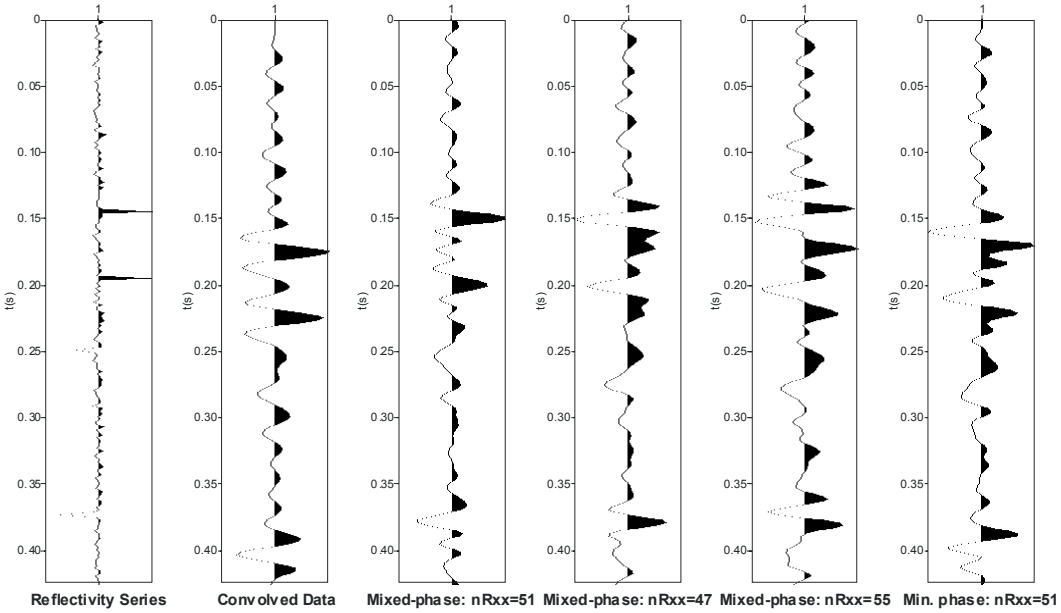


Figure 2 – Synthetic data testing. The figure shows the Reflectivity Series, Convolved with 36Hz Gabor wavelet, its mixed-phase deconvolution with nRxx values of 51, 47 and 55 and the minimum-phase decnvolution.

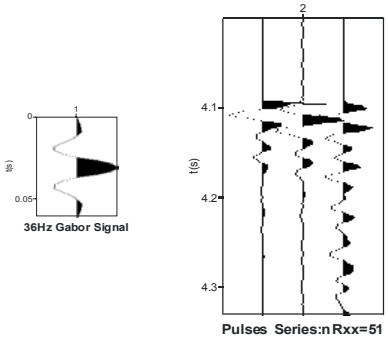


Figure 3 – Input pulse and estimated components and pulse (middle).

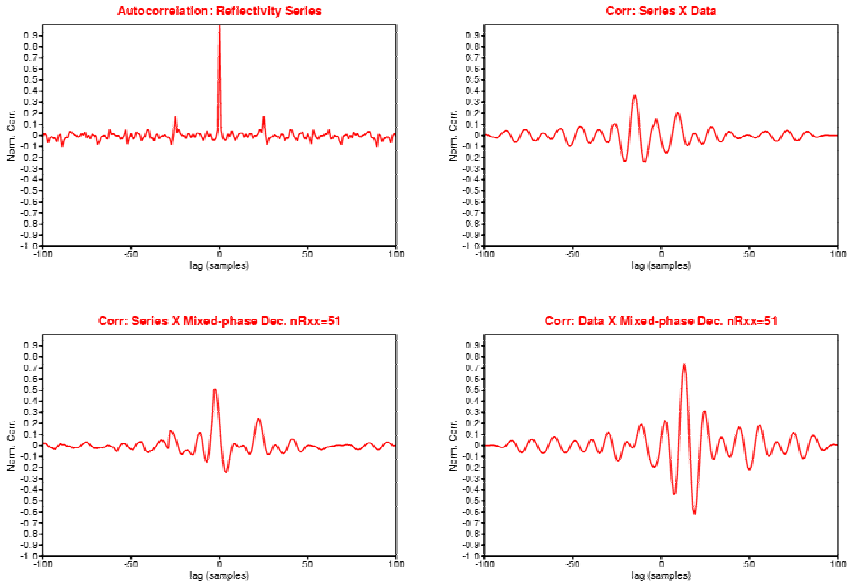


Figure 4 – Cross-correlation plots showing the behavior of the deconvolution output with the optimum value of nRxx

