ACADEMY OF FINLAND

RESOLVE CENTRE OF EXCELLENCE

UNIVERSITY of OULU
OULUN YLIOPISTO

FINNISH METEOROLOGICAL INSTITUTE

Aalto University

Long-term homogeneity and intercalibration of datasets

Timo Asikainen
ReSoLVE Centre of Excellence,
Astronomy and Space Physics Unit,
University of Oulu
timo.asikainen@oulu.fi

Contents

- Why do we need to homogenize and intercalibrate?
- Differences/changes in instrumentation, location of measurements, method of measurement etc.

- Basic methods of forming homogeneous data and inter-calibrating between datasets

- Homogenization = making sure your data describes the same thing in same way at all times

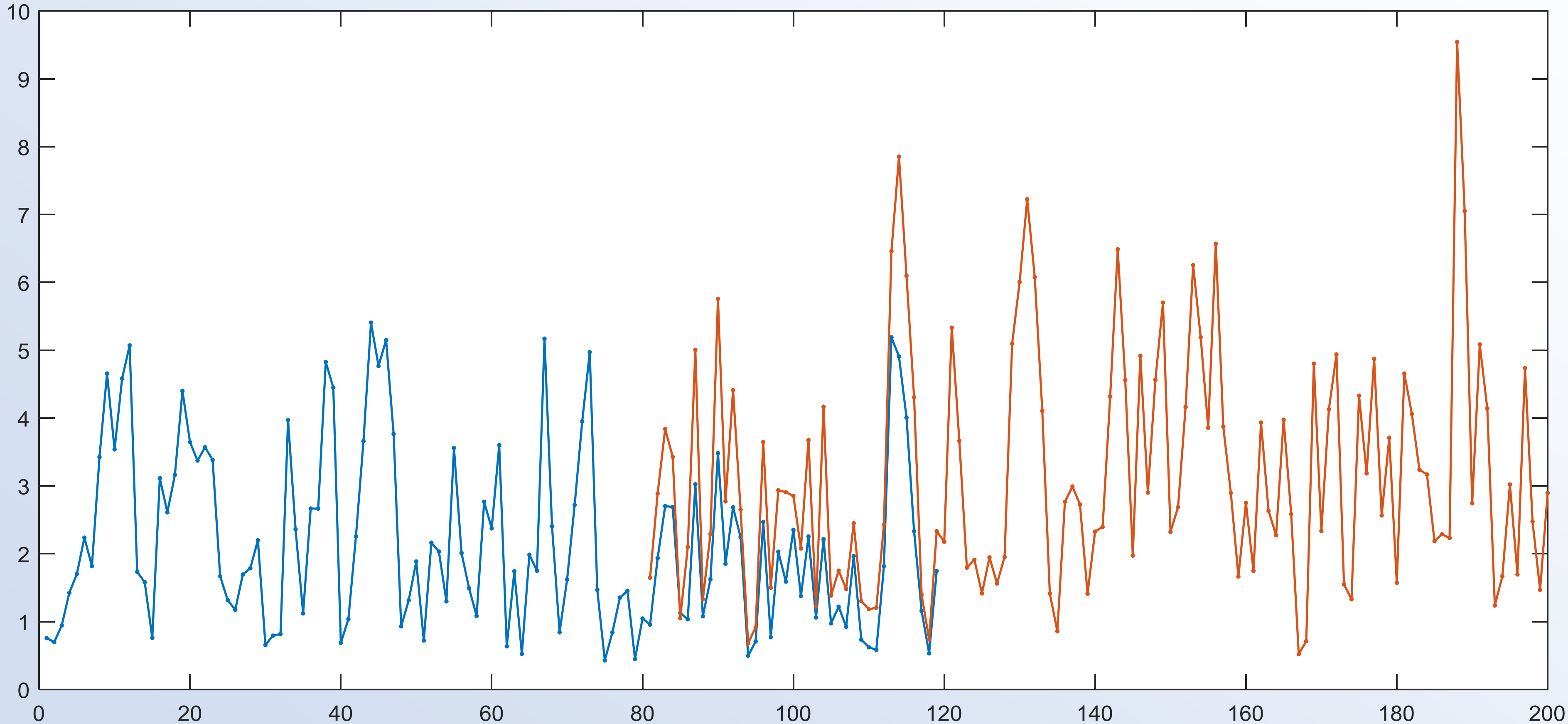
- Intercalibration = making sure two (or more) data series describe the same thing in same way at all times

- Foundation: understand your data and measurements!

- Examples:
- Geomagnetic indices
- Particle measurements from satellites

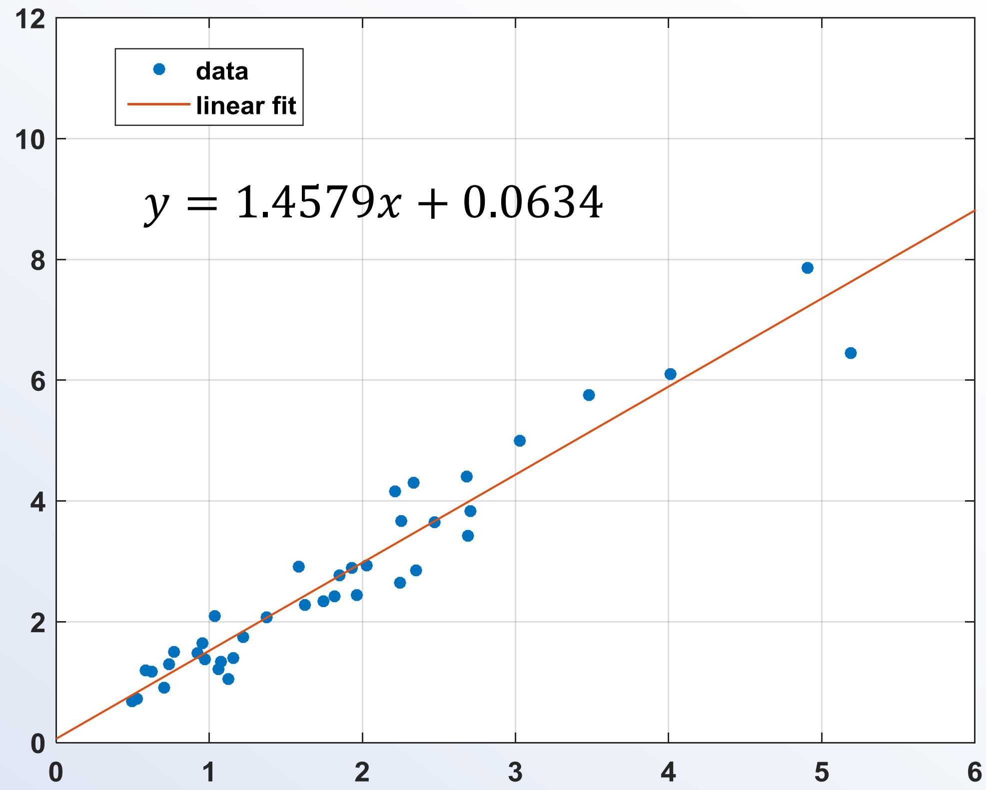
Compositing measurements: inter-calibration

- Blue series ends and red series begins
- Overlapping data differs ➔ two series systematically different ➔ Compare overlapping part



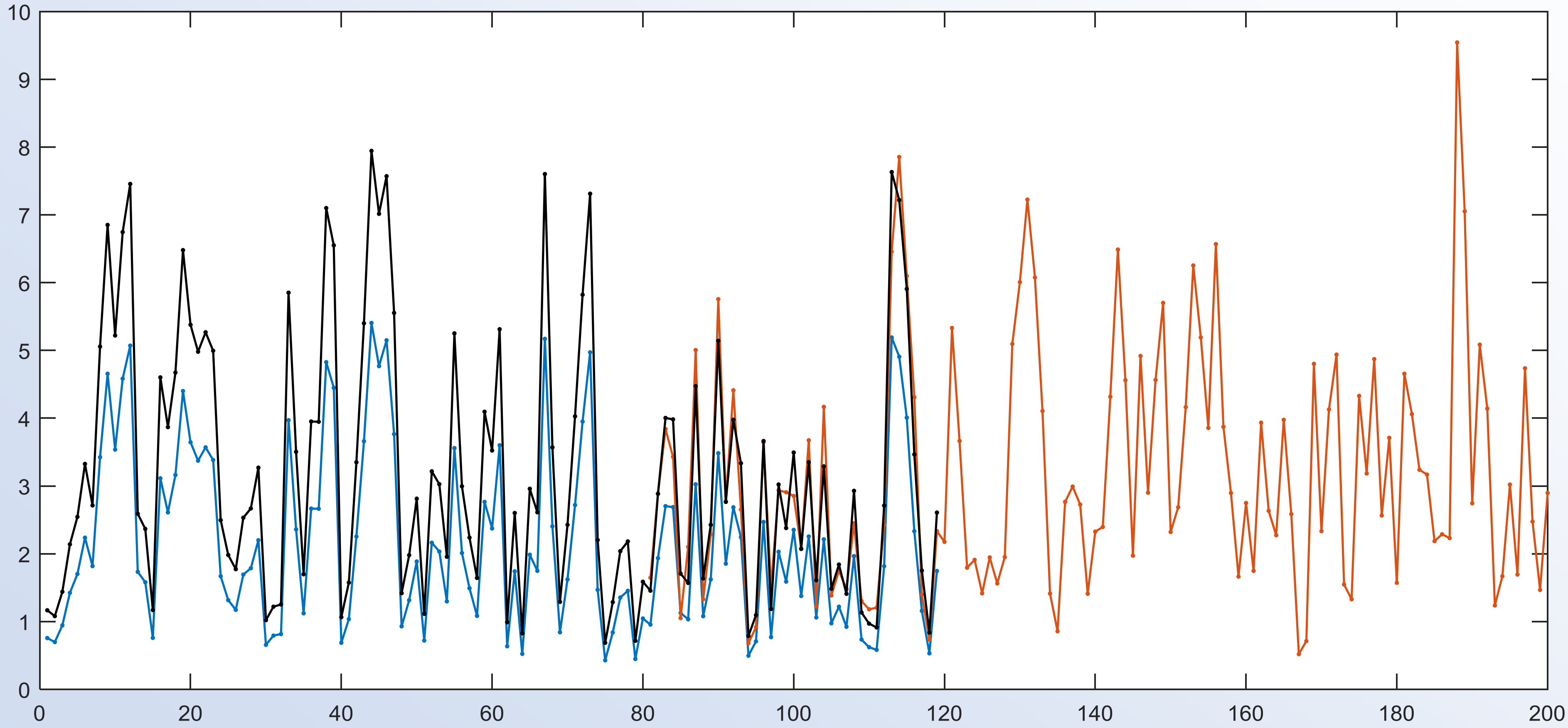
Compositing measurements: inter-calibration

- First make sure data are comparable!
- We assume here that individual series are homogeneous during their individual periodss
- Comparison by scatterplot ➔ fit a suitable curve (e.g., a line)



Compositing measurements: inter-calibration

- Estimate from blue series what would red series be if it had been measured in the past.

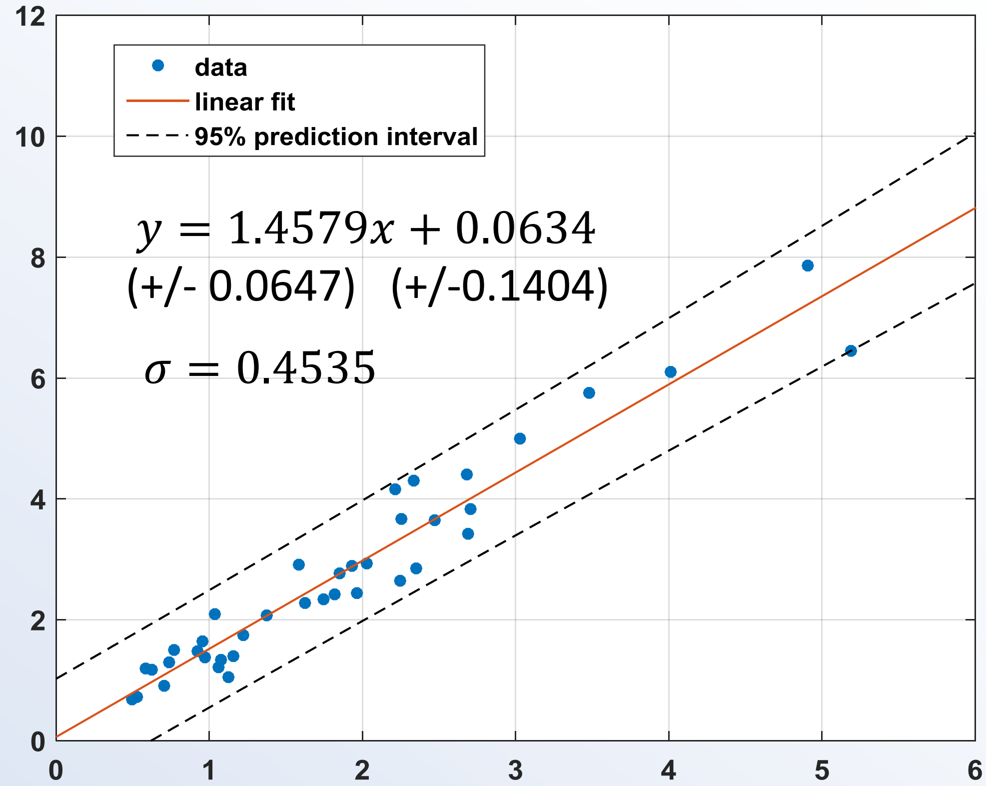


Compositing measurements: inter-calibration

- However, important to estimate error!
- Regression

$$y = \beta_0 + \beta_1 x + \varepsilon$$

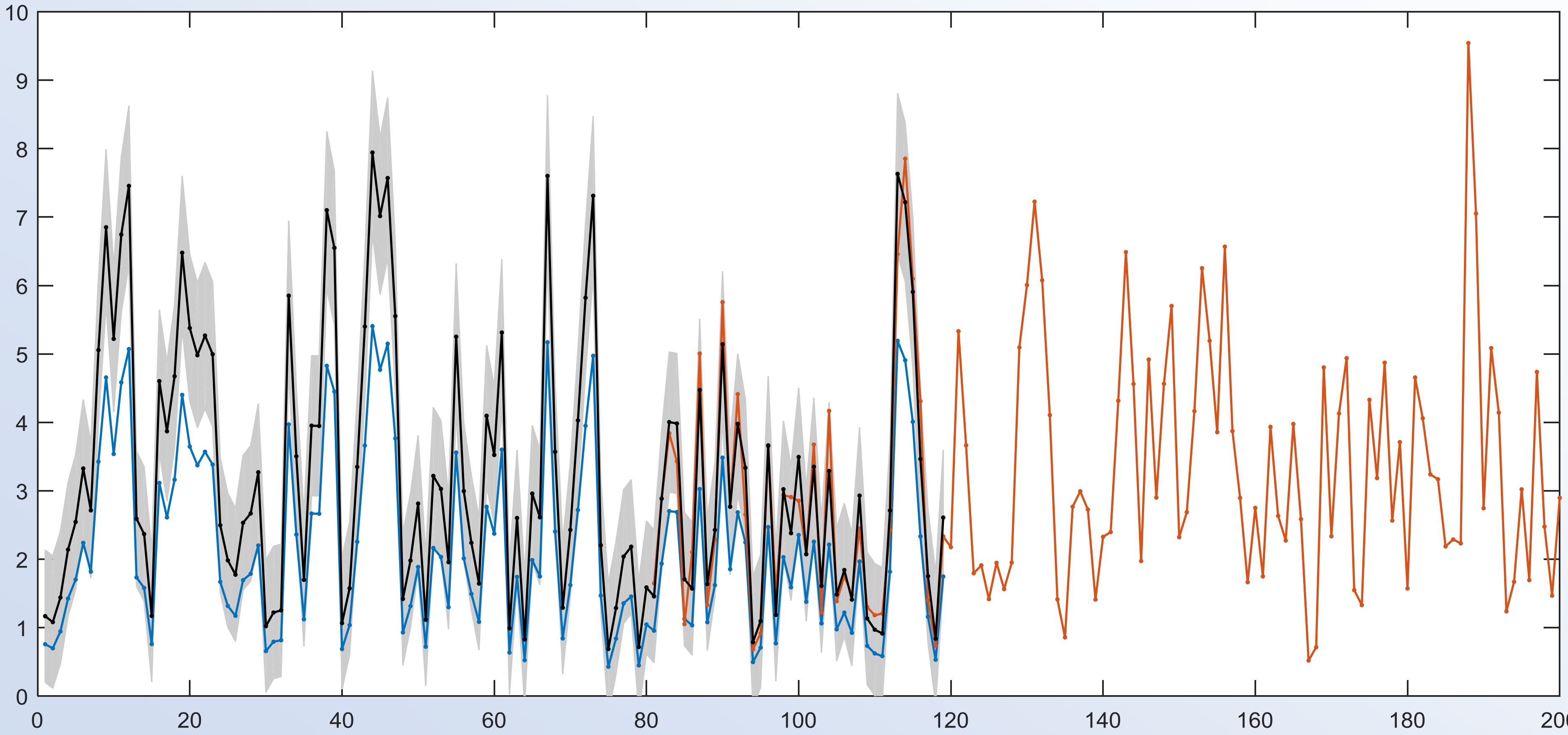
- $\varepsilon = N(0, \sigma)$
- Since σ is not known sampling distribution of ε is Student T-distr. With n-2 DF
- 95% prediction interval:



