Lecture 5 Data processing 1

Acknowledgments to

- Prof HP Schmid, Indiana University
- Dr John Finnigan, CSIRO
- Dr Eva van Gorsel, CSIRO
- Dr Vanessa Haverd, Univ. Wollongong
- Dr Helen Cleugh, CSIRO

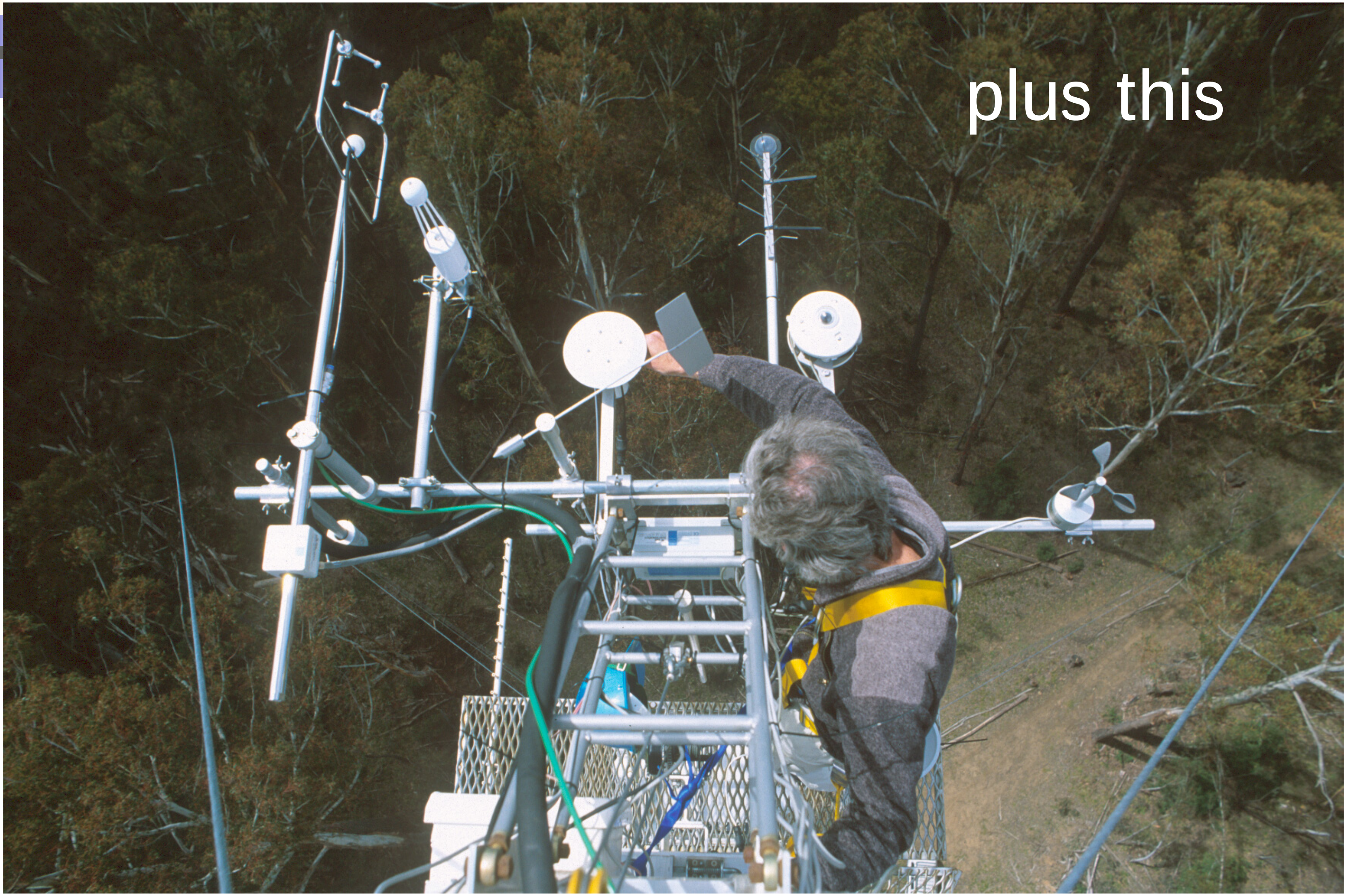
Final Flux Calculations and Validity of the Results

- Need to process raw data to produce final flux results before we address
- Quality control
  - Eliminate bad data
  - Fill the gaps
- Interpreting the data
  - Ecosystem processes
  - Micrometeorology
  - Extrapolation?

