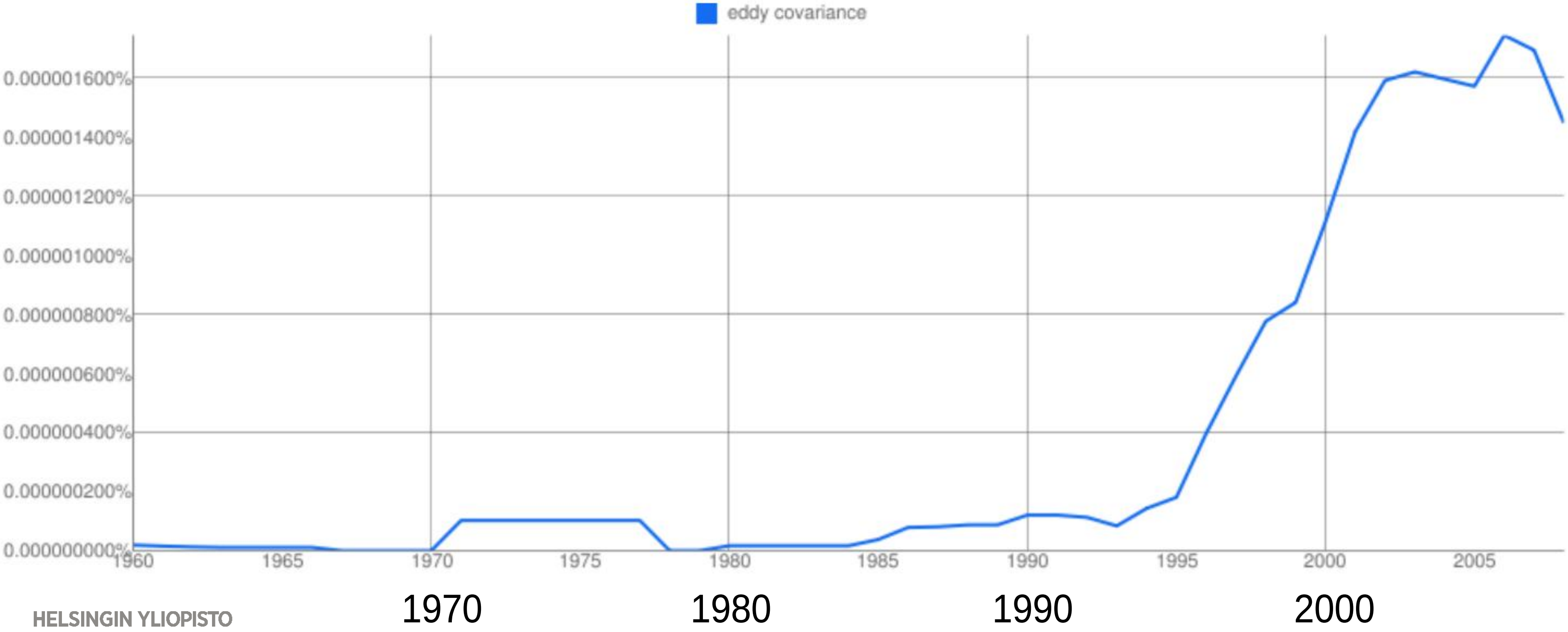
EC raw data post processing

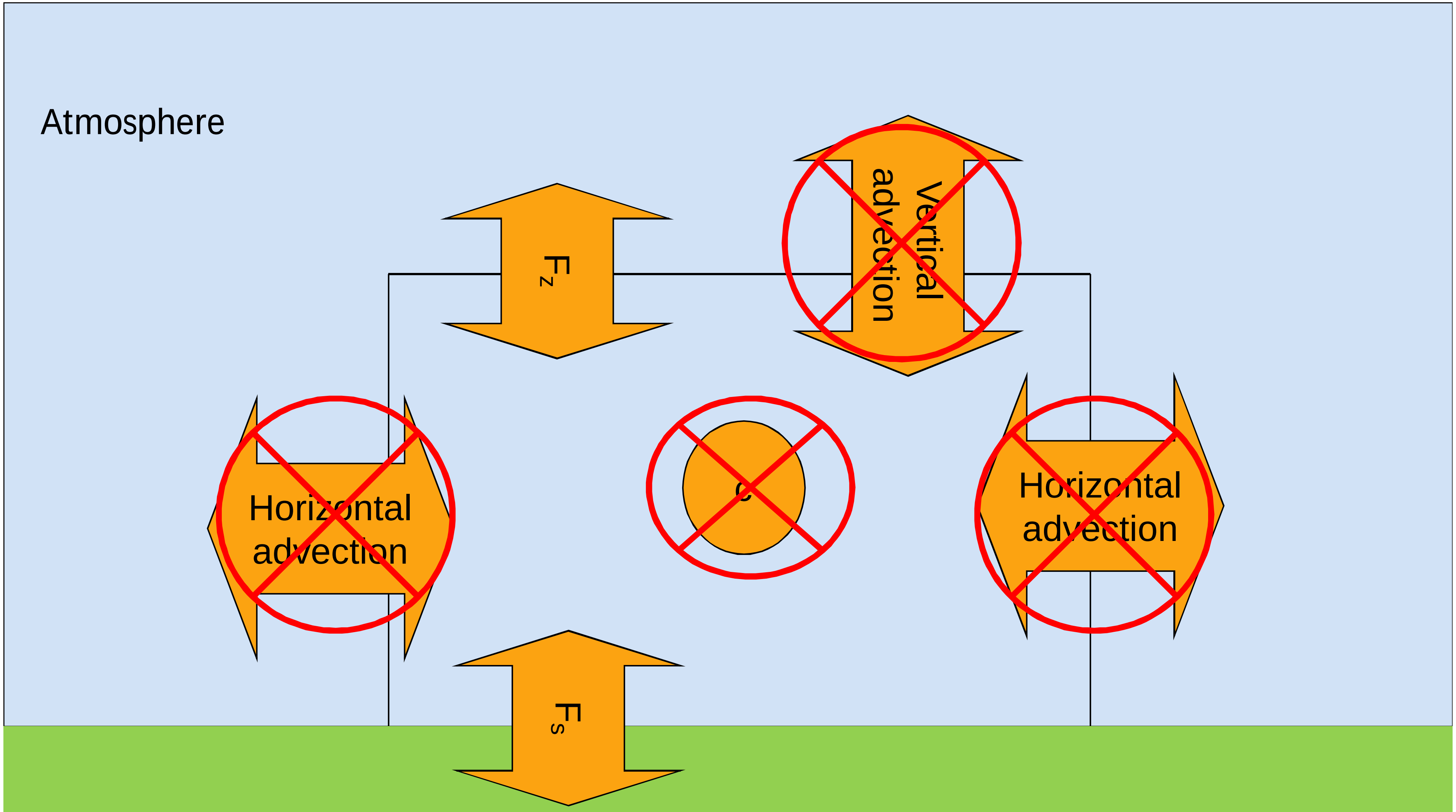
Training course/workshop on
EddyUH: a software for eddy covariance flux calculations
Jan 21- 25 2013, University of Helsinki, Finland

Annika Nordbo

Intro

How often does the phrase ’eddy covariance’ occur in books (books.google.com/ngrams)





Atmosphere

F_z

~~Vertical advection~~

~~Horizontal advection~~

~~c~~

~~Horizontal advection~~

F_s

Surface

~~Change in concentration~~ =

surface flux + ~~vertical advection~~ + ~~horizontal advection~~ – flux through the top

Adapted from Janne Rinne

F_z=F_s if

Horizontally homogenous site

→ no net advection

Chemical reactions slow

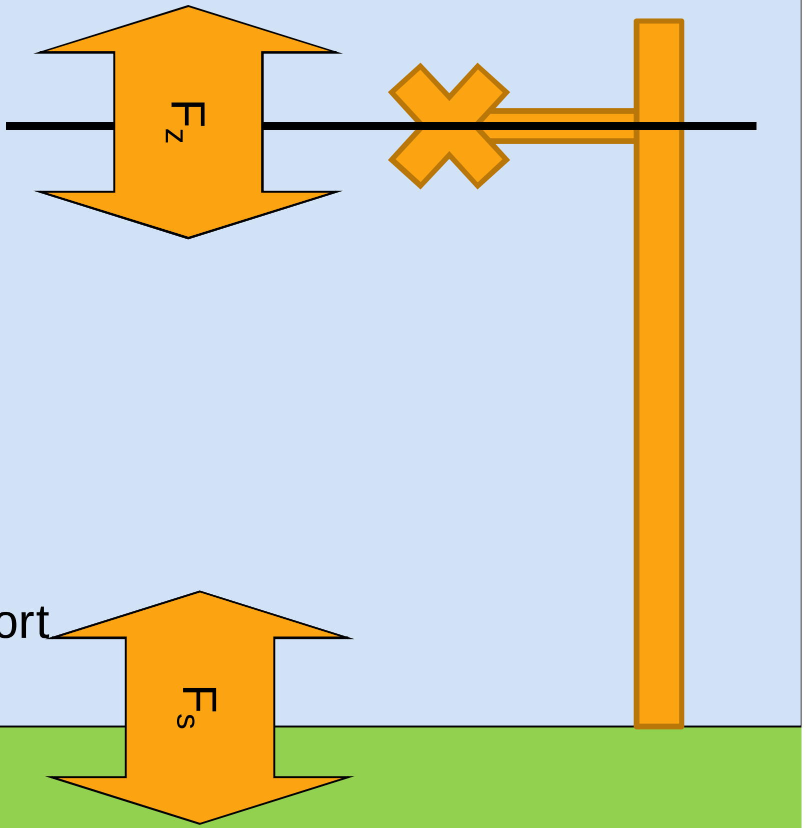
Stationary situation

No storage inside canopy

} dC/dt = 0

Vertical velocity = 0

→ vertical flux consists only of turbulent transport



~~Change in concentration~~ =

surface flux + ~~vertical advection~~ + ~~horizontal advection~~ – flux through the top