$$\Delta y = t_{0.975,n-2}\sqrt{var(\beta_0) + x^2 var(\beta_1) + \sigma^2}$$

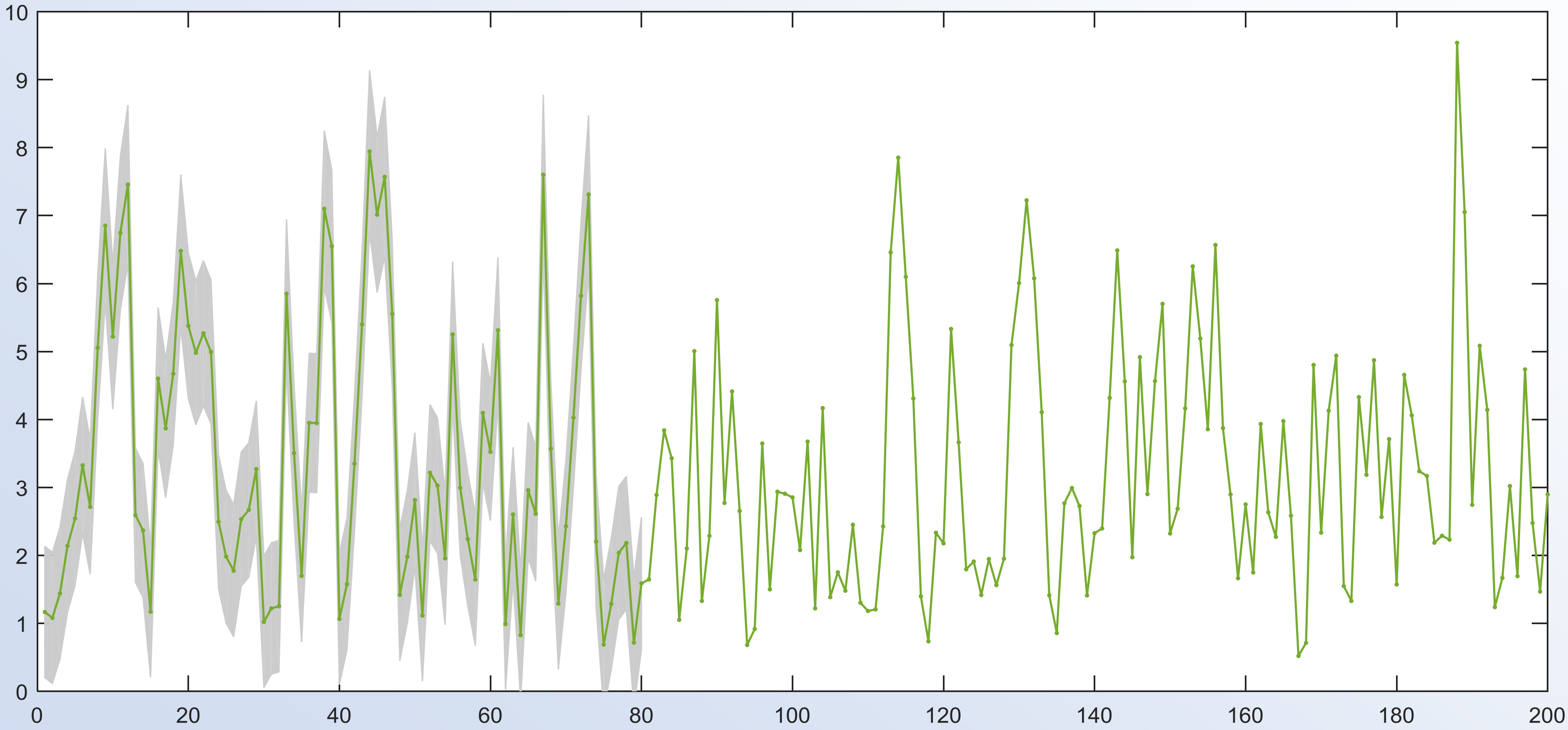
Compositing measurements: inter-calibration

- Composited data is now black and red series



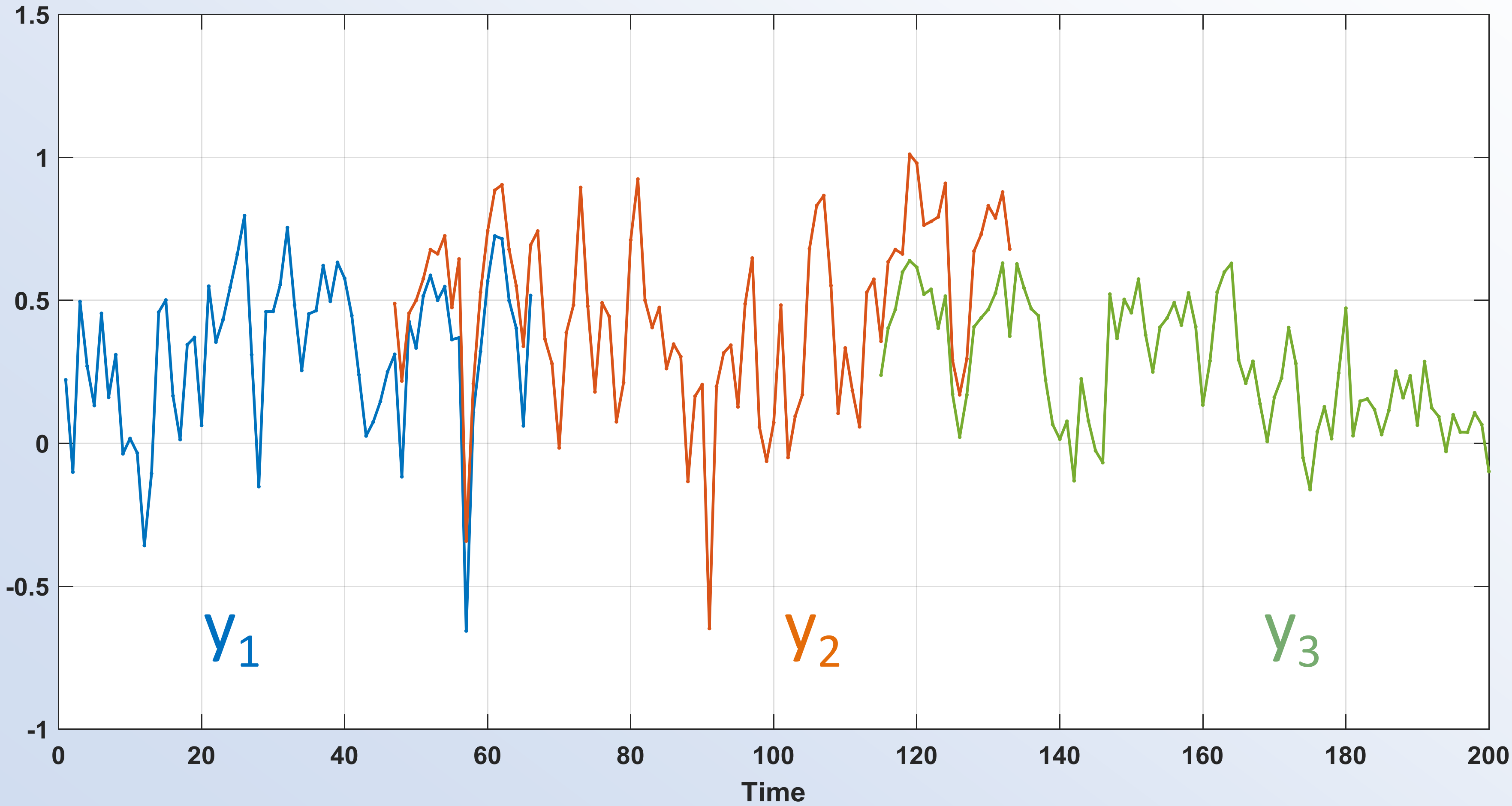
Compositing measurements: inter-calibration

- Composited data is now green series



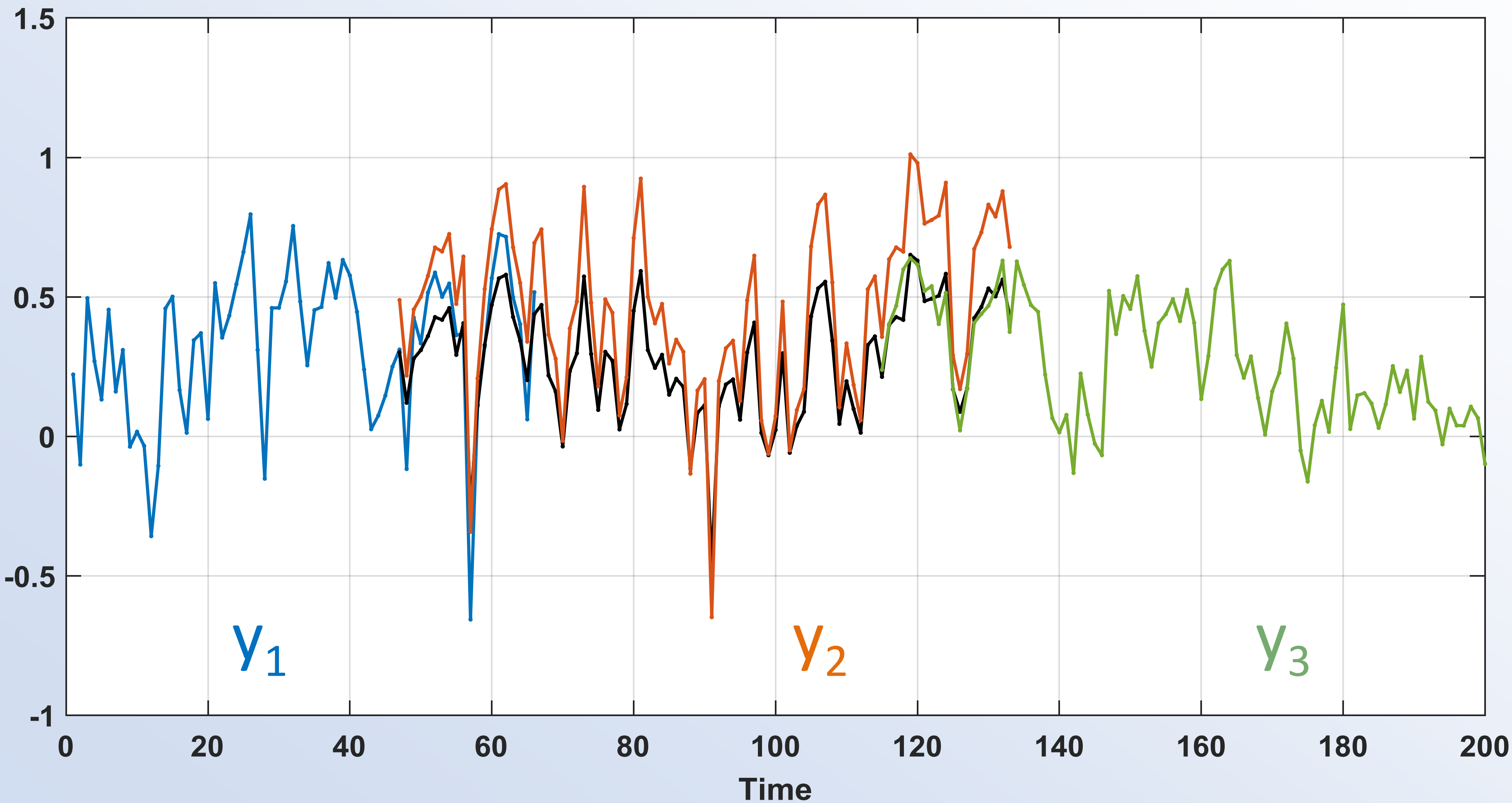
Stitching together multiple datasets

- The above method can be generalized to multiple overlapping datasets. We select y_3 as the basis



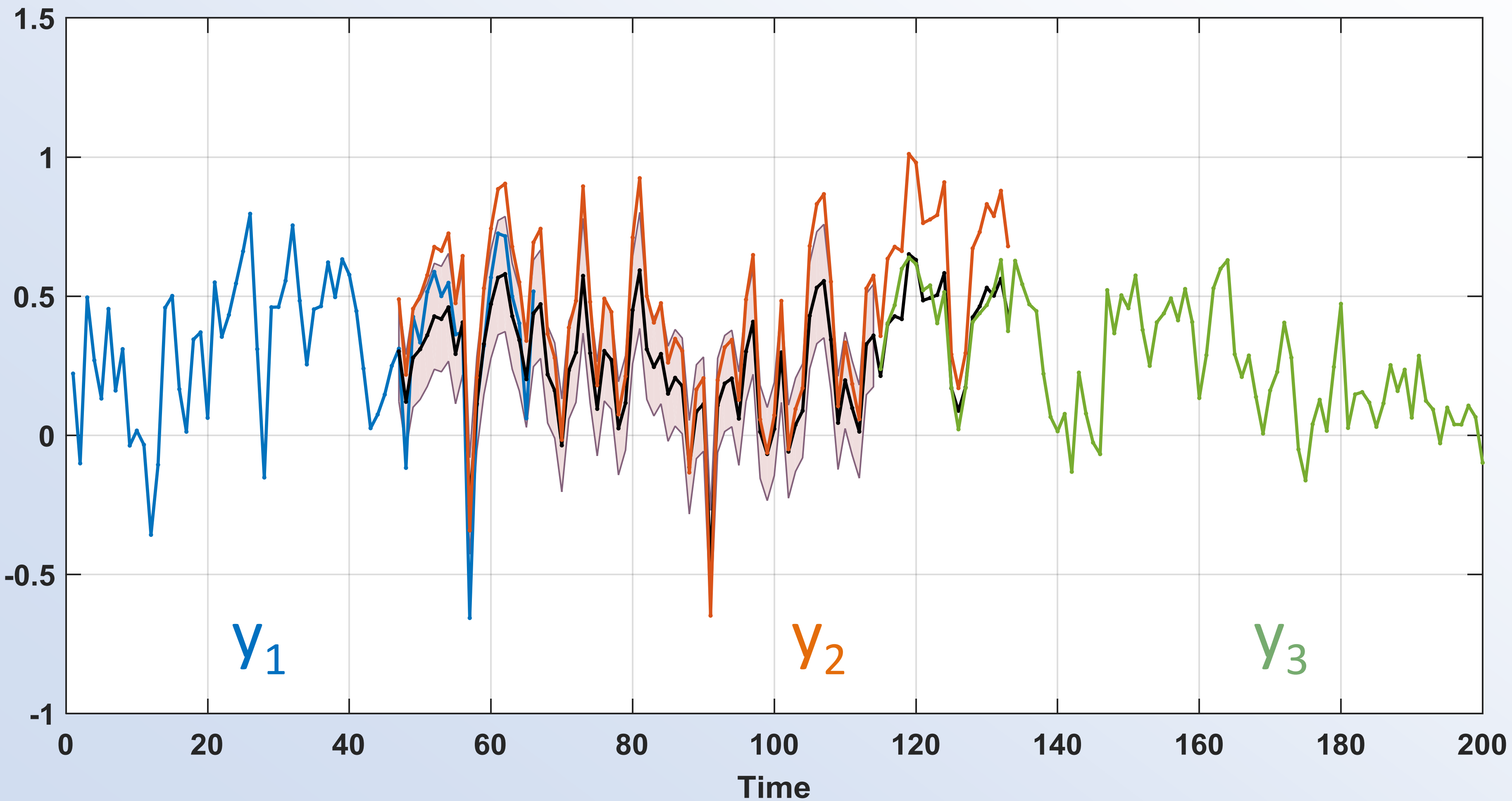
Stitching together multiple datasets

- First normalize y_2 to level of y_3



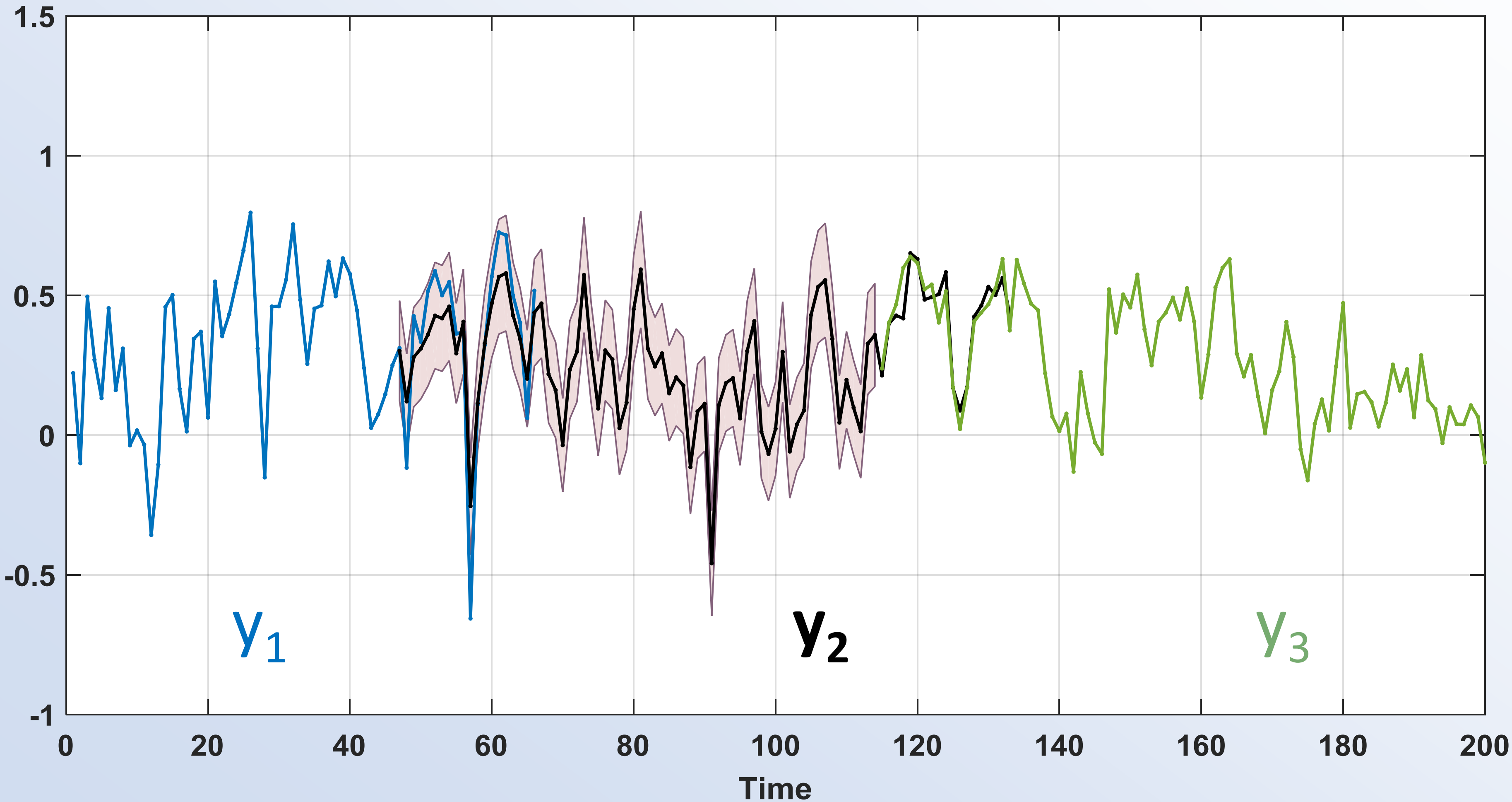
Stitching together multiple datasets

- First normalize y_2 to level of y_3
- This normalization has some statistical error



