Sea Level Trend Estimation

Paul Eilers

Erasmus Medical Center

Richard Duin

Dutch Ministry of Public Works

Douwe Dillingh

Deltares

Groningen, March 2011

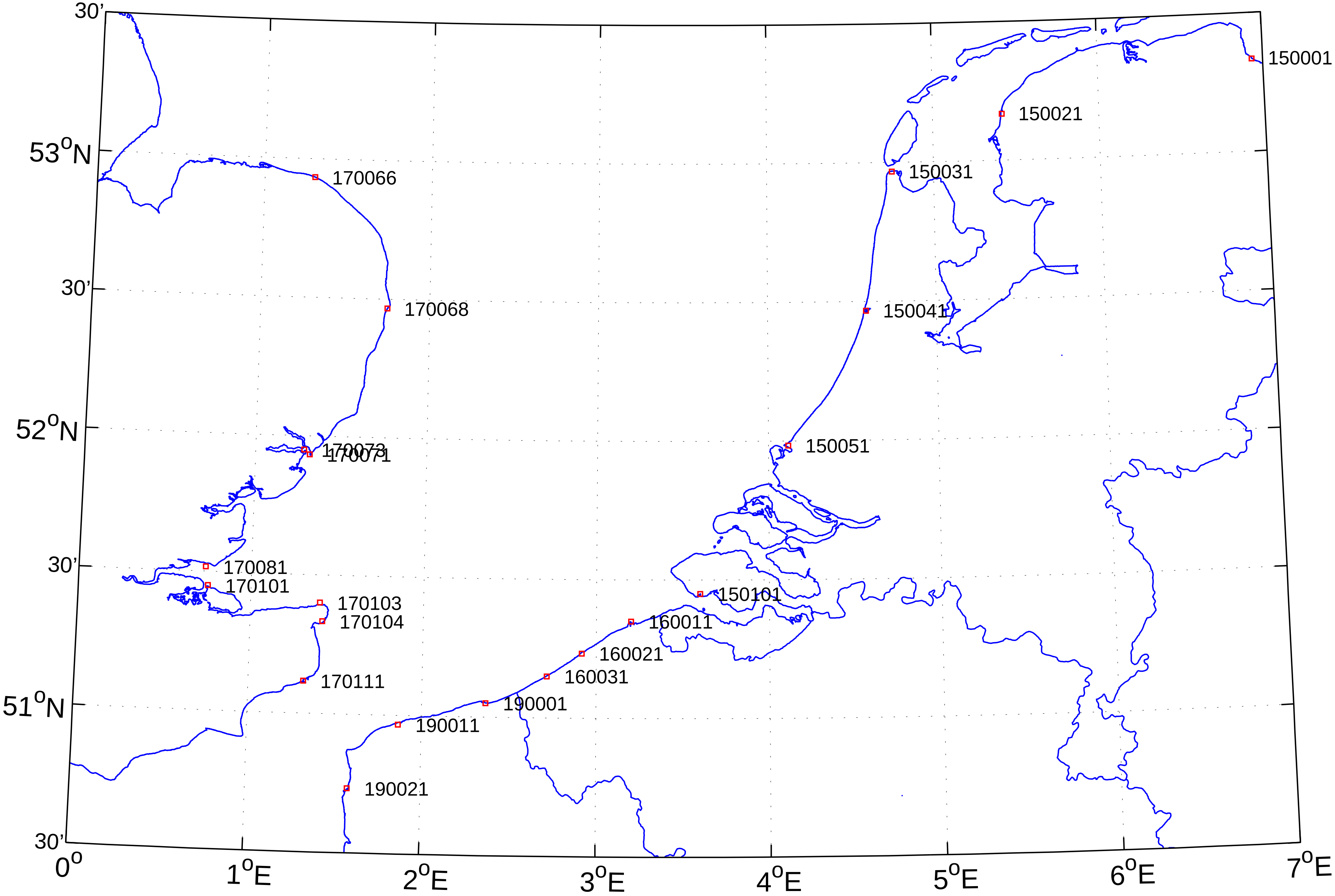
The old country and the sea

- Approximately half of the Netherlands lies below sea level
- But luckily we are protected
- By nature: dunes, and by our own work, strong dikes and dams
- Global warning is a serious concern
- How fast are sea levels rising?
- We estimate trends from observations
- We have lots of data

