Multi-taper measurements for quantifying differences between two time series

Vala Hjörleifsdóttir, Jeroen Tromp
Seismological Laboratory, California Institute of Technology, Pasadena, California 91125, USA
October 24, 2006 (last compilation October 26, 2009)

`This file : /home/carltape/latex/notes/tomo/multitaper_vala.tex`

Summary

Some notes on multi-taper measurements. Based on *Laske and Masters* (1996) and *Zhou et al.* (2004). Also includes notes on codes on /opt/seismo-util that do multi-taper measurements. Another useful discussion of mtm-measurements can be found in *Percival and Walden* (1993), page 333–347.

1 Estimating the differences between two time series

In this note we wish to quantify the differences between two time series. We start with a data trace $d(t)$ and a corresponding synthetic trace $s(t)$. We want to compare individual phases in the seismograms, more specifically, we want to estimate the time-shift and amplitude ratio between the traces, as a function of frequency, within a given time window. We quantify the difference between the data and synthetics in terms of a transfer function, $T(f)$, that satisfies

$$[d(f) - T(f)s(f)]^2 = \text{minimum} \tag{1}$$

In this case the solution is just:

$$T(f) = \frac{d(f)}{s(f)} \,. \tag{2}$$

Writing the synthetic as $s(f) = A(f)e^{-i\tau(f)}$ and the data as $d(f) = [A(f) + \delta A(f)]e^{-i[\tau(f)+\delta\tau(f)]}$ we can write:

$$d(f) = T(f)s(f) = s(f)[1 + \delta A(f)/A(f)]e^{-i\delta\tau(f)} \tag{3}$$

with $T(f) = [1 + \delta \ln A]e^{-i\delta\tau(f)}$. In the Born approximation $d(f) = s(f) + \delta s(f)$, and thus $T(f) = 1 + \delta T(f)$ where $\delta T(f) = \delta s(f)/s(f)$. For small $\tau(f)$ we have $e^{-i\delta\tau(f)} \approx 1 - i\delta\tau(f)$. Correct to first order in small perturbations:

$$\delta T(f) = T(f) - 1 = [1 + \delta \ln A(f)][1 - i\delta\tau(f)] - 1 \approx \delta \ln A(f) - i\delta\tau(f) \tag{4}$$

and finally

$$\delta\tau(f) = -\mathrm{Im}\left(\frac{\delta s}{s}\right), \quad \delta \ln A(f) = \mathrm{Re}\left(\frac{\delta s}{s}\right) \tag{5}$$

1.1 Multi-taper measurements

In the discussion above we didn't specify the type of window to use. Care has to be taken when windowing though, as the type of window can affect the measurement, due to spectral leakage. This is a well known problem in signal processing, as windowing in the time domain corresponds to convolution of the fourier transform of the windowing function in the frequency domain. Denoting the window in the time-domain by $h(t)$ and in the frequency domain by $h(f)$ the windowed data becomes

$$d_w(t) = h(t)d(t) \tag{6}$$

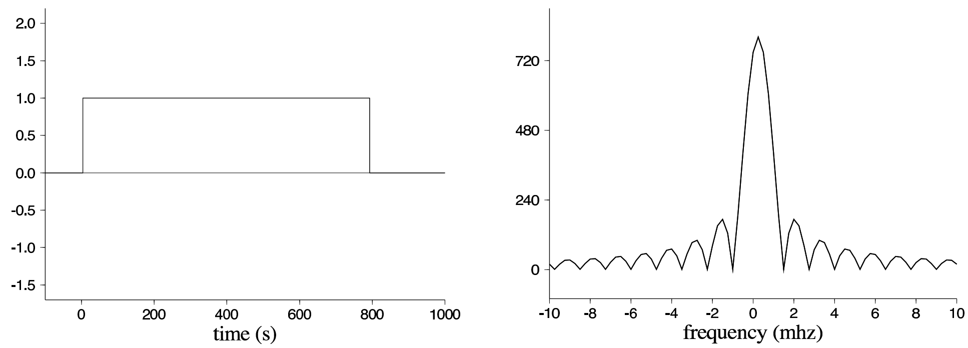

Figure 1: The time and frequency versions of a boxcar.