Stitching together multiple datasets

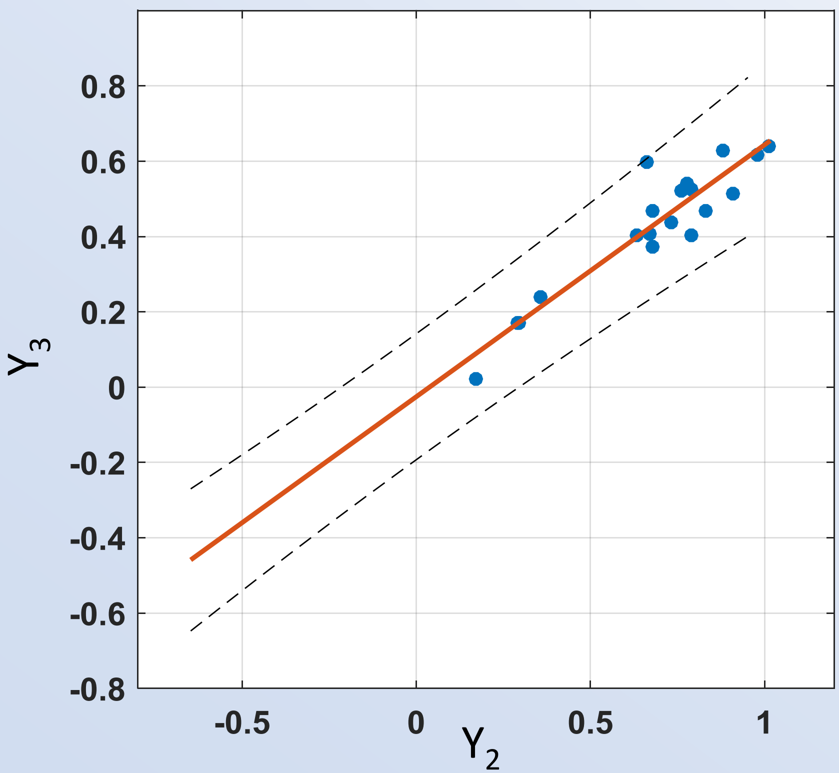
- Next normalize y_1 to new $\mathbf{y_2}$

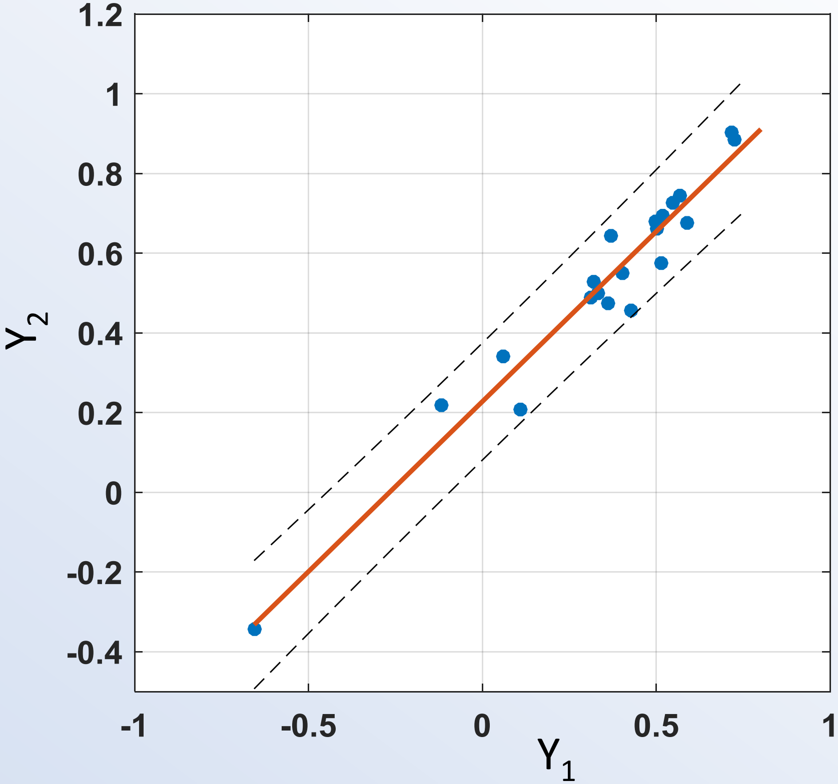


Stitching together multiple datasets

- From the two regressions we have
- 1st norm: $y_3 = \beta_1 + \beta_2 y_2$
- 2nd norm: $y_2 = \alpha_1 + \alpha_2 y_1$
- 2nd norm: $y_3 = \beta_1 + \beta_2(\alpha_1 + \alpha_2 y_1)$

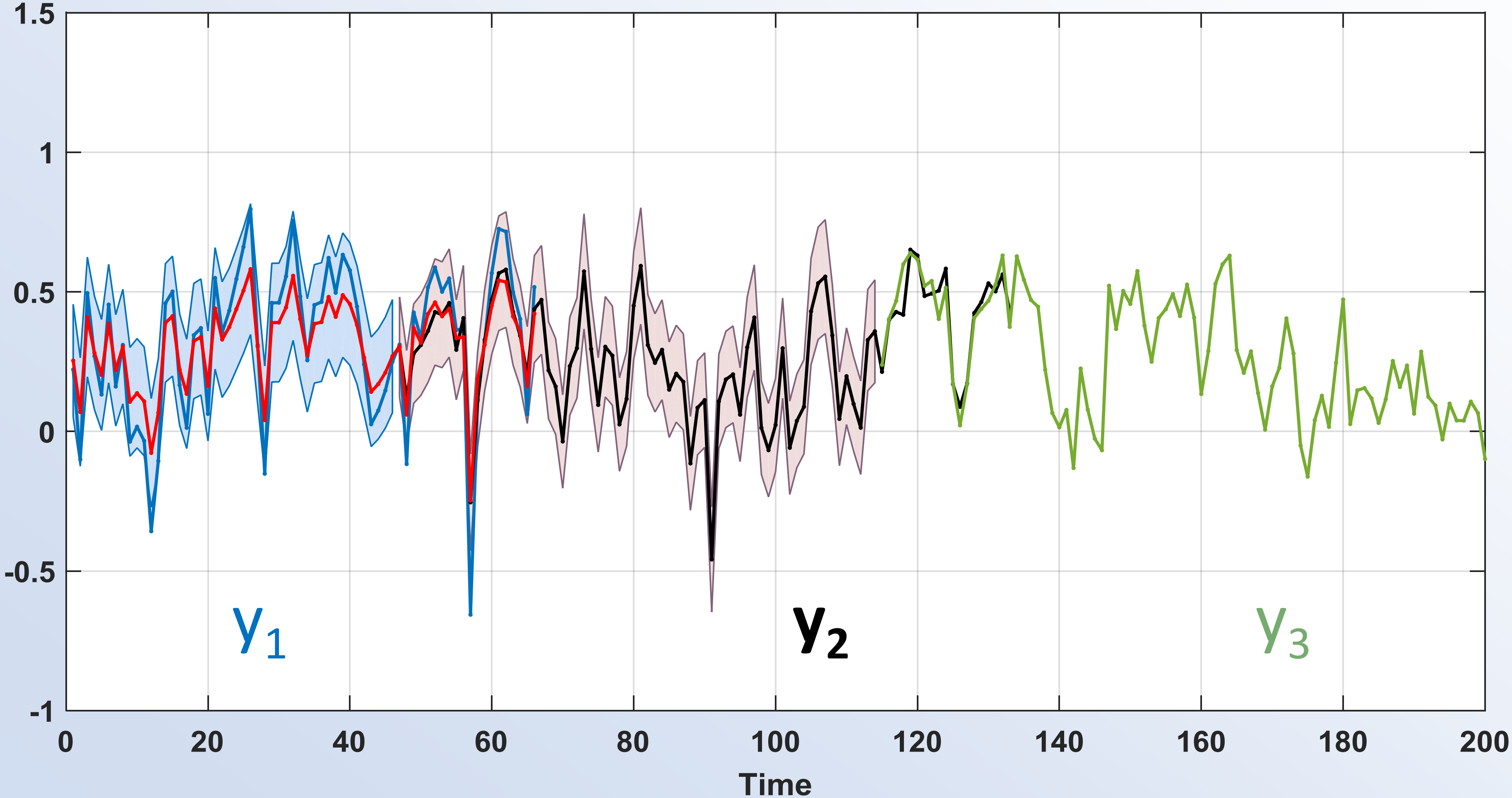
- $var(y_2) = var(\alpha_1) + y_1^2 var(\alpha_2) + var(\varepsilon_1)$
- $var(y_3) = var(\beta_1) + y_2^2 var(\beta_2) + var(y_2)\beta_1^2 + var(\beta_2)var(y_2) + var(\varepsilon_2)$
- 95% Prediction interval is given by:
- $\Delta y = t_{0.975,n-4}\sqrt{var(y_3)}$





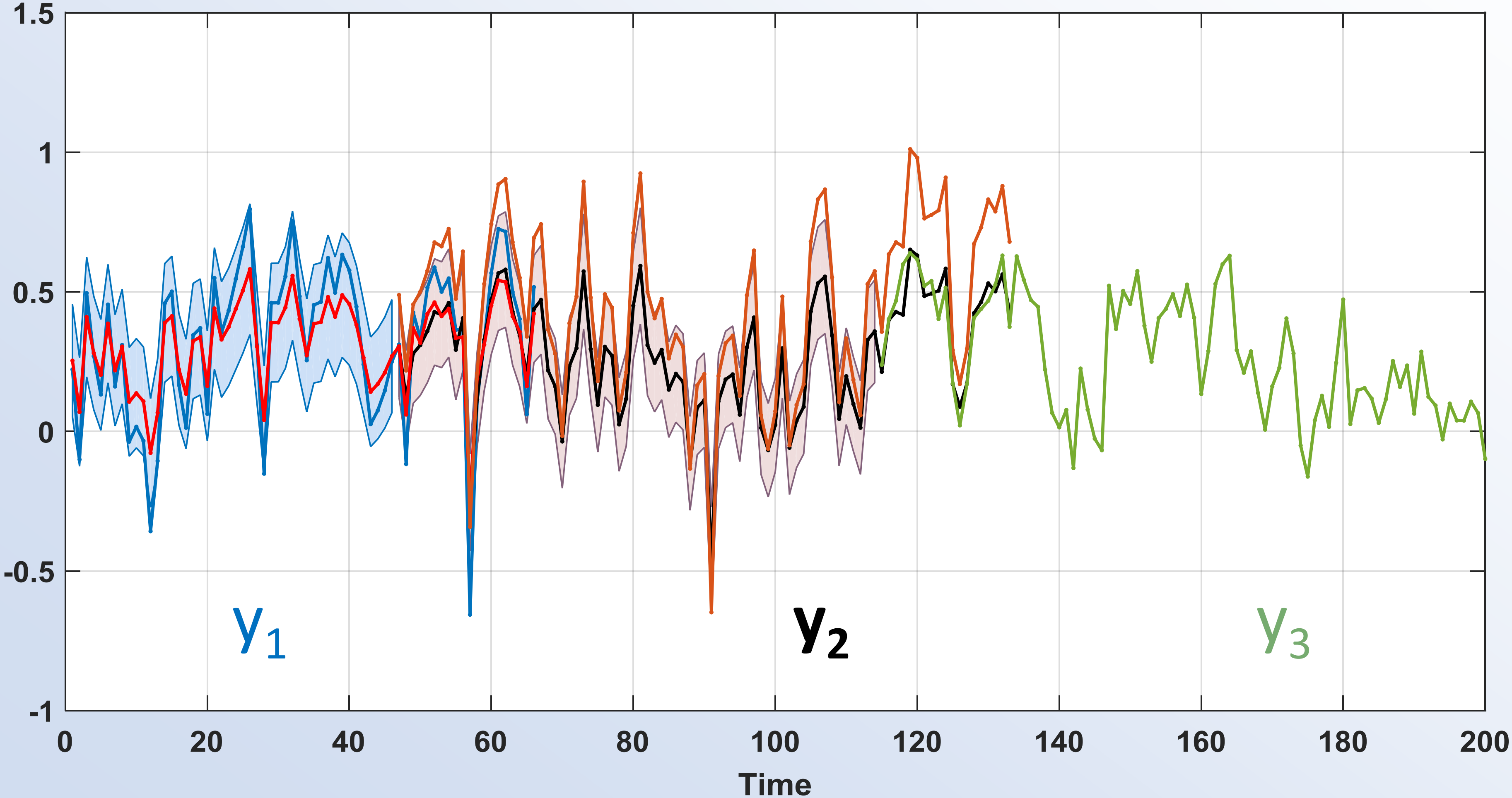
Stitching together multiple datasets

- Next normalize y_1 to new $\mathbf{y_2}$
- And estimate error limits



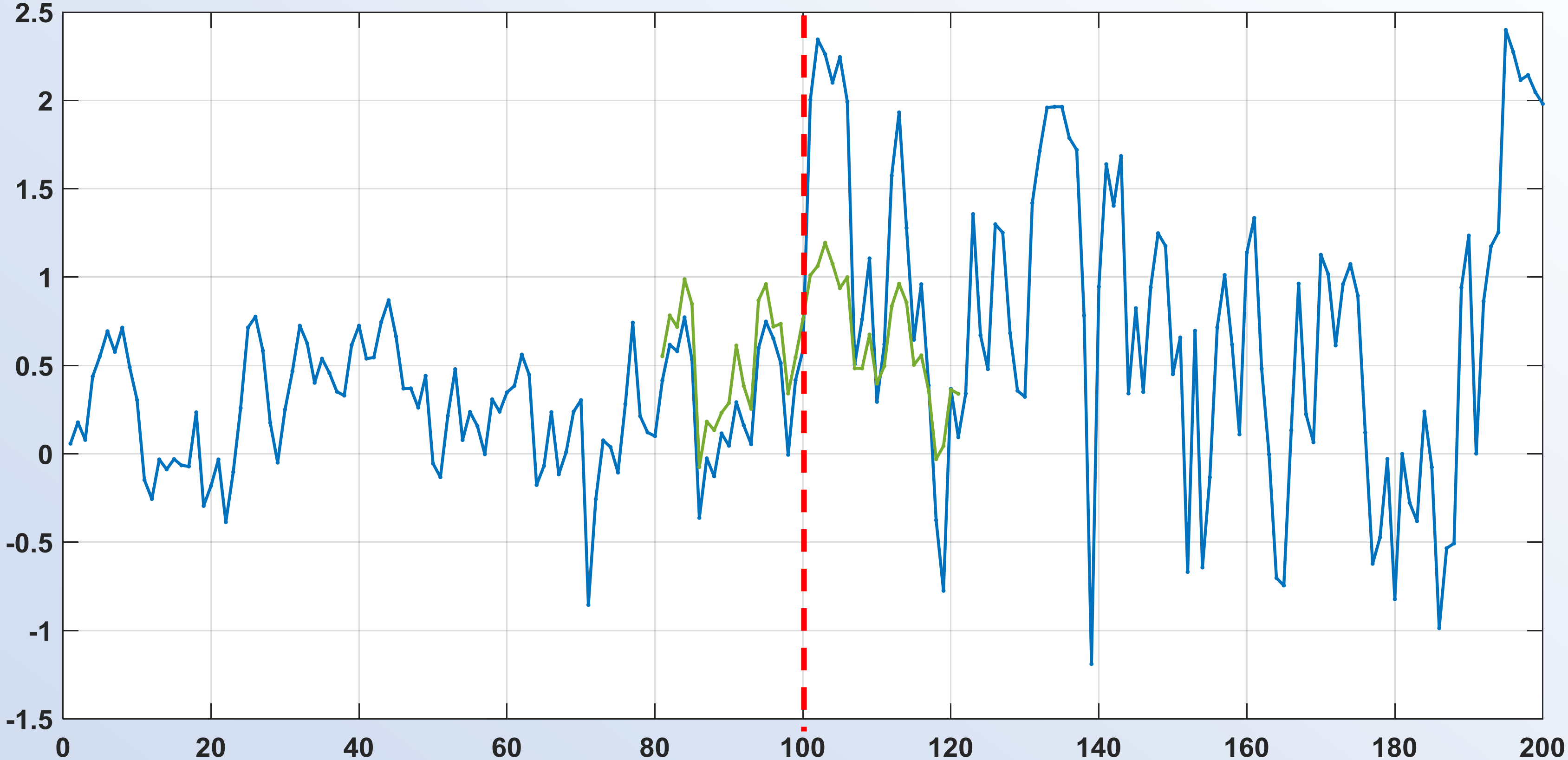
Stitching together multiple datasets

- Finally we have the composited data series and its error limits.
- Note: Errors are accumulated. In this example the errors in the first time series are larger than the difference between the original and the composite!