Adapted from Janne Rinne

Outline

1. Detrending
2. Despiking
3. Coordinate rotations
4. Different sonic coordinate systems
5. Inclination correction (moving platforms)
6. Lag time and covariance
7. Calculating power spectra and co-spectra
8. Choosing an averaging period

Example data:

CH_4 from Siikaneva, July 19 2010 13:00-13:30 (Los Gatos instrument)

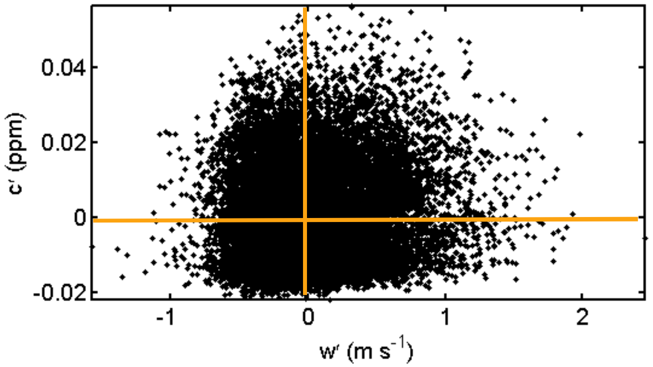
1. Detrending

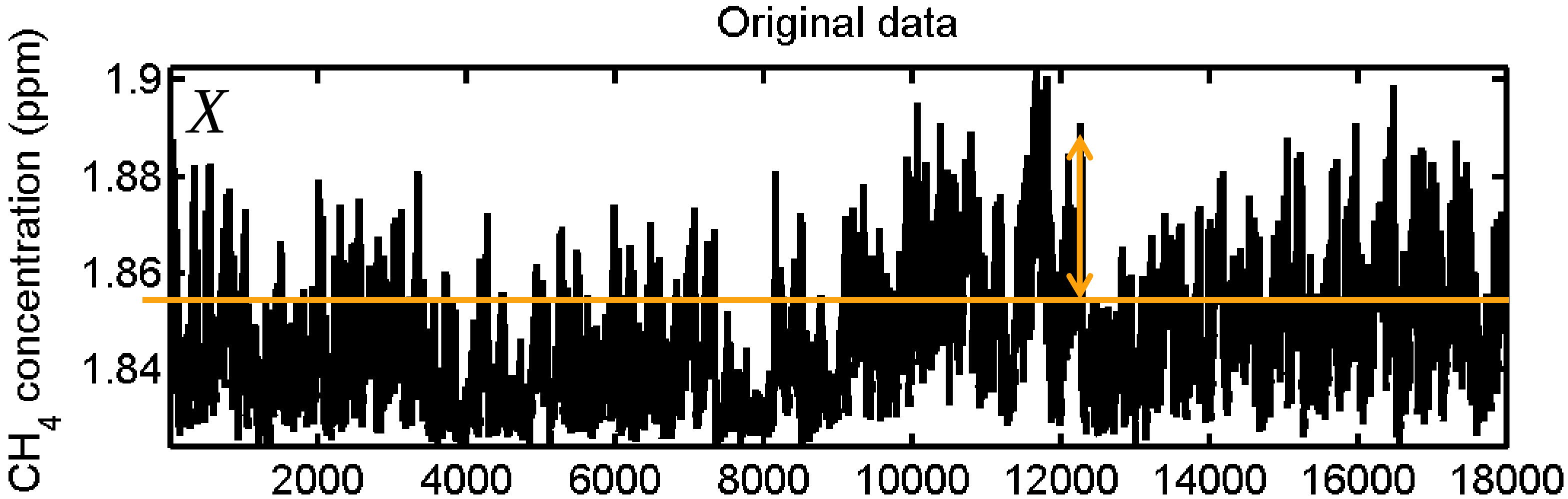
A signal consists of a mean and perturbations around the mean $X = \overline{X} + X'$

Variance $\sigma_X{}^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \overline{X})^2 = \overline{X'^2}$

Covariance $\mathrm{cov}_{X_1X_2} = \frac{1}{N}\sum_{i=1}^{N}(X_{1i} - \overline{X_1})(X_{2i} - \overline{X_2}) = \overline{X_1{}'X_2{}'}$

Correlation $cor_{X_1X_2} = \overline{X_1{}'X_2{}'} / \sigma_{X1}\sigma_{X2}$





1.1 Mean removal

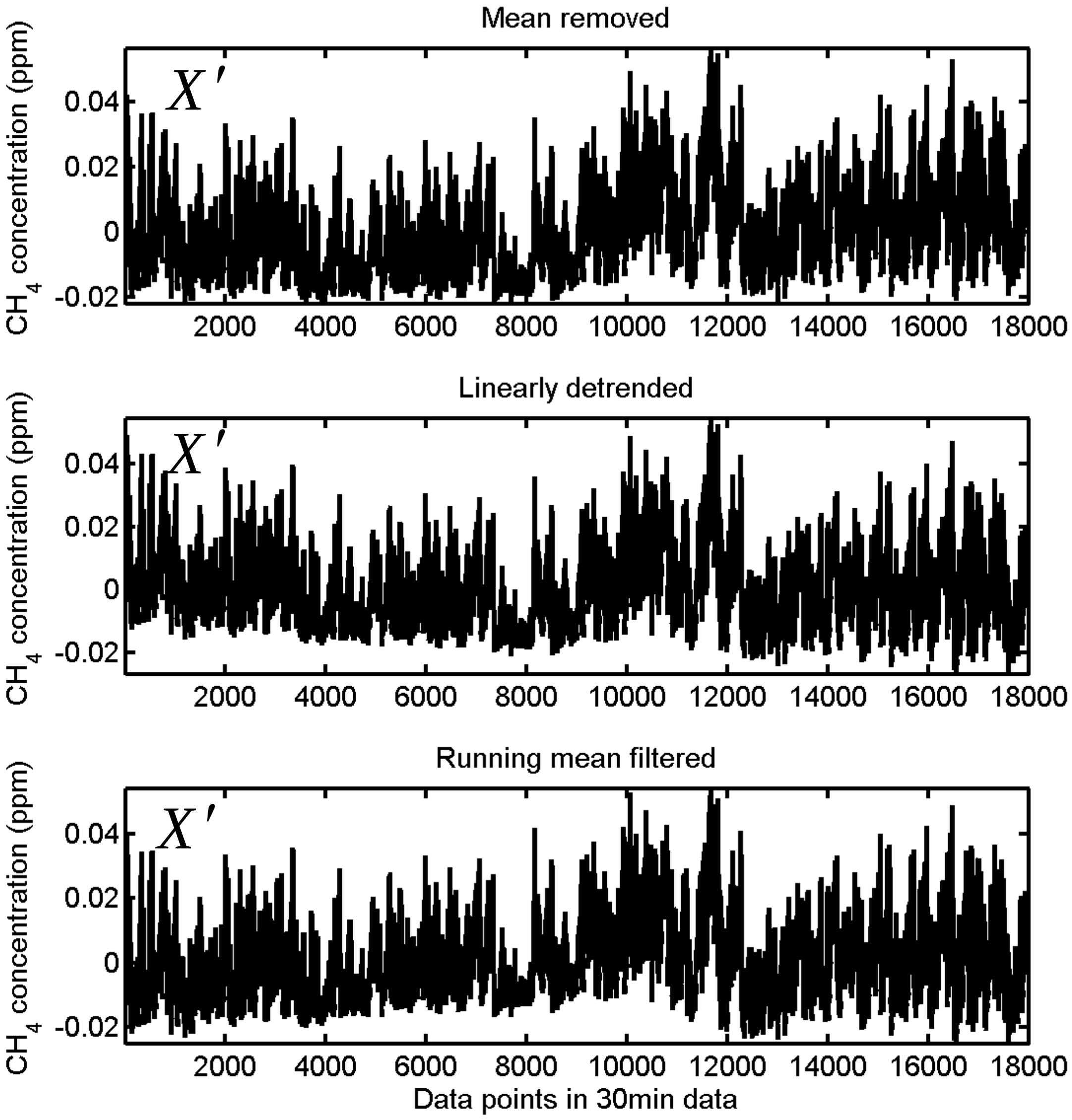
- Simplest way, many prefer
- may yield artificial flux