or in the frequency domain

$$d_w(f) = h(f) \otimes d(f) \tag{7}$$

where $\otimes$ denotes convolution. To get an accurate estimate $d_w(f)$ of $d(f)$ we want $h(f)$ to be as close to a delta function as possible. If we choose $h(t)$ to be a box-car, then $h(f)$ is a sinc-function. When we convolve $d(f)$ with the sinc-function, its side-lobes cause the spectral values away from f to contribute to the estimated value at f. This example suggests we should choose a windowing function without side lobes in the frequency domain, such as a cosine taper. This choice reduces the spectral leakage, but it creates a new problem, as the signal in the middle of the window is weighted more heavily than the signal at the edges, biasing the measurement. In order to minimize the spectral leakage, while keeping the bias at a minimum, we use a multi-taper measurement technique (*Thomson*, 1982).This technique uses several tapers, $h_j(t)$, all concentrated within a small window in the frequency domain, without sidelobes, to window the data and the synthetics. We denote the data windowed by the jth taper by $d_j(f)$ and the windowed synthetic by $s_j(f)$. The basic idea is that even though the spectra from each of the windowed traces is biased, then by using orthogonal tapers and averaging the spectra, one gets a less biased final spectra. An added benefit is that since we get several estimates for each spectral measurement we can compute the error in the estimate, in addition to the average.

1.2 Prolate Spheroidal Eigentapers

Now we focus our efforts on finding the ideal tapers, that have a compact support in the frequency domain while still sampling a large part of the trace in the time domain. Suppose we have chosen a measurement window with width

$$L = N\Delta t, \tag{8}$$

where L denotes the length of the time window, Δt the sampling rate, and N the number of time samples contained in the window. The Rayleigh frequency is then

$$f_R = \frac{1}{L} = \frac{1}{N\Delta t}\,. \tag{9}$$

This is the lowest frequency, i.e., the longest period, that we can hope to resolve with a window length L. The highest frequency, i.e., shortest period, that we can resolve is determined by the Nyquist frequency:

$$f_c = \frac{1}{2\Delta t}\,. \tag{10}$$

The frequency content of our time window $[0, L]$ lies between $[-f_c, f_c]$. The frequency spacing Δf is equal to the Rayleigh frequency:

$$\Delta f = \frac{2f_c}{N} = \frac{1}{N\Delta t} = f_R. \tag{11}$$

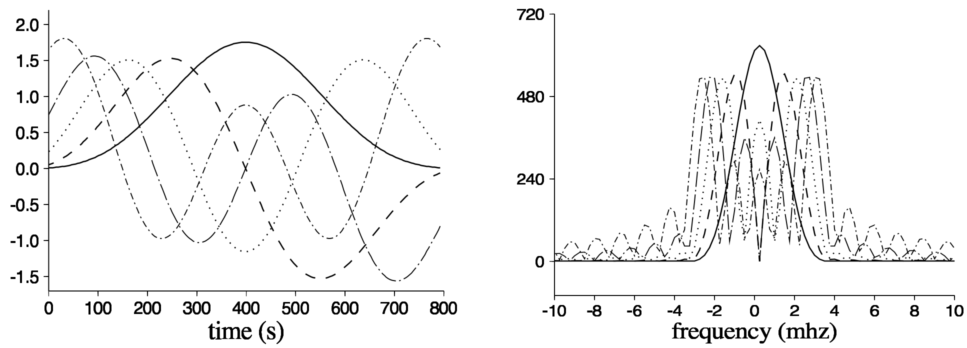

Figure 2: The first five 2.5 pi tapers.

