Stat 565

Bivariate Series

March 1 2016

Charlotte Wickham

stat565.cwick.co.nz

Announcements

Project Proposals all look good!
Presentations?

Thursday this week

Tuesday next week

Thursday next week

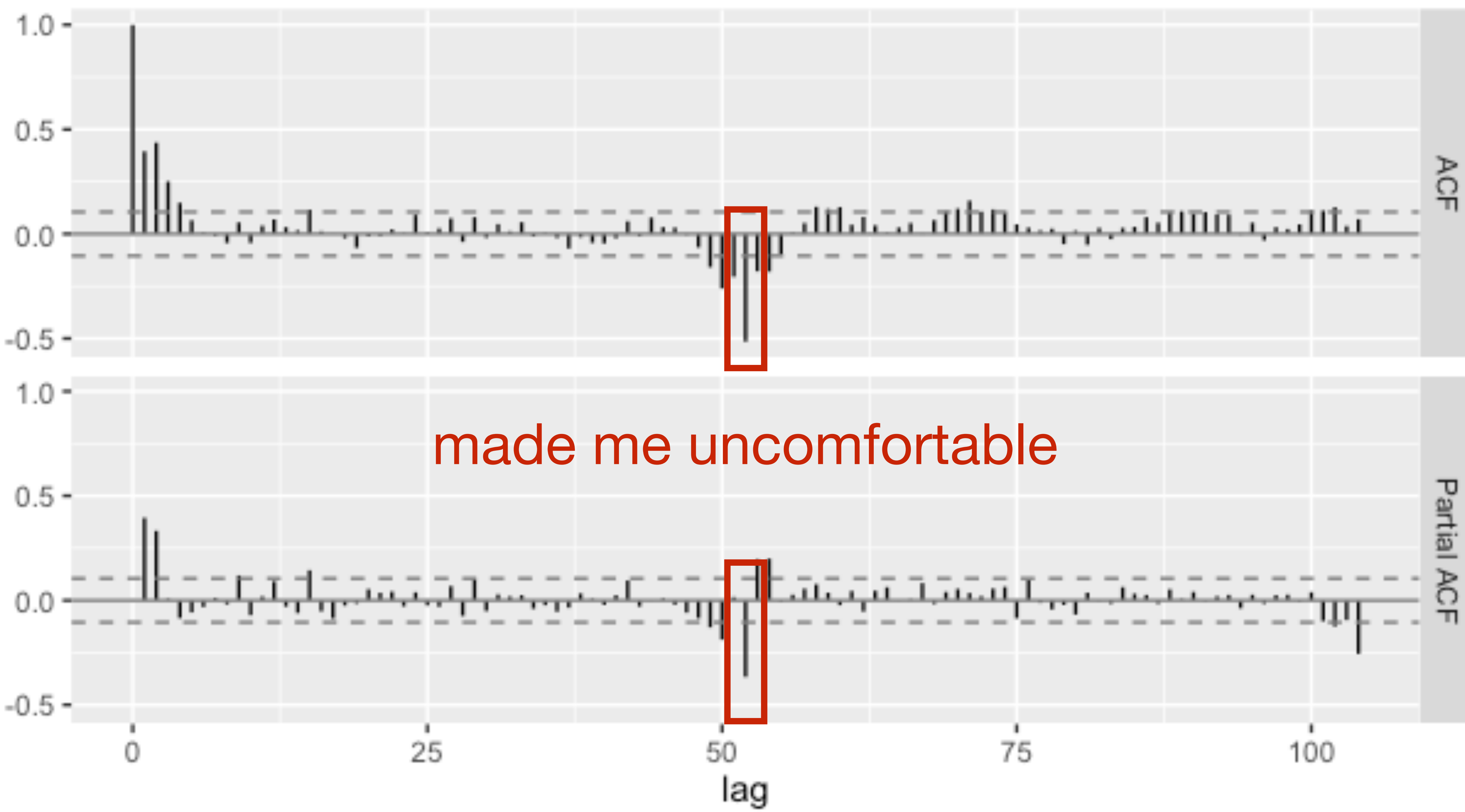
Today

HW #6 Return to mortality example

Bivariate series

Project time

Mortality Regression Example



Over differencing

Bivariate series

Imagine we observe two stationary time series

$$(X_t, Y_t), \quad t = 1, ..., N$$

We call the pairs a bivariate time series

$$E(X_t) = \mu_X, \quad Cov(X_t, X_{t+h}) = \gamma_X(h)$$

$$E(Y_t) = \mu_Y, \quad Cov(Y_t, Y_{t+h}) = \gamma_Y(h)$$

We've already considered this case when doing regression, but now we'll let the two series stand on equal footing.

