# On the use of dynamic time warping for multicomponent seismic data registration

José Fernando Caparica Jr. and Fernando Sérgio de Moraes, Universidade Estadual do Norte Fluminense, Brazil.



## Abstract

This work presents a method for multicomponent seismic data registration using initial estimated velocities and a subsequent fine-tune adjustment using dynamic time warping. While the first part can be solved purely in an analytical manner, the time warping process requires a dynamic approach. Tests were conducted using synthetic data modeled using Zoeppritz equations, with different approximations for velocities and different levels of random noise. The results show dynamic time warping as a promising tool for data registration, as long as a first approximation can bring the events on both datasets reasonably close together, as demonstrated by our synthetic data examples.

## Introduction

For the last couple of decades, multicomponent seismic data using converted waves has been used to improve AVO analysis and reservoir monitoring. Although being object of study for some time (STEWART, 1990; LARSEN, 1999; WANG, 1999; HAMPSON et al., 2005) multicomponent AVO joint inversion still presents us with technical challenges, from which we highlight the registration of PP and PS events. Compressional waves travel faster than shear waves, a reflector shown at a particular time on PP data will be registered at a later time on the PS data. Since the simultaneous AVO joint-inversion is done at each particular time (HAMPSON et al., 2005), it is essential that the same event is shown at the exact same time on both data sets.

The multicomponent data registration usually is done with the interpreter's experience on recognizing the corresponding horizons on both data sets. Hardage et al. (2011) list various techniques to properly tie the converted wave data to the compressional wave data. One of them consists on a method which employs the compressional and shear wave velocities, $\alpha$ and $\beta$, to make a first-order adjustment, then a second step must take place, where the interpreter must recognize the similar events on both data and tie them together. Or one could use a numerical method for fine-tuning image registration, such as the one proposed by Fomel and Backus (2003). Hale (2012) proposed a method called *dynamic image warping*, based on a well-known speech recognition algorithm called *dynamic time warping* (BERNDT and CLIFFORD, 1994), suggesting that it could indeed be used for fine-tuning PP and PS data registration.

In this work we describe a necessary sequence of steps to put both PP and PS data on the same time domain (PP time), aligning the events on both data sets, placing them at the exact same time. We shall present the methods and the application on a synthetic data, based on a real log data, modeled using Zoeppritz equations, with different levels and types of random noise.

### Time domain conversion at zero-offset

Given a horizontally layered half-space, and each layer being homogeneous and isotropic, one can easily find the relationship between interval times ($\Delta t_{PP}$ and $\Delta t_{PS}$) and interval velocities ($\alpha^{(\text{int})}$ and $\beta^{(\text{int})}$) at a layer of index $j$:

$$\Delta t_{PS_j} = \Delta t_{PP_j}\left(\frac{\alpha_j^{(\text{int})}+\beta_j^{(\text{int})}}{2\beta_j^{(\text{int})}}\right). \qquad (1)$$

We can rewrite equation 1 in terms of the interval velocities ratio, $\gamma_j^{(\text{int})}=\alpha_j^{(\text{int})}/\beta_j^{(\text{int})}$, as show equation 2:

$$\Delta t_{PS_j} = \Delta t_{PP_j}\left(\frac{\gamma_j^{(\text{int})}+1}{2}\right). \qquad (2)$$

Ursenbach et al. (2013) have shown that we can use either interval velocities ratio (equation 2) or the average velocities ratio

$$t_{PS_j} = t_{PP_j}\left(\frac{\gamma_j^{(\text{avg})}+1}{2}\right), \qquad (3)$$

which is more convenient, because allows us to convert the times directly, instead of having to work with time intervals. Since $\gamma^{(\text{avg})}$ can be computed directly from $\gamma^{(\text{int})}$ (GAISER, 1996), equation 3 allows us to directly convert the PS data, originally recorded in PS time domain, into PP time domain.

This first-order adjustment heavily depends on the quality of the velocities, and often only processing velocities are available. Since processing velocities are not necessarily equal to the real velocities (AL-CHALABI, 1994), using this kind of velocities to convert PS data into PP time domain most likely will not produce a satisfactory alignment. That's why a fine-tuning algorithm is important, and the one we chose will be presented next.