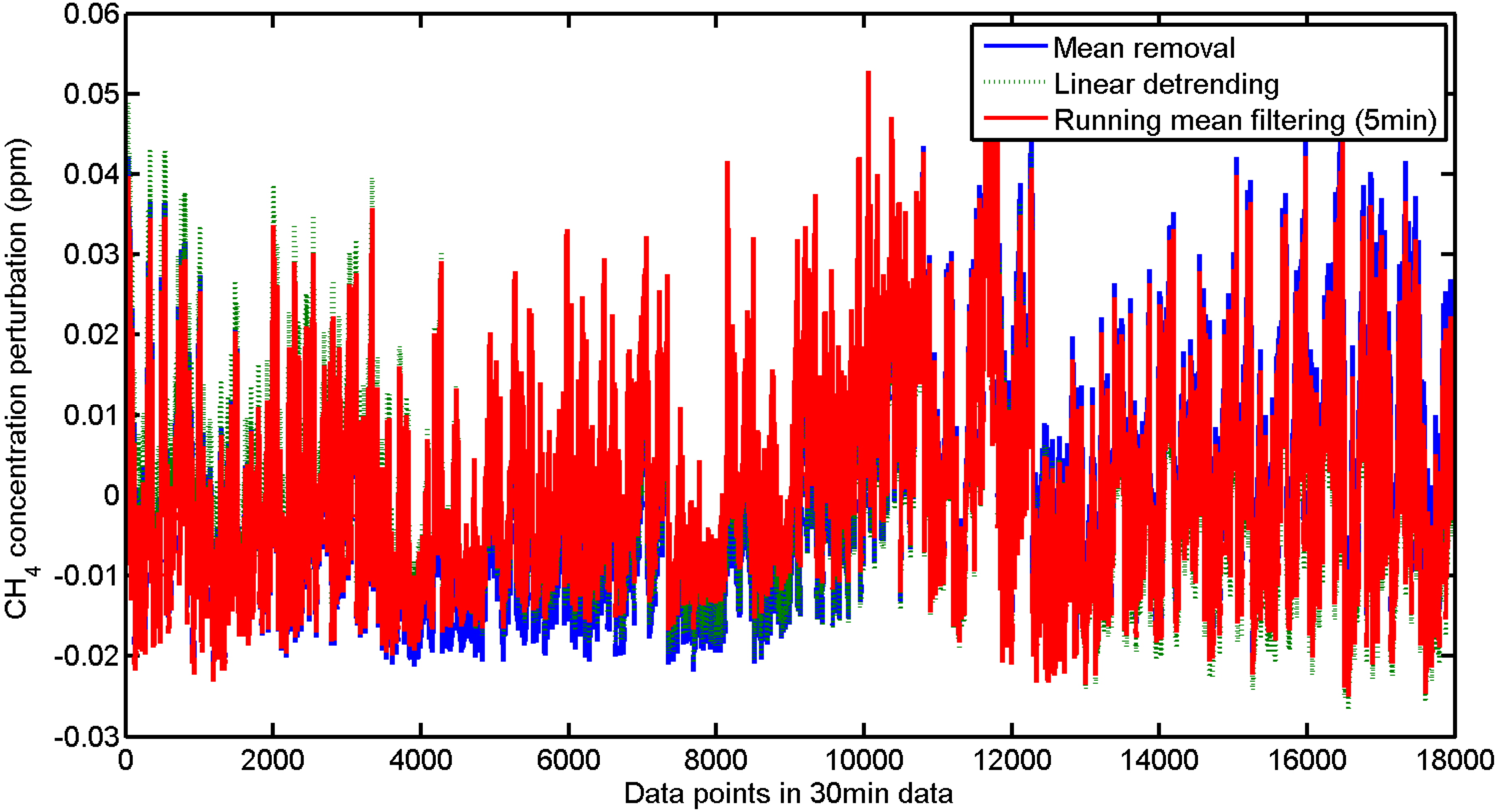
1.2 Linear detrending

- perhaps most common
- may lose some flux (here 5%)
- Good when rapid changes
  (e.g. diurnal variation)

1.3 Running mean detrending

- May lose a lot of flux
- Good if very rapid changes
  (e.g. < 5min)




Mean removal
Linear detrending
Running mean filtering (5min)
CH_4 concentration perturbation (ppm)
Data points in 30min data
0.06
0.05
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
0
2000
4000
6000
8000
10000
12000
14000
16000
18000

2. Despiking

Why?

- Spikes (= unreasonably high/low values) are caused by electronic and physical noise
- Might cause an error in the flux calculations. The error is hard to assess.

2.1 adjacent points

— Spike: adjacent CH_4 data points depart more than 1ppm (for example)

— Solution: spike data replaced by the value of the previous point

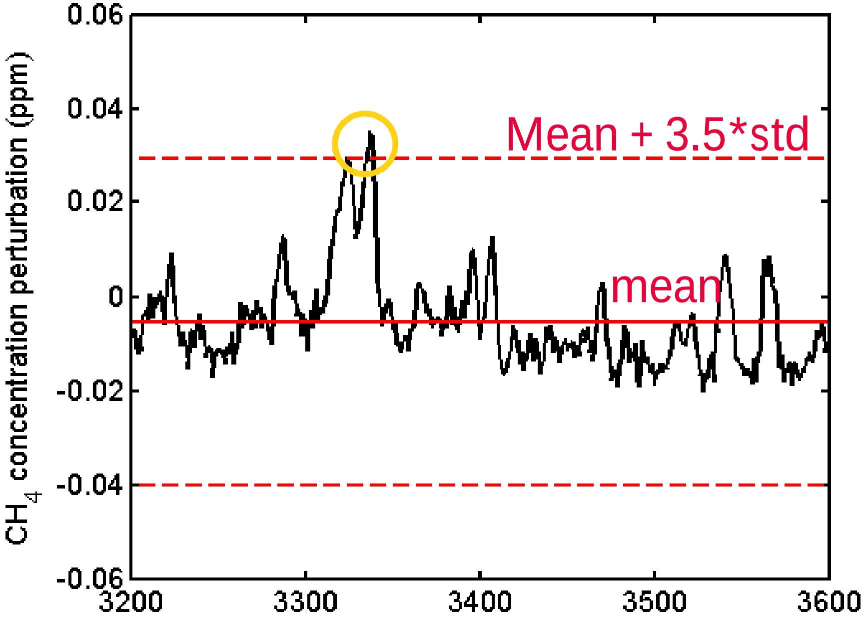
2.2 upper and lower limits

— Spike: data that exceed an upper or a lower limit

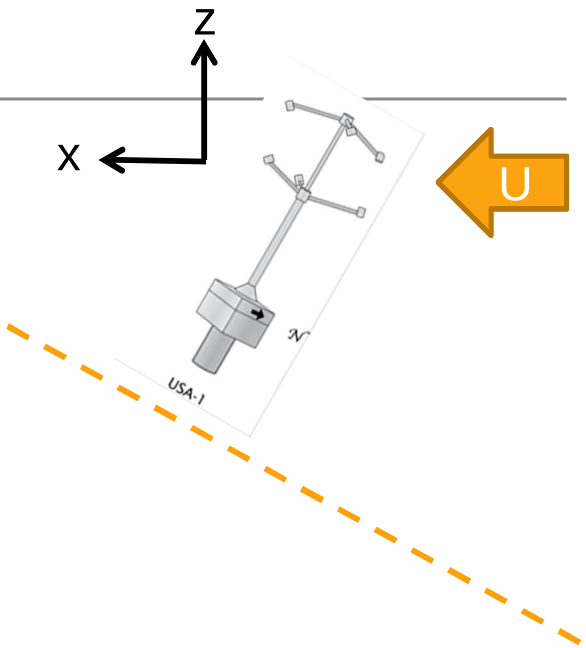
— Solution: spike data replaced by the value of the previous point

2.3 According to std (Vickers & Mahrt 1997)

- Means and standard deviations for a series of moving data windows are calculated (window moves 400 points)
- Spikes: points which are an optional amount of std away from the mean (you can choose the std amount in EddyUH, e.g. 3.5*std)
- Solution: spike data are replaced by linear interpolation
- The process is repeated until no spikes are left, the std level is increased by 0.3 every time



3. Coordinate rotations

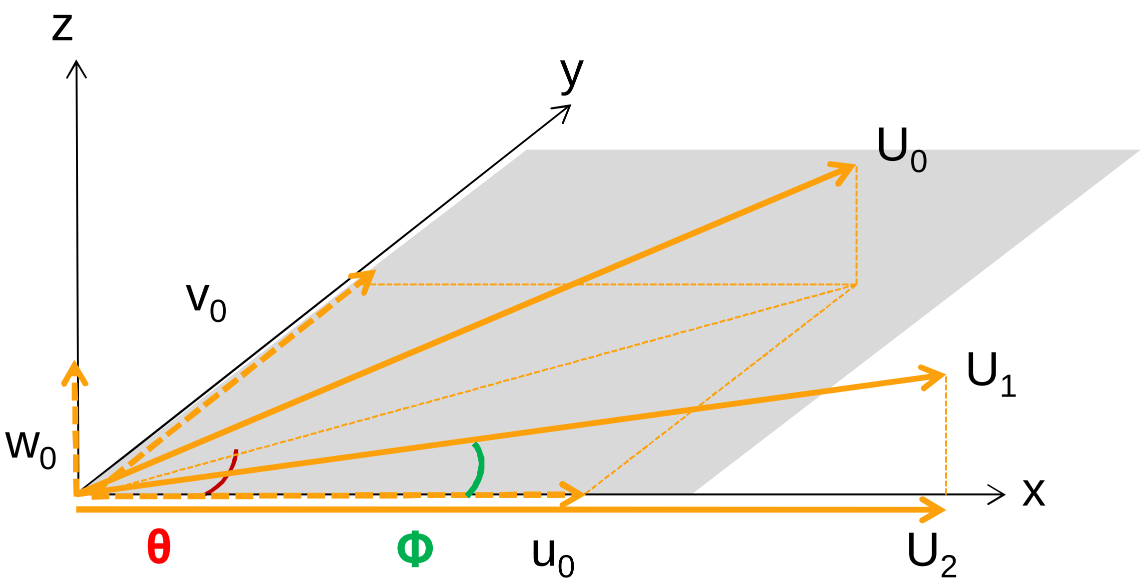


Why?