The cross covariance

$$\text{Cov}(X_t, Y_{t+k}) = E\left[(X_t - \mu_X)(Y_{t+k} - \mu_Y)\right] = \gamma_{XY}(k)$$

measures the covariance between one series and leading values of the other

$$\gamma_{XY}(k) \neq \gamma_{XY}(-k)$$

it isn't an even function

$$\gamma_{XY}(k) = \gamma_{YX}(-k)$$

$$\rho_{XY}(k) = \frac{\gamma_{XY}(k)}{\sqrt{\gamma_X(0)\gamma_Y(0)}}$$

the cross correlation

$$\rho_{XY}(0) \neq 1$$

The sample cross covariance

$$c_{XY}(k) = \begin{cases} \sum_{t=1}^{N-k}(x_t - \bar{x})(y_{t+k} - \bar{y})/N & k = 0, 1, \ldots, N-1 \\ \sum_{t=1-k}^{N}(x_t - \bar{x})(y_{t+k} - \bar{y})/N & k = -1, \ldots, -(N-1) \end{cases}$$

$r_{XY}(k) = c_{XY}(k)/s_X s_Y$ sample cross correlation

asymptotically consistent and unbiased estimate of the cross correlation, **but**

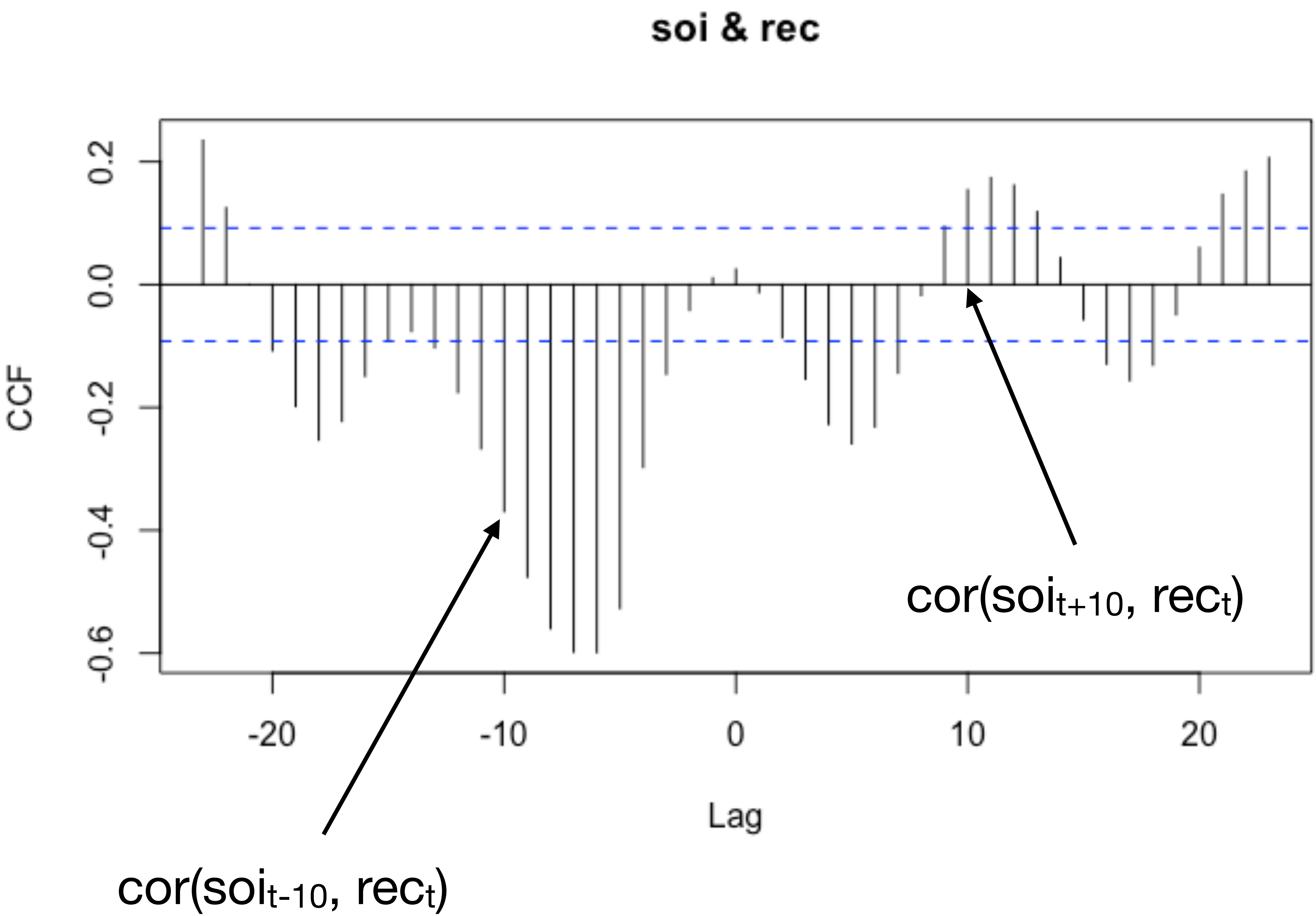
$r_{XY}(k)$ correlated with $r_{XY}(k+1)$,
and the variance of $r_{XY}(k)$ depends on autocorrelation of X and Y