Our objective is to find to find functions (tapers) that are optimally concentrated within the window W in the frequency domain. For convenience we define the window width in terms of the Rayleigh frequency f_R, such that $W = kf_R$. Now our objective is accomplished by optimizing the quantity

$$\lambda = \frac{\int_{-W}^{W} |h(f)|^2 \, \mathrm{d}f}{\int_{-f_c}^{f_c} |h(f)|^2 \, \mathrm{d}f}. \tag{12}$$

This leads to an eigenvalue problem with eigenvalues λ_j and associated eigenfunctions ('prolate multi-tapers') $h_j(f)$ (*Slepian*, 1978). A remarkable property of the eigenvalues λ_j is that the first $2k = 2LW$ values are ~ 1, and the remaining eigenvalues quickly drop off to zero. The implication is that only the first $2k$ eigentapers are optimally concentrated in the window W. So for small k the window is narrow, and for large k it is wider. Similarly, for a long time window L the frequency window W is narrower, and for a short time window L the frequency window W is wider. Effectively you are choosing the width W around the target frequency of interest over which you are going to average the measurement. Frequently the tapers are referred to in terms of their k in the form "k π tapers" ('π' for prolate).

Next, one uses the multi-tapers as windowing functions. Suppose we have a time series $s(t)$ with a corresponding spectrum $s(f)$. Now, rather than working with the time series directly, we multiply it by the $2k$ multi-tapers to get $2k$ versions of the time series:

$$s_j(t) = h_j(t)s(t), \qquad j = 1, \ldots, 2k. \tag{13}$$

In the frequency domain, this corresponds to a convolution with the frequency-version of the taper, leading to $2k$ spectral estimates:

$$s_j(f) = s(f) \otimes h_j(f) = \int_{-f_c}^{f_c} s(f')h_j(f - f') \, \mathrm{d}f', \qquad j = 1, \ldots, 2k. \tag{14}$$

Here one can really see how the spectrum $s(f')$ is convolved with the taper centered on f', $h_j(f - f')$. So the wider the bandwidth of h_j, i.e., the wider W, the more we average over neighboring frequencies. In the limit $k \to 0$ we get a delta function, which corresponds to a boxcar taper in the time domain.

1.3 Transfer function for multitaper measurements

Now that we have made a choice of windowing function, we can go back to estimating the transfer function $T(f)$. In this case we wish to find $T(f)$ such that:

$$||\mathbf{d}(f) - T(f)\mathbf{s}(f)||^2 = \text{minimum} \tag{15}$$

where $\mathbf{d}(f) = [d_1(f), ..., d_j(f), ..., d_{2k}(f)]^T$ and $d_j(f) = d(f) \otimes h_j(f)$. The solution is given by $\mathbf{s}^T[\mathbf{d}(f) - T(f)\mathbf{s}(f)] = 0$. Now we have $T(f) = \mathbf{s}^T\mathbf{d}/\mathbf{s}^T\mathbf{s} = \mathbf{s}^T(\mathbf{s} + \delta\mathbf{s})/\mathbf{s}^T\mathbf{s}$ or:

$$T(f) = \frac{\sum_{j=1}^{2k} d_j(f)s_j^*(f)}{\sum_{j=1}^{2k} s_j(f)s_j^*(f)} = 1 + \delta T(f) \tag{16}$$

where

$$\delta T(f) = \frac{\sum_{j=1}^{2k} \delta s_j(f)s_j^*(f)}{\sum_{j=1}^{2k} s_j(f)s_j^*(f)} \tag{17}$$

Rembembering eq. 4 we get the expressions:

$$\delta\tau(f) = -\mathrm{Im}\left[\frac{\sum_{j=1}^{2k} \delta s_j(f)s_j^*(f)}{\sum_{j=1}^{2k} s_j(f)s_j^*(f)}\right], \quad \delta\ln A(f) = \mathrm{Re}\left[\frac{\sum_{j=1}^{2k} \delta s_j(f)s_j^*(f)}{\sum_{j=1}^{2k} s_j(f)s_j^*(f)}\right] \tag{18}$$

1.4 Effect of taper parameters

The main parameter controlling the behavior of the taper is k. The following figures illustrate the effect of k: We use the first five 2.5π prolate spheroidal tapers, to estimate the spectra. These spectra of these tapers are localized within $2.5/L$ where L is the length of the time-series. Remembering equation (7) this leads to independent estimates of the true spectra every $2.5/L$ Hz. For a window-length of 800 s this corresponds to independent estimates every $2.5/800 = 0.003125$ Hz.