Tumbarumba mast

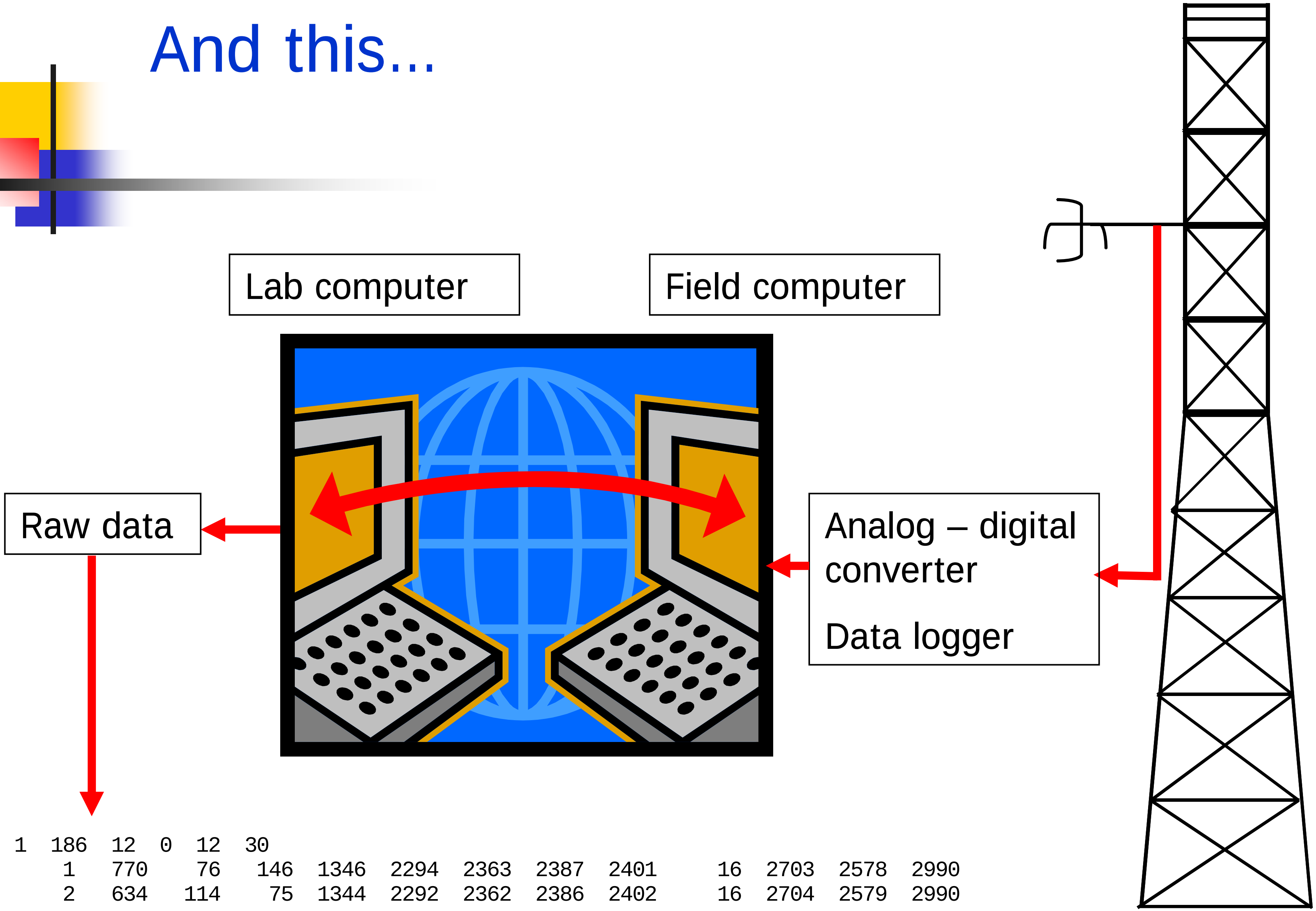
How do we get from this

Tumbarumba mast instruments



plus this

And this...



Lab computer

Field computer

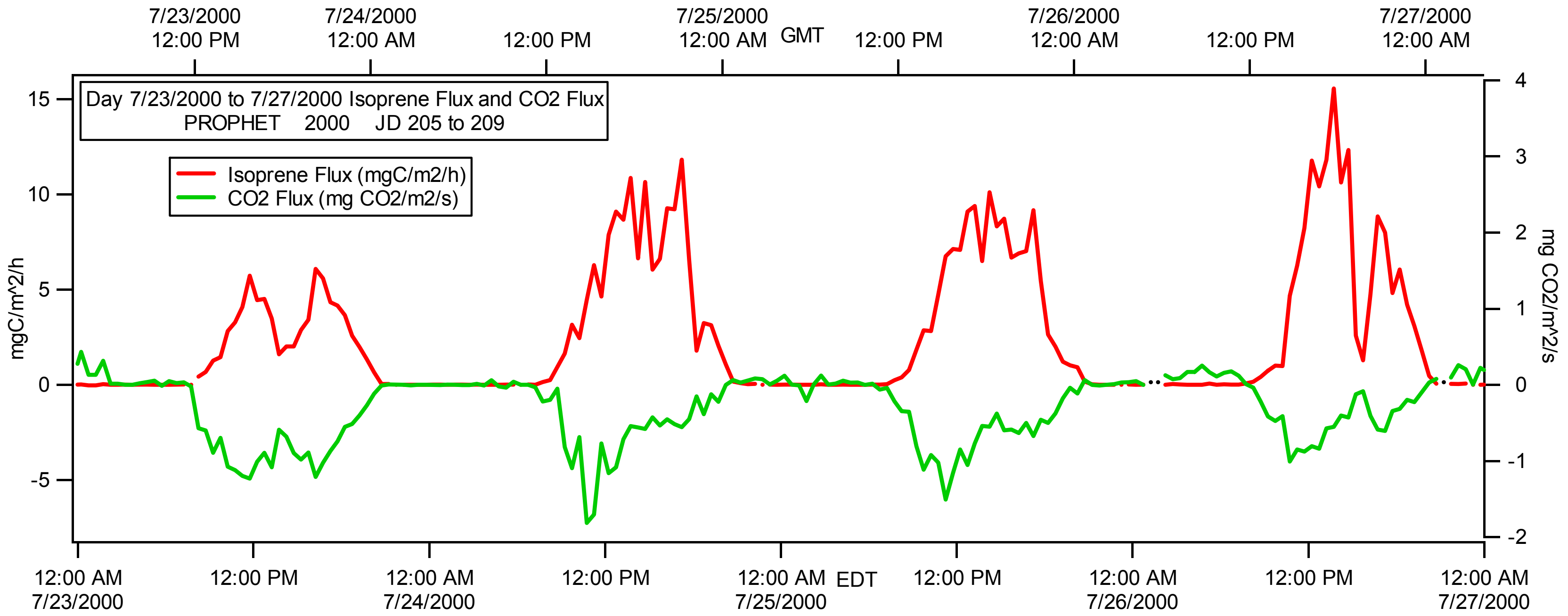
Raw data

Analog – digital converter

Data logger

```
1  186  12  0  12  30
   1  770   76   146  1346  2294  2363  2387  2401     16  2703  2578  2990
   2  634  114    75  1344  2292  2362  2386  2402     16  2704  2579  2990
   3  530  179    52  1343  2290  2362  2386  2401     16  2703  2578  2990
   4  464  230    53  1337  2289  2362  2386  2401     12  2704  2578  2990
   5  478  261    15  1341  2294  2363  2386  2401     12  2703  2579  2990
```

To Fluxes….such as these?



And all the steps in between

Steps along the way

- Despiking (quality of raw data)
- Calibration and conversion to real units (see lecture 4)
- Coordinate rotation
- Averaging/detrending/filtering
- Determining lag times
- Frequency response corrections
- WPL Corrections (see lecture 3)

Despiking...

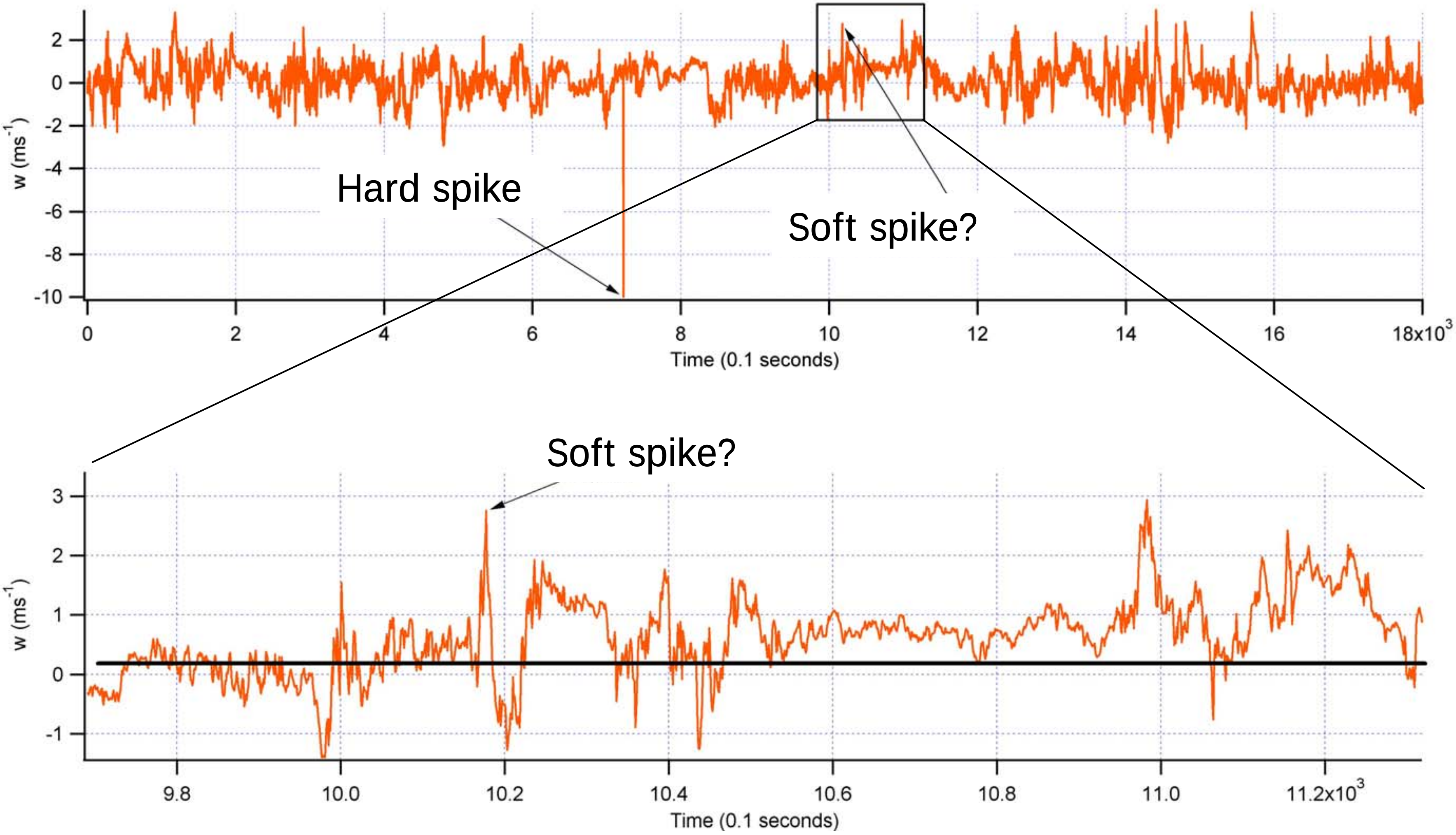
- Removing spikes from the data that are caused by:
  - ✓ Blocking the path of the sensor (i.e. precipitation, spider webs, bird droppings)
  - ✓ Large short-lived departures from the mean, usually caused by instrument errors