**Dynamic time warping (*DTW*)**

Let's consider two distinct time series - $A$ and $B$ - and assume they have similar events, but they are registered in different times. Our task is to warp one of these series in order to align its events with the events from the other series. Considering that $A$ and $B$ series have $n$ and $m$ elements respectively, such that

$$A = a_1, a_2, ..., a_i, ..., a_n \,,$$
$$B = b_1, b_2, ..., b_j, ..., b_m \,. \qquad (4)$$

We can compute the absolute difference between each term of both series (equation 5), resulting in a matrix of $n \times m$ dimensions.

$$\delta(a_i, b_j) = |a_i - b_j| \,. \qquad (5)$$

This $\delta$ function quantifies the proximity of both series at all times. There can be variations of this equation, such as the squared difference between elements, which sometimes is called as the distance between them.

We can formulate time warping problem as the being the minimum path over the cumulative differences (BERNDT and CLIFFORD, 1994):

$$DTW(A,B) = \min_w \sum_{k=1}^{p} \delta(w_k), \qquad (6)$$

where $w_k$ represents a coordinate $(i, j)_k$ from the warping path, as shows Figure 1.

The warping path typically has conditions and constraints, such as continuity and monotonicity (KEOGH and RATANAMAHATANA, 2005), meaning that there cannot be gaps on the warping path, and it can only move forward in time. To improve the algorithm performance and avoid extreme distortions, we can also impose a warping window constraint (BERNDT and CLIFFORD, 1994), which is represented in Figure 1.

In order to find the path which minimizes the difference between both series, usually a dynamic programming approach is used, evaluating the following recursion:

$$\lambda(i, j) = \delta(a_i, b_j) + \min[\lambda(i-1, j), \lambda(i-1, j-1), \lambda(i, j-1)] \,, \qquad (7)$$

where $\delta(a_i, b_j)$ is the absolute difference between $a_i$ and $b_j$, and $\lambda(i, j)$ is the sum of the difference between both elements and the minimum cumulative difference from the 3 adjacent elements. There are only three possibilities because of the continuity and monotonicity constraints.

The dynamic image warping method, proposed by Hale (2012), is a slightly modified version of the dynamic time warping algorithm, applied to an entire seismic section, dealing with each trace individually. This slightly modified version of *DTW* allows only smoother distortions by limiting the strain over the warping path (stretch or

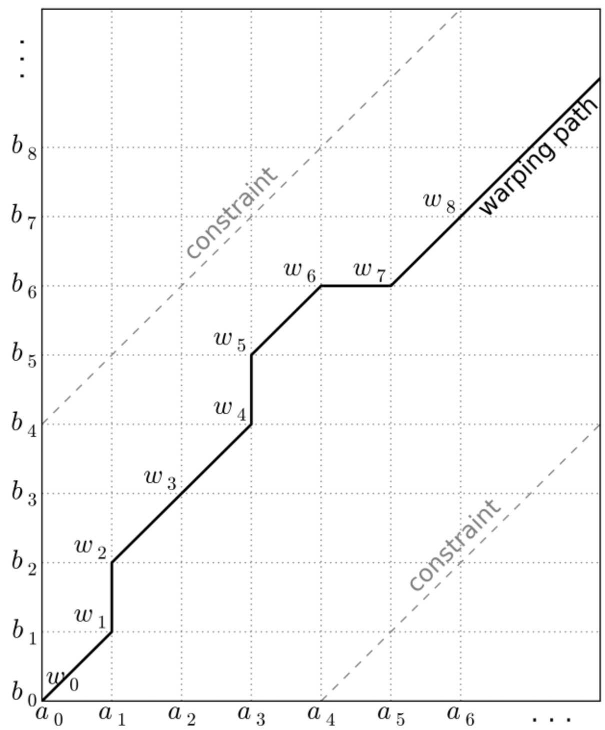


Figure 1: Warping path defined by the smallest absolute difference between series $A$ and $B$.

squeeze). For each horizontal or vertical segment on the warping path we have a stretch or squeeze of 100% on one of the series, which is usually unreasonable. The proposed algorithm uses a ratio factor which imposes that, for each horizontal or vertical distortion, there will have to be $x$ segments without distortion (i.e. diagonal segments). The proposed method also optimizes the process by computing the differences only inside a range of time lags, which is equivalent to compute the differences only in between the constraints shown in Figure 1. The smaller the range of lags, the smaller is the possibility of distortions. This particular feature is really attractive since our series are expected to be quite similar after the first-order corrections made with the velocities ratio $\gamma$. Considering that *DTW* is a purely mathematical method, with no regards for the medium's physical properties, restraining the amount of allowed distortion is also a safer choice.