generally filter series to remove correlation first

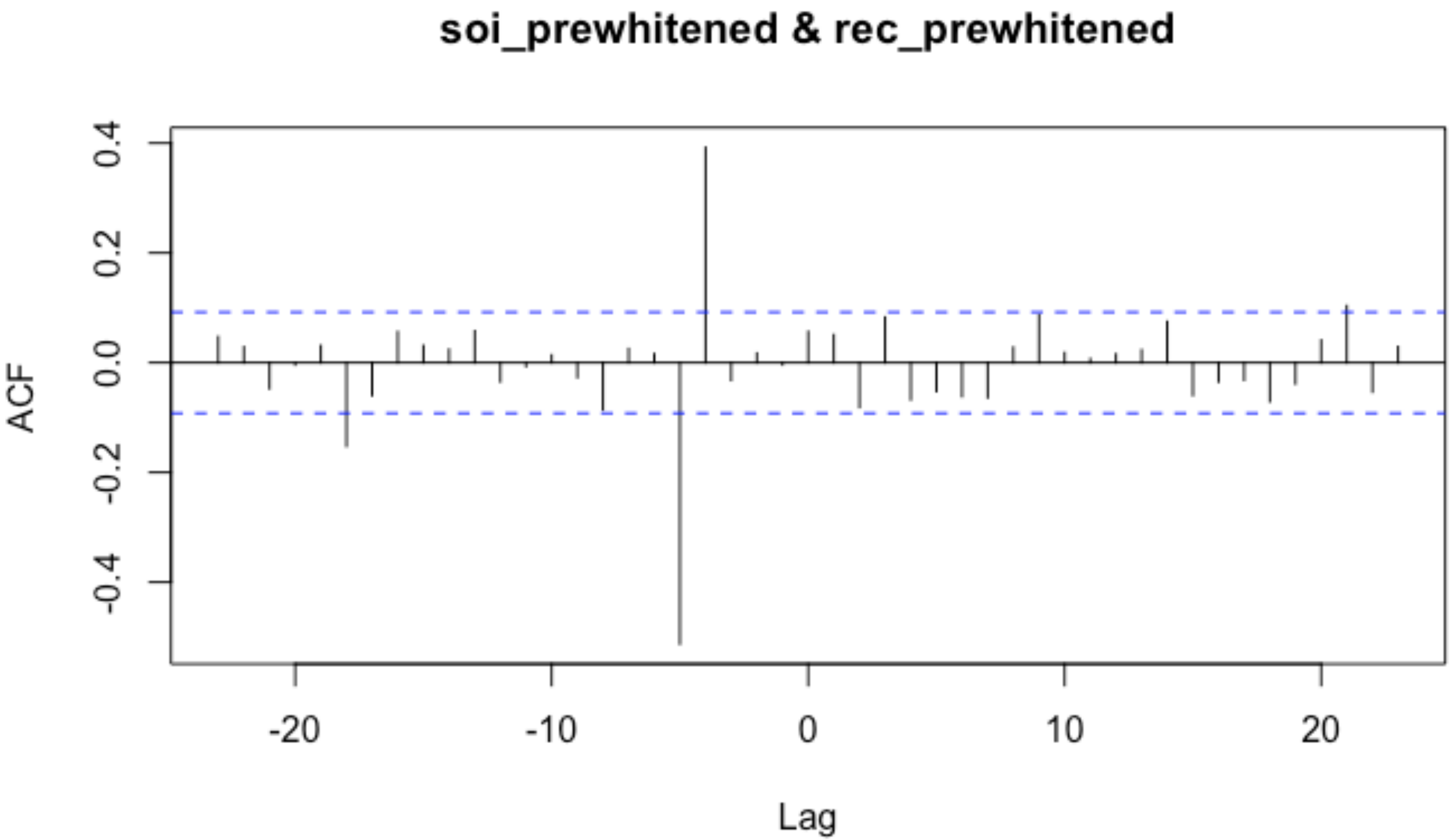
Example

Southern Oscillation Index (proxy of El Nino)

Recruitment series - that’s the number of new fish to some area



soi & rec
CCF
0.2
0.0
-0.2
-0.4
-0.6
-20
-10
0
10
20
Lag
cor(soi$_{t+10}$, rec$_t$)
cor(soi$_{t-10}$, rec$_t$)



soi_prewhitened & rec_prewhitened
ACF
0.4
0.2
0.0
-0.2
-0.4
-20
-10
0
10
20
Lag

The cross spectrum

$$f_{XY}(\omega) = \frac{1}{\pi}\left[\sum_{k=-\infty}^{\infty} \gamma_{XY}(k) e^{-i\omega k}\right] \quad 0 < \omega < \pi$$

complex valued, i.e. it has real and imaginary parts

$$c(\omega) = \frac{1}{\pi}\left\{\gamma_{XY}(0) + \sum_{k=1}^{\infty} [\gamma_{XY}(k) + \gamma_{YX}(k)] \cos(\omega k)\right\}$$

co-spectrum, real part

$$q(\omega) = \frac{1}{\pi}\left\{\sum_{k=1}^{\infty} [\gamma_{XY}(k) - \gamma_{YX}(k)] \sin(\omega k)\right\}$$

quadrature, imaginary part