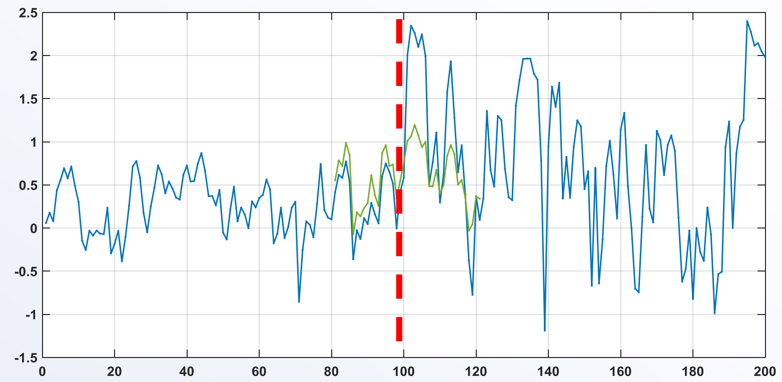
Removing inhomogeneity with a reference series

- Often a single long data series can be inhomogeneous due to changes in instrument settings, location etc.
- Sudden changes can be calibrated with a reference time series, which is homogeneous over the same time interval



Removing inhomogeneity with a reference series

- x_1=data before jump, x_2=data after the jump, y= reference series
- After the jump:
- $x_2 = \beta_1 + \beta_2 y + \varepsilon_2$ n_2 points
- Before the jump:
- $y = \alpha_1 + \alpha_2 x_1 + \varepsilon_1$ n_1 points



- Estimate what x_2 would be BEFORE the jump
- ➔ $x_2 = \beta_1 + \beta_2(\alpha_1 + \alpha_2 x_1 + \varepsilon_1) + \varepsilon_2$

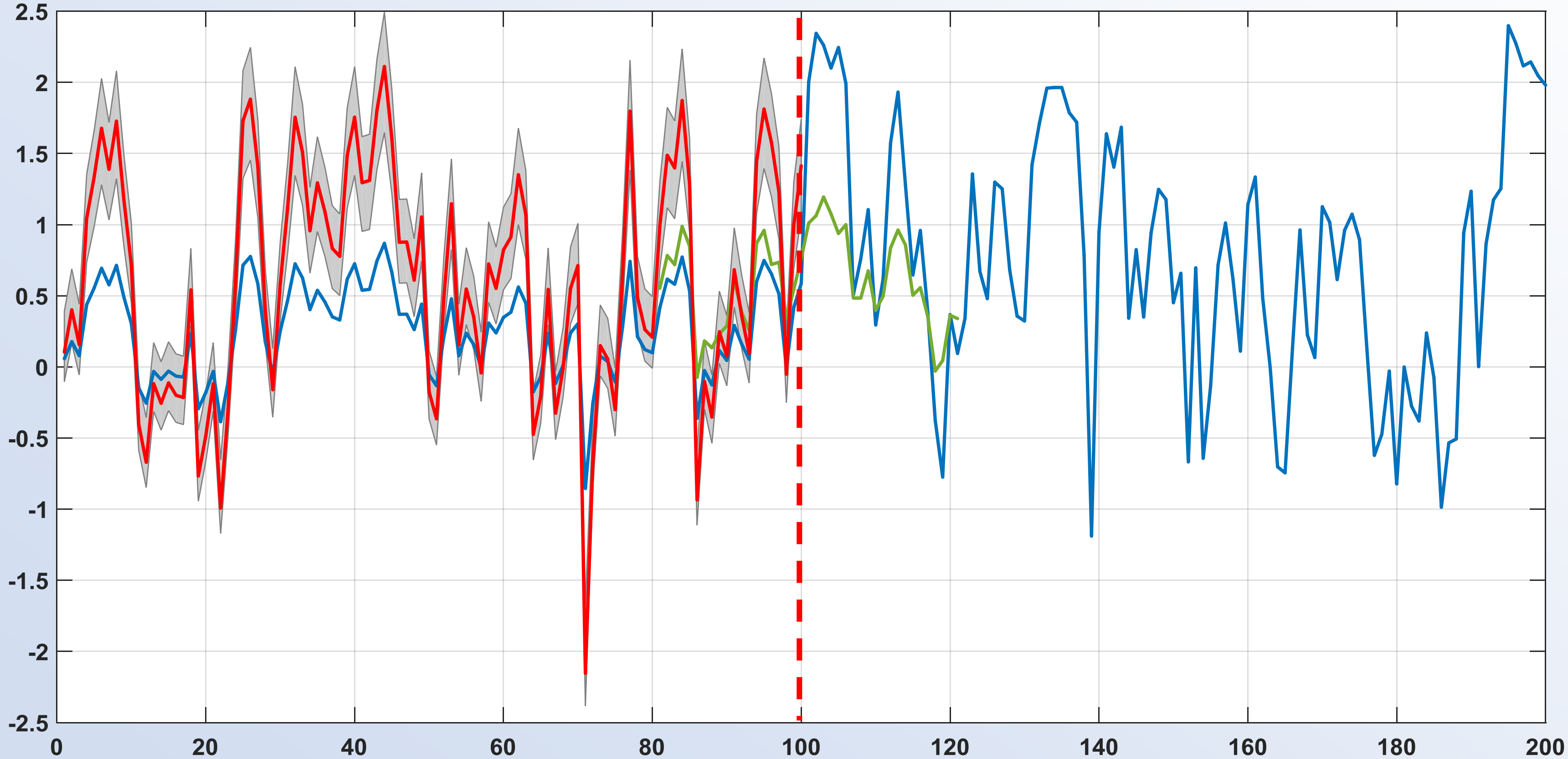
- From summation of variance formula we get for the time period before the jump

- $var(y) = var(\alpha_1) + x_1^2 var(\alpha_2) + var(\varepsilon_1)$
- $var(x_2) = var(\beta_1) + y^2 var(\beta_2) + \beta_2^2 var(y) + var(\beta_2) var(y) + var(\varepsilon_2)$

- And the 95% prediction interval would be
- $\Delta x_2 = t_{0.975, n_1+n_2-4} \sqrt{var(x_2)}$

Removing inhomogeneity with a reference series

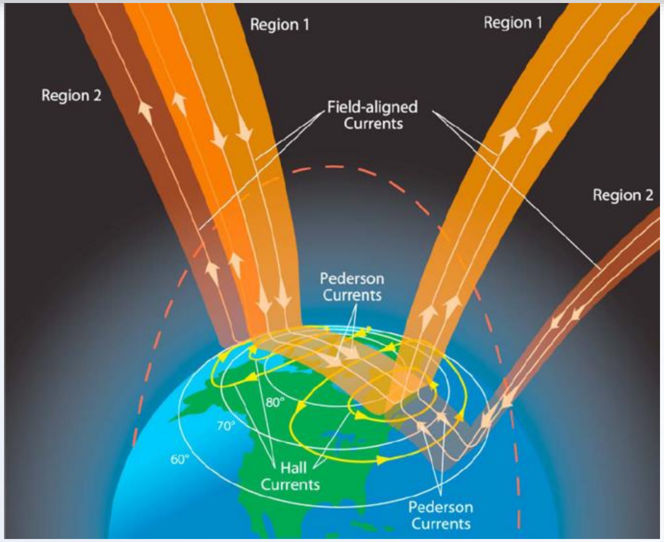
- Here the inhomogeneity has been corrected

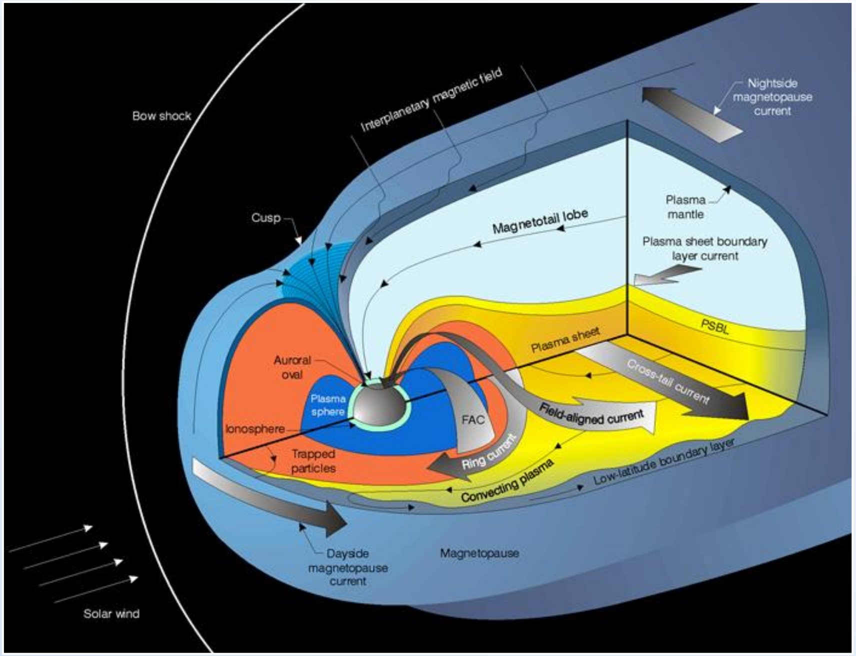


Example 1.
Geomagnetic indices

Geomagnetic activity

- Geomagnetic activity ⇔ Variations of magnetic field on ground
- Results from variations of electric currents in magnetosphere and ionosphere

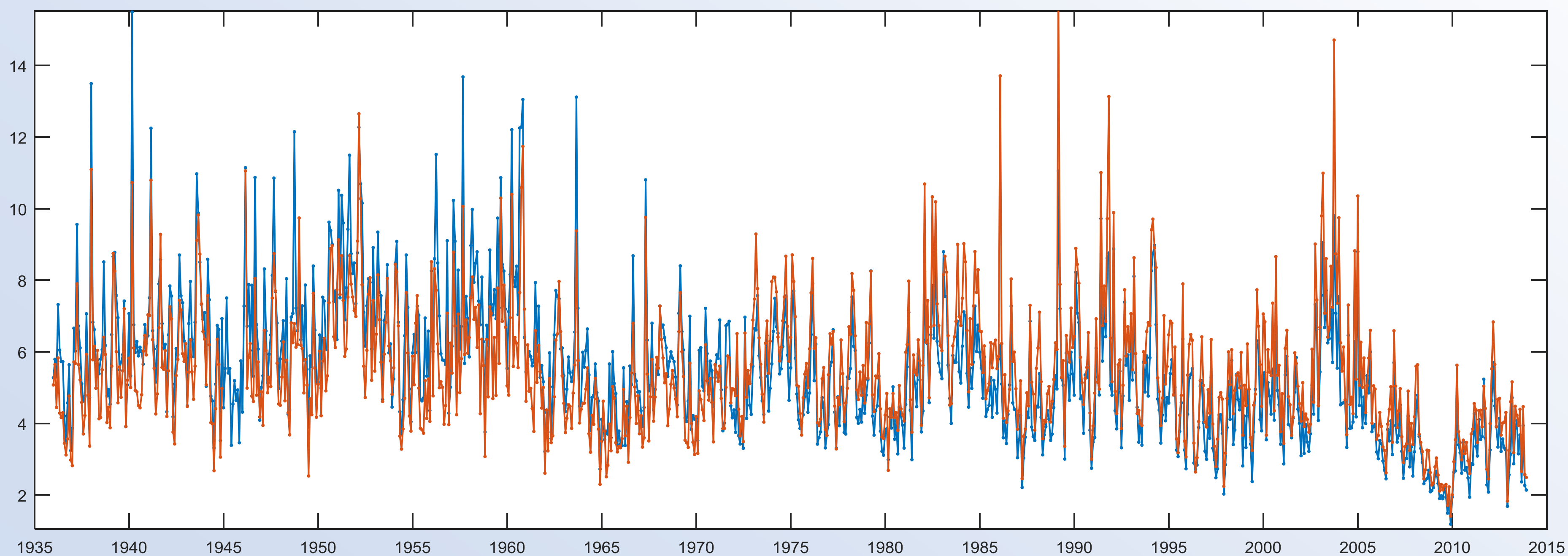




IHV indices

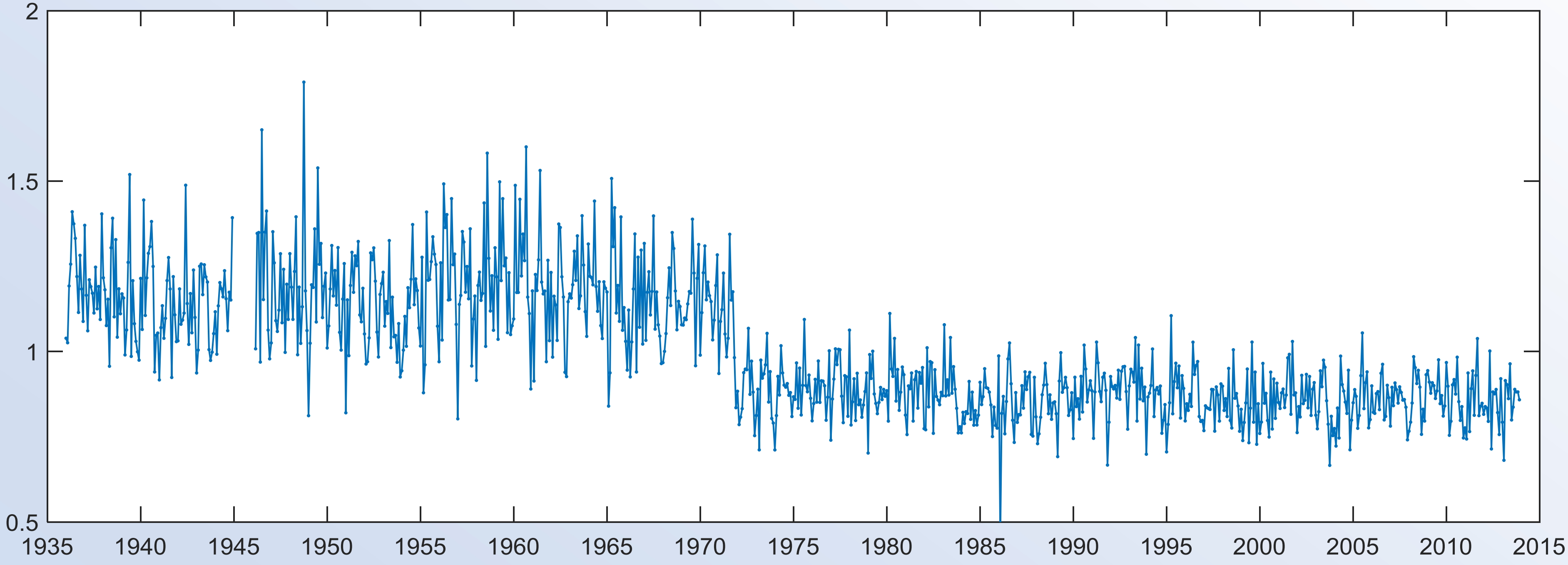
- IHV indices from two different mid-latitude stations (NGK/Germany, CLF/France)
- CLF changed from spot sampling to hourly means in 1972

$$IHV = \frac{1}{6} \sum_{t=21LT}^{02LT} |H(t+1) - H(t)|$$



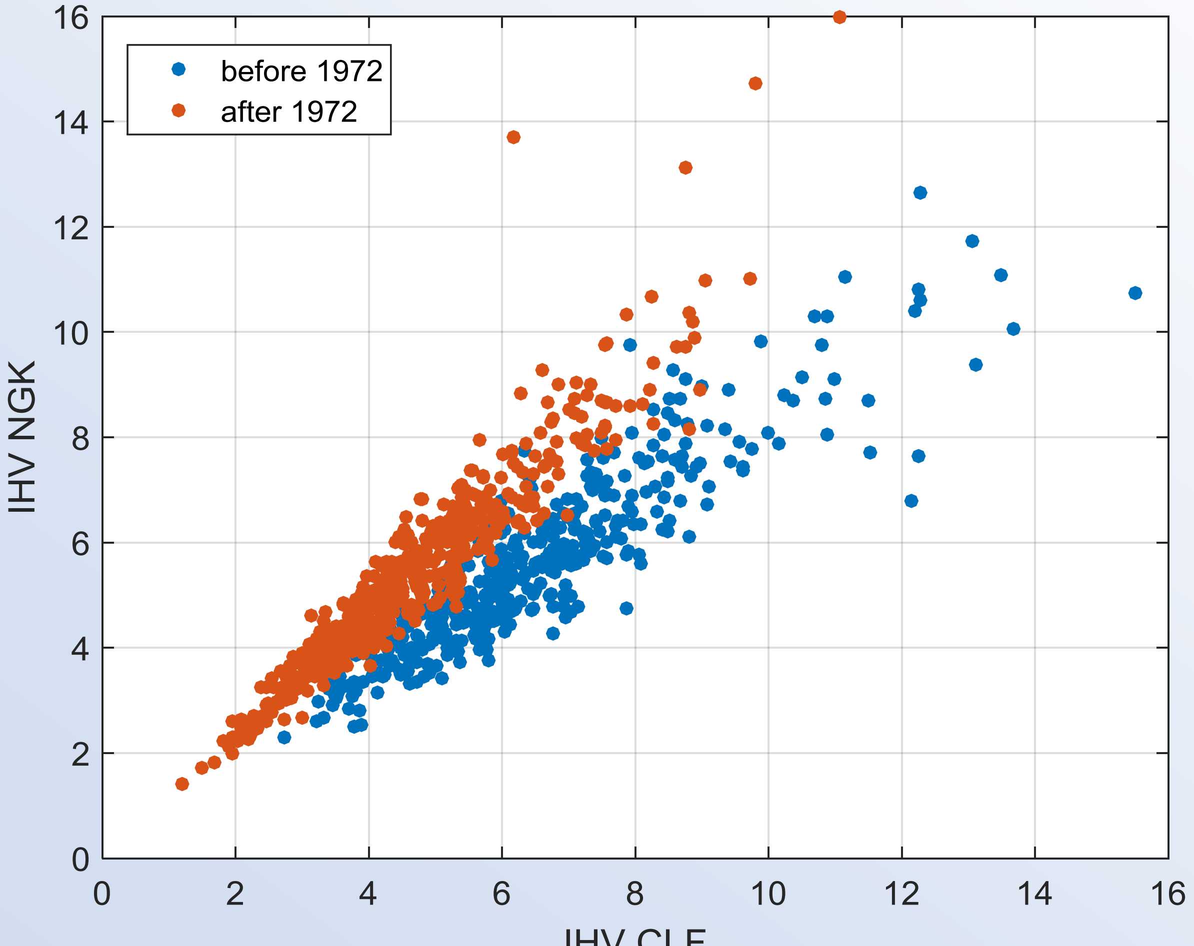
Effect of sampling change to IHV

- Effect of sampling change is clearly seen in CLF/NGK ratio
- Before 1972 CLF sees systematically larger values than NGK compared to period after 1972
- Note! Ratio after 1972 is not 1 ➔ Real difference in the indices.
- We are not trying to make the indices identical!!