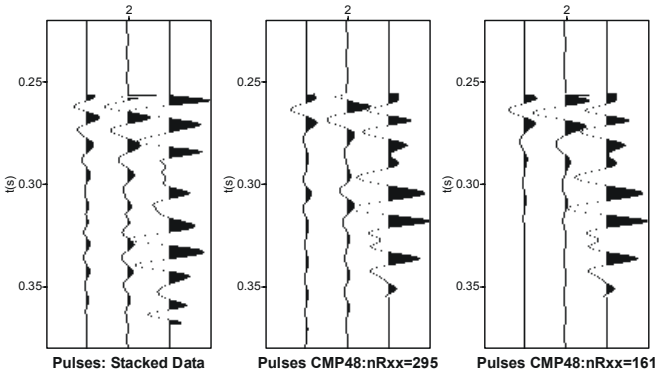


Figure 6 – Estimated pulses and their components: over the stacked data with nRxx=555 (left), over CMP 48 with nRxx=295 (middle) and with nRxx=161 (right)

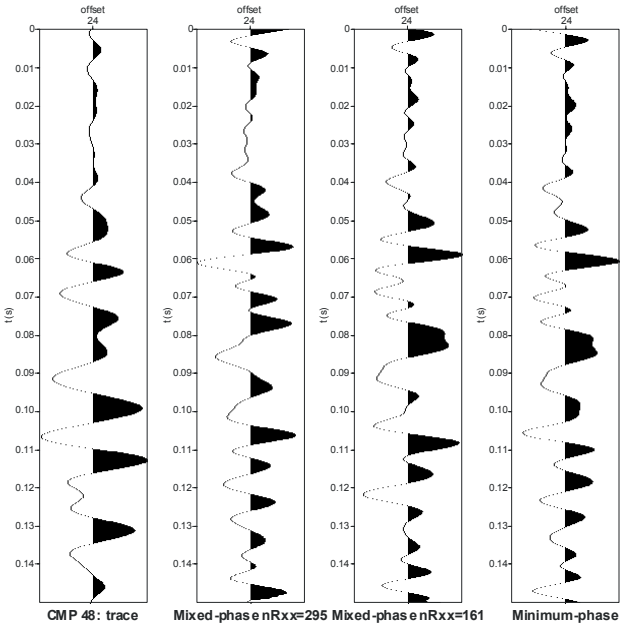


Figure 5 – From left to right: trace with offset 24m from CMP 48, mixed-phase deconvolution with nRxx values of 295 and 161, and minimum-phase deconvolution.

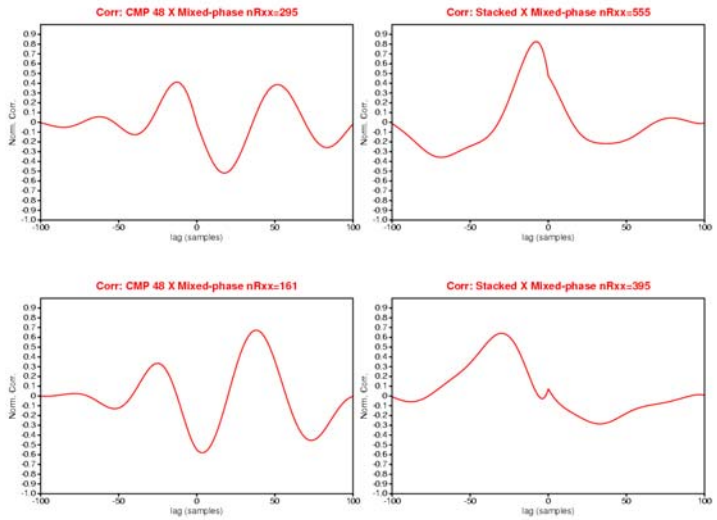


Figure 7 – Cross-correlation of I/O of the deconvolution, for the CMP 48 (plots on the left), and for the stacked section (pots on the right)

