Stabilizing time-shift estimates in coda wave interferometry with the dynamic time warping method

T. Dylan Mikesell, Alison Malcolm and Di Yang, Earth Resources Laboratory, MIT; Matt M. Haney, Alaska Volcano Observatory, USGS

SUMMARY

Accurate time-shift estimation between arrivals in two seismic traces before and after a small velocity change is crucial for estimating the location and amplitude of velocity change. Windowed crosscorrelation and trace stretching are two techniques commonly used to estimate local time shifts in multiply scattered coda signals. These methods both suffer when the induced changes in the scattered wavefield are not simple shifts. Cycle skipping is an example of one such obstacle that neither method is able to overcome. A common approach to mitigate such problems is to choose only part of the coda to analyze. In the work presented here, we implement Dynamic Time Warping (DTW) to search for the time shift at each time sample that globally minimizes the misfit between two seismic traces. We show that DTW is not as susceptible to errors in time-shift estimates caused by cycle skipping or disappearance of coda phases due to changes in scattering. Our approach provides a new tool to estimate small time shifts in coda and has wide application across all disciplines of seismic monitoring with coda waves.

INTRODUCTION

Unraveling the multiply scattered wavefield, commonly called coda, in the Earth is a difficult task. Over decades seismologists have developed different approaches to understand and use the coda signal, e.g., to determine earthquake magnitude (see Chapter 3.2 in Sato et al., 2012, and references therein). Numerous authors have demonstrated the use of coda waves to monitor small velocity changes in different geologic settings (see Poupinet et al., 2008, and references therein). The reason that coda waves are often used is because they travel through the subsurface over larger distances than the ballistic waves. Therefore, coda waves sample the medium much more; in effect making them very sensitive to the velocity structure. The difficulty however, comes in using coda waves to understand the spatial distribution of the velocity heterogeneity.

Researchers across many fields have suggested ways to analyze and interpret complex coda wave signals. In seismology the active doublet (Poupinet et al., 1984; Roberts et al., 1992) and coda wave interferometry (CWI) (Snieder, 2006) methods are approaches to quantify small velocity changes using coda waves. The methods are based on the assumption that a velocity change induces measurable phase shifts in the coda. Pacheco and Snieder (2005) developed a framework for the spatial sensitivity of coda time shifts ($\Delta t$) to velocity perturbations ($\Delta V$) using a diffusion based kernel to represent the coda wavefield with time:

$$\frac{\Delta t}{t} = -K(x_i, x_r, t) \frac{\Delta V}{V},$$

where the kernel ($K$) is a function of the source and receiver positions $x_i$ and $x_r$, respectively. The time shifts can be estimated using windowed crosscorrelation or crosscoherency (e.g. Poupinet et al., 1984; Grêt et al., 2006; Haney et al., 2009) or trace stretching (e.g. Sens-Schönfelder and Wegler, 2006; Hadzioannou et al., 2009), and the kernels can also be derived using radiative transfer theory (e.g. Planès et al., 2014).

If only isolated parts of the medium velocity are perturbed, then only isolated parts of the coda should change (Pacheco and Snieder, 2006). For example, if only a single reservoir velocity changes, then only parts of the coda that contain waves sampling this reservoir will change (e.g. Khatiwada et al., 2012). When the velocity change $\Delta V$ is not homogeneous, windowed crosscorrelation is known to under estimate $\Delta t$, especially at large lag times. This underestimation can be due to the crosscorrelation detecting the lag time that maximizes the correlation within the window, which may contain waves that did not sample the velocity change. Another cause is cycle skipping due to the time shift exceeding the dominant period (e.g. McGuire et al., 2012) or due to changes in the waveform caused by changes in scattering.