**Synthetic data examples**

Using real set of well data we have modeled synthetic CDP gathers for both PP and PS data, using Zoeppritz equations. Their CDP stacks are shown in Figure 2, with the same time scale. Notice the reverse polarity of PS data when compared with PP data, which means we will have to revert it's polarity before applying the time warping algorithm.

The $\alpha/\beta$ ratio $\gamma$ used to create the model is displayed in Figure 3, as are the approximations used in the first-order adjustment. These approximations were created using a low pass filter by a truncated series of cosines, one of them using 10 terms, and the other using only 3 terms.

These two approximations are used to simulate the ratio of interval processing velocities, to be used as a first-order adjustment.

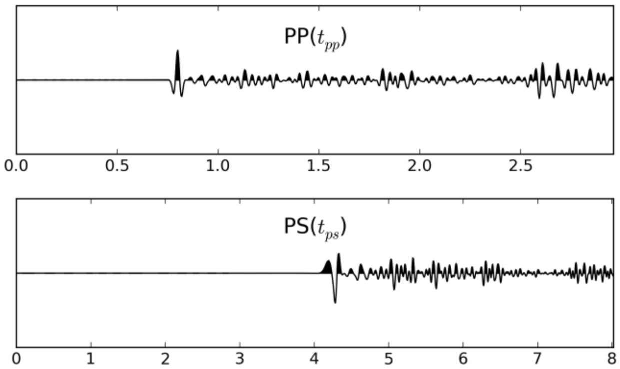


Figure 2: Synthetic PP and PS traces registered in their respective time domains.

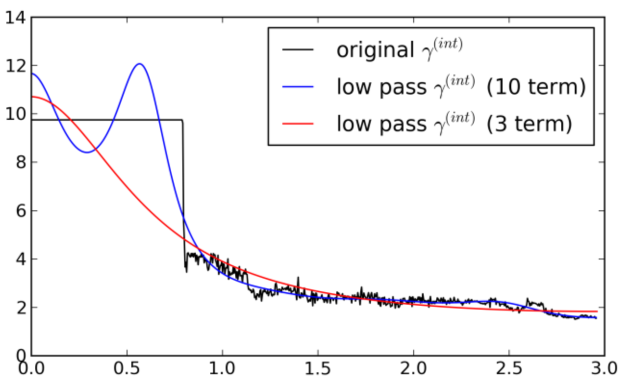


Figure 3: Original $\gamma^{(\text{int})}$ computed from well logs (black), and two low pass filtered versions, one using 10 terms (blue) and the other 3 terms (red), for simulating a $\gamma^{(\text{int})}$ ratio obtained from processing velocities, used for the first-order adjustment.

## Results

Using the models described in the previous section, we first converted the synthetic PS data from PS time domain to PP time domain using the 10 term $\gamma$ (Figure 3), applying it into equation 3. Figure 4 shows the original PP data compared to this first-order approximation of PS data, called PS ($t_{pp}$). Notice the misaligned events shown in detail, as indicated by the red and green dotted lines.

We then applied the time warping algorithm, resulting in the third trace, called PS($w(t_{pp})$). As Figure 4 show, the marked events were then aligned by the warping algorithm. In fact, the correlation coefficient between the two series improved considerably, going from 33.35% between PP and PS($t_{pp}$), to 94.12% between PP and PS($w(t_{pp})$).

Using the 3 term approximation for $\gamma$ to simulate poorer processing velocities, we've obtained the results shown in Figure 5. The time domain conversion did not succeed to bring the corresponding events of the PS data to a time sufficiently close to the ones in PP data. With a poor first-order conversion, the subsequent application of the dynamic time warping method could not align the events correctly.