1.5 Combining measurements

Each multi-taper measurement gives us an estimate of the time shift, $\delta\tau_i(f)$, and the amplitude anomaly, $\delta\ln A_i(f)$ at station i and frequency f. This provides us with a slue of measurements for each run. In order to visualize the results we combine the measurements, either integrating over all frequencies at a given station and see the variation with receiver location or summing all the measurements at a given frequency over stations to see the variation with frequency. We define the average time shift at a given frequency as:

$$< \delta\tau(f) >^2 = \frac{1}{N}\sum_{i=1}^{N} |\delta\tau_i(f)|^2 \tag{19}$$

and the average time-shift at station i:

$$< \delta\tau_i >^2 = \frac{1}{f_1 - f_0}\int_{f_0}^{f_1} |\delta\tau_i(f)|^2 df \tag{20}$$

finally, the average over all measurements is given by:

$$< \delta\tau >^2 = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{f_1 - f_0}\int_{f_0}^{f_1} |\delta\tau_i(f)|^2 df \tag{21}$$

For the amplitudes, we define $< \delta\ln A(f) >$, $< \delta\ln A_i >$ and $< \delta\ln A >$ in the same manner.

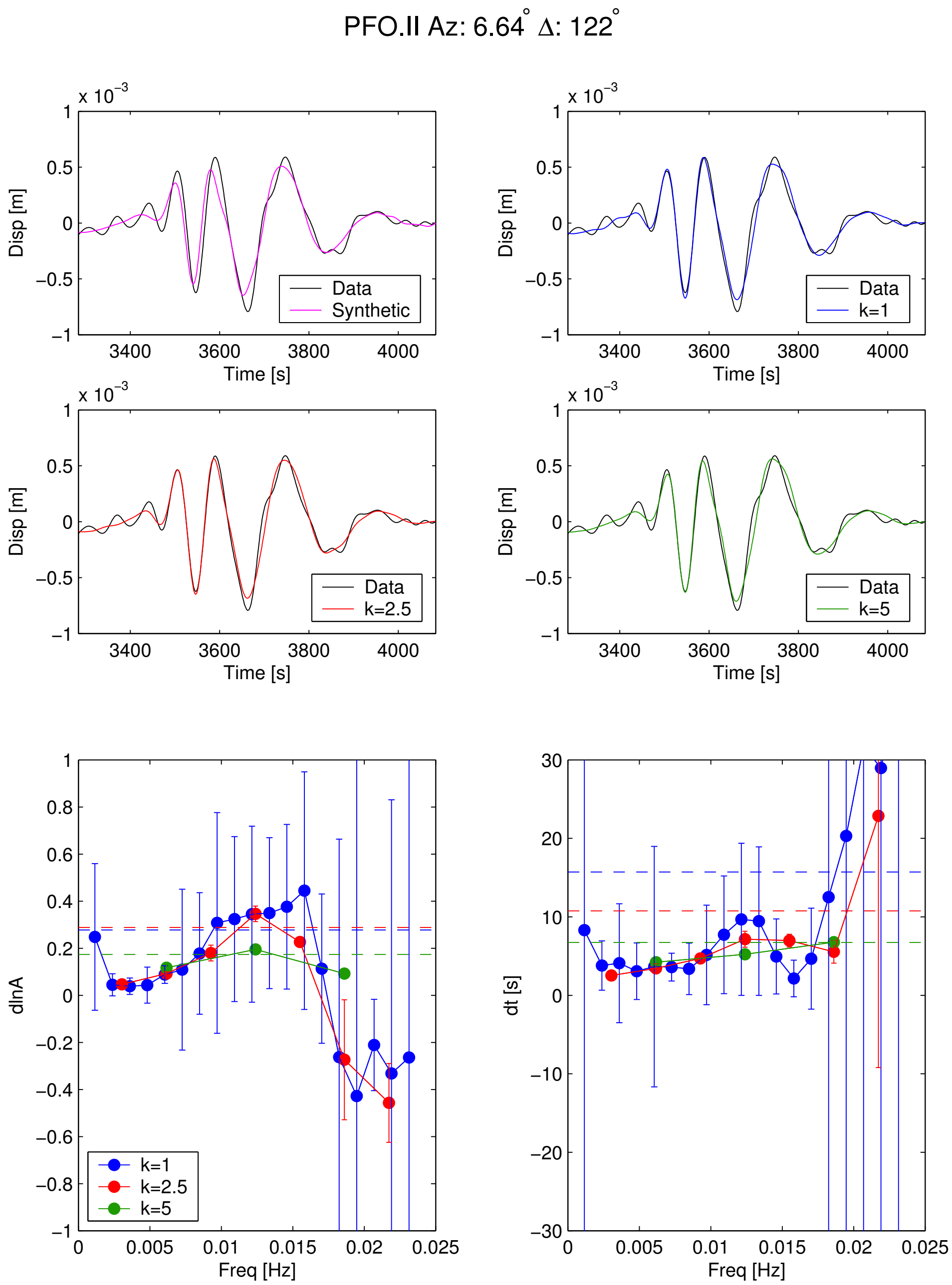

Figure 3: The effect of changing k for a fixed window length. The observed and synthetic traces are shown at the top left, data is black and synthetic is magenta. The other waveform plots compare the data to the reconstructed seismograms for various k. The The multitaper measurements, with errorbars, are shown on the left, the square of the measurements is shown on the right. The colored lines represent measurements made using different k. The number of tapers is set to $2k$. We only show measurements that are "independent", or k/L apart in the frequency domain. Notice how small values of k give many measurements with high variance, and large values of k give few measurements (in the case of k=5, only three measurements), but small variance. Fig. 4 shows the same, for k=2,2.5,3.