Hale (2013) demonstrates the advantages of the Dynamic Time Warping (DTW) method over crosscorrelation to measure variations in time-lapse seismic images. Here we compare the DTW estimated shifts to time shifts estimated by the windowed crosscorrelation method. DTW has been shown to out perform crosscorrelation when time and frequency shifts are nonlinear and strong noise is present (Hale, 2013). Recent examples of DTW use in the seismic exploration industry can be found in well-tie experiments (e.g. Muñoz and Hale, 2012; Herrera and van der Baan, 2012). With this in mind, we investigate the use of DTW to measure small delays in coda wave signals caused by isolated velocity perturbations. Using DTW we can impose constraints on how rapidly shifts vary with time to suppress cycle skipping, but we need not constrain the linearity of these shifts. This is an advantage over current methods.

THEORY

As evident by equation 1, if the $\Delta t$ value is not correctly estimated, the corresponding velocity perturbation $\Delta V$ is not correctly estimated, even if an accurate kernel is available. Dynamic time warping (DTW) is a nonlinear optimization approach (Sakoe and Chiba, 1978) which we use to measure time shifts between coda arrivals before and after a velocity change has occurred in the subsurface. We implement DTW following Hale (2013) and quantify the magnitude of time shifts in coda waves. Below we give a short description of the method; however, we suggest the reader see Müller (2007) or Hale (2013) for a more complete description of the DTW algorithm and the underlying details and assumptions.

Consider the seismogram $g(t_i)$, where $t_i$ is the time sample in-
A NUMERICAL EXPERIMENT

We investigate the use of DTW to estimate time shifts in the coda using a 2-dimensional (2D) acoustic model with a random Gaussian slowness distribution. We use the 2D background velocity model shown in Figure 1(a). The inner circle velocity is 6 km/s and remains constant throughout all of the experiments. We vary the outer circle velocity from -3% to +3% of the inner circle velocity at 1% increments (i.e., ∆V = -180, -120, -60, 0, +60, +120, +180 m/s, respectively). In Figure 1(a) we show the +1% velocity model. The model is 32 km × 32 km with absorbing boundaries. The grid spacing is 20 m in the simulations.

We use the SeisUnix 2D finite-difference code *sufdmod2* to model the acoustic wavefield for 10 s after the source impact. The source is a 10 Hz (center frequency) Ricker wavelet; therefore, at 6 km/s the dominant wavelength is 600 m. After simulation, we resample the data to an interval of 1 ms. The receiver (gray reverse triangle in Figure 1(a)) and source (gray star in Figure 1(b)) are co-located at the center of model. To suppress reflections from the inner-outer boundary we taper the background velocity model from the inner to the outer circle. The taper width is 2 km and starts 6 km from the center of the model.

To create the multiply scattered coda wavefield, we generate a random Gaussian slowness distribution and overlay this on the various background velocity models. The random slowness model is shown in Figure 1(b) as velocity. The average size of the perturbations are on the order of 400 m. For each of the various background velocity models we show the individual waveforms in Figure 2(a). The wiggle traces are underlain by the grey boxes in certain places to help visualize the wavefield changes in the different models. For instance, looking at arrivals around 8 and 9 s, we see the appearance or disappearance of coda arrivals, as well as changes in the frequency of arrivals.

TIME-SHIFT ESTIMATION

We compare the DTW method to the windowed crosscorrelation method. The window length is 0.8 s – roughly 8T, where T = 0.1 s is the dominant period. For the reference trace, to which we compare all other traces, we use the model that does not differ across the inner and outer boundary (i.e., ∆V = 0, gray trace in Figure 2(a)). For each of the different traces in Figure 2(a), we slide the crosscorrelation window 1 sample at a time, recording the maximum correlation coefficient and the lag at that maximum correlation. The correlation coefficients and lag times are plotted in Figure 2(b) and 2(c), respectively.