Considering once again the adjustment made with the 10 term $\gamma$, we inserted different levels of random noise and then have observed the correlation coefficients to drop as the noise level was raised. Figure 6 shows four different situations, each one with a different level of noise. The level of noise represents the percentage of the maximum value of each particular time series. Random values within this noise range were then added to the series to observe how it would affect the correlation coefficient.

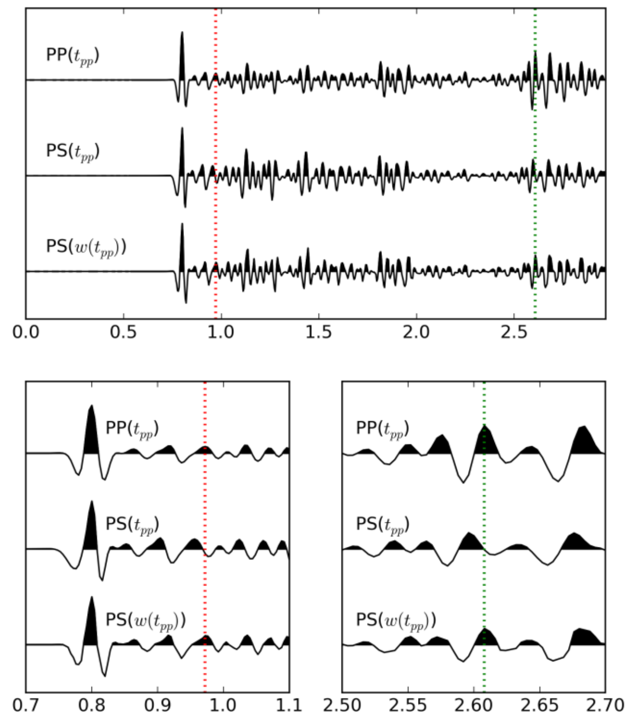


Figure 4: PP and PS data after the first-order adjustment made using the 10 terms approximation of $\gamma$, and then after *DTW*. In the detailed view (Figure 4 bottom) we can see the two events marked in red and green corrected by the warping technique. The polarity of PS data was already reverted to match the PP data polarity.

## Conclusions

The results show that dynamic time warping is a promising technique to perform fine-tuning adjustments in the multicomponent seismic data registration, provided that the first-order PS to PP time conversion using the velocities ratio has already been performed, and the events on both series are sufficiently close to each other. As Figure 5 has shown, a poor first-order approximation will most likely ruin any chance that *DTW* have to correctly align events on both datasets. When trying to use the time warping method directly, before any adjustment with $\gamma$, the results became unpredictable.

Preliminary studies have shown that this problem could be avoided by including an intermediate step of manual event picking, which would be used to tie together horizons according to the interpreter's recognition of corresponding events. This step should be done before

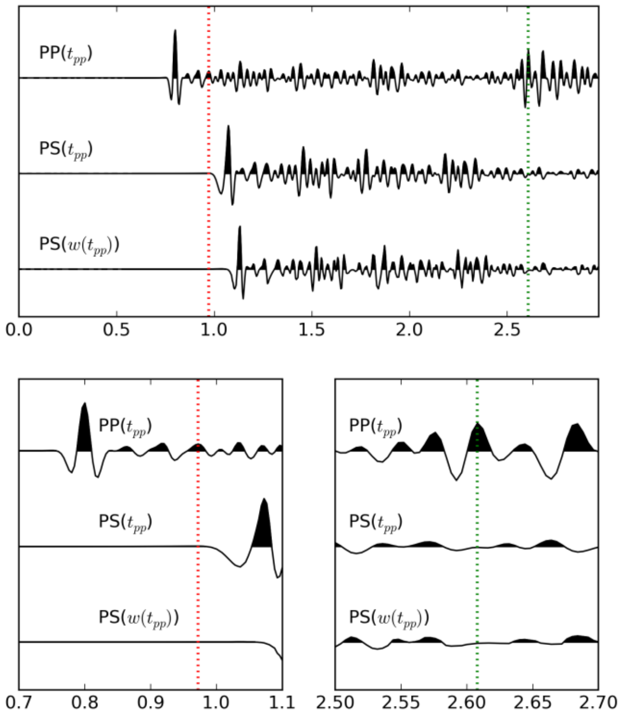


Figure 5: PP and PS data after the first-order adjustment made using the 3 terms approximation of $\gamma$, and then adjusted using *DTW*. Notice the poor first-order adjustment and it's reflection on the following time warping adjustment.