- The z-axis should be perpendicular to the mean flow (U). This is not the case if the anemometer is not levelled perfectly and the terrain might be sloping.
- Hence, the w-signal might be contaminated by the u- and v-components

3.1 Traditional 1D, 2D, 3D (additive)

- Applied to each averaging period separately
- 1D: mean v=0, u aligned with x
- 2D: mean w=0, w aligned with z (1D made first, MOST COMMON!)
- 3D: mean w’v’=0



$$\theta = \tan^{-1}\frac{v_0}{u_0}$$

$$u_1 = u_0\cos\theta + \mathrm{v}_0\sin\theta$$

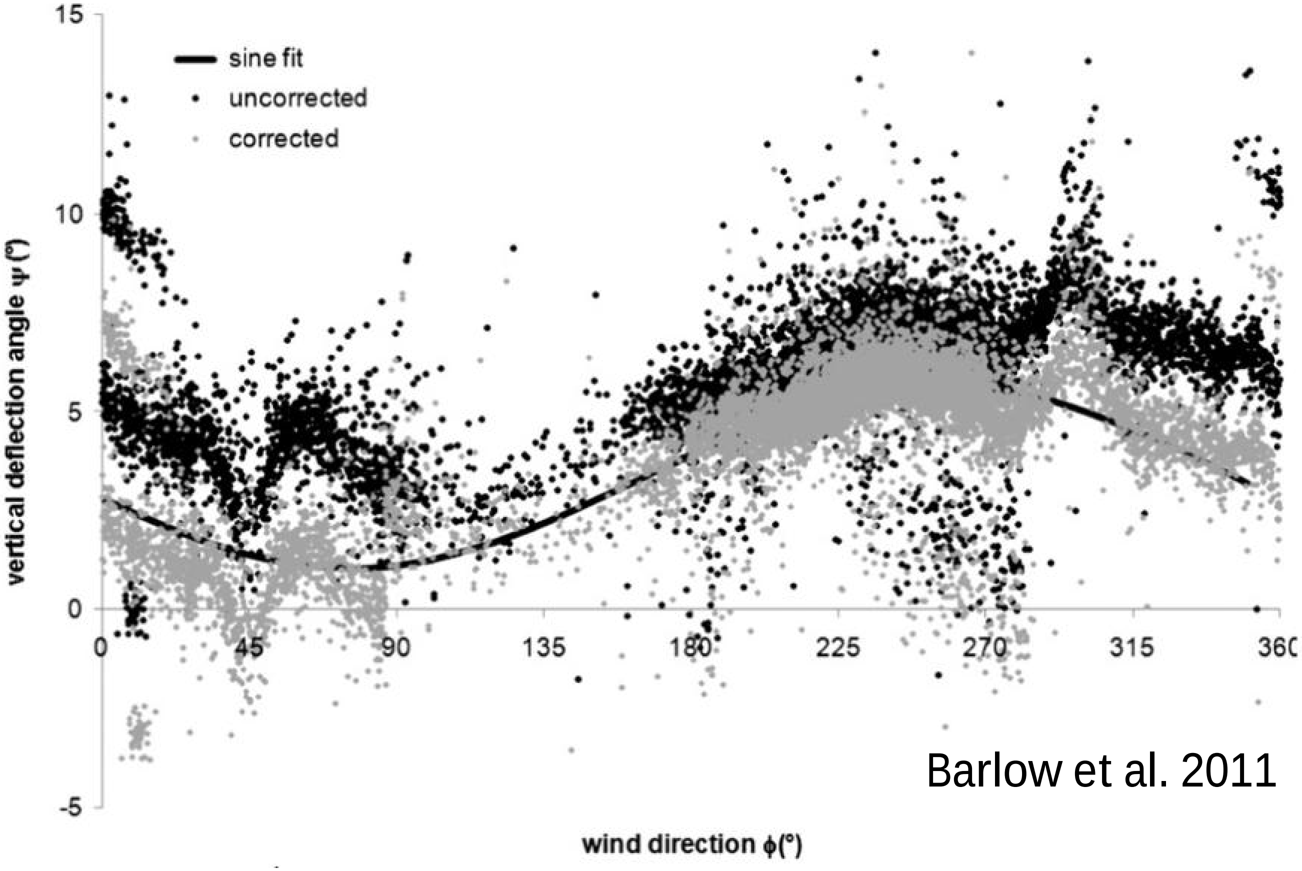
$$v_1 = -u_0\sin\theta + \mathrm{v}_0\cos\theta = 0$$

$$\phi = \tan^{-1}\frac{w}{u_1}$$

$$u_2 = u_1\cos\phi + \mathrm{w}_0\sin\phi$$

$$w_1 = -u_1\sin\phi + \mathrm{w}_0\cos\phi = 0$$

You know that your tower is tilted when…



Barlow et al. 2011

3.2 Planar fitting (Wilczak et al. 2001)

- For stationary set-up conditions
- A hypothetical plane with constant tilt angles for a set of data runs (2 months or so) is subtracted from the data

  → Non-zero mean vertical velocities of averaging periods (Influence of large eddies)

- Tilt-angles might depend on wind direction in heterogeneous surroundings

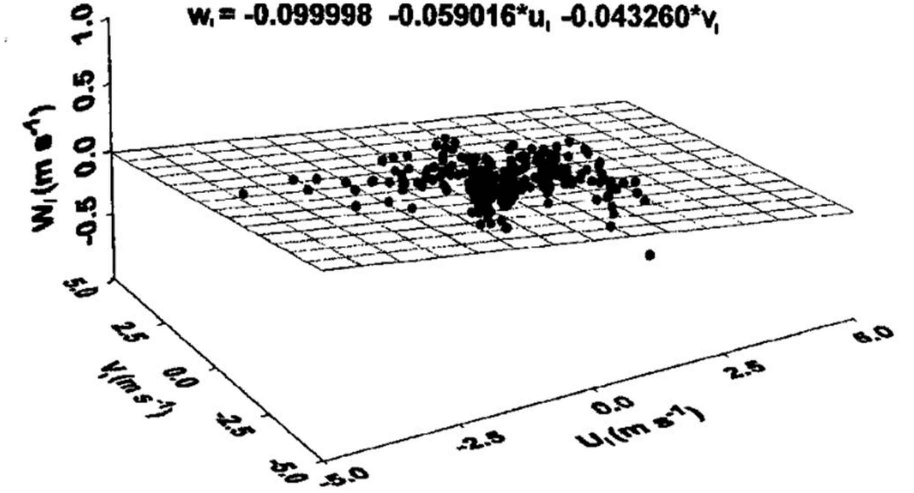


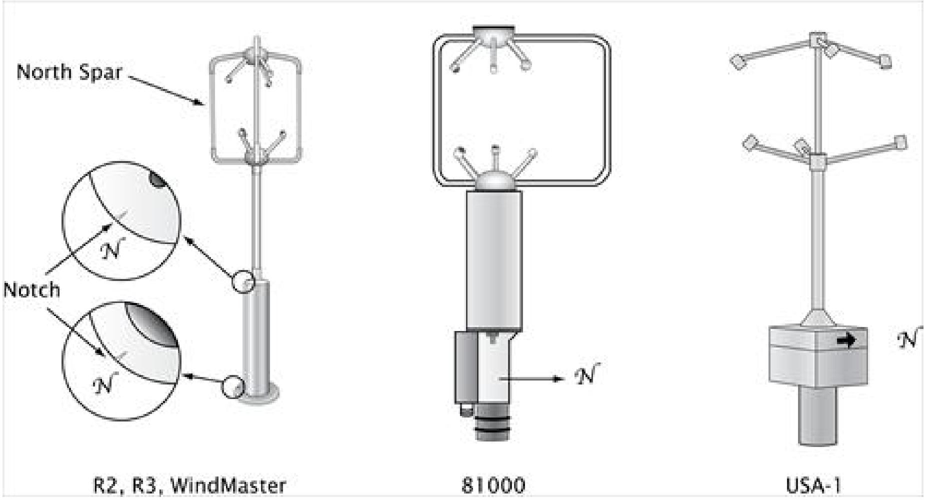
Figure 9.5. An example of planar fit regression with wind data over a maize canopy in Davis, California (Paw U et al. 2000).