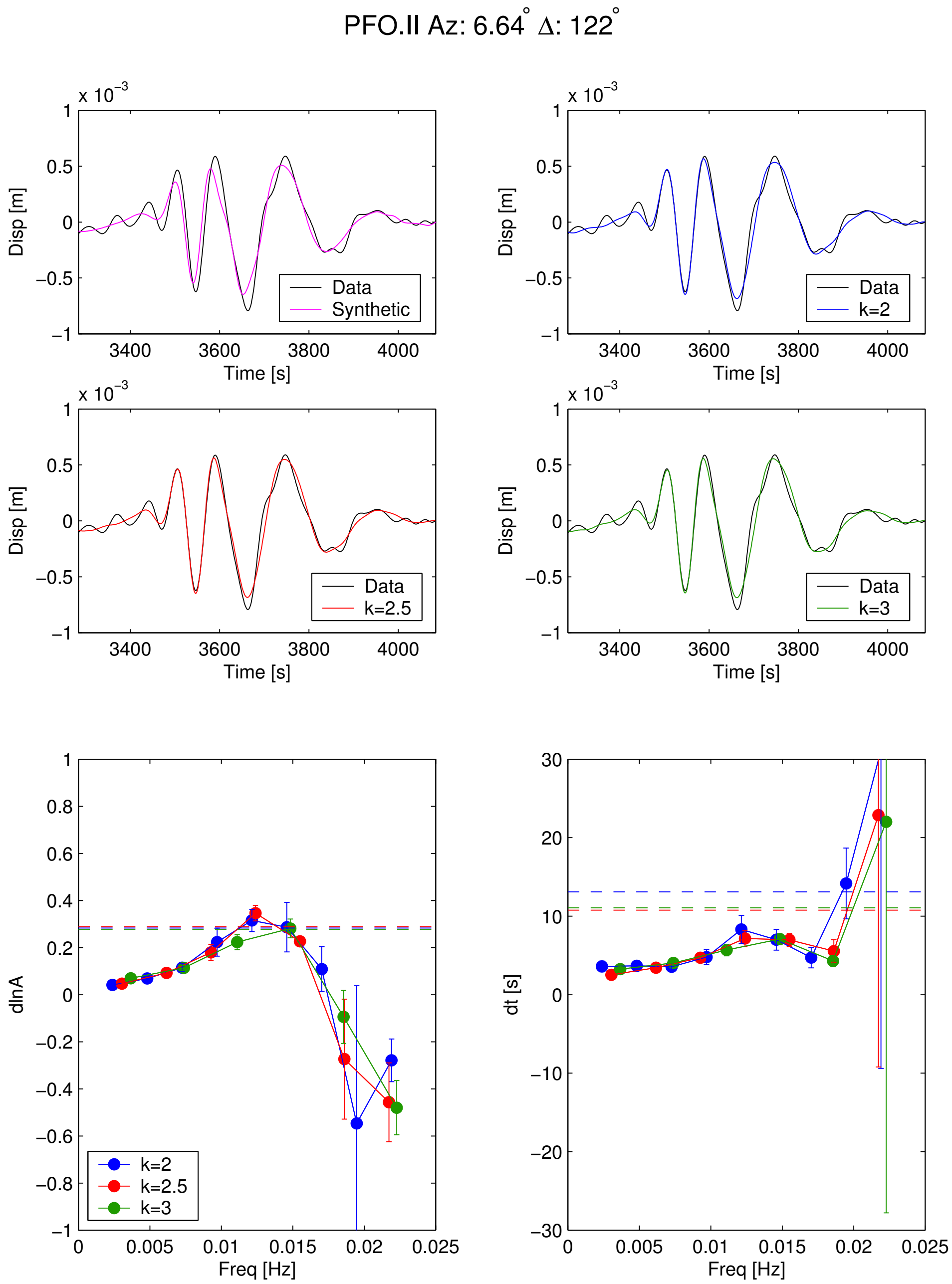

Figure 4: The effect of changing k for a fixed window length. Same as Fig. 3 except other values of k.

2 Programs

In /opt/seismo-util/bin there are several programs to simplify making multitaper measurements. For the perl-scripts, simply type the name of the code and it will give you a usage statement.

2.1 mtm-measure

The fortran code from Ying Zhou (modified by Alessia Maggi, Carl Tape and myself). Reads from standard input (run the code and it will ask you for the parameters). The output is in several files:

obs.win The part of the observed seismogram that is used for the measurement (sac-file).

syn.win The part of the synthetic that is used for the measurements (sac-file).

obs.win.new The reconstructed seismogram (sac-file).

trans.am.mtm Amplitude ratio as a function of frequency (two columns).

trans.dlnA.mtm Same, but dlnA.

trans.dt.mtm Time shift as a function of frequency.

trans.ph.cor.mtm Phase shift as a function of frequency.

err.am, err.dlnA, err.dt,err.ph Standard deviation of amplitude ratio, dlnA, time shift and phase shift respectively.

rh_quality.dat Before and after quality check. Outputs F1, F2, cc-max and dlnA (see top of program for definitions)

The code uses a jack-knife routine to estimate the errors.

2.2 mtm.pl

This script is a wrapper for mtm-measure.

```