Some derived functions

$$f_{XY}(\omega) = c(\omega) + iq(\omega)$$

$$f_{XY}(\omega) = \alpha_{XY}(\omega)e^{i\phi_{XY}(\omega)}$$

$$\alpha_{XY}(\omega) = \sqrt{c^2(\omega) + q^2(\omega)}$$ cross-amplitude

$$\phi_{XY}(\omega) = \tan^{-1}[-q(\omega)/c(\omega)]$$ **phase**

$$C(\omega) = \alpha^2_{XY}(\omega)/f_X(\omega)f_Y(\omega)$$ **coherence**

squared linear correlation at frequency ω

$$G_{XY}(\omega) = \alpha_{XY}(\omega)/f_X(\omega)$$ gain

regression coefficient of Y_t on X_t at frequency ω

Estimating the cross spectrum

The cross periodogram is a function of the coefficients from the Fourier expansion of X_t and Y_t

$$I_{XY}(\omega_p) = N(a_{px}a_{py} + b_{px}b_{py})/4\pi + iN(a_{px}b_{py} - a_{py}b_{px})/4\pi$$

smooth to get consistent estimates of $\hat{c}(\omega_p)$ and $\hat{q}(\omega_p)$

No extra calculation beyond individual series periodograms.

Southern Oscillation Index

Back to example in R