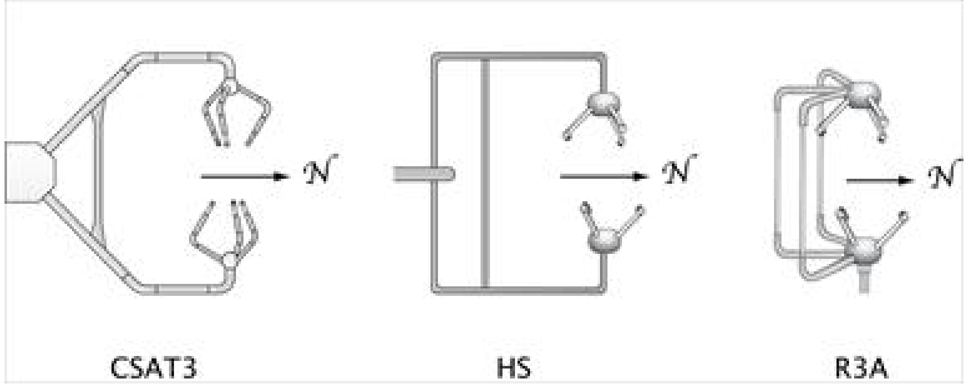
- Subject to debate (flow depends on stability)

4. Different sonic coordinate systems

All sonics have a marking which should point towards north

The deviation from north is as ‘boom_correction’ in EddyUH





http://envsupport.licor.com/help/EddyPro3/Content/Topics/Adjusting_Anemometer_Coordinate_System.htm

Output u and v are defined differently in different sonics

EddyUH calculates wind direction so that N = 0°

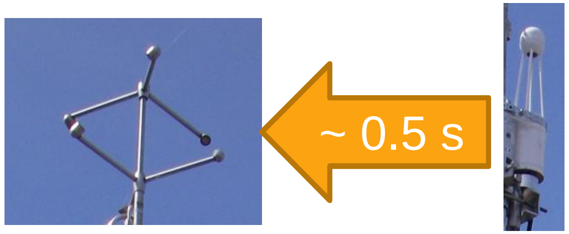
Anemometer	u	v
Gill R3-50	N	W
Gill R2	S-30°	S+60°
Gill HS	W	N
Metek USA-1	N	E
Cambell CSAT3	S	E

5. Inclination correction (moving platforms)

Lake Valkea-Kotinen, Nordbo et al. 2011

- If the EC mast is moving itself, inclination corrections must be applied to the wind data
- Rafts may move fast with waves and tall masts may move slower with the wind
- Requires a fast response inclinometer
- Optional in EddyUH

6. Lag time and covariance calculations



Why is a lag time needed?

- the gas analyzer and the sonic might be far apart
- time differences in data collection due to computers

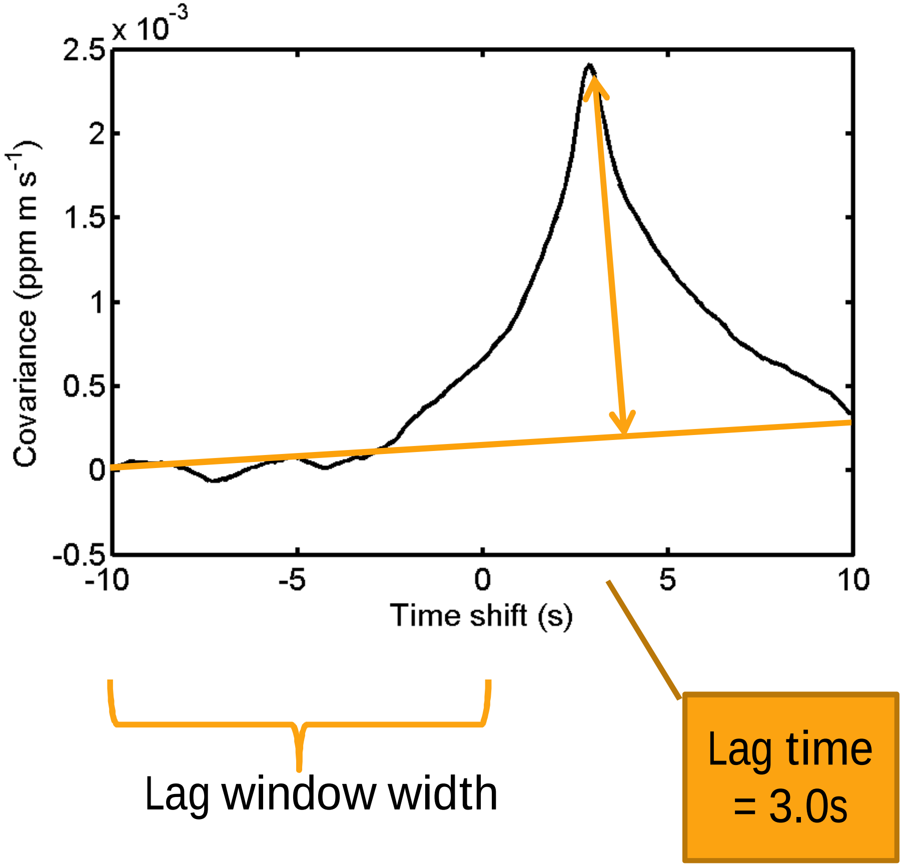
6.1 Constant lag time

- a constant value for the time difference is given
- lag window width = 0 in EddyUH
- closed-path analyzers: can be approximated from the tube length and flow rate, t_{lag}=U/L
- this is not recommended since even tube flows are not always steady

$\overline{w'c'}=2.41\cdot10^{-3}\text{ppm m s}^{-2}$

6.2 Varying lag time

- cross-correlation function gives the covariance as a function of the time shift between w’ and c’
- Lag time is the point where the function has its maximum
- EddyUH needs: lag window centre & width (here 0 and 10s)
- ‘linear detrending’ of the cross-correlation function in EddyUH (optional in future)
- this method is recommended



Lag window width

Lag time = 3.0s

Lag time of H2O depends on relative humidity due to sorption effects on tube walls/filters (may depend also on temperature!)
Example from SMEAR III, Helsinki, Nordbo et al. 2012a:

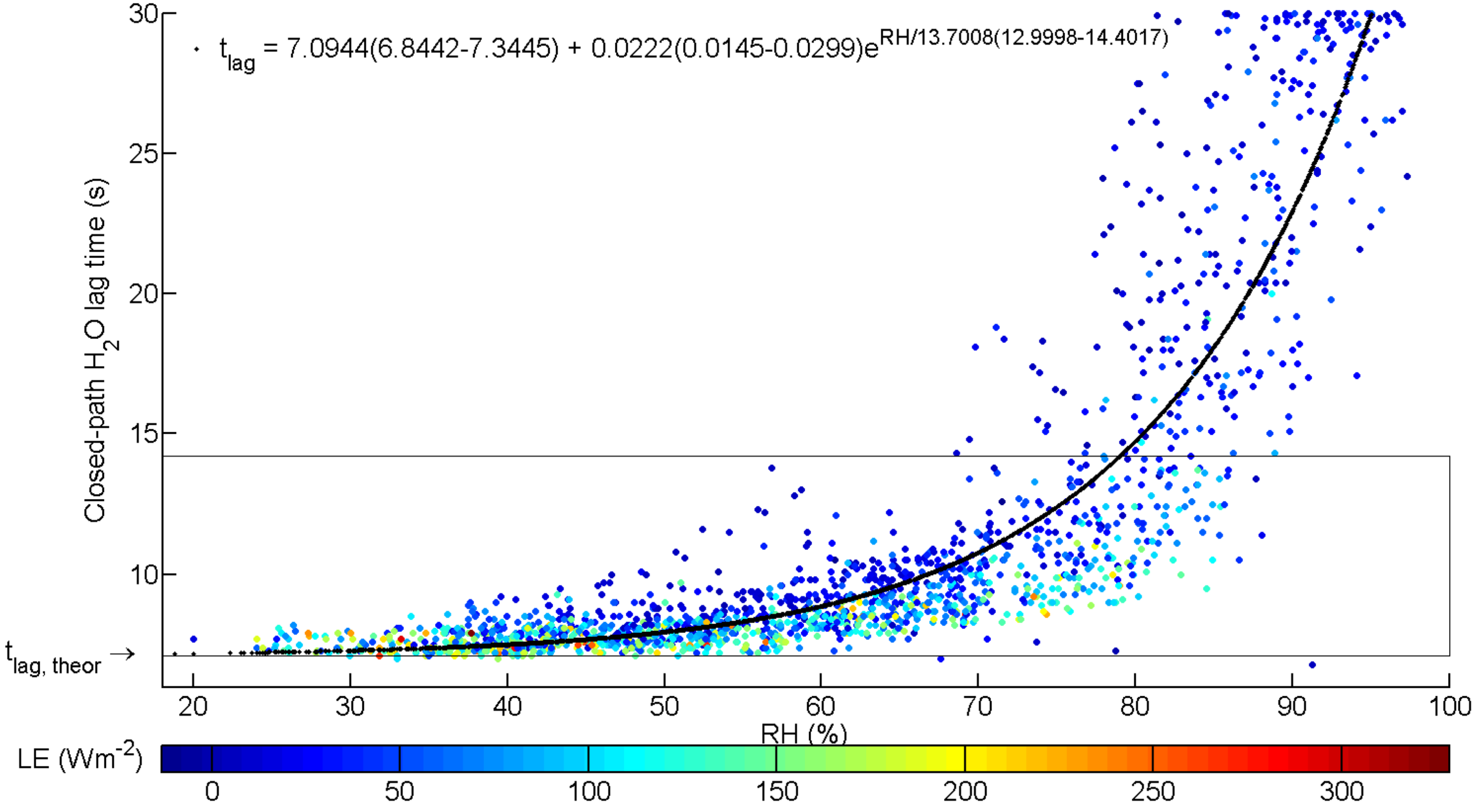


Fig. 7. Lag time (s) of closed-path H_2O measurements as a function of relative humidity (%) and an exponential with 95% confidence intervals. The theoretical lag time ($t_{lag,theor}$) and the theoretical lag window (rectangle) are also given. The colorbar gives the latent heat flux magnitude (LE, Wm^{-2}). Data are from May and June 2010.

7. Calculating power spectra and co-spectra

- Why?