mtm.pl -b -s syn_dir,syn_suffix -o outputdir -c cut0/cut1
-w wave -m ntapers/waterlevel/npi -V sac_files
```

2.3 plot_mtm.pl

Plots dlnA and dt for individual pairs of data and synthetics, given the output from mtm.pl. Also plots the output from the quality check.

```
plot_mtm.pl mtmfiles.dlnA
```

2.4 combine_mtm_measurements.pl

Combines the output from several measurements into averages over frequency for a given station, or averages over all stations for a given frequency. Also gives the "average" misfit for the whole dataset. Accepts three types of averages, as described above ("regular", "χ^2" or "weighted average").

```
combine_mtm.pl -i mtmfiledir -T min_per -X -A r|x|w
```

References

Laske, G., and G. Masters (1996), Constraints on global phase velocity maps from long-period polarization data, *J. Geophys. Res.*, *101*(B7), 16,059–16,075.

Percival, D., and A. Walden (1993), *Spectral Analysis for Physical Applications*, Cambridge U. Press, Cambridge, UK.

Slepian, D. (1978), Prolate spheroidal wave functions, Fourier analysis, and uncertainty. V: The discrete case, *Bell Syst. Tech.*, *57*, 1371–1430.

Thomson, D. J. (1982), Spectrum estimation and harmonic analysis, *IEEE Proc.*, *70*, 1055–1096.

Zhou, Y., F. A. Dahlen, and G. Nolet (2004), Three-dimensional sensitivity kernels for surface wave observables, *Geophys. J. Int.*, *158*, 142–168.