Lecture 19:

Recap

Key themes from last time focused on multi-taper methods for computing spectra when identifying peaks and energy levels in the peaks is important. We talked about definitions of decibels. And we considered the transfer function.

Transfer functions: a proper example

Last time I gave you a very fleeting and inadequate view of a transfer function example, so I want to work through this more carefully today, before taking a more holistic view at a case study.

Salinity spiking examples

The slides show some examples of 'salinity spiking', which results from the temperature and conductivity sensors having different response times. If we assume that simultaneous measurements represent the same water sample, when in reality they don't, we end up computing:

$$S = S(T(t), C(t - \Delta t)). \tag{1}$$

If temperature and salinity are constant, this doesn't pose a problem, but when T or C change abruptly, this can lead to very odd results. How can we use the transfer function (or coherence) approach to examine this (even if we don't know the actual response characteristics of the sensor). Here's a basic procedure.

1. Identify a segment of the water column in which $\hat{T}$ and $\hat{C}$ (the measured T and C) should be well behaved, with fluctations due to temperature only.

2. Collect a lot of profiles of data.

3. Now treat this as a linear system:

$$\hat{T}(k) = H_T(k)T(k) \tag{2}$$
$$\hat{C}(k) = H_C(k)T(k) \tag{3}$$
$$\tag{4}$$

   Here $H_T(k)$ is the spatial/frequency response of the temperature sensor and $H_C(k)$ is the spatial/frequency response of the conductivity sensor

4. Compute cross spectra:

$$\hat{G}_{TT}(k) = \frac{\langle \hat{T}^*(k)\hat{T}(k)\rangle}{\Delta k} \tag{5}$$

$$\hat{G}_{CC}(k) = \frac{\langle \hat{C}^*(k)\hat{C}(k)\rangle}{\Delta k} \tag{6}$$

$$\hat{G}_{TC}(k) = \frac{\langle \hat{T}^*(k)\hat{C}(k)\rangle}{\Delta k} \tag{7}$$

$$\tag{8}$$

Then if we substitute in the expessions linking the observed values to the true values we obtain:

$$\frac{\langle \hat{T}^*(k)\hat{C}(k)\rangle}{\Delta k} = \frac{\langle (H_T(k)T(k))^* H_C(k)T(k)\rangle}{\Delta k} \tag{9}$$

$$= [H_T^*(k)H_C(k)]\frac{\langle T^*(k)T(k)\rangle}{\Delta k} \tag{10}$$

$$\hat{G}_{TC}(k) = [H_T^*(k)H_C(k)]G_{TT}(k) \tag{11}$$

The same applies for the temperature spectrum:

$$\hat{G}_{TT}(k) = \frac{\langle \hat{T}^*(k)\hat{T}(k)\rangle}{\Delta k} \tag{12}$$

$$= [H_T^*(k)H_T(k)]G_{TT}(k) \tag{13}$$

So the ratio of these becomes:

$$\frac{\hat{G}_{TC}(k)}{\hat{G}_{TT}(k)} = \frac{H_T^*(k)H_C(k)}{|H_T|^2} = \frac{H_C(k)}{H_T(k)} \tag{14}$$

This means that even without knowing the response function H, we can compute the ratio of the response functions from the transfer function, the ratio of the cross-spectrum to the spectrum.

An analogous relationship also holds:

$$\frac{\hat{G}_{TC}(k)}{\hat{G}_{CC}(k)} = \frac{H_T^*(k)H_C(k)}{|H_C|^2} = \frac{H_T(k)}{H_C(k)} \tag{15}$$

5. Now use this information to correct the conductivity sensor to have the same response as the temperature sensor. Here we'll define our corrected conductivity as $\hat{\hat{C}}(k)$, and we want to understand its relationship with the observed temperature $\hat{T}(k)$ and the true temperature $T(k)$.

$$\hat{\hat{C}}(k) = \alpha\hat{T}(k) = \alpha H_T(k)T(k) \tag{16}$$

This means we need a correction of the form:

$$\hat{\hat{C}}(k) = \hat{C}(k) \cdot P(k) = \alpha\hat{T}(k), \tag{17}$$

and our task is to figure out $P(k)$. We can also write:

$$\hat{C}(k) \cdot P(k) = H_C(k)T(k)P(k) \tag{18}$$

so putting this together:

$$\alpha H_T(k)T(k) = H_C(k)T(k)P(k) \tag{19}$$

Thus

$$P(k) = \frac{\alpha H_T(k)}{H_C(k)} = \alpha\frac{\hat{G}_{TC}(k)}{\hat{G}_{CC}(k)} \tag{20}$$

where α is real. So we do a bit of curve fitting to optimize our correction.