— one wants to know what sized turbulence transports the flux (not just only one flux value)

— good way to check that measurements are ok

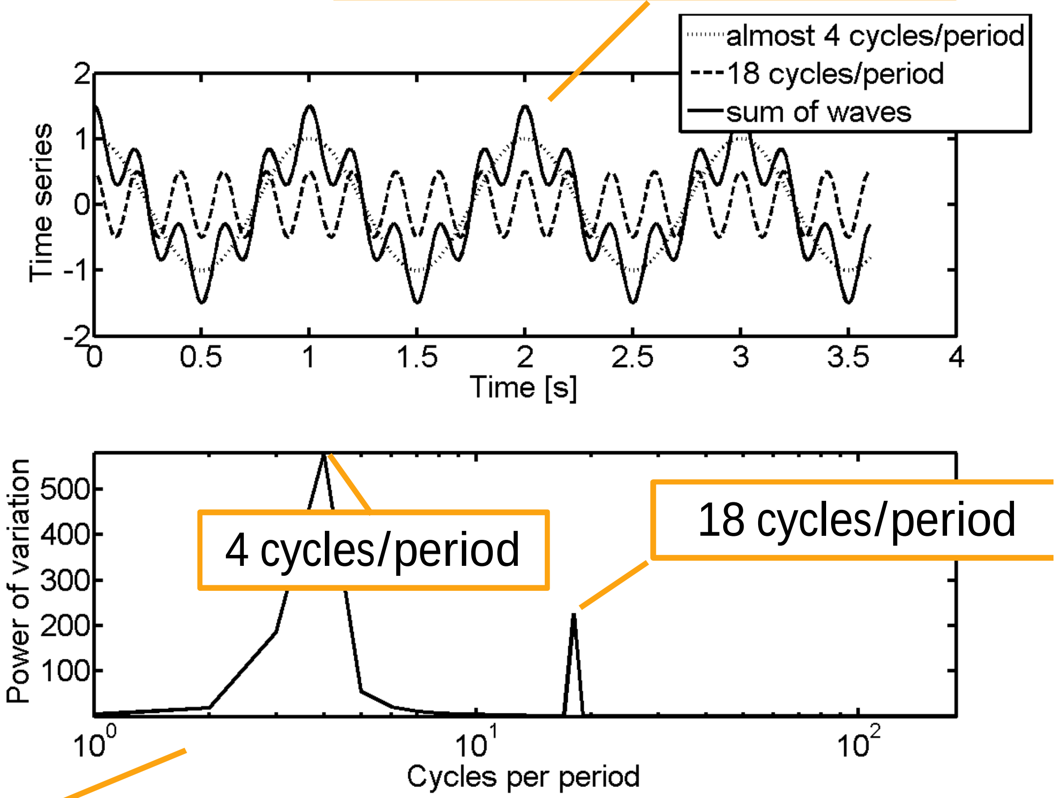
— interesting information can be drawn from spectra

- Calculated using mathematical methods such as Fast Fourier Transform or Wavelet analysis (FFT in EddyUH)
- FFT gives the amplitudes of variation for different frequencies (f, Hz)
- Normalized frequency ($n = f(z-d)/U$) takes into account the measurement height (z), displacement height (d) and mean wind speed (U); makes easier to compare between sites

7.1 Basics of discrete Fourier analysis

- A numerical tool for decomposing a time series to a sum of time series consisting of cosine and sine waves
- for continuous periodic time series
- can be used for converting a time series to a function of frequency (does a certain thing repeat itself every 2s?)
- The amplitudes of each variation period (e.g. 2s) tell something about turbulence with that time period
- Variation periods can be converted to frequencies or turbulence diameters

Simple example: time series consists of two cosine waves



4 cycles/period

18 cycles/period

The fourier transform reveals two main periods

7.2 Power spectra

— power spectra tell how much different sized eddies contribute to the variance

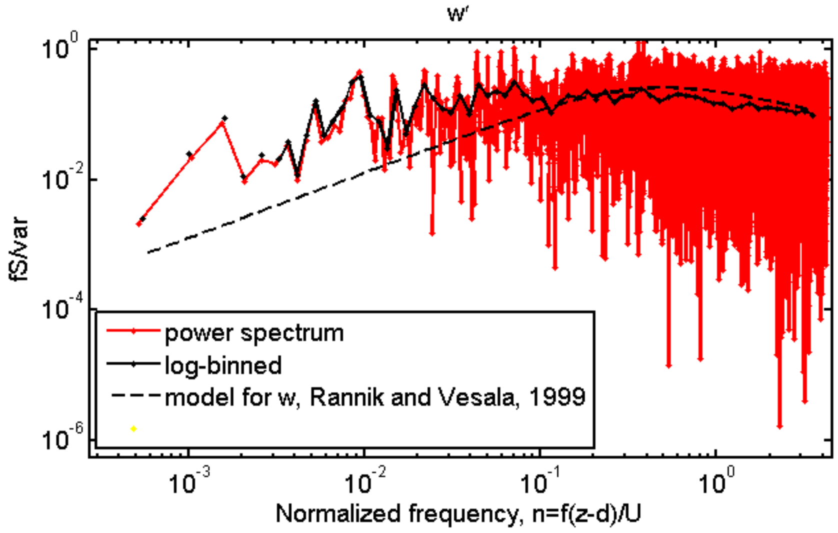
$A = \mathtt{fft}(x)$ amplitudes come from the Fast Fourier Transform, has a real (cosine) and imaginary (sine) part

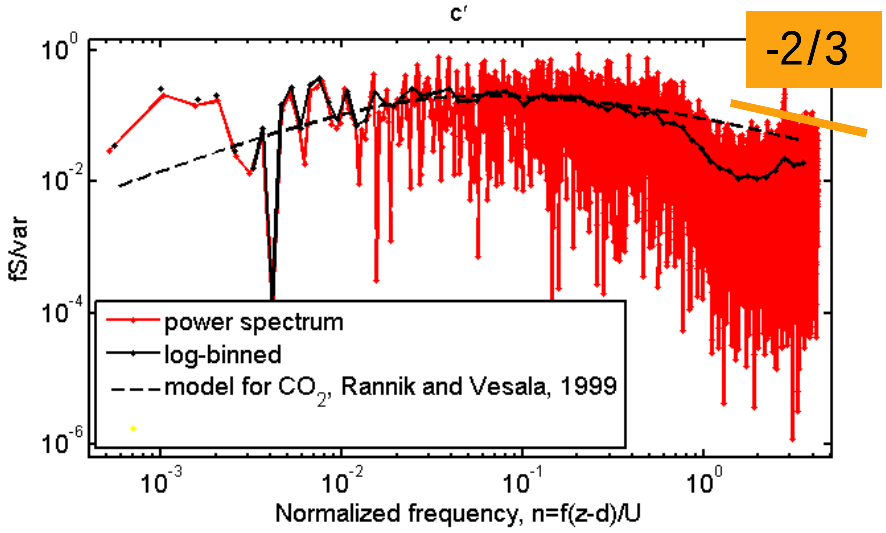
$P = |A|^2$ power is the square of the absolute value of the amplitude = $\mathrm{Re}(A)^2 + \mathrm{Im}(A)^2 = A \cdot A*$

$E = \dfrac{2P}{N^2}$ spectral energy is twice the power devided by the length of x squared (N^2), Stull 1988, p.313

$S = \dfrac{E}{\Delta f}$ spectral density (units of variance per frequency interval)

$\dfrac{fS}{\sigma^2}$ is spectral density * frequency / variance (dimensionless), always positive!





7.3 Co-spectra

— how much different sized eddies contribute to the Covariance of two variables

A_x and A_y the amplitudes of variables x and y

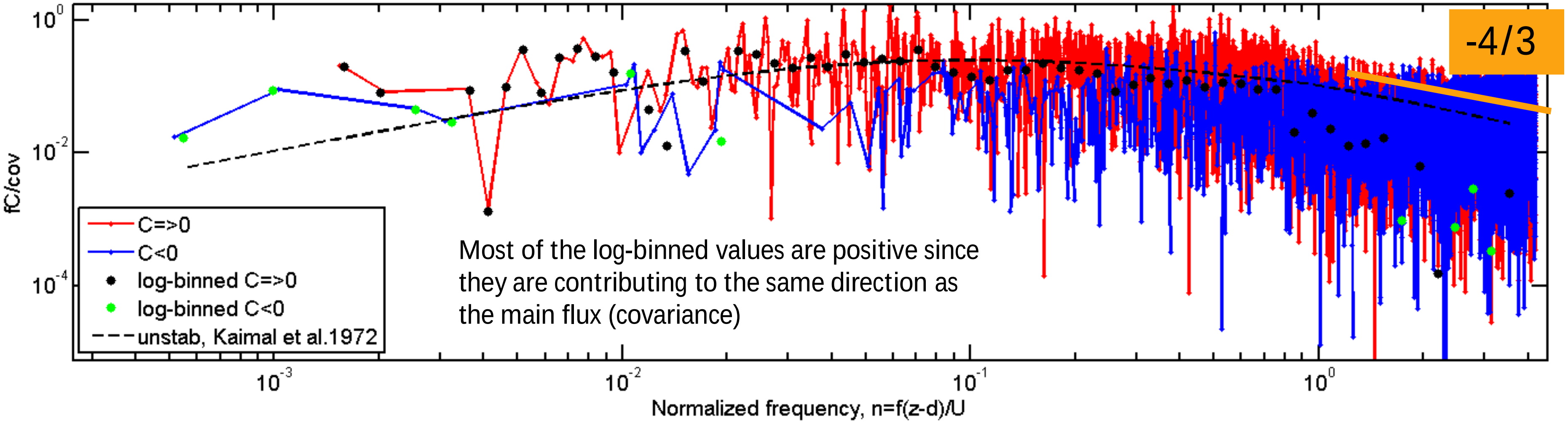
$$P_{xy} = \mathrm{Re}(A_x)\cdot\mathrm{Re}(A_y) + \mathrm{Im}(A_x)\cdot\mathrm{Im}(A_y)$$

co-spectral power is the product of the amplitudes of the cosine and sine waves

$$E_{xy} = \frac{2P_{xy}}{N^2}$$ co-spectral energy is twice the co-spectral power devided by the length of x squared (N^2)

$$C = \frac{E_{xy}}{\Delta f}$$ co-spectral density (units of co-variance per frequency interval)

$$\frac{fC}{\overline{w'c'}}$$ is co-spectral density · frequency / co-variance (dimensionless), positive or negative!