These are sometimes called "Hard" spikes

These are called "Soft" spikes

(Schmid, et al. 2000; Vickers and Mahrt, 1997)

Determining spikes...



Hard spike

Soft spike?

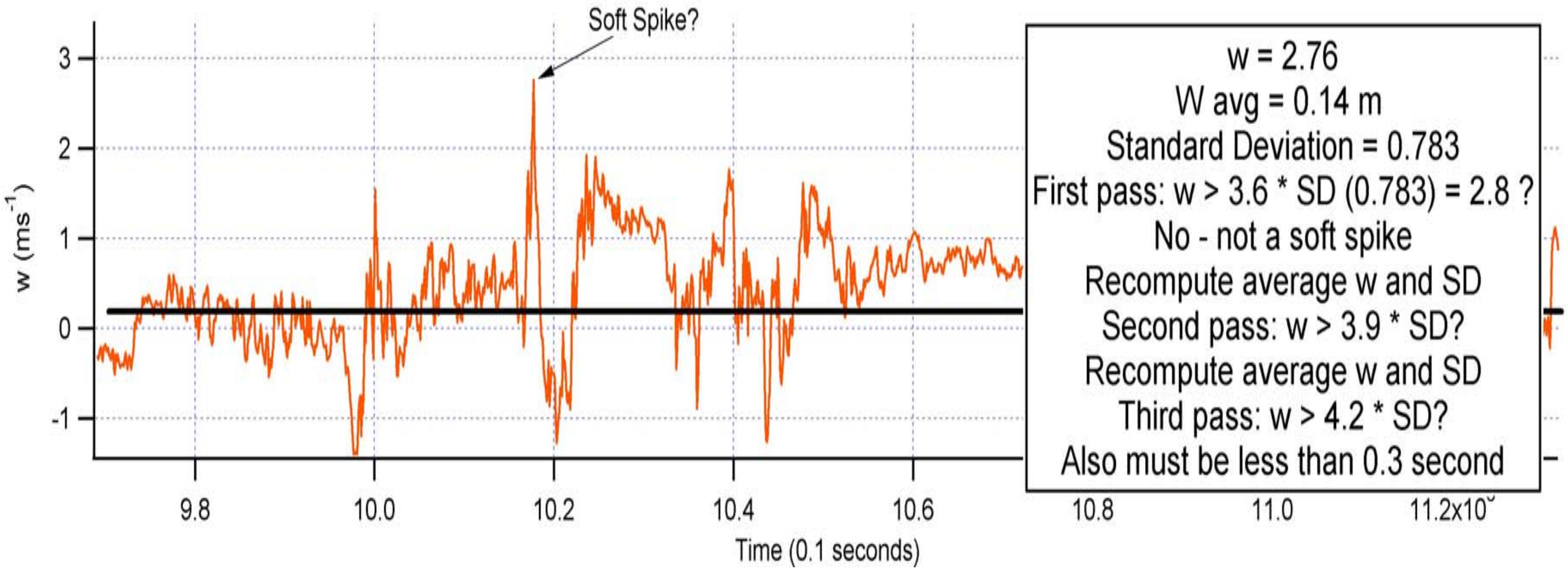
Soft spike?

Despiking con't...

- Hard spikes easily detected and rejected
  - A flag with the sonic software
  - An extreme digital signal (power failure)
- Soft spikes detected by an iterative process
  - Calculate mean & standard deviation, s.d. for averaging period
  - Spike threshold = 3.6*s.d. initially, increased by 0.3 after each pass.
  - Set soft spike flag when signal > threshold and < 0.3 s

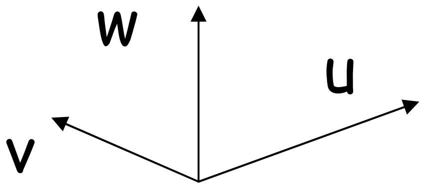
(Schmid, et al. 2000)

No soft spike here!



w = 2.76
W avg = 0.14 m
Standard Deviation = 0.783
First pass: w > 3.6 * SD (0.783) = 2.8 ?
No - not a soft spike
Recompute average w and SD
Second pass: w > 3.9 * SD?
Recompute average w and SD
Third pass: w > 4.2 * SD?
Also must be less than 0.3 second

Coordinate rotation (Chapter 3)



- Forces $\overline{w} = 0$
- Orients $\overline{u}$ along the mean wind
- Maximizes gradients normal to surface
- Removes anemometer tilt errors
- Keep in mind – it is also a high pass filter!

(Handbook of Micrometeorology, 2004)

(Kaimal and Finnigan, 1994)

Coordinate rotation ...

- Instrument Coordinate
  - Orthogonal coordinate frame employed by the sonic anemometer
  - Absolute one, and independent of the flow field
  - Always archive the data!

(Handbook of Micrometeorology, 2004)

Coordinate rotation ...

- Planar Fit Coordinate (Chapter 3, Handbook)
  - z-axis perpendicular to the mean streamline plane
  - y-axis perpendicular to the plane of the short-term u and z axis
  - Use multiple linear regression of *w* vs *u* and *v* using *long-term* measurements to obtain planar fit

$$\overline{w} = \overline{w_m} - (a + b\overline{u} + c\overline{v})$$

current $\overline{\mathrm{w}}$

long-term $\overline{\mathrm{w}}$

Regression coefficients

(Handbook of Micrometeorology, 2004)

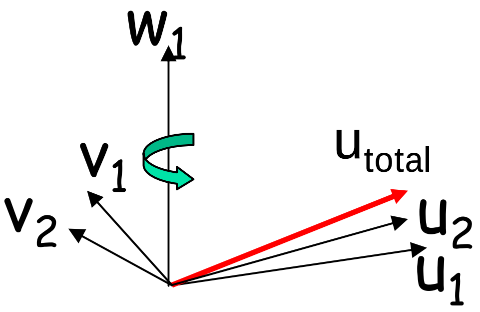
Coordinate rotation...