A typical correction might allow for errors both in the response time and a direct time lag:

$$\tau \frac{d\hat{T}(t)}{dt} + \hat{T}(t) = T(t - L) \tag{21}$$

(from Giles and McDougall, Deep-Sea Research, 1986) and this suggests corrections both in the frequency and time domain, either by minimizing phase differences or by maximizing correlation.

Last week we looked at noise in coherence, but let's take a quick look at what noise does to the transfer function. Suppose:

$$Y(\sigma) = X(\sigma)H(\sigma) + N(\sigma) \tag{22}$$

If we multiply through by X^*:

$$\langle X^*Y \rangle = \langle X^*X \rangle H(\sigma) + \langle X^*N \rangle \tag{23}$$

Since the signal is uncorrelated with noise, this still gives us

$$H(\sigma) = \frac{G_{xy}(\sigma)}{/} G_{xx}(\sigma), \tag{24}$$

so noise appears to have no impact on the results, but is it all so rosy?

Alternatively, we might imagine that our forcing x is noisy, so that

$$Y(\sigma) = [X(\sigma) + N(\sigma)] H(\sigma) \tag{25}$$

Then:

$$\langle (X + N)^*Y \rangle = \langle (X + N)^*(X + N) \rangle H(\sigma) \tag{26}$$

And

$$\hat{H}(\sigma) = \frac{G_{xy}(\sigma)}{/} G_{xx}(\sigma) + G_{nn}(\sigma), \tag{27}$$

which is biased relative to the true response. This is analogou to the noise-related formulation that we looked at for coherence.

Putting it all together

In this course, we've looked at a broad range of strategies for analyzing time series. Can we make some decisions about how we might plan an experiment and analyze our data?

Let's consider a central problem of physical oceanography. Can we evaluate the ocean response to wind:

$$\frac{\partial u}{\partial t} + u \cdot \nabla \mathbf{u} - fv = -\frac{1}{\rho}\frac{\partial p}{\partial x} + \frac{1}{\rho}\frac{\partial \tau^x}{\partial z} \tag{28}$$

$$\frac{\partial v}{\partial t} + v \cdot \nabla \mathbf{u} + fu = -\frac{1}{\rho}\frac{\partial p}{\partial y} + \frac{1}{\rho}\frac{\partial \tau^y}{\partial z} \tag{29}$$

Suppose we cross out a few terms. What do we need to measure to evaluate whether we've crossed out the right terms? Can we show that one of these is a reasonable approximation?

$$\frac{\partial u}{\partial t} = \frac{1}{\rho}\frac{\partial \tau^x}{\partial z} \tag{30}$$

$$-fv = \frac{1}{\rho}\frac{\partial \tau^x}{\partial z} \tag{31}$$

$$-fv + -\frac{1}{\rho}\frac{\partial p}{\partial x} = \frac{1}{\rho}\frac{\partial \tau^x}{\partial z} \tag{32}$$

$$\tag{33}$$

Methods we've explored in this class include:

1. Basic statistics: means, standard deviation, variance, standard error

2. Probability density function

3. Least-squares fitting

4. One-dimensional spectra (with windowing and uncertainties)

5. Two-dimensional spectra

6. Monte Carlo methods for evaluating confidence limits

7. Coherence

Intangibles

Besides the formal topics for which you've done problem sets, this class has aimed to start you thinking more like a data analyst. This has thrown you into the thorny world of real data problems, and I'm immensely grateful to you for your persistence. Some life lessons from this class:

1. In science, we favor evidence-based decision making over shoot-from-the-hip opinons. Data analysis gives you a set of tools for this.

2. Your good judgement matters in deciding how to approach a data analysis problem. You should always ask yourself how your understanding of the physics can inform your approach.

3. Even the rigors of the peer review process cannot guarantee the fidelity of published sources of information. Be skeptical and inquisitive.

4. You have the tools at your disposal to address your skepticism. Fake data and Monte Carlo methods are always an option.

5. Many questions have not been answered carefully, and there is room for you to make significant contributions. (I think the pier station staff are still looking for the perfect means to validate the methodology used for the pier samples.)

6. Just because something appears to be significant at the 95% level doesn't guarantee that it is a robust signal.

7. If your results are wildly dependent on the details of your methodology, that might mean that paying attention to methodology matters, but it also could be a warning sign that you're trying to identify a signal that is more wishful thinking than real signal.

8. The methods that we apply to real data can be exactly analogous to problems that we've done in this class, or they can be surprisingly divergent.