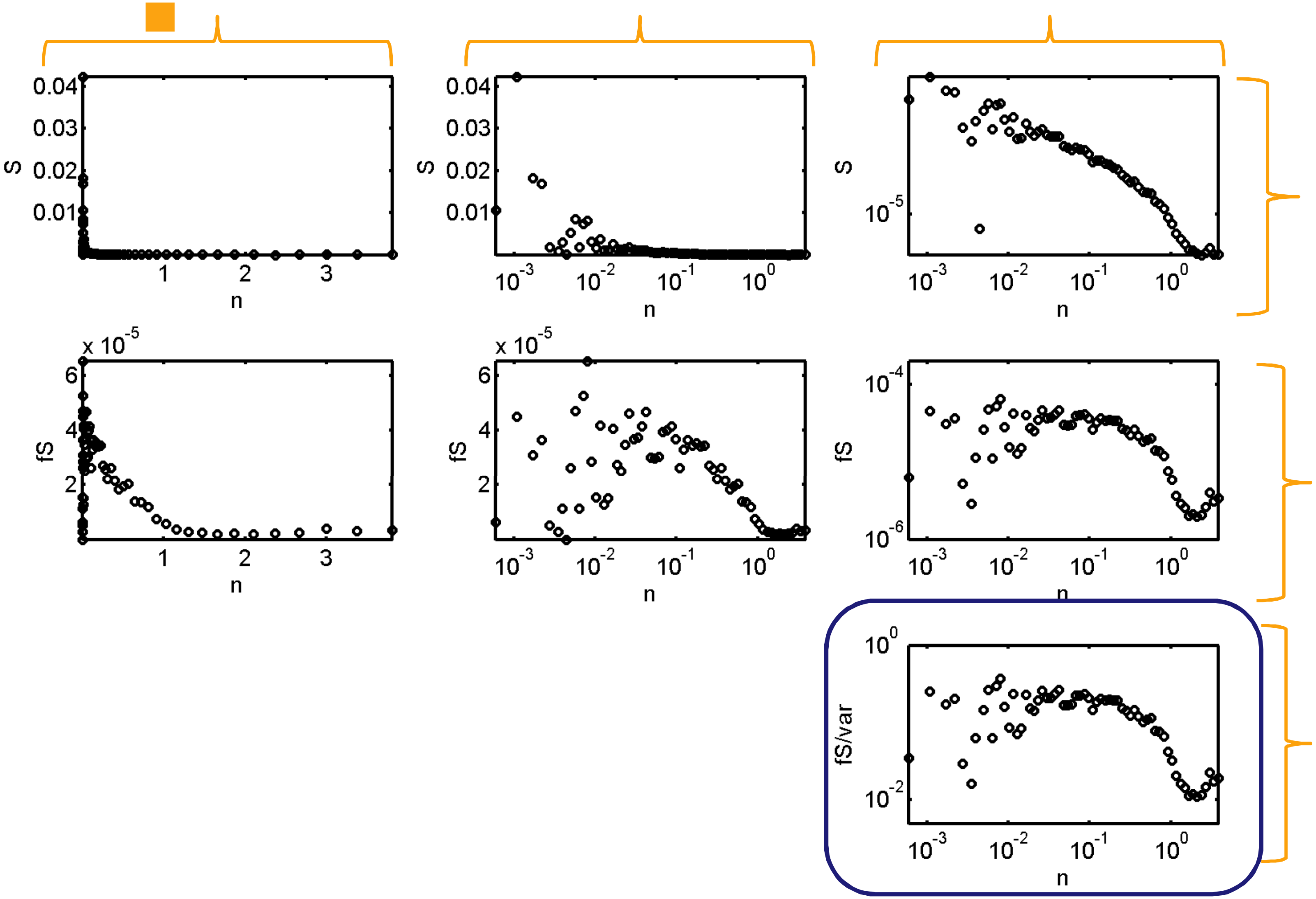
Most of the log-binned values are positive since they are contributing to the same direction as the main flux (covariance)

7.3 Different ways to display power spectra (applies for co-spectra, too)

linlin

semilogx

loglog



Units of variance per frequency interval

Units of variance

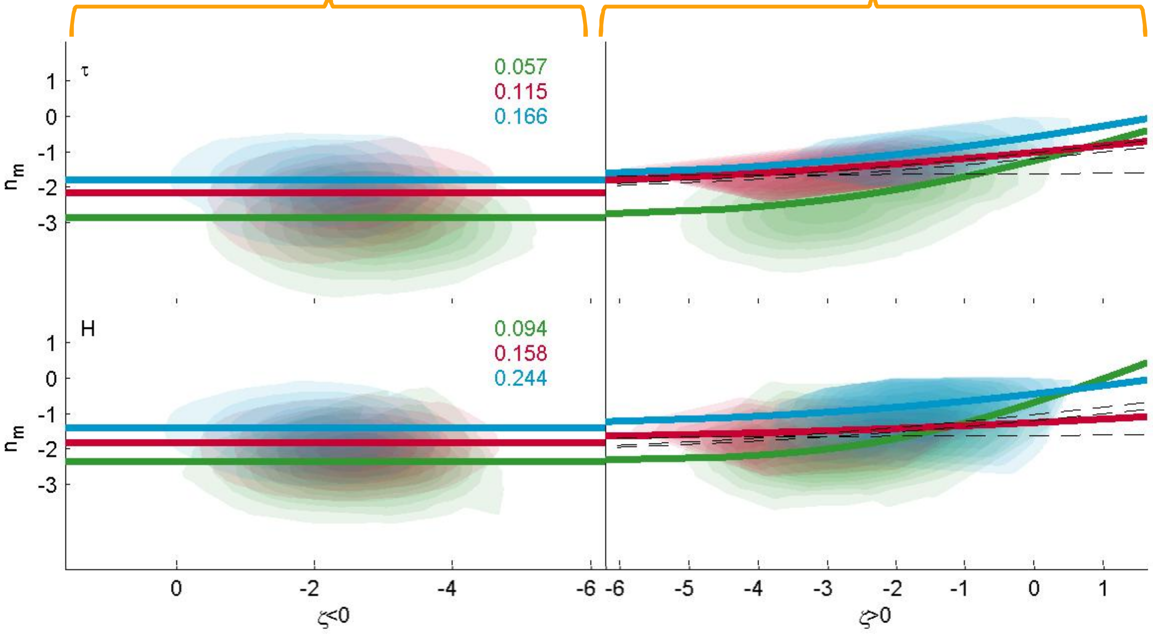
Dimensionless

7.4 Spectral peak location depends on stability

- example with 1 year of data from 3 sites in Helsinki
- the peak frequency is the higher the rougher the surface

Peak frequency (n_m) is constant in unstable conditions

Peak frequency grows as a function of stability (small turbulence is more important in stable conditions)



Momentum co-spectral peak frequency

Sensible heat co-spectral peak frequency

7.5 Engineer’s approximations for co-spectra

- Scalar co-spectra for unstable (L<0) and stable (L>0) situations

$$\frac{fC_{ws}}{\overline{w's'}} = \begin{cases} \begin{cases} \dfrac{1.05\, n/n_m}{(1+1.33\, n/n_m)^{7/4}}, & n \le 1.0 \qquad n_m = 0.085 \\ \dfrac{0.387\, n/n_m}{(1+38\, n/n_m)^{7/3}}, & n \ge 1.0 \qquad n_m = 0.085 \end{cases} & L \le 0, \quad \text{Horst 1997, eq. 8} \\ \dfrac{0.637\, n/n_m}{1+0.91(n/n_m)^{2.1}}, \qquad n_m = 2.0 - 1.915/(1+0.5\zeta) & L \ge 0, \quad \text{Horst 1997, eq. 10} \end{cases}$$

- We use a different type of equation $n_m = \beta_1 (1+\beta_2 \zeta)^{\beta_3}$
- Similar equations also for momentum co-spectra and all power spectra.

8. Choosing an averaging period

Why?

When the averaging period is too short, flux contribution from large eddies might be lost.

When the averaging period is too long, stationarity might become a problem (diurnal cycle might show).

How to choose the averaging period

1. Fixed arbitrary: 30min, 1h
2. Empirical: calculate the flux (for example) for 5min, 15min, 30min, 1h and choose the one with the highest flux (you can choose the averaging period in EddyUH)
3. Ogive method: plot the cumulative co-spectrum and find the point after which there is no flux contribution any more. Averaging period could be different for all time points. Nobody has ever used a varying averaging period!