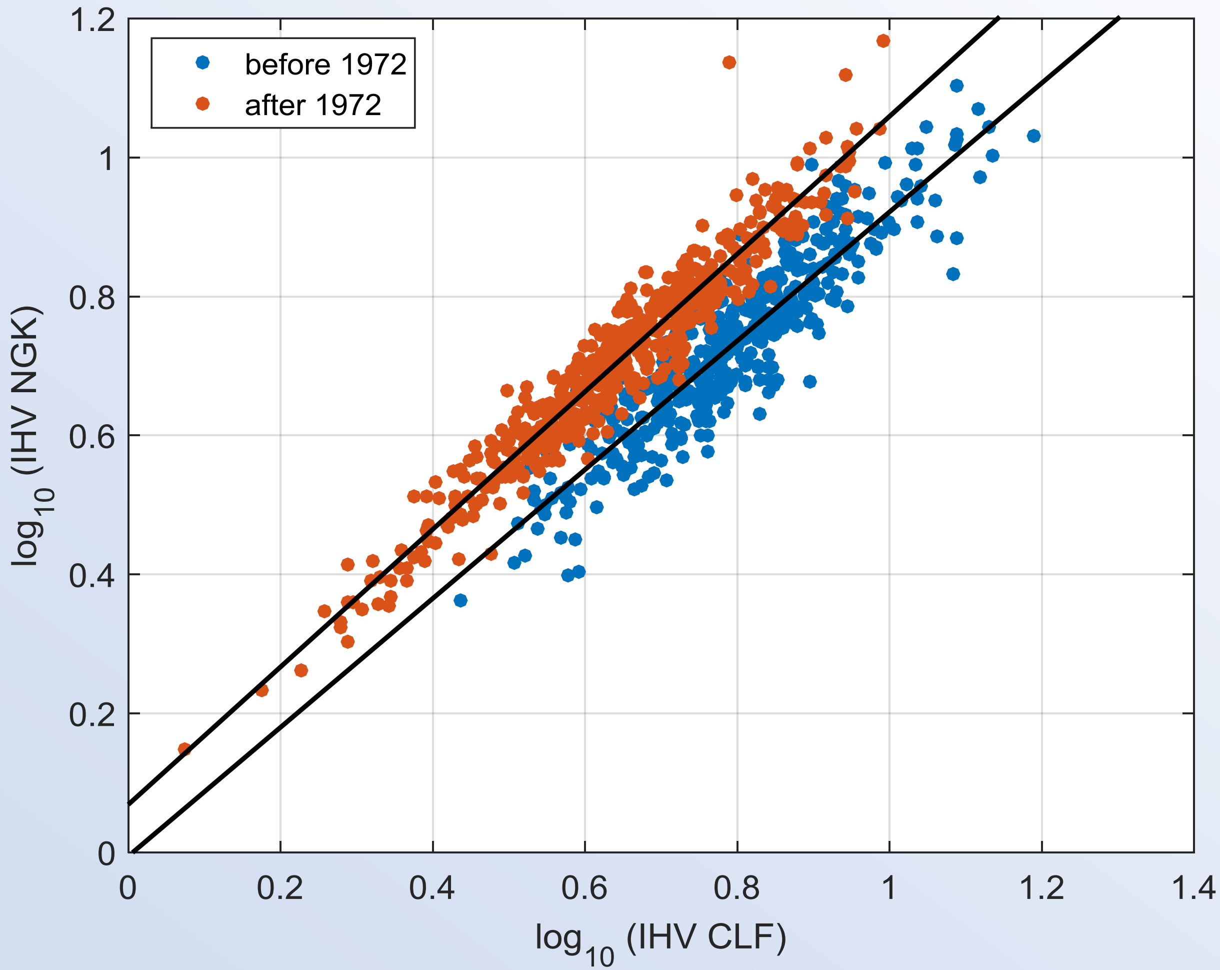
Effect of sampling change to IHV

- Clear difference in two time periods
- But linear fit is not good (not a constant variance)



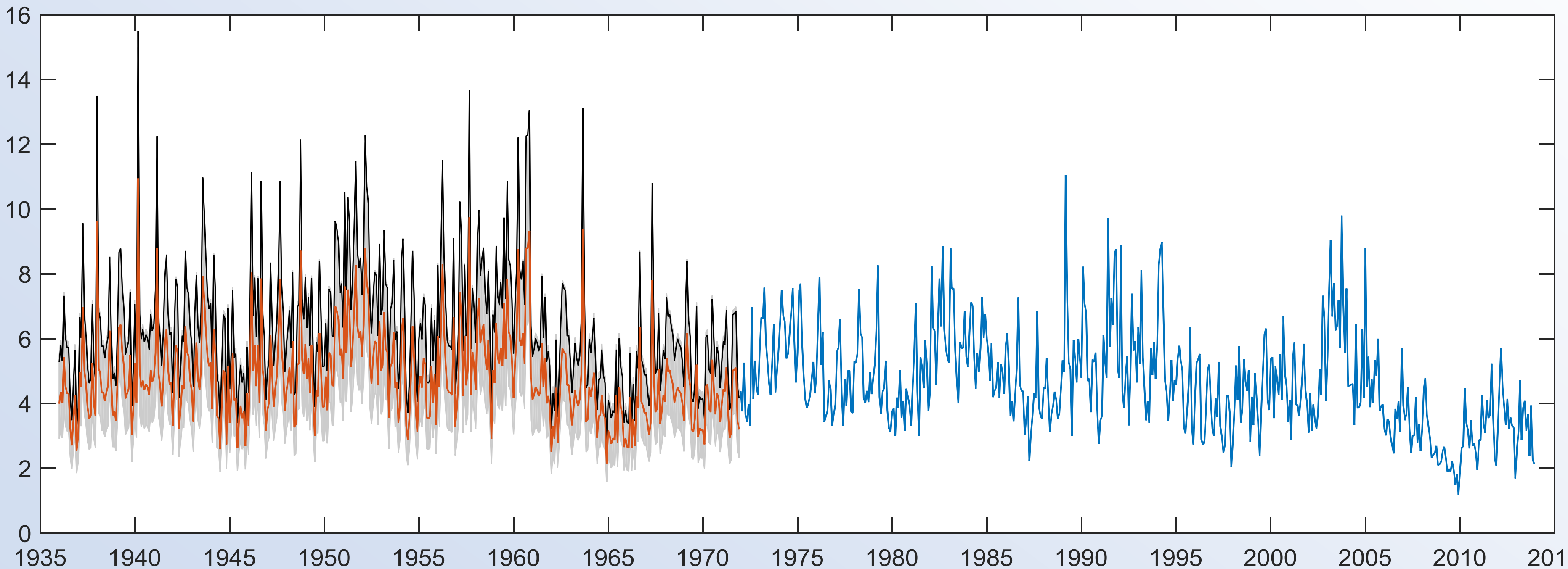
Effect of sampling change to IHV

- Log-log scale much better



Result of homogenization

- After calibration the data before 1972 is in correct level and we have also uncertainty limits



More problematic inhomogeneities

- In many cases inhomogeneity of a measurement series changes with time

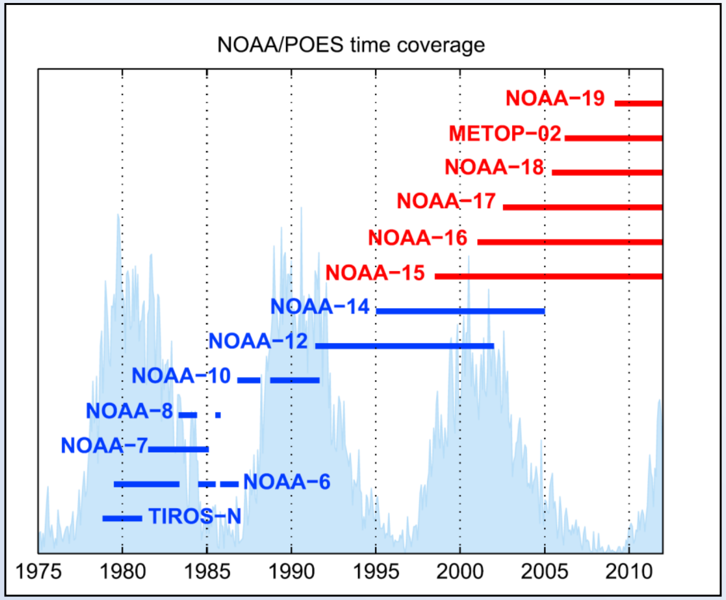
- Examples:
  - Instrument efficiency changes continuously with time
  - Instrument degrades with time

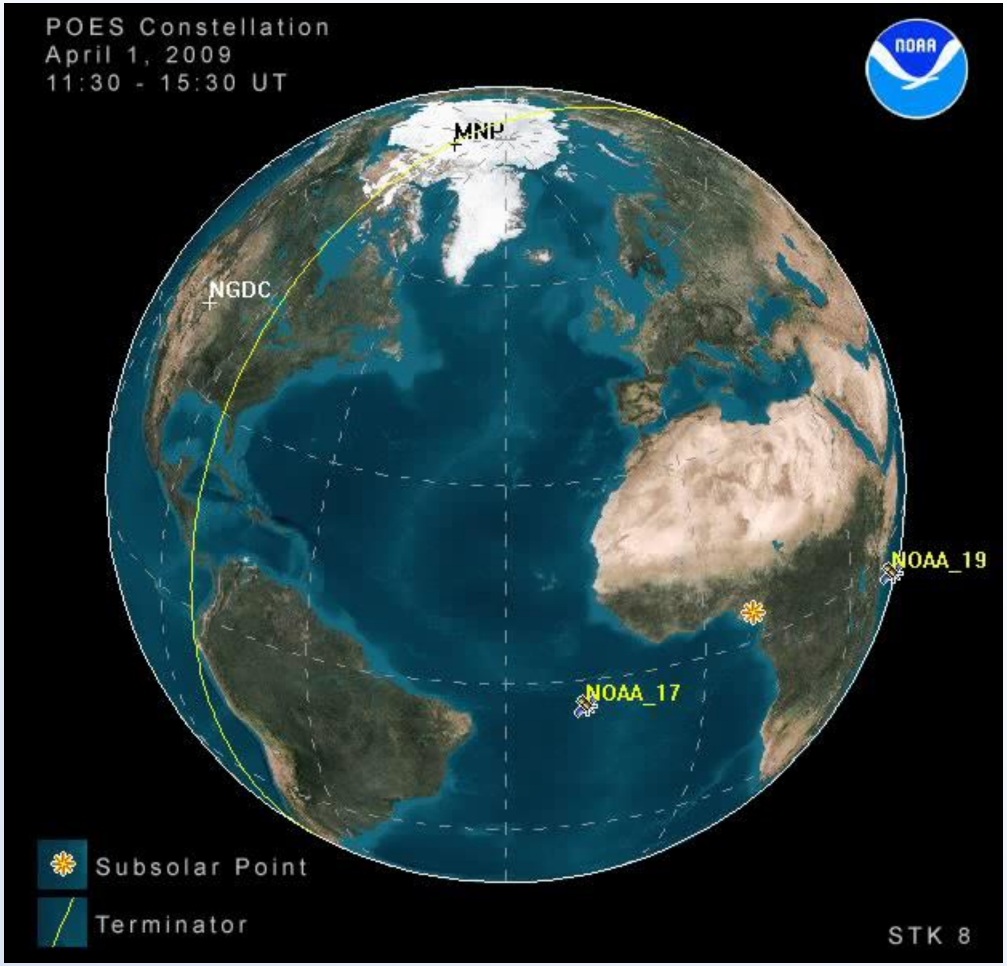
- Often these cannot be corrected by comparing with some other time series at one interval of time

- ➔ Need to understand the cause of changing data homogeneity and how to compensate it

Example 2.
Particle data from satellites

NOAA/POES satellites: long dataset

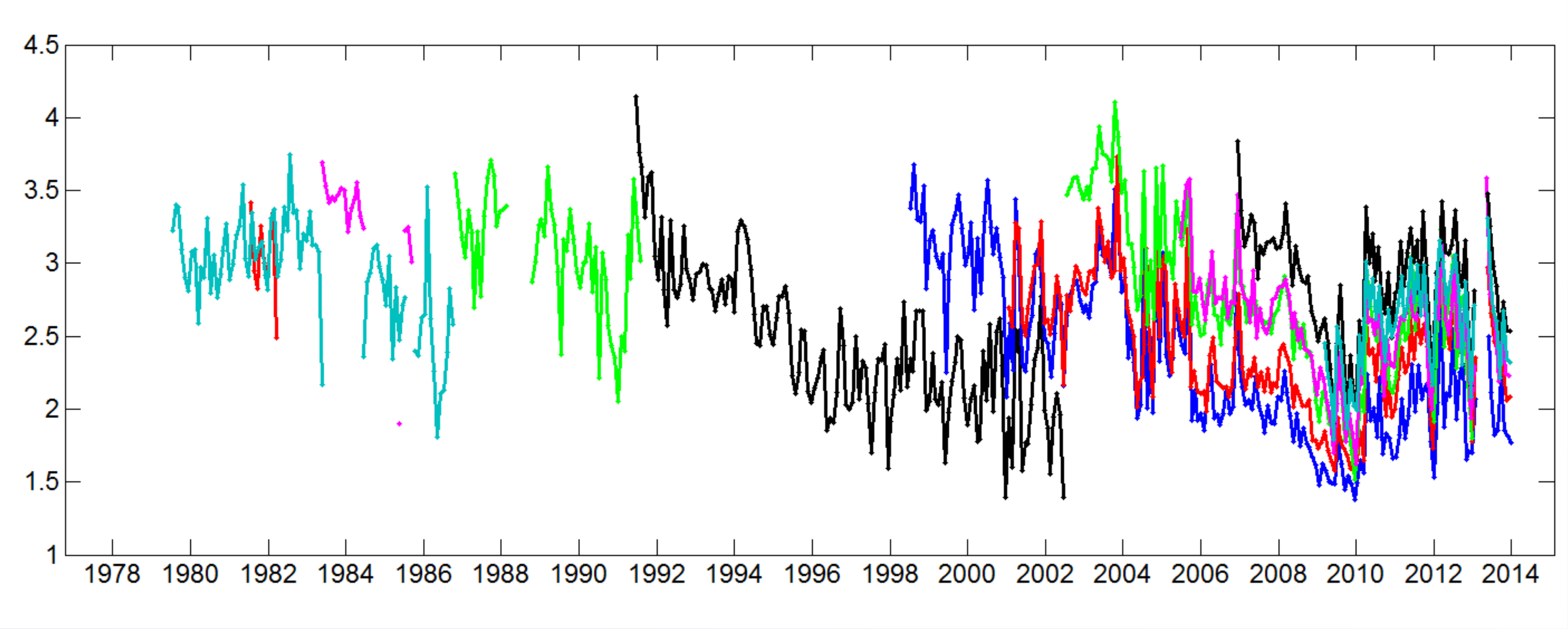




- The dataset has been plagued by significant problems related to the MEPED instrument.

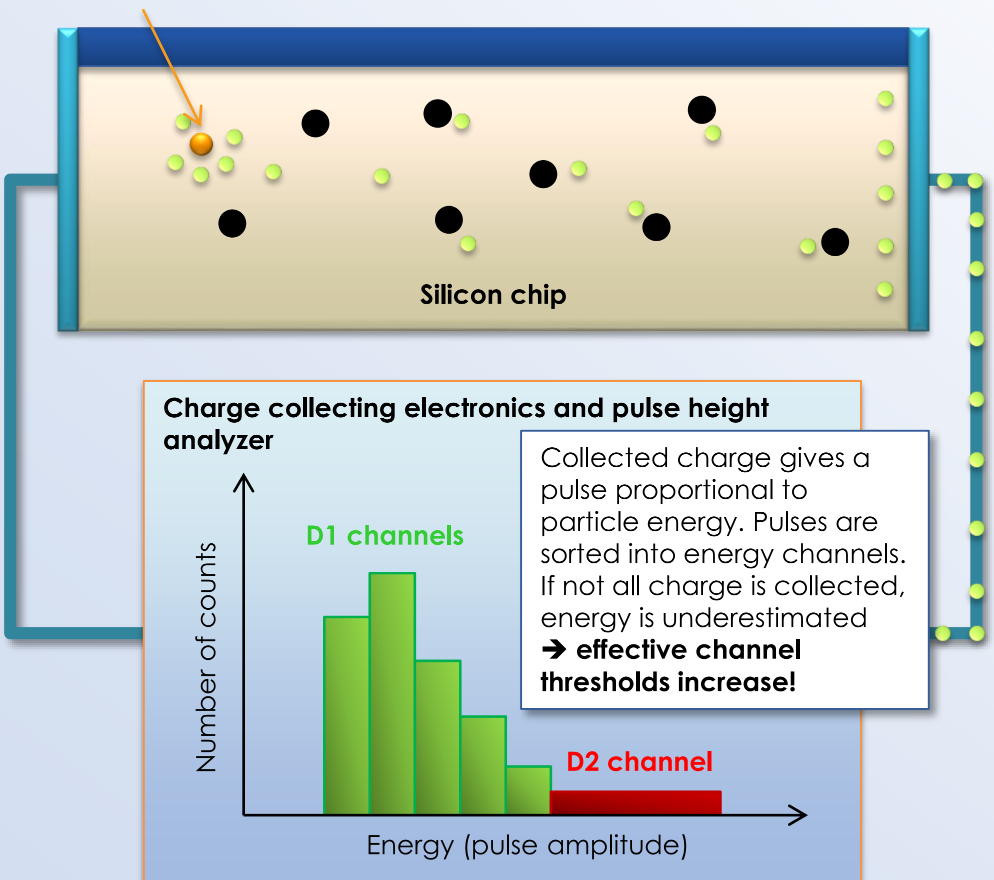
Raw proton data is obviously erroneus

- Steady decrease in all satellites ➔ degradation of instrument
- Large spread between simultaneous satellites
- Simple stitching of overlapping data series does not work.