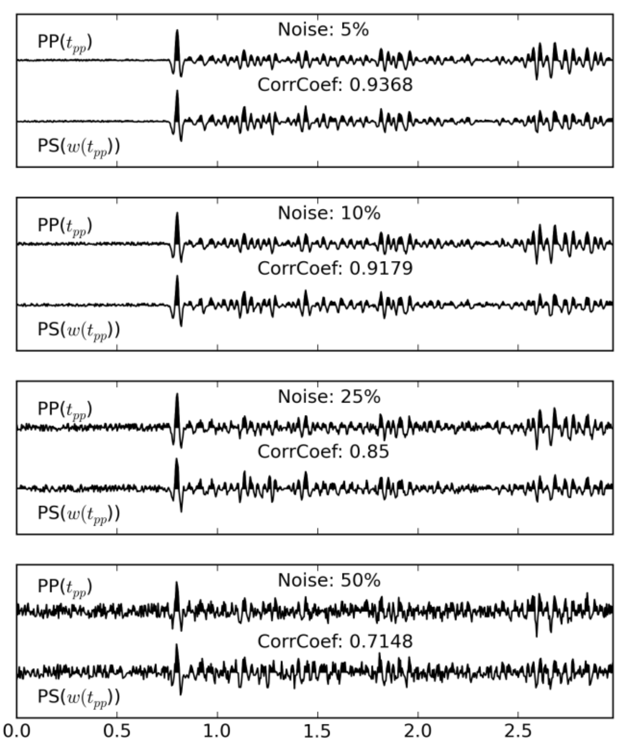


Figure 6: Different noise levels generate different correlation coefficient. As expected, the correlation coefficients get worse as the noise levels increase.

the application of the time warping method, which would be used just as a fine-tuning adjustment tool.

**Acknowledgments**

The authors wish to thank PETROBRAS for financial support and for providing the well data used in this work.

**References**

Al-Chalabi, M., 1994, Seismic velocities - a critique: First Break, December, 12(12), pp. 589-596.

Berndt, D. J. and J. Clifford, 1994, Using Dynamic Time Warping to Find Patterns in Time Series: KDD workshop, pp. 359-370.

Fomel, S., and M. M. Backus, 2003, Multicomponent seismic data registration by least squares: SEG Technical Program Expanded Abstracts, 22(1), pp. 781-784.

Gaiser, J., 1996. Multicomponent correlation analysis: Geophysics, 61(4), pp. 1137-1149.

Hale, D., 2012, Dynamic warping of seismic images: CWP research report, 723, Colorado School of Mines.

Hampson, D. P., B. H. Russell, and B. Bankhead, 2005, Simultaneous inversion of pre-stack seismic data: SEG Technical Program Expanded Abstracts, 24(1), pp. 1633-1637.

Hardage, B. A., M.V. DeAngelo, P. E. Murray, and D. Sava, 2011, Multicomponent Seismic Technology: Society of Exploration Geophysicists.

Keogh, E., and C. A. Ratanamahatana, 2005, Exact indexing of dynamic time warping: Knowledge and Information Systems, Springer London, v. 7, p. 358-386.

Larsen, J. A., 1999, AVO Inversion by Simultaneous P-P and P-S Inversion: M.Sc. thesis, University of Calgary.

Stewart, R. R., 1990, Joint P and P-SV Inversion: CREWES research report, University of Calgary.

Ursenbach, C., P. Cary, and M. Perz, 2013, Limits on resolution enhancement for PS data mapped to PP time: The Leading Edge, 32(1), pp. 64-71.

Wang, Y., 1999, Approximations to the Zoeppritz equations and their use in AVO analysis: Geophysics, 64(6), pp. 1920-1927.