- Natural Wind Coordinate (short term)
  - x-axis is parallel to the (60-min) mean flow
  - z-axis is normal to surface
  - Each period is processed individually

(Handbook of Micrometeorology, 2004)

Natural wind (short term) coordinate rotation

- First rotation—x_1 and y_1 coordinates around z_1
- Mean $v_2 = 0$
- θ = mean wind direction during 60 min period



$$u_2 = u_1 \cos\theta + v_1 \sin\theta$$

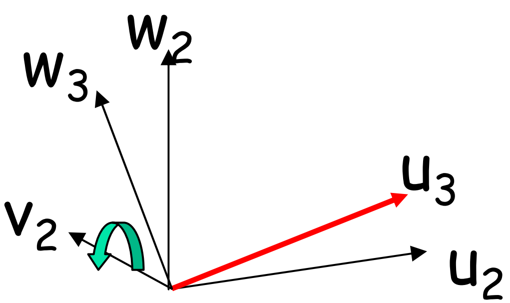
$$v_2 = -u_1 \sin\theta + v_1 \cos\theta = 0$$

$$w_2 = w_1$$

$$\theta = \tan^{-1}\left(\frac{\overline{v_1}}{\overline{u_1}}\right)$$

Coordinate rotation

- Second rotation
  - x_2 and z_2 coordinates around y_2
- Mean $v_3=0$
- Mean $w_3=0$
- Mean $u_3=U_{total}$ aligned along the mean wind direction



$$u_3 = u_2 \cos\phi + w_2 \sin\phi$$

$$w_3 = -u_2 \sin\phi + w_2 \cos\phi$$

$$v_3 = v_2$$

$$\phi = \tan^{-1}\left(\frac{\overline{w_2}}{\overline{u_2}}\right)$$

Natural wind vs. Planar fit

- Planar fit overcomes problems associated with the natural wind coordinate system
  - over-rotation,
  - loss of information
  - degradation of data quality
- Planar fit (or related method)
  - requires data for several weeks with no movement of sonic anemometer
  - to determine 'tilted plane' (pitch, roll and yaw angles)
- Sample dataset comparison indicates that the natural wind system underestimates the flux by ~4% (Schmid)

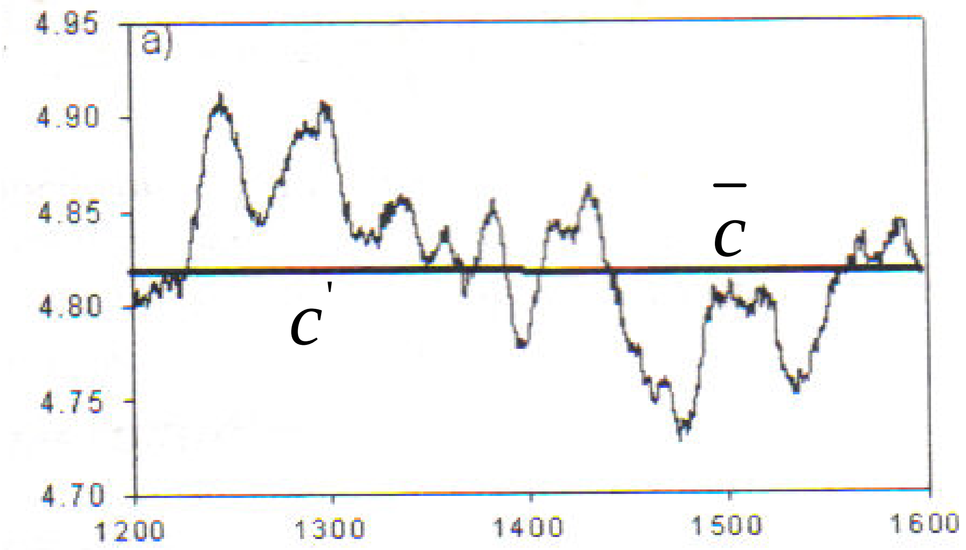
(Handbook of Micrometeorology, 2004)

Averaging and filtering (Chapter 2)

- Used to separate the turbulent signals from low frequency components caused by
  - Instrumental drift
  - Changes in meteorological conditions
- Three main types of operations (time averaging, detrending, and filtering)

(Handbook of Micrometeorology, 2004)

Time averaging



- Obeys Reynolds averaging and is simple
- Well approximated by running mean filter if averaging time T >> period of any fluctuations

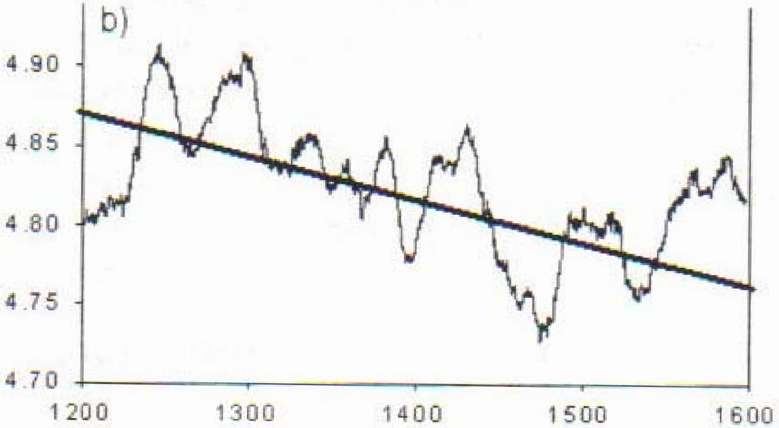
$$w = w' - \bar{w}, \quad c = \bar{c} + c'$$

$$\overline{wc} = \overline{(\bar{w} + w')(\bar{c} + c')}$$

$$= \bar{w}\bar{c} + \overline{w'c'}$$

(Handbook of Micrometeorology, 2004)