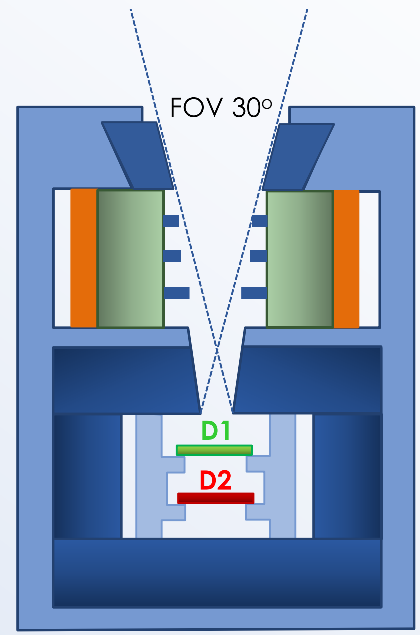


MEPED proton instrument

Zoom-up of a single silicon detector chip

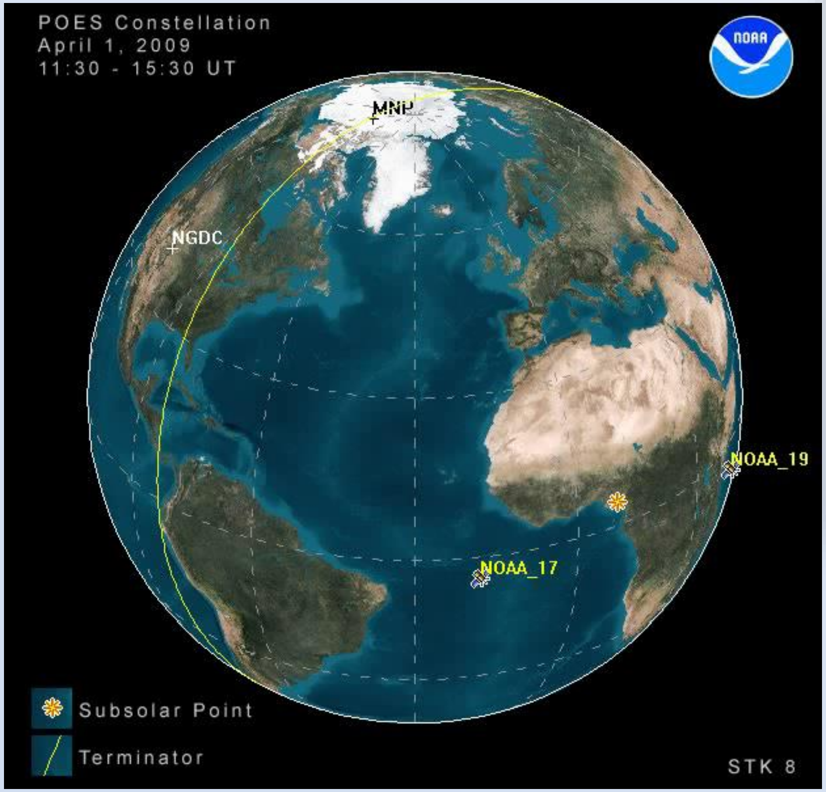


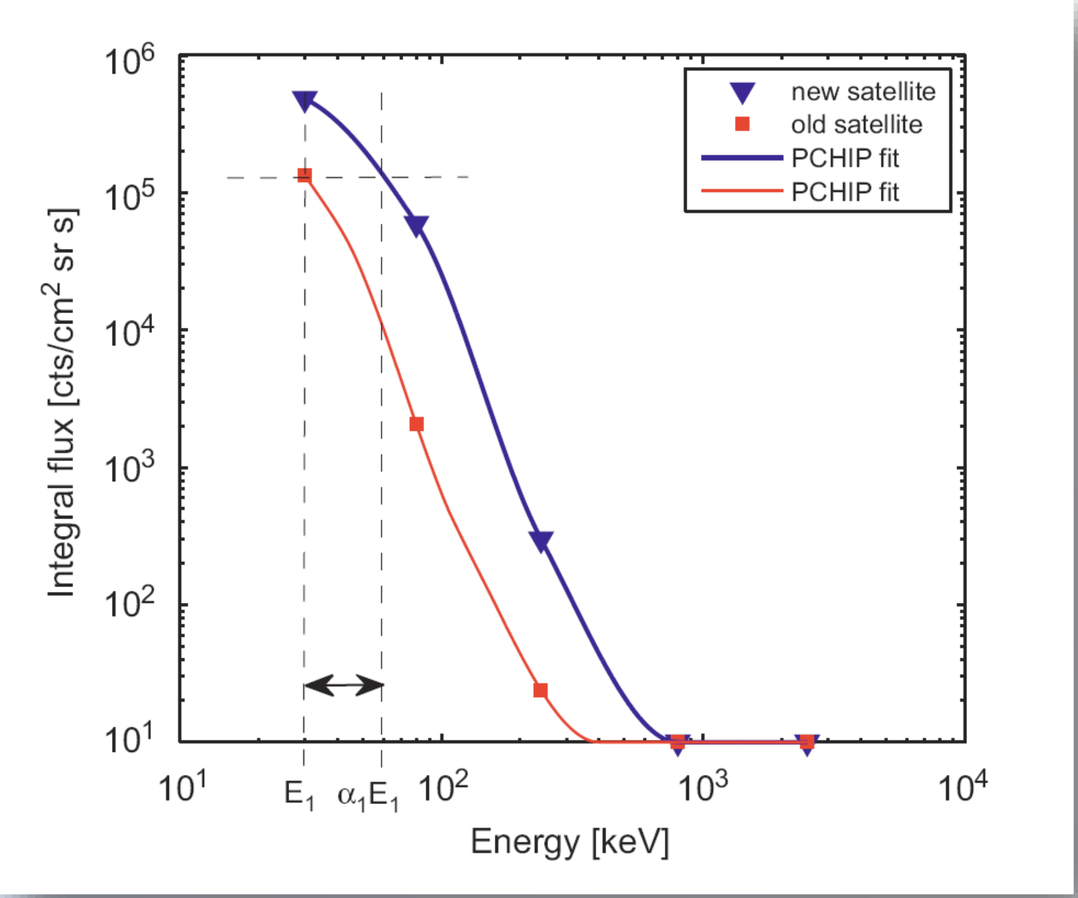
Cross-cut of the MEPED proton instrument



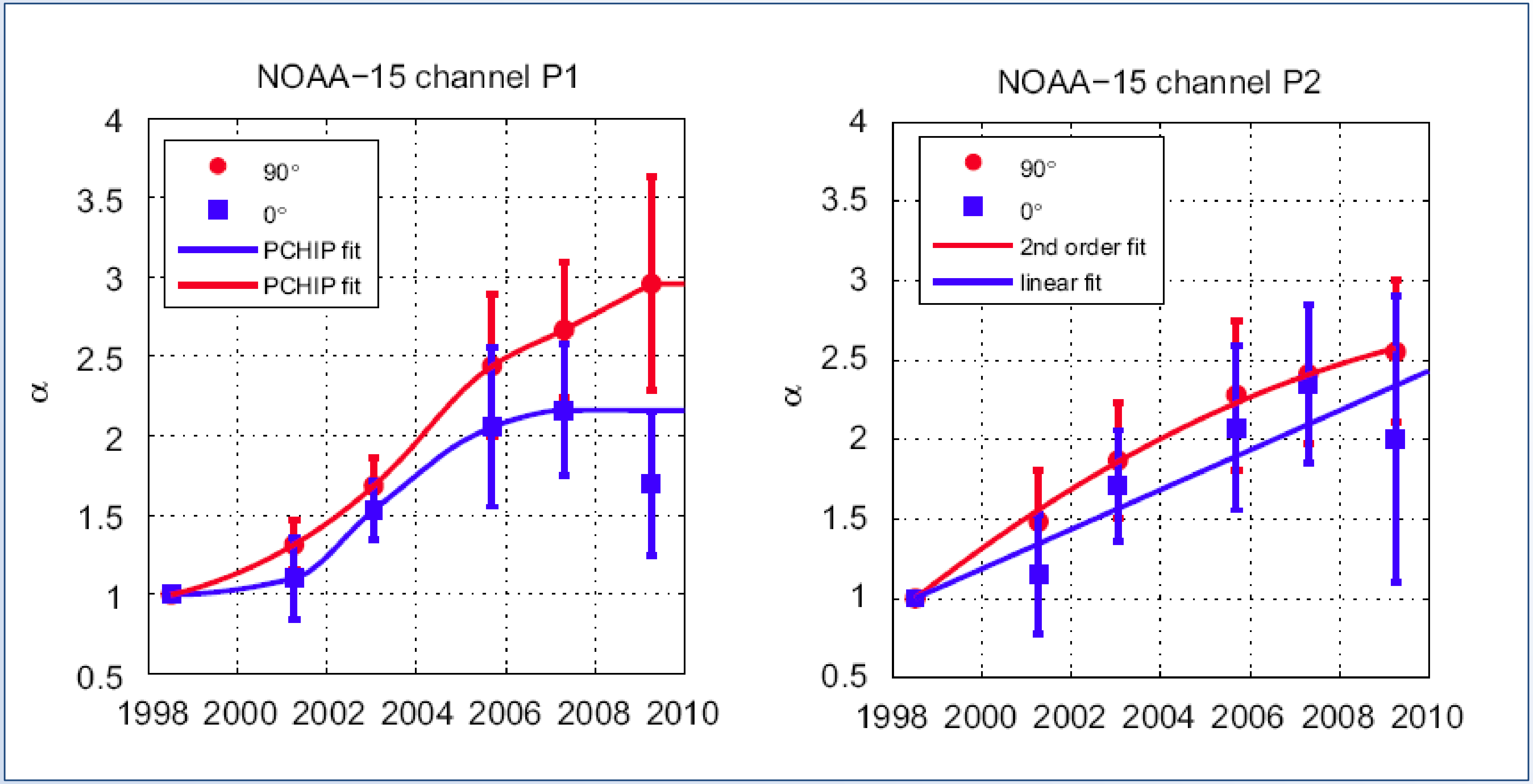
Anti-coincidence logic between **front detector (D1)** and **back detector (D2)**
➔ a noise(false) count in D2 erases a coincident real count from D1.

Estimate instrument energy thresholds by comparing an old satellite to new



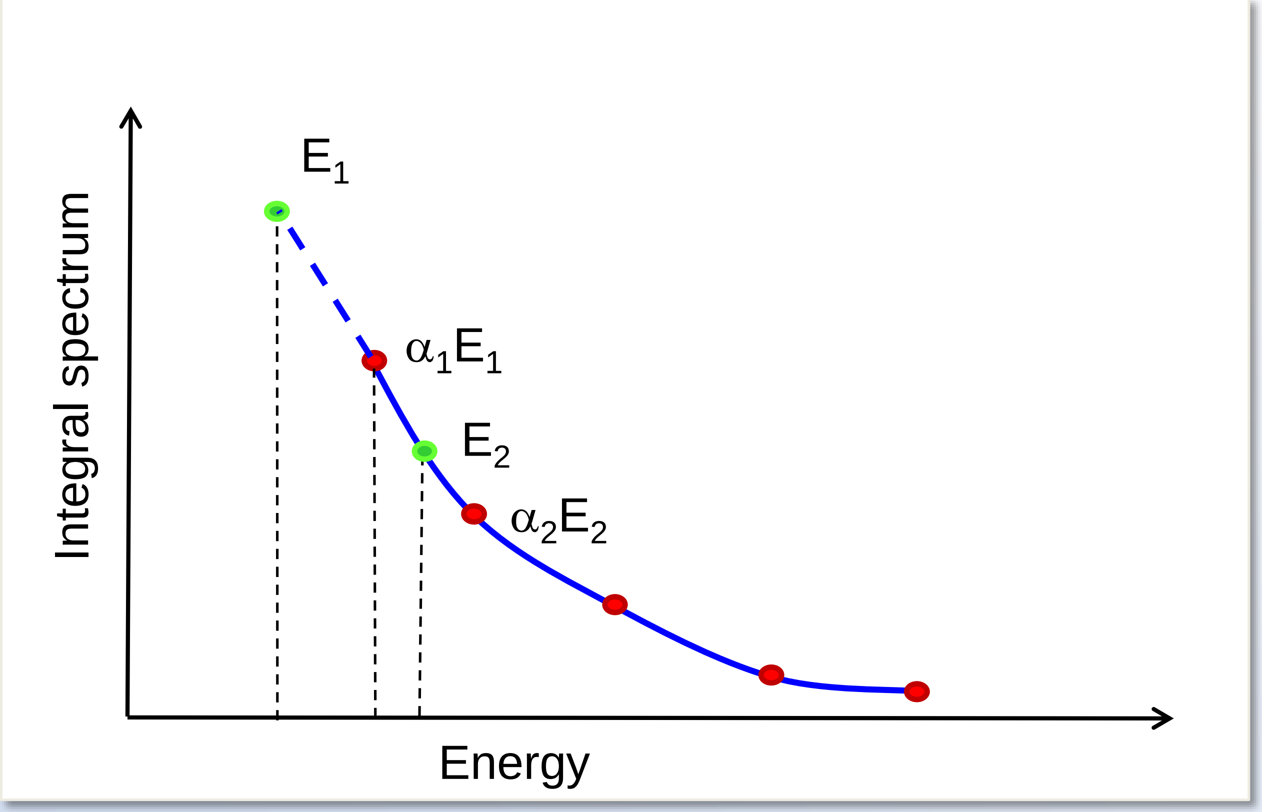


We get real energy thresholds as a function of time



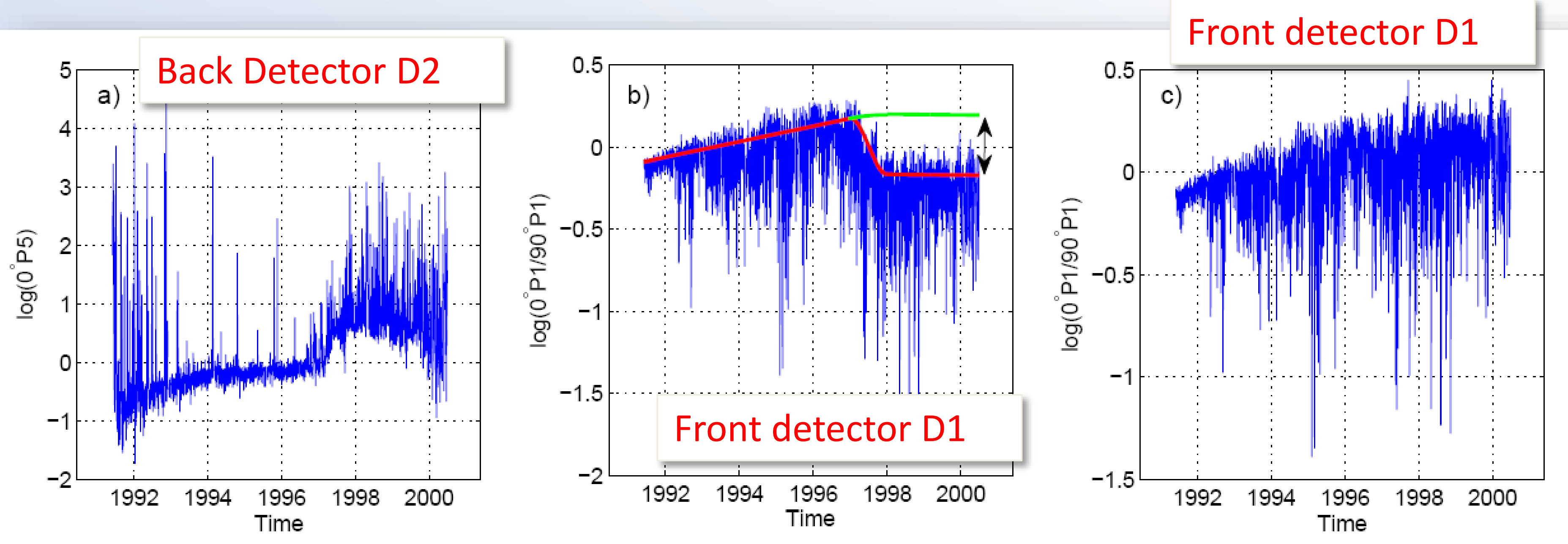
Compute fluxes at nominal energies from spectrum fitted to actual energies

- Note a caveat! The P1 energy channel cannot be reliably corrected, since it requires extrapolation to energies lower than the instrument can measure at the time.