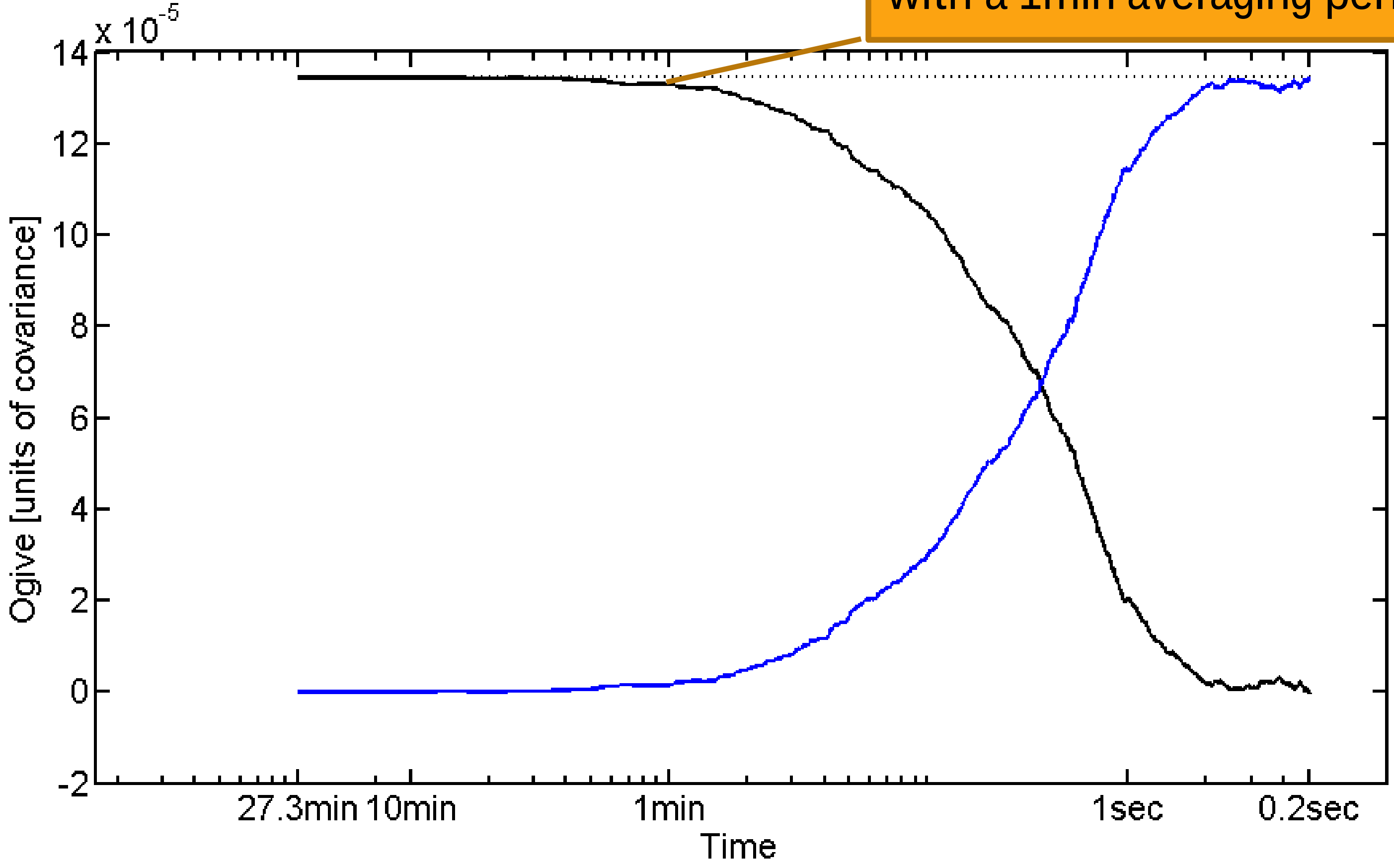
Ogive = cumulatice co-spectrum

Usually used only sometimes when trying to decide for the adequate averaging period

---

The adequate averaging period depends on surface roughness, measurement height, atmospheric stability…

Most of the flux is already reached with a 1min averaging period!



References

Aubinet M, Grelle A, Ibrom A, Rannik Ü, Moncrieff J, Foken T, Kowalski AS, Martin PH, Berbigier P, Bernhofer C, Clement R, Elbers J, Granier A, Grunwald T, Morgenstern K, Pilegaard K, Rebmann C, Snijders W, Valentini R, Vesala T (2000) Estimates of the annual net carbon and water exchange of forests: The EUROFLUX methodology. Adv.Ecol.Res. 30:113-175.

Barlow JF, Harrison J, Robins AG, Wood CR (2011) A wind-tunnel study of flow distortion at a meteorological sensor on top of the BT tower, london, UK. J.Wind Eng.Ind.Aerodyn. 99:899-907.

Horst TW (1997) A simple formula for attenuation of eddy fluxes measured with first-order-response scalar sensors. Bound.-Layer Meteorol. 82:219-233.

Nordbo A, Järvi L, Vesala T (2012) Revised eddy covariance flux calculation methodologies - effect on urban energy balance. Tellus Ser. B-Chem. Phys. Meteorol. 64:18184. http://dx.doi.org/10.3402/tellusb.v64i0.18184

Nordbo A, Launiainen S, Mammarella I, Leppäranta M, Huotari J, Ojala A, Vesala T (2011) Long-term energy flux measurements and energy balance over a small boreal lake using eddy covariance technique. J. Geophys. res. 116:1-17. doi:10.1029/2010JD014542

Vickers D and Mahrt L (1997) Quality control and flux sampling problems for tower and aircraft data. J.Atmos.Ocean.Technol. 14:512-526.

Wilczak JM, Oncley SP, Stage SA (2001) Sonic anemometer tilt correction algorithms. Bound.-Layer Meteorol. 99:127-150.

Thanks!

Annika Nordbo

Department of physics

P.O.Box 48

FI – 00014 University of Helsinki

annika.nordbo@helsinki.fi

+358 50 415 4830

www.mv.helsinki.fi/home/nordbo/index.html

Some unit conversions

$$C_{air} = \frac{\rho}{M_{air}} = \frac{1.25\ \mathrm{kg\ m^{-3}}}{28.965\cdot 10^{-3}\ \mathrm{kg\ mol_{air}^{-1}}} \approx 43.16\ \mathrm{mol_{air}\ m^{-3}} \quad \text{molar density of air}$$

$$F_{CH4} = \overline{w'c'}\cdot C_{air} = 2.41\cdot 10^{-3}\ \frac{\mathrm{mol_{CH4}}}{\mathrm{mol_{air}}\cdot 10^{6}}\ \mathrm{m\ s^{-2}}\cdot 43.16\ \mathrm{mol_{air}\ m^{-3}} \approx 1.04\times 10^{-10}\ \frac{\mathrm{mol_{CH4}}}{\mathrm{m^{-2}s}} \quad \text{methane flux}$$

$$F_{CH4} = \overline{w'c'}\cdot C_{air}\cdot M_{CH4} = 2.41\cdot 10^{-3}\ \frac{\mathrm{mol_{CH4}}}{\mathrm{mol_{air}}\cdot 10^{6}}\ \mathrm{m\ s^{-2}}\cdot 43.16\ \mathrm{mol_{air}\ m^{-3}}\times 16.04\times 10^{-3}\ \mathrm{kg\ mol_{CH4}^{-1}}\cdot 3600\ \frac{\mathrm{s}}{\mathrm{h}}$$

$$\approx 6.01\times 10^{-9}\ \frac{\mathrm{g_{CH4}}}{\mathrm{m^2h}} = 6.01\ \frac{\mathrm{\mu g_{CH4}}}{\mathrm{m^2h}}$$

$$E = \frac{\overline{w'q'}}{M_{water}} = \frac{5\cdot 10^{-5}\ \mathrm{kg\ m^{-2}\ s^{-1}}}{18.0153\cdot 10^{-3}\ \mathrm{kg\ mol_{water}^{-1}}} = 2.78\times 10^{-3}\ \frac{\mathrm{mol_{water}}}{\mathrm{m^2\ s}} \quad \text{water vapor flux}$$

$$LE = L\overline{w'q'} = (3147.5 - 2.37\cdot \mathrm{T_{air}})\cdot 1000\cdot 5\cdot 10^{-5}\ \mathrm{kg\ m^{-2}\ s^{-1}} = 125\ \mathrm{W\ m^{-2}} \quad \text{latent heat flux}$$

Matlab code for making fig on slide 21

```
Fs =10;                                     % sampling frequency [Hz]
L = 6^2;                                    % length of signal
t = (0:0.1:L)./Fs;                          % time vector
x = cos(2.*pi.*t)  +  0.5.*cos(10.*pi.*t);  % artificial data consisting of two cosine waves
f = Fs.*(0:0.1:L/2);                        % frequency
```

```
% Fast fourier transform for the first 6^2 elements gives a vector with a real and an imaginary part.
% Each point corresponds to a number of cycles/period.
fftx = fft(x);
% the firts value is useless
fftx(1)=fftx(2);
% the second half is just the mirror of the first half so it can also be deleted
fftx(length(fftx)/2+1:end)=[];
% the power is the square of the absolute value of the fourier amplitudes. abs(fftx).^2 = real(fftx).^2+imag(fftx).^2
Power = abs(fftx).^2./L;
figure(1)
subplot(2,1,1)
plot(t,cos(2.*pi.*t),':k',t,0.5.*cos(10.*pi.*t),'--k',t,x,'-k'); xlabel('Time [s]'); ylabel('Time series'); legend('almost 4 cycles/period','18 cycles/period','sum of waves')
subplot(2,1,2)
semilogx(f,Power,'-k'); xlabel('Cycles per period'); axis tight; ylabel('Power of variation')


% figure has a clear spike at 18 cycles/period and a less clear spike at 4
% cycles/period as should be expected from x, line 4
```