Linear detrending

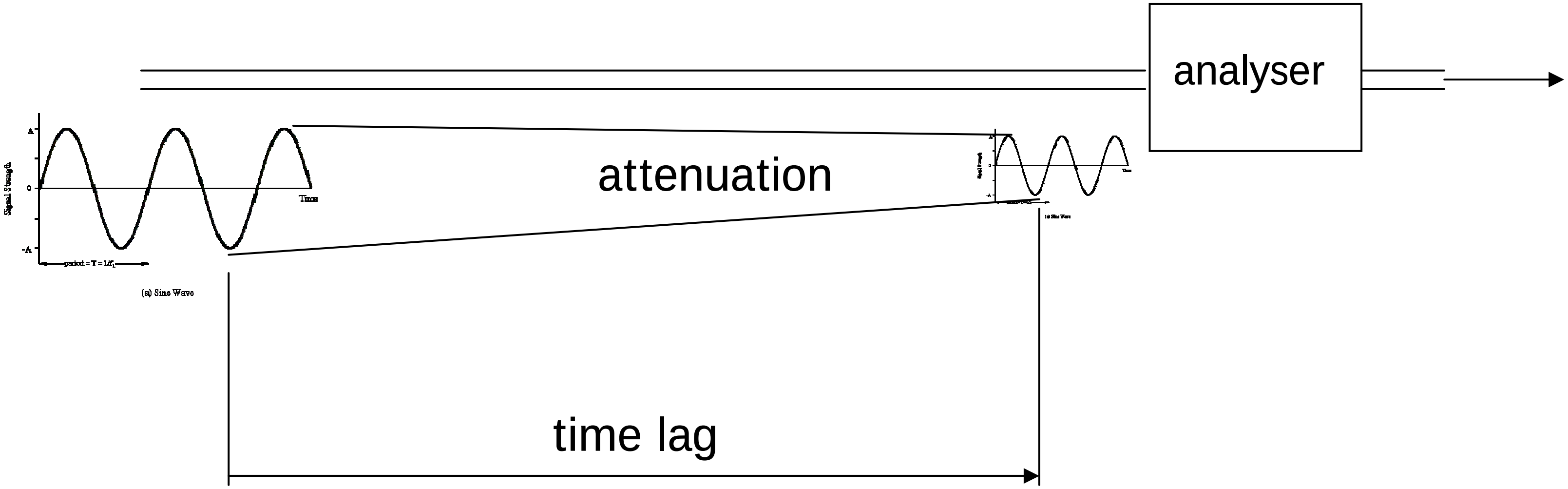


- Find the best linear fit over the period and subtract that from each value

- Mainly affects the low frequency part of the signal, but it affects all frequencies
- Does not obey Reynolds averaging
- **Not recommended!**

(Handbook of Micrometeorology, 2004)