For each model, the inner background velocity does not change. Therefore, we do not expect any change in the waveforms until after the time when waves could have reached the inner-outer boundary and returned to the receiver. This time is approximately 2.5 s and the correlation coefficients in Figure 2(b) demonstrate this; the traces are identical up to ~ 2.5 s, after which the correlation coefficient begins to decrease. Note that the models with larger velocity perturbations start to lose correlation more rapidly than those with smaller velocity perturbations.

The time shifts estimated by the crosscorrelation method in Figure 2(c) also show that ∆t = 0 before 2.5 s and starts to decrease or increase more-or-less linearly until the first cycle skip occurs. The first cycle skip time is different for each trace, but the skipping begins when ∆t approaches 1 period, T = 0.1 s. Therefore, the skipping begins earlier for the larger velocity perturbation models (e.g., the blue line in Figure 2(c) for ∆V = -180 m/s).

Once cycle skipping occurs, the correlation coefficients start to decay and oscillate, and it is difficult to trust the ∆t estimates based correlation coefficients, even as they overcome the cycle skipping at later times. It is worth noting that when the crosscorrelation window is 8T, the cycle skipping remains a problem and only worsens as the window size decreases. This is because the correlation function only identifies the local lag time between the two windows and is sensitive to changes in scattering (e.g., Larose et al., 2010). We can directly see this sensitivity by considering the ∆t estimates around 8 s for the ∆V = 180 m/s trace. We see that the arrival just after 8 s in the ∆V = 0 trace disappears in the ∆V = 180 m/s trace. This causes a cycle skip in the ∆t estimate (Figure 2(c)).

We apply DTW to the same reference and perturbed traces to estimate time shifts (Figure 2(d)). The maximum possible lag is 0.5 s and the lag interval is 0.1 ms. Right away we see that cycle skipping is no longer a problem. This is due to search-
Dynamic time warping does not impose constraints on the linearity of shifts, only on the slope of the $\Delta t$ function. Therefore, DTW has the ability to more accurately represent the time variant $\Delta t$ as demonstrated in Figure 2(d).

**CONCLUSION**

We apply the Dynamic Time Warping (DTW) method to two seismic traces in order to estimate time shifts in the multiply scattered coda wavefield. These time shifts are caused by perturbations in the background velocity of a random Gaussian slowness model. We compare the DTW method to the often used windowed crosscorrelation method. The velocity model perturbations range from -3% to +3% at a ±1% interval. We show that the DTW method is considerably more stable than crosscorrelation, even when the correlation window length is 8 times the period. In all of the tested velocity models, the DTW $\Delta t$ estimate does not cycle skip when we set the lag interval to one-tenth the sample interval. Furthermore, the DTW estimate shows more fine scale structure than the crosscorrelation estimate. This is because the crosscorrelation averages over the entire window. This new application has the potential for widespread use across all disciplines using coda waves to monitor changes in velocity over time.

**ACKNOWLEDGMENTS**

We are grateful to Oleg Poliannikov for providing the code to generate the random Gaussian slowness model, as well as the developers of SeisUnix, which we used for the finite difference modeling in this work. T. D. M. thanks Kasper van Wijk and Guust Nolet for useful discussions about changes in coda waves and acknowledges financial support from the National Science Foundation under Award No. 1144883.
Figure 2: (a) Seismic wiggle traces; the large amplitude direct wave arrival ($t < 0.25$ s) is muted because it has significantly larger amplitude than the scattered waves. Underlying grey boxes indicate notable changes in the coda waveforms, not just time shifts. From bottom to top, the outer velocity for each trace changes by -180, -120, -60, 0, +60, +120, & +180 m/s, respectively. The middle gray trace ($\Delta V = 0$) is the reference trace to which all other traces are correlated or warped. (b) Correlation coefficients from the sliding window crosscorrelation method; the window length is 0.8 s. (c) $\Delta t$ estimates from sliding window crosscorrelation. (d) $\Delta t$ estimates from dynamic time warping. The maximum lag is set to 0.5 s and the lag interval is 0.1 ms.
REFERENCES