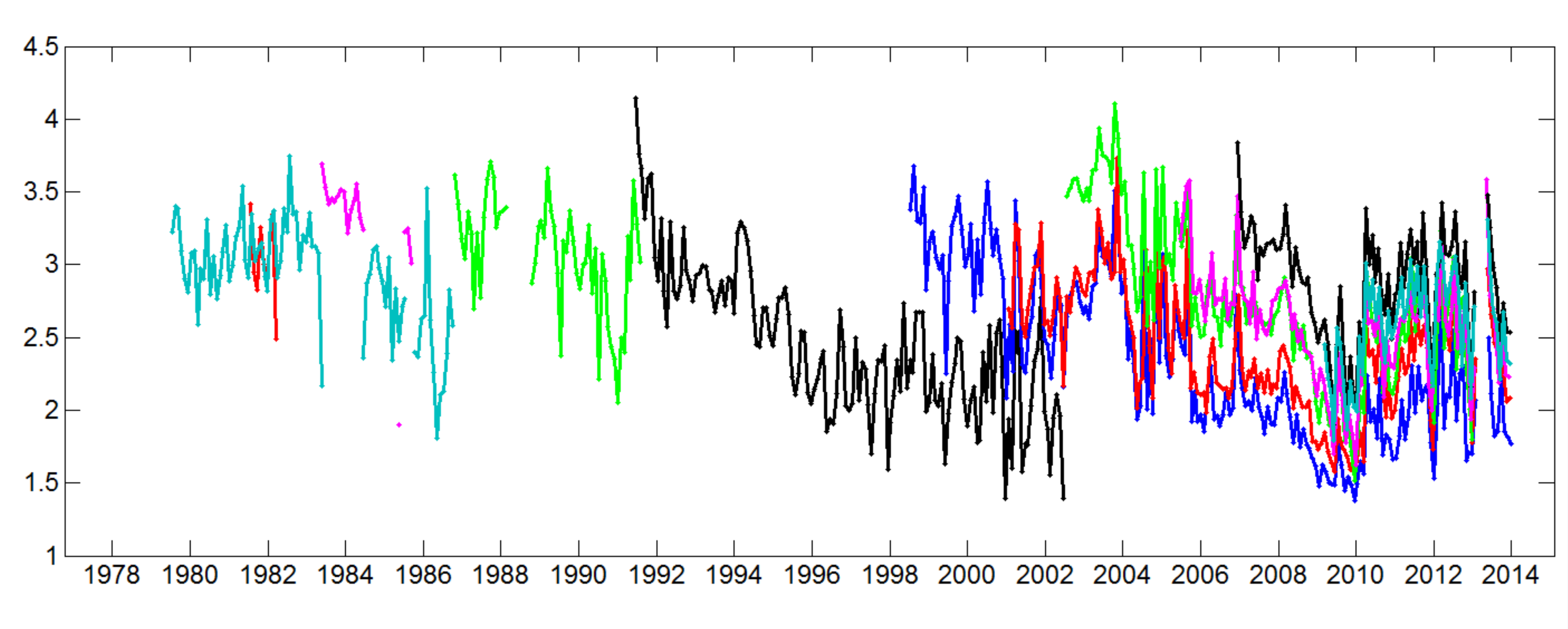
Electronic back detector noise in NOAA-08 and NOAA-12

- Back and front detectors work in anti-coincidence logic
- False counts (noise) in back detector erase real counts from front detector
- Modify the front detector baseline to correct for the noise



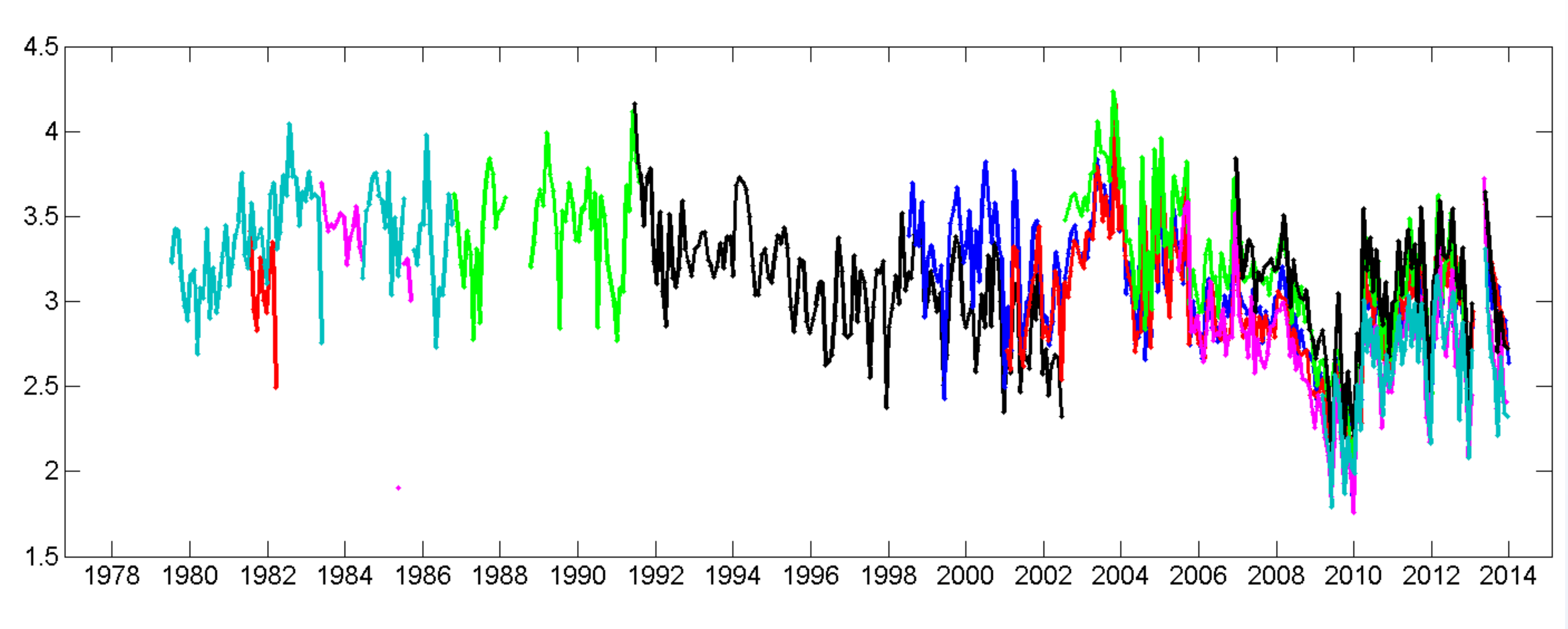
Raw proton data once more

- Steady decrease in all satellites ➔ degradation of instrument
- Large spread between simultaneous satellites



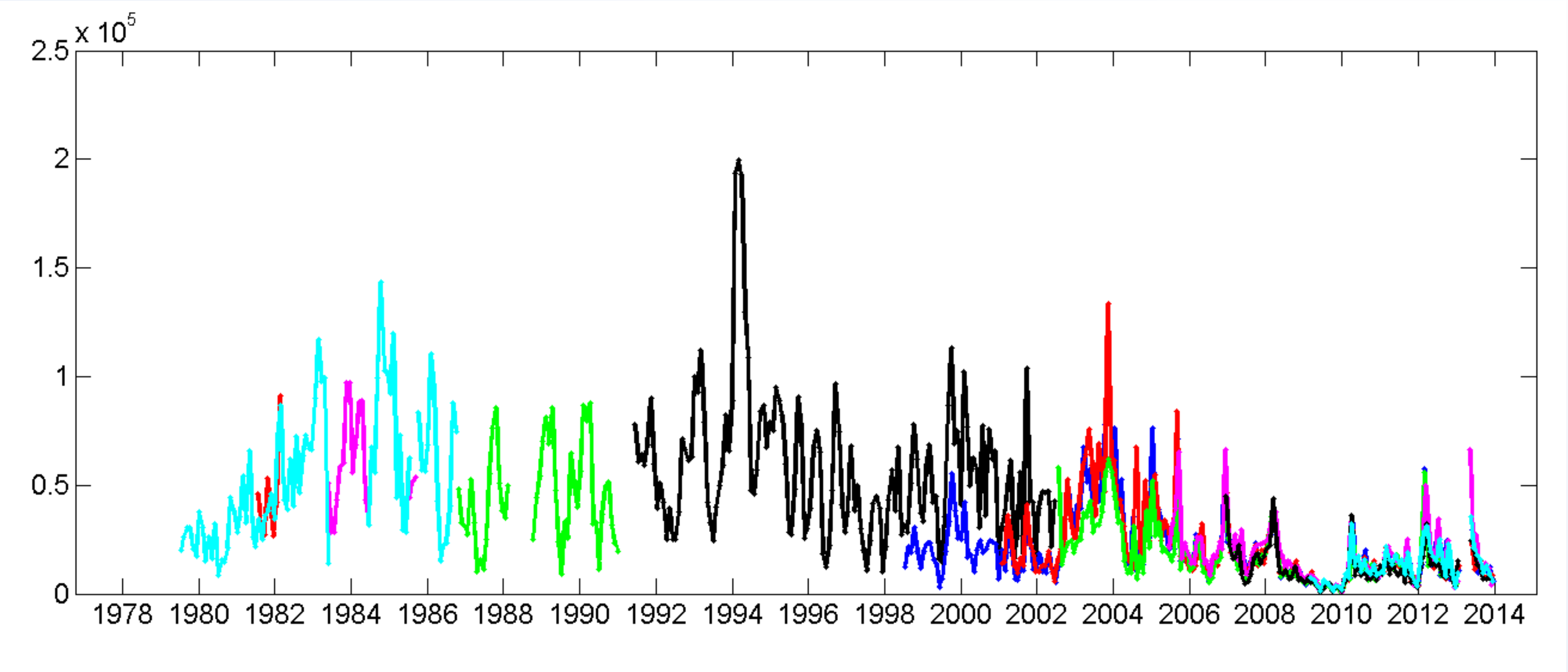
Corrected protons

- Continuous series from different satellites
- Satellite differences greatly reduced



Electrons >30 keV uncorrected

- Newer SEM-2 satellites (since mid-1998) systematically smaller than SEM-1 satellites



Problems in electrons

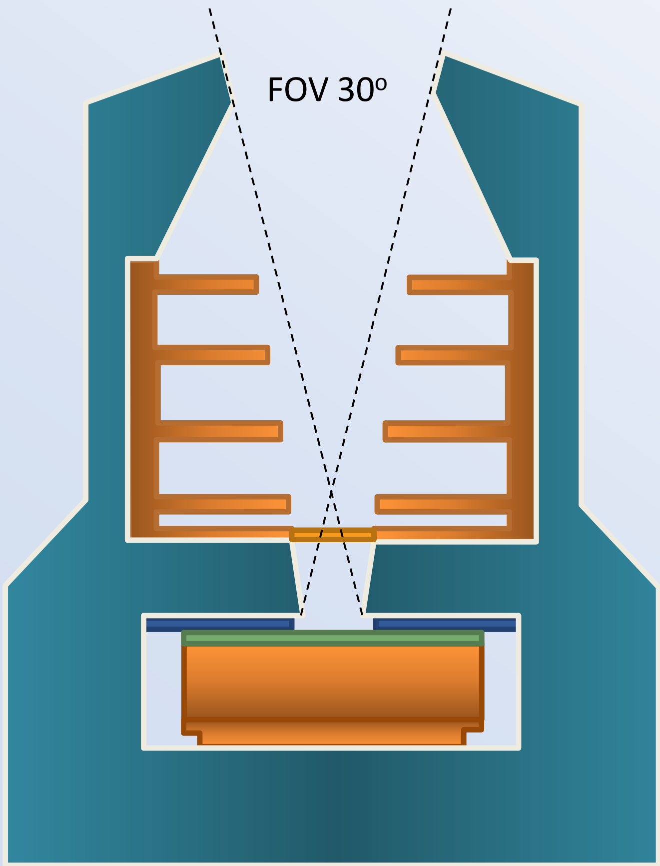
- Satellites up to NOAA-14 carry an older SEM-1 instrument. Starting with NOAA-15 (mid 1998) the satellites carry a newer SEM-2 instrument.

- Problems in MEPED electron detectors:
  - Contamination by energetic protons
  - Significantly non-ideal detector response

- These problems lead to significant errors and long-term inhomogeneities in the electron data, and require correction.

Modeling the electron detector response with Geant4 Monte Carlo 1D simulation

Schematic electron instrument design:



1D detector models (note that the layers are not in scale):

SEM-1 model

Ni-foil, thickness 0.511 mm

Si detector layer thickness 700 mm

SEM-2 model (note the thicker shielding layers)

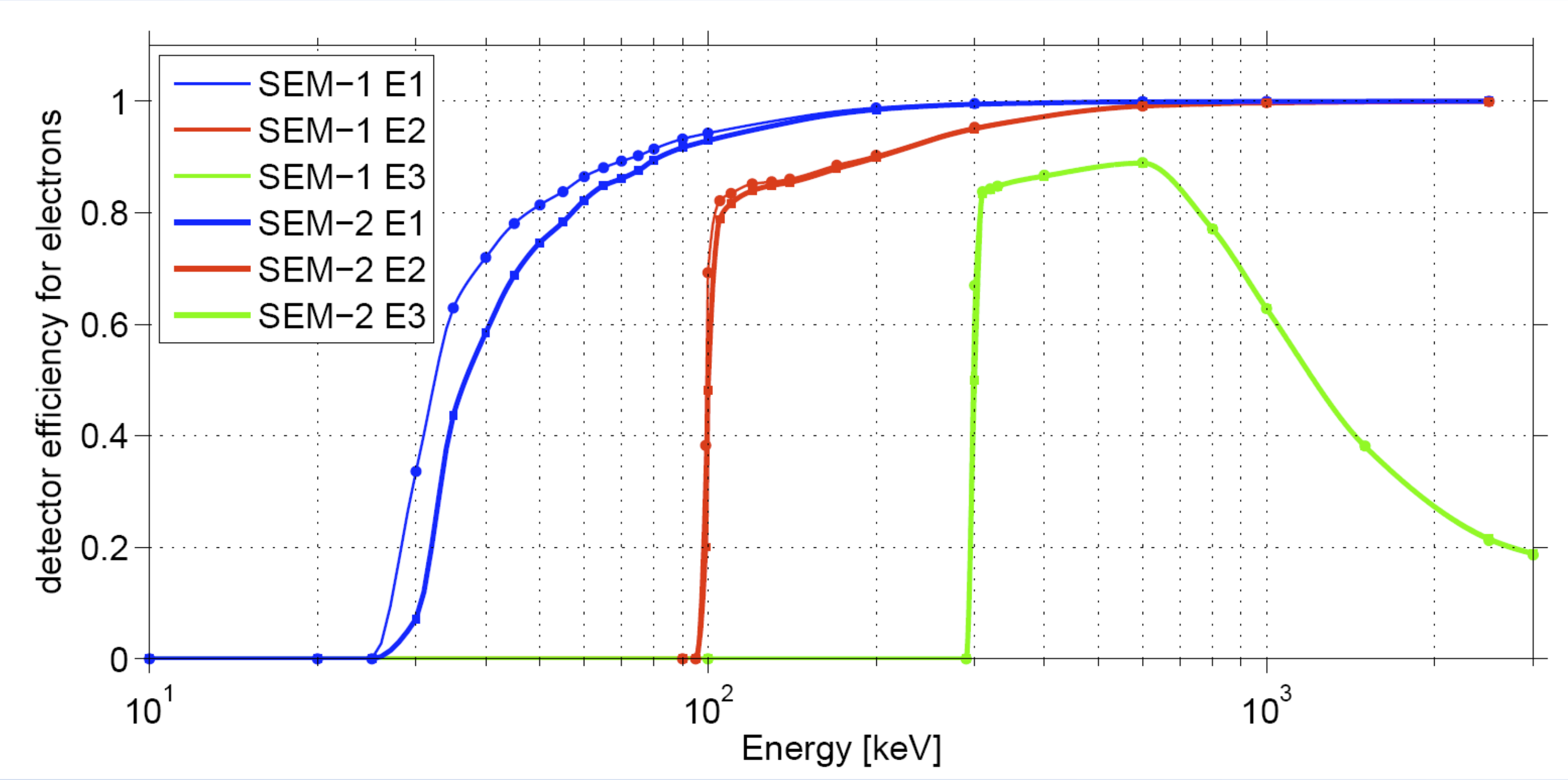
Ni-foil, thickness 0.76 mm

Al layer, thickness 7.41 mm

Si detector layer thickness 700 mm

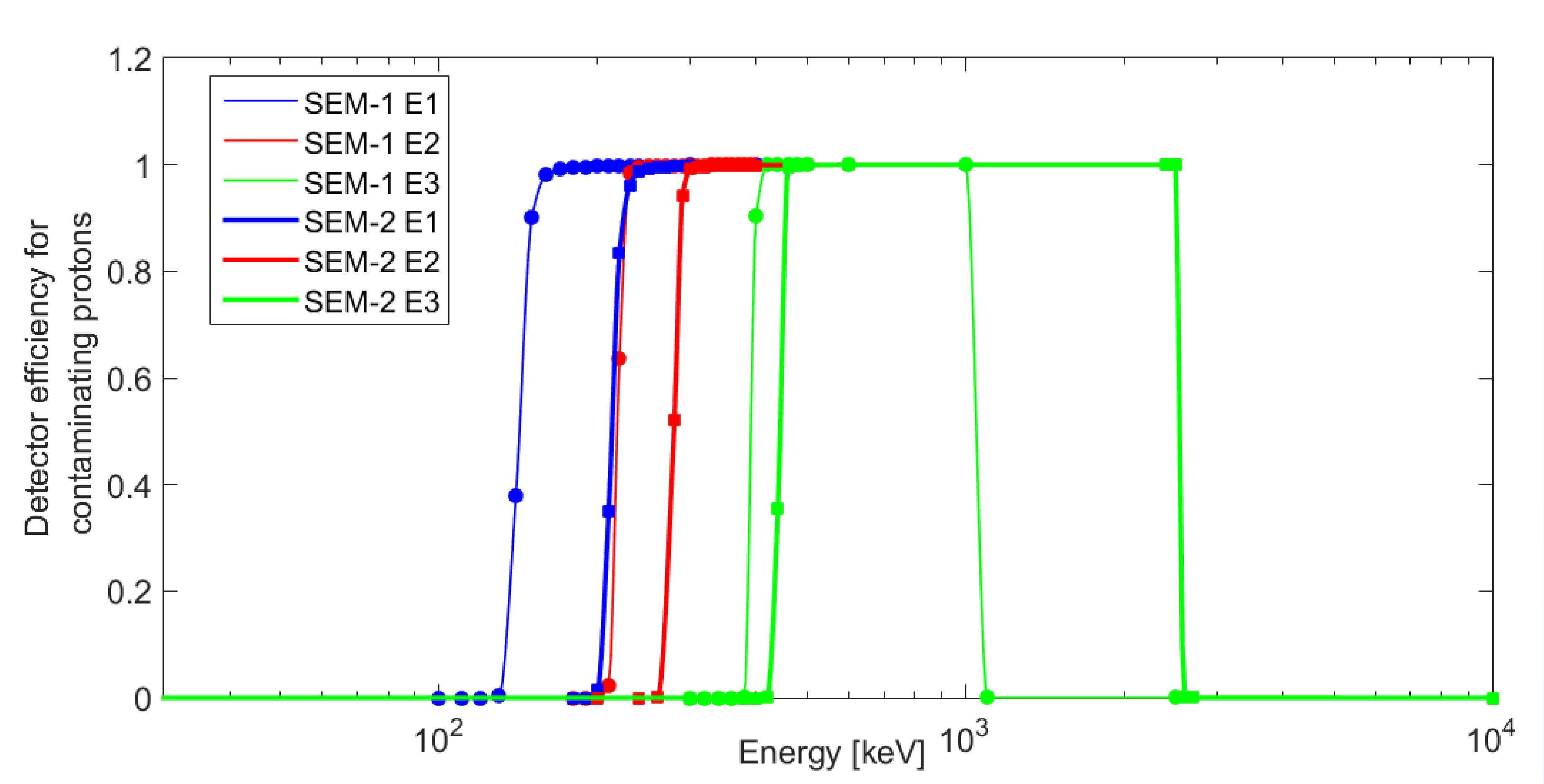
Modeled electron efficiencies

- Efficiencies for all channels deviate from ideal
- Large differences esp. in E1 channel between SEM-1 and SEM-2 instruments.
- The differences in efficiencies result from the thicker Ni-foil in SEM-2



Modeled proton efficiencies

- Efficiencies for all channels deviate from ideal
- Large differences in all channels
- Differences again due to different shielding thicknesses.