Closed-path gas sampling



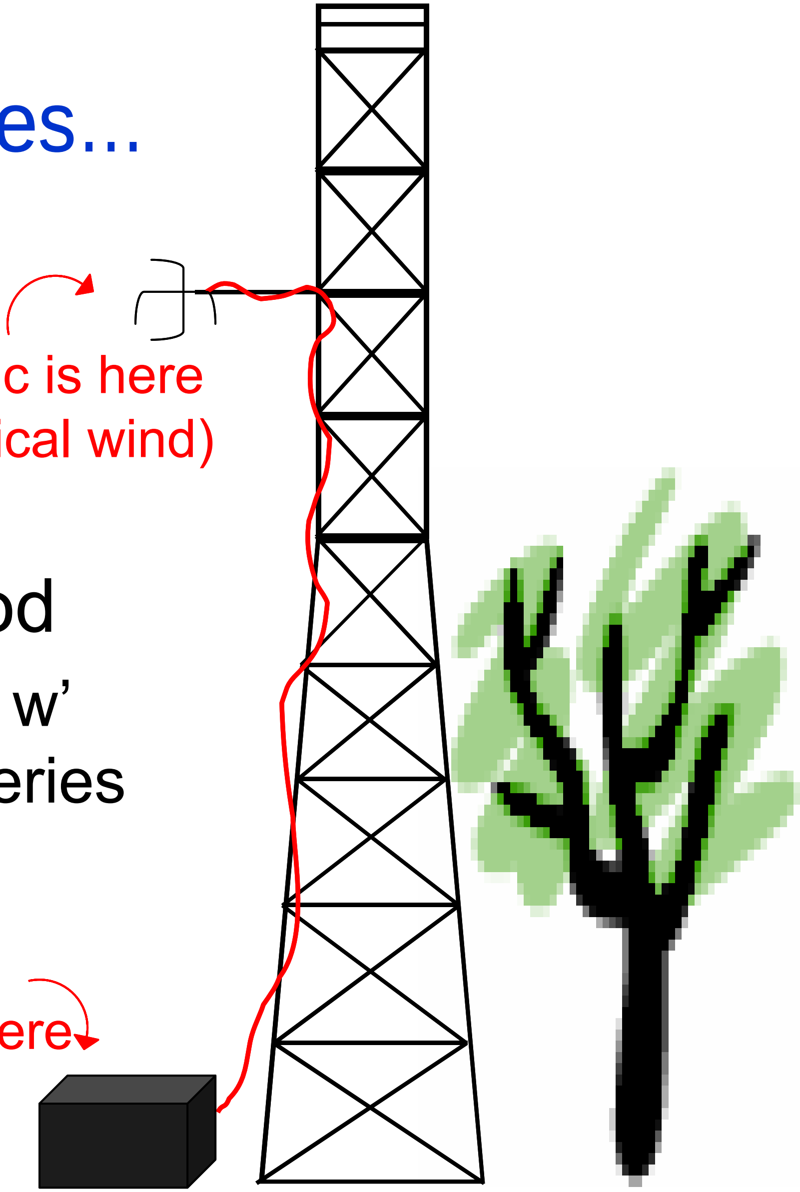
analyser

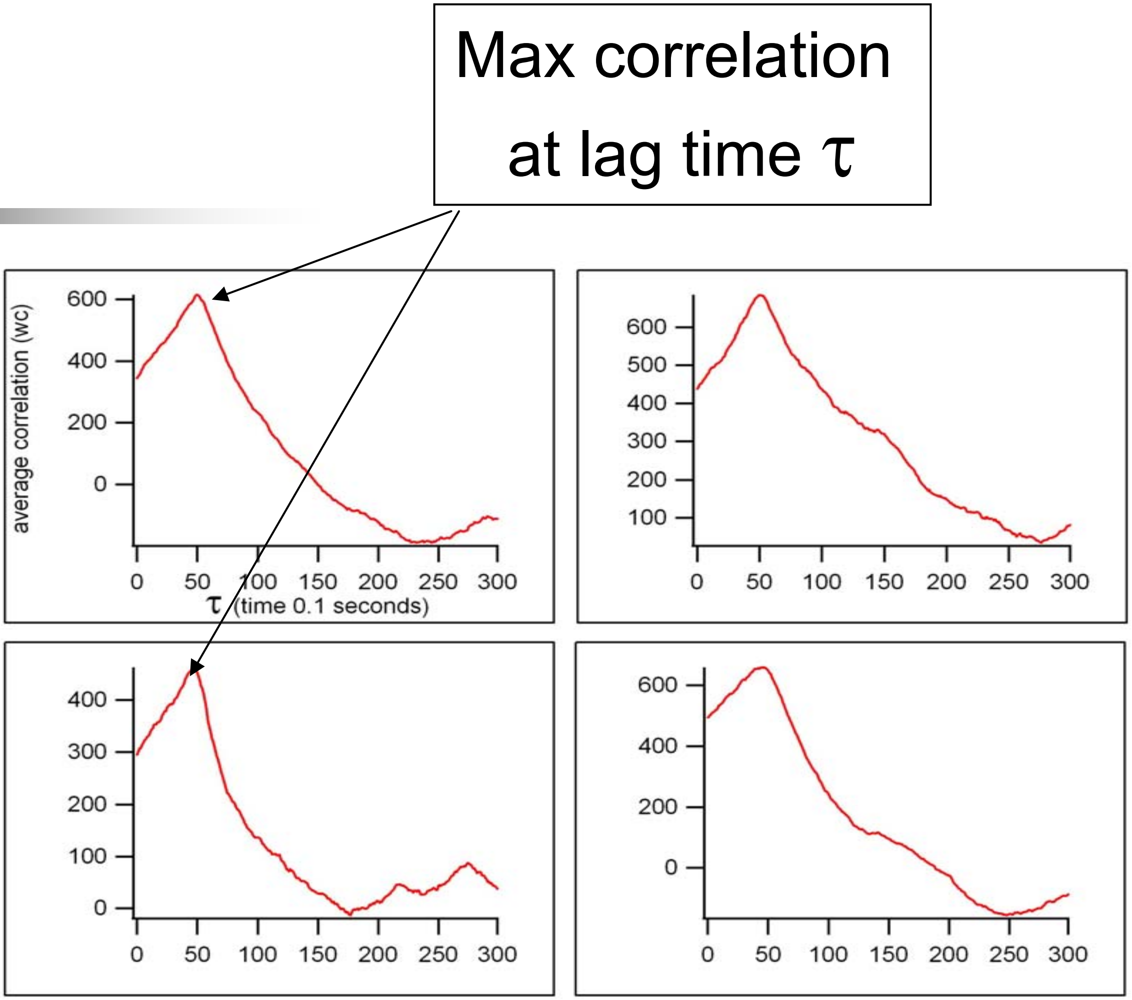
attenuation

time lag

Determining lag times...

- Maximum correlation method
  - determines lag time between w' time series and scalar time series




Max correlation at lag time τ

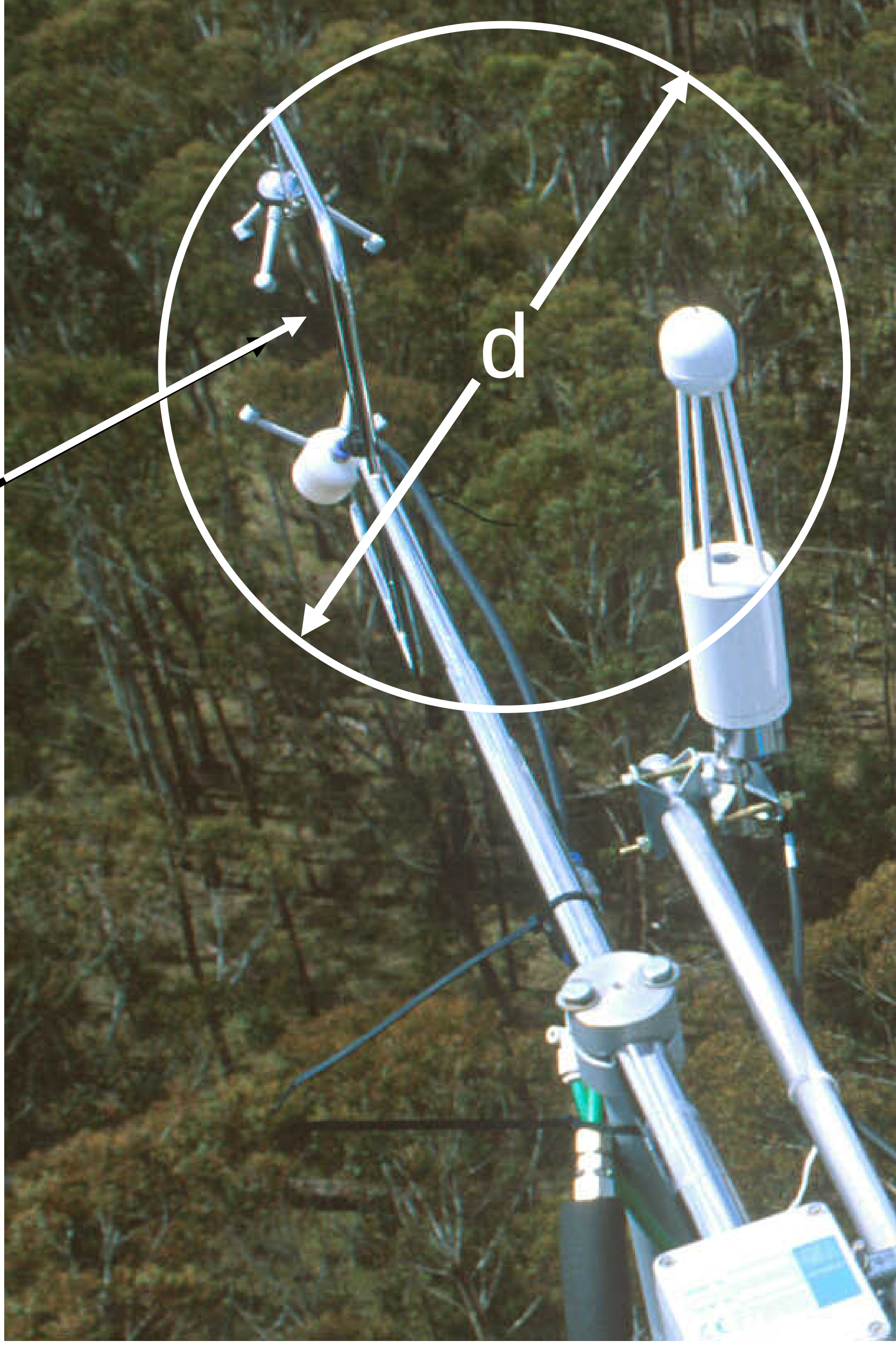
Average time lag = 47.25 or 4.7 seconds

Causes of High Frequency Attenuation

- Slow response of scalar sensors
  - Time constants > 0.1 s
- Errors largest at low wind speeds

(Su, et al. 2004; Massman, 2000)

High Frequency Attenuation - Open path



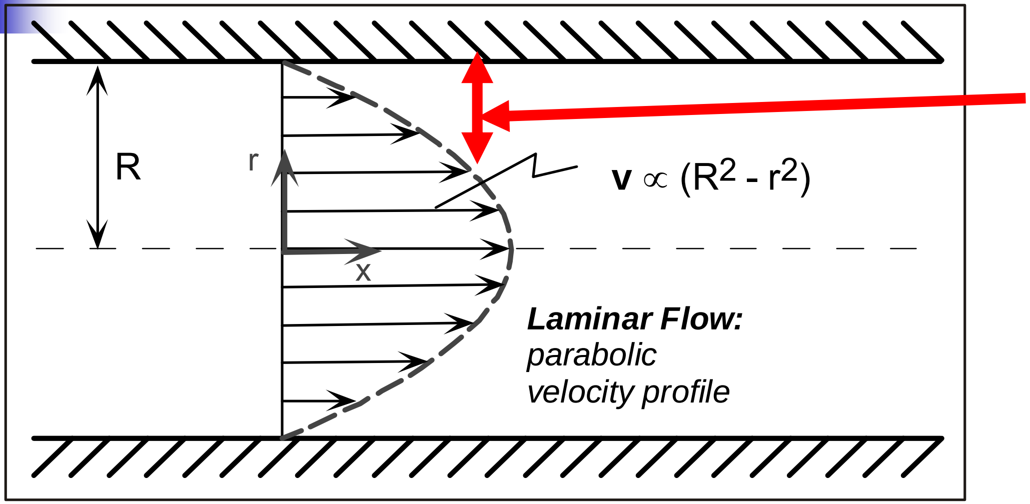
- Line-averaging along instrument path
- Spatial separation between instruments
- Samples eddies > ~2d