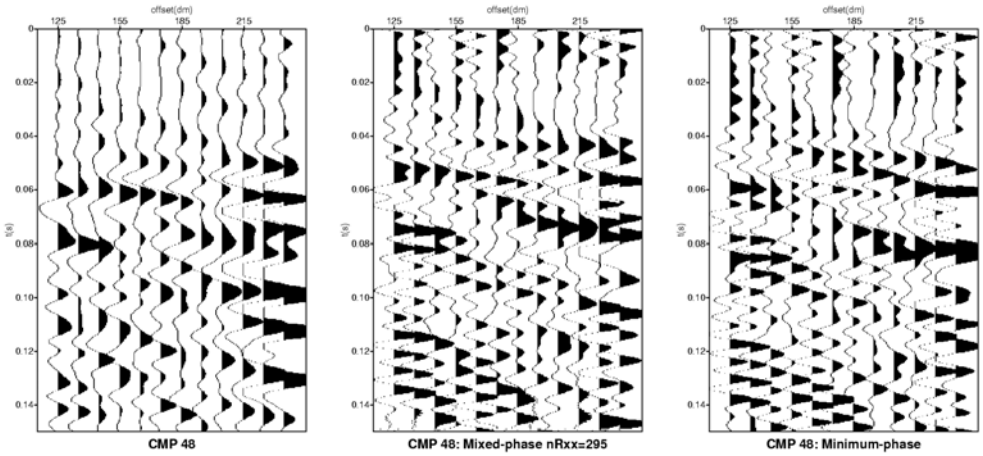


Figure 8 –12 fold CMP48 (left), its mixed-phase deconvolution with nRxx=295 (middle), and minimum-phase deconvolution.

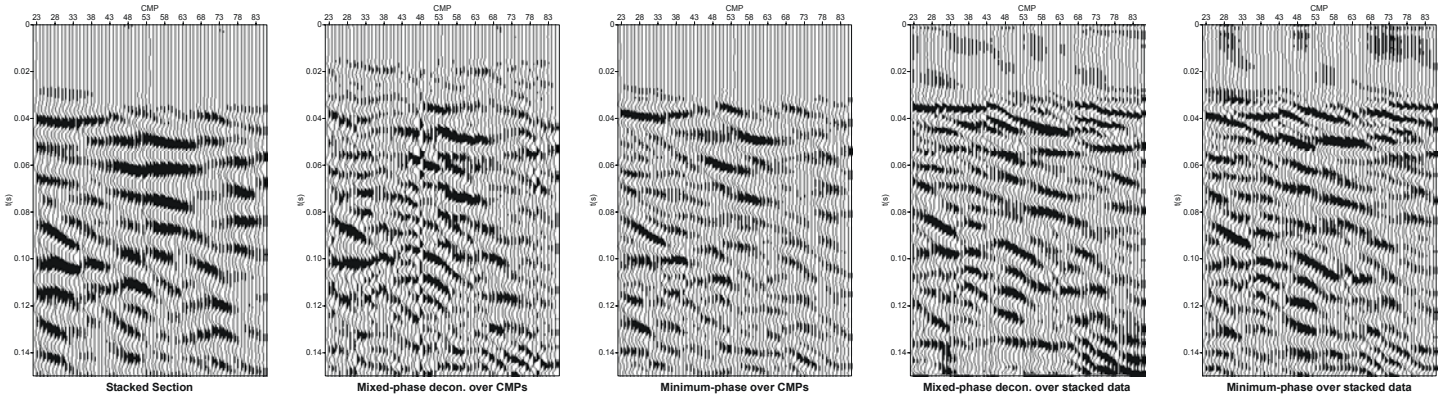


Figure 9 – From left to right: Stacked section, stacked section of mixed-phase deconvolved CMP's, stacked section of minimum-phase deconvolved CMP's, mixed-phase deconvolved stacked section and minimum-phase deconvolved stacked section