Modeled proton efficiencies

- We assume that the differential electron spectrum is piecewise continuous power law:

$$f_e(E) = AE_{xo}^{\gamma_1-\gamma_2}E^{-\gamma_1}, E < E_{xo}$$
$$f_e(E) = AE^{-\gamma_2}, E \geq E_{xo}$$

where E_{xo}=95 keV.

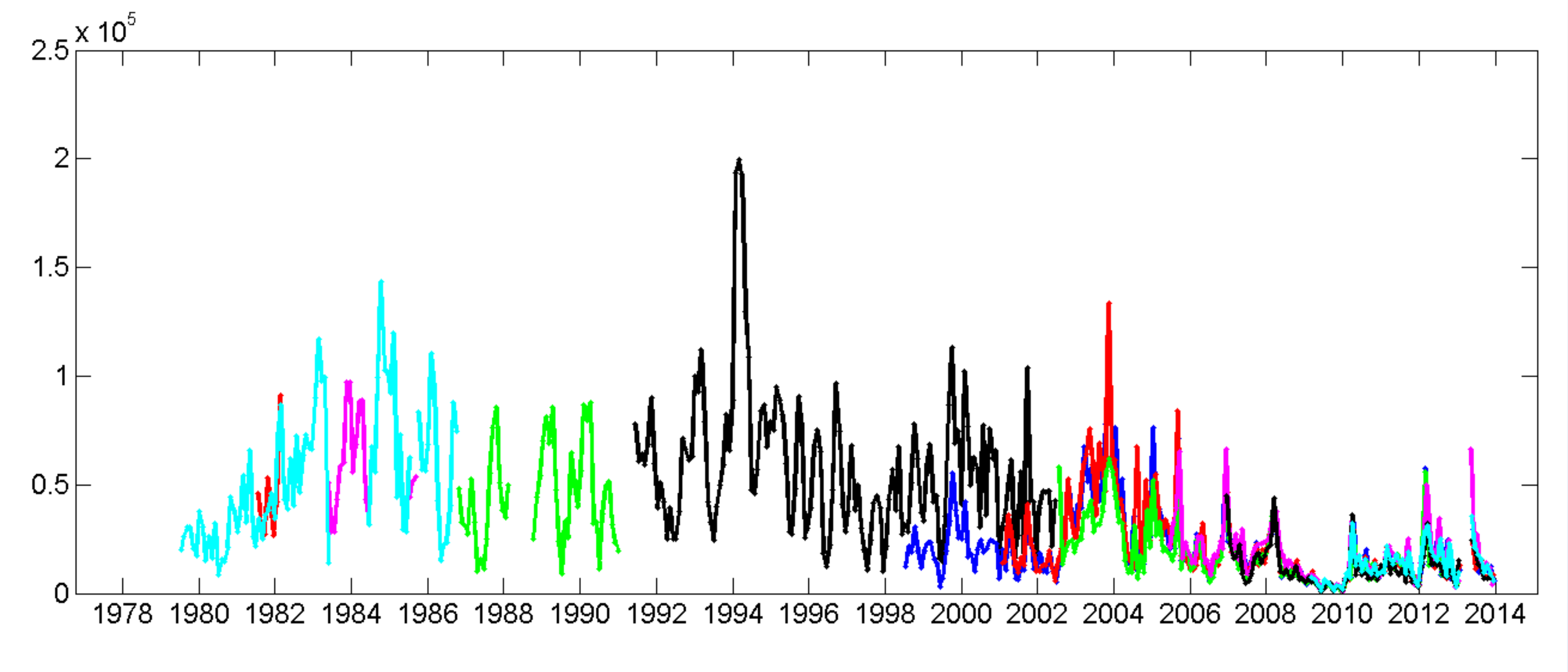
- In discrete form the measured fluxes in i:th channel (i=1,2,3) can be written as

$$j_i = \sum_{E_k=0}^{E_{\max}} \varepsilon_i(E_k) f_e(E_k)\Delta E + \sum_{E_k=0}^{E_{\max}} \rho_i(E_k) f_p(E_k)\Delta E$$

- where $\varepsilon_i(E)$ and $\rho_i(E)$ are the electron and proton efficiency functions for the i:th channel, E_{max}=2.5 MeV, ΔE=1 keV, and the differential proton spectrum $f_p(E)$ is obtained from corrected proton measurements.
- ➔ Numerically solve for A, γ_1 and γ_2 and use them to compute the fluxes at nominal channel energies (>30 keV, >100 keV and >300 keV) by integrating the differential spectrum.

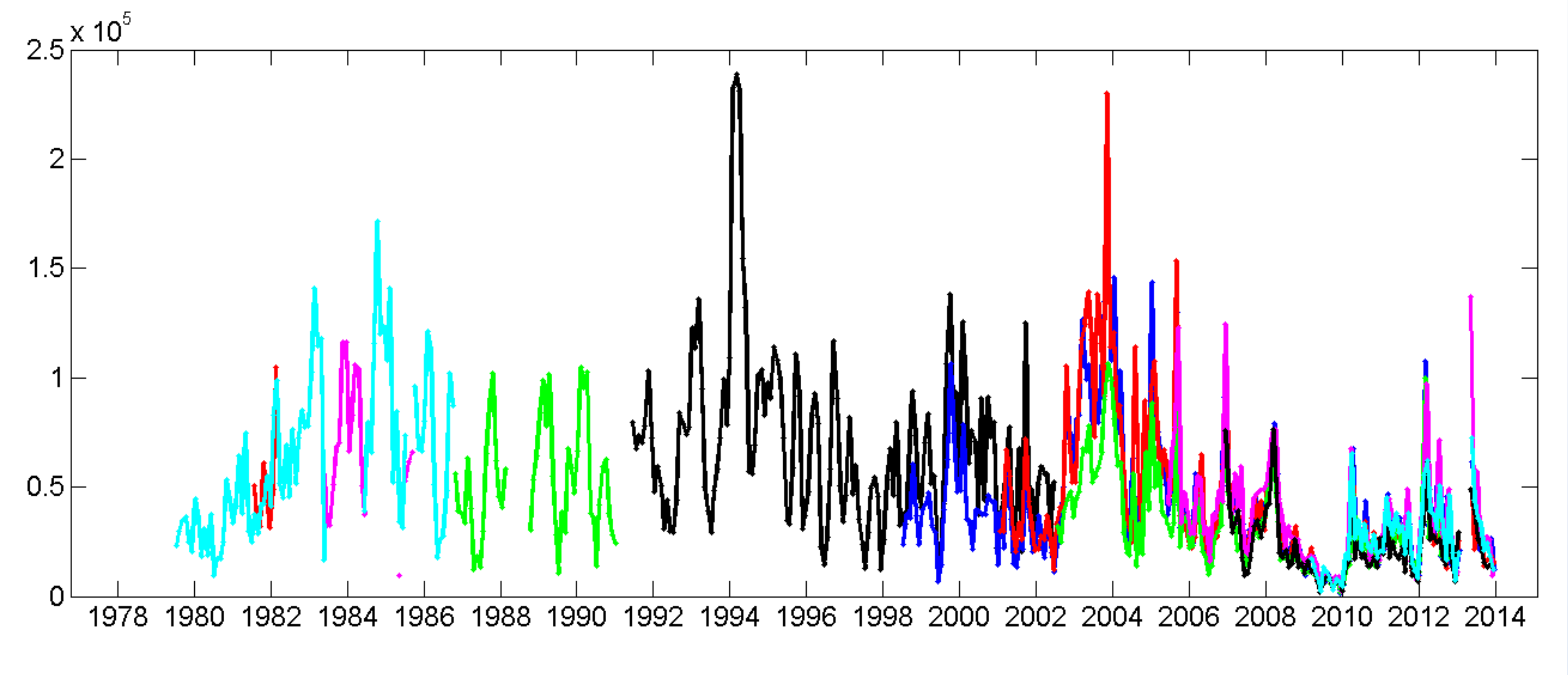
Electrons >30 keV uncorrected

- Newer SEM-2 satellites (since mid-1998) systematically smaller than SEM-1 satellites



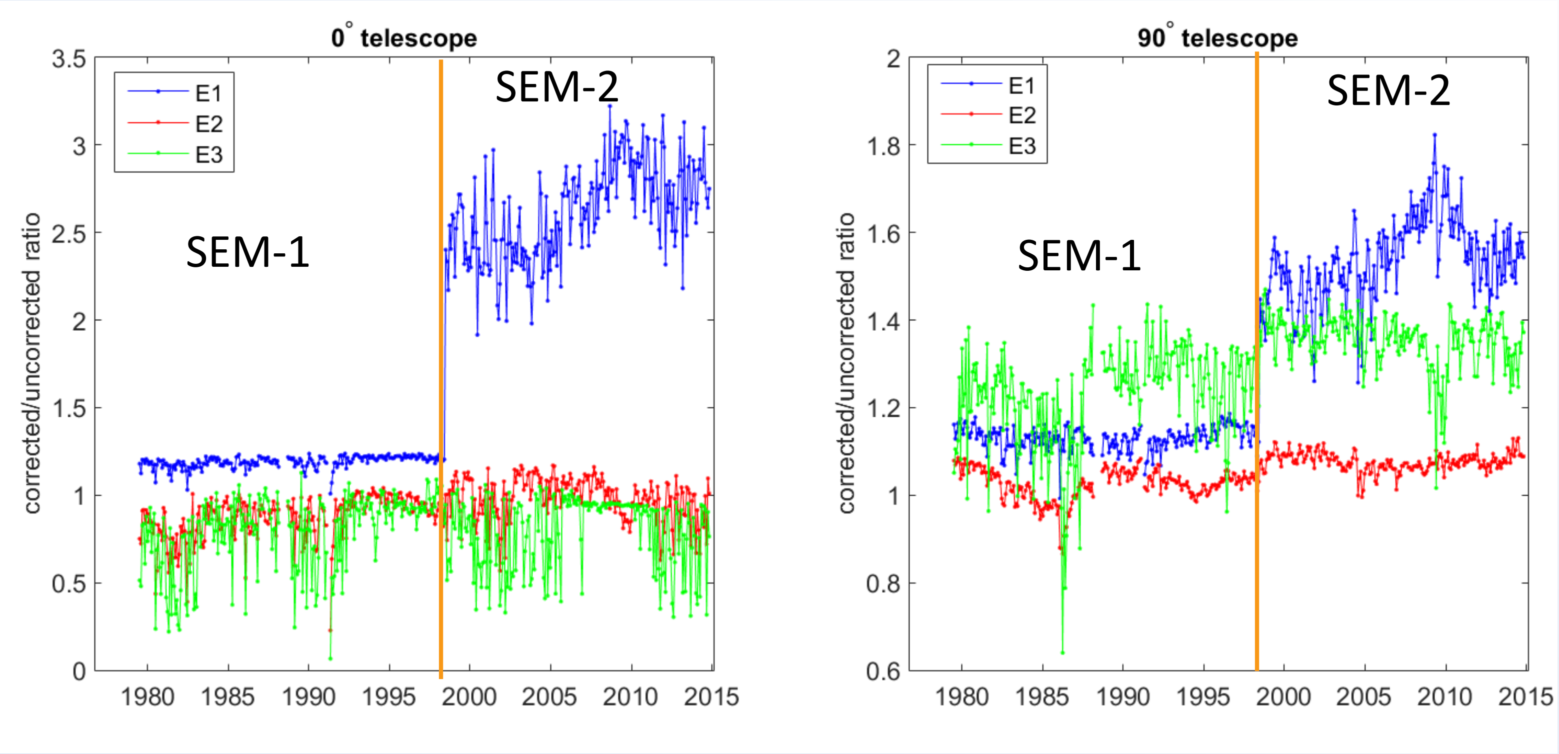
Electrons >30 keV corrected

- Correction raises all fluxes, but SEM-2 fluxes are raised more than SEM-1
- ➔ Correction removes most of the systematic differences between SEM-1 and SEM-2.
- Note: Some systematic difference is still there, because of the different orientation of SEM-1 and SEM-2 particle telescopes.



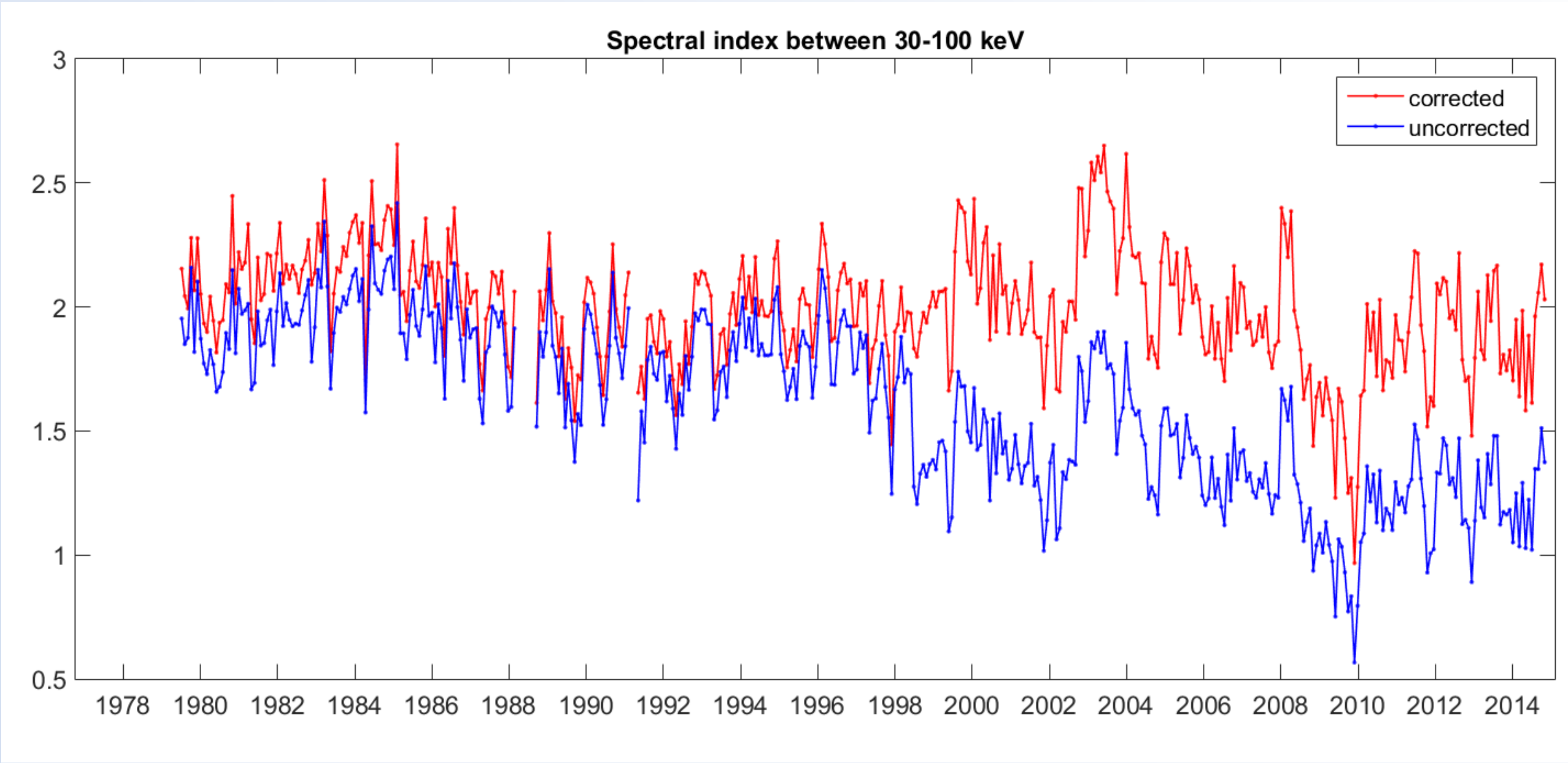
Correction affects different channels differently

- Correction has two parts: removing proton contamination and correcting for detector sensitivity.
- ➔Overall effect of correction is different in different energy channels, telescopes and instrument versions (SEM-1/SEM-2)



Energy spectrum is severely distorted without correction

- Erroneous long-term evolution of spectral index
- Energy spectrum too hard without correction



Summary

- Data from different sources is typically always inhomogeneous
- Constant difference or temporally changing?
- ➔ Understand your data!!

- Calibrate by comparing simulataneous and comparable measurements
- More refined calibrations and homogenization require deep understanding of instruments and how their properties change with time
- Last but not least: Estimate the uncertainty!

THE END