High Frequency Attenuation - Closed path

- Tubing acts like a low-pass filter by mixing the air
- Higher frequencies strongly attenuated – depends on:
  - Flow rate through tube
  - Tube diameter, length and material

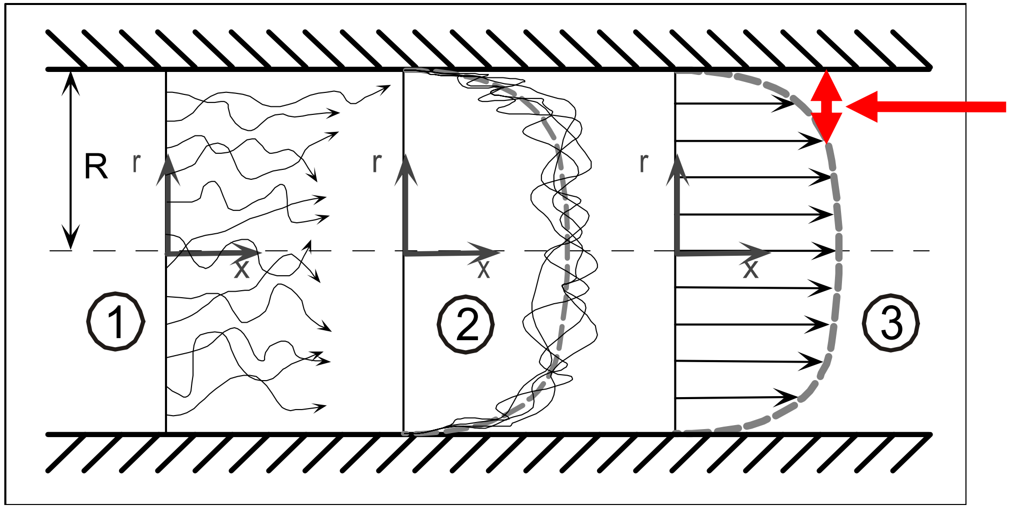
(Leuning and Moncrieff, 1990; Leuning & Judd 1996)

Flow in pipes



Mixing

Laminar



Mixing

Turbulent

Frequency Response Corrections...con’t

Define correction factor

$$C_F = \frac{\int_0^\infty C_{wc}(f))df}{\int_0^\infty G_{wc}(f)C_{wc}(f))df}$$

‘true’ cospectrum, eg w’T’

filtered cospectrum

filter function

(Leuning and Moncrieff, 1990; Leuning & Judd 1996)

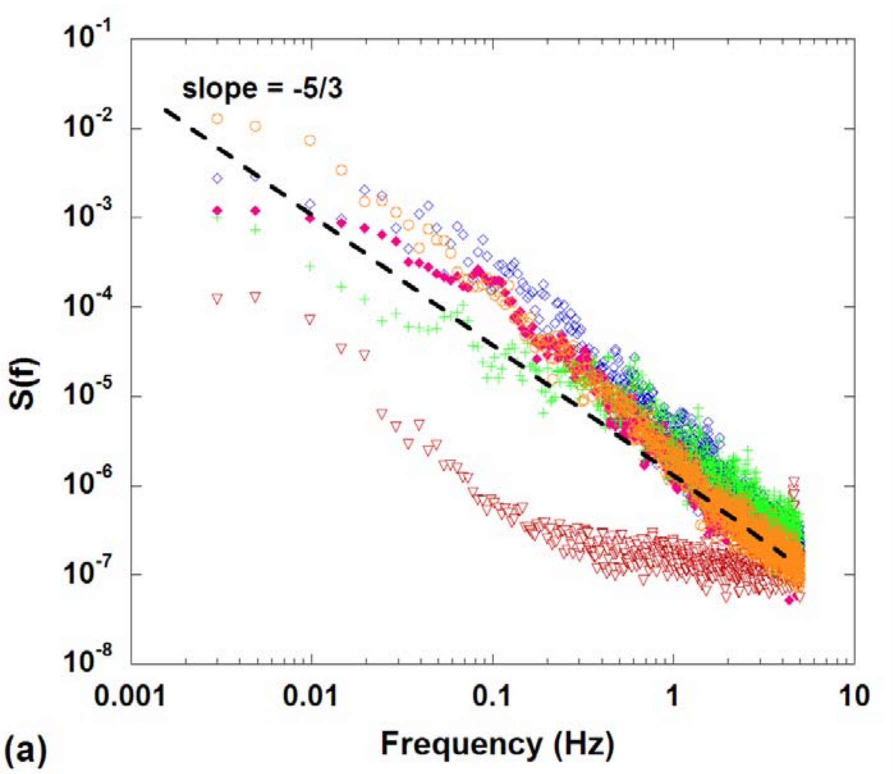
Magnitude of Corrections...

- Losses depend on
  - Ratio of separation distance to measurement height (d_{xy}/h_{m})
  - Atmospheric stability (h_{m}/L)
  - Windspeed
- Losses
  - close to ground (>10%)
  - over forests (< 1-2%)

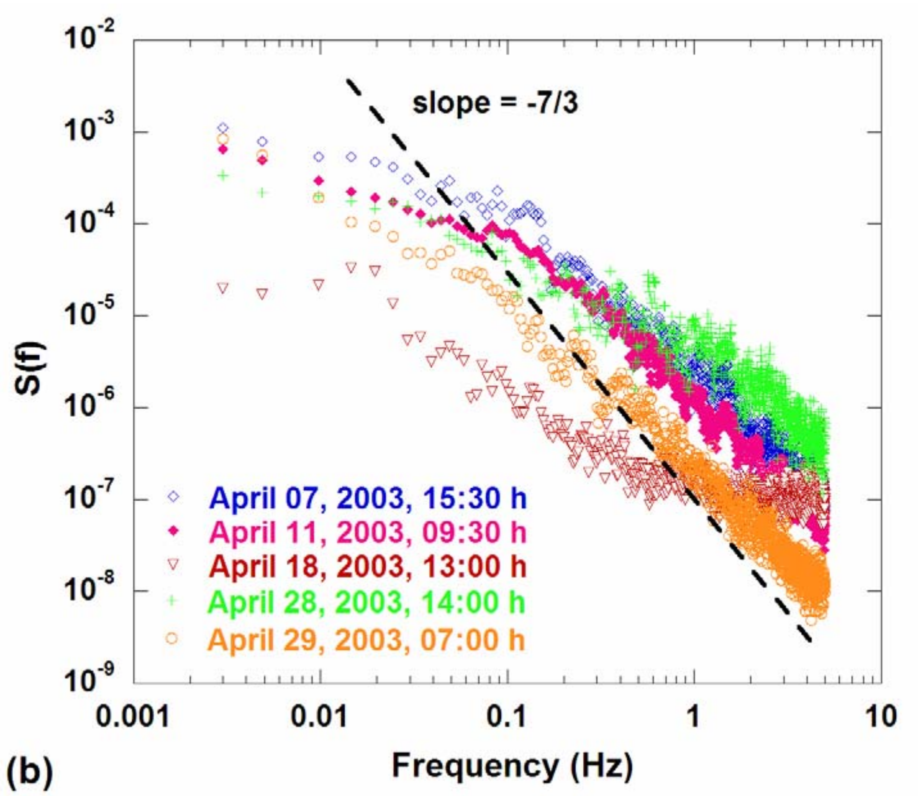
(Su et. al, 2004; Webb Pearman and Leuning, 1980)

Spectral and Co-spectral Analysis

Spectral and co-spectral analyses demonstrate the expected -5/3 and -7/3 slopes in the inertial subrange



Spectra for CO$_2$ mixing ratio



Cospectra for CO$_2$ mixing ratio and vertical wind speed.

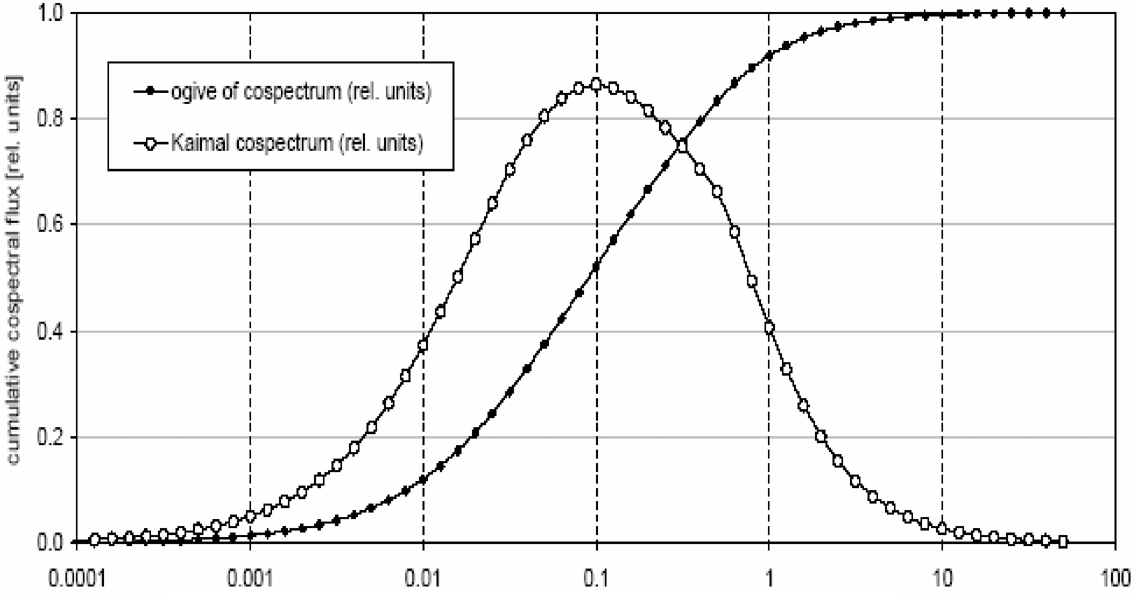
Ogive Analysis

- Statistics: cumulative frequency distribution curve
- Atmospheric turbulence: cumulative cospectrum (or power spectrum)
- Ogive analysis is the integral of the spectral analysis

(Oncley, et al. 1996; Ammann and Neftel ??)

Ogive analysis

Cumulative cospectrum



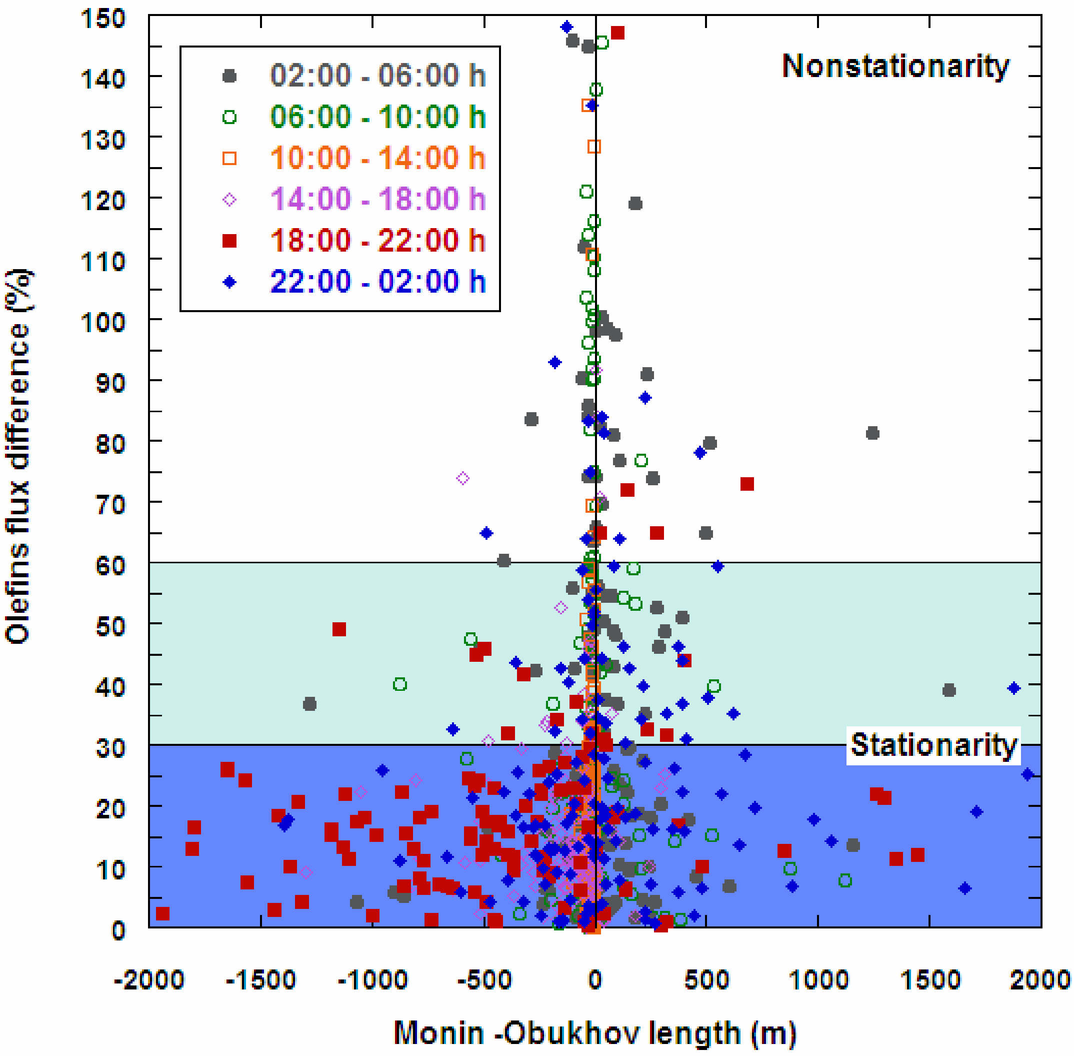
Normalized frequency n= fz/u

Provide visualization of spectral information

- Proportional to the flux contribution
- With integral smoothing effect

(Oncley, et al. 1996)

Stationarity



One criterion for stationarity is if the average flux from 6 continuous subperiods of 5 min is within 60% of the flux obtained from a 30 min average.

In study by Schmid the stationarity condition was fulfilled in 82% of the half hour periods for olefin fluxes and 70% for CO_2 fluxes. Conditions of non-stationarity were related to very unstable or stable atmospheric conditions.

Foken et al Chapter 9

Summary: data processing

- This lecture has discussed the following data processing issues
- Despiking
- Coordinate rotation
- Averaging/detrending/filtering
- Determining lag times
- Frequency response corrections
- Statistical stationarity