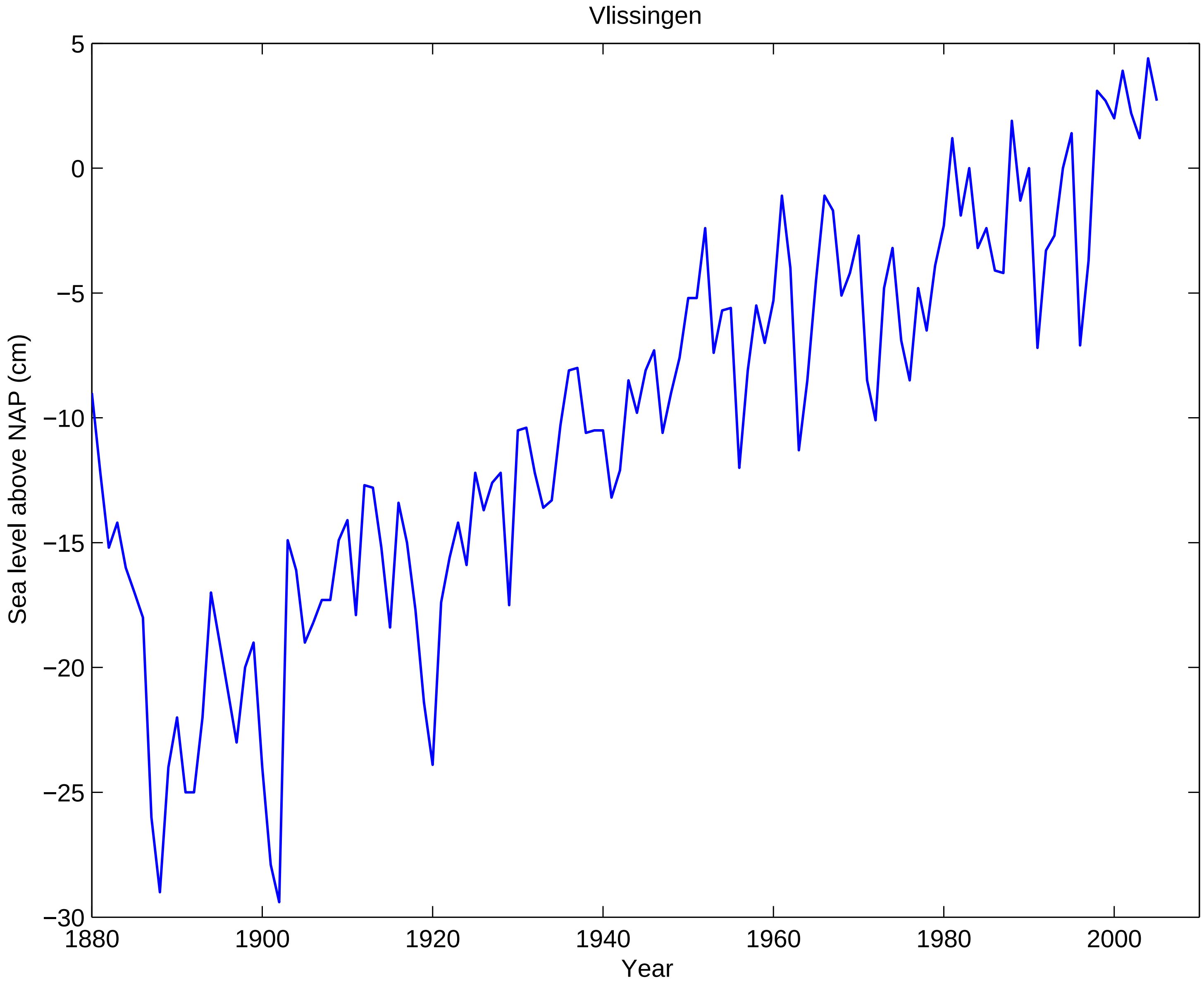
The Dutch sea level monitoring network

- Started in almost 200 years ago (Vlissingen)
- Gradually extended (now over 30 stations)
- Six so-called principal stations
- Continuous measurements, translated to averages
- Monthly and yearly averages of daily high, low, mean levels
- Available in a global database (in the UK)
- Permanent Service for Mean Sea Level (www.psmsl.org)

The stations on the map



A typical record of yearly averages



Trend estimation with the Whittaker smoother

- Time series y_i and trend to estimate z_i, for $i = 1, \ldots, m$
- Two competing goals for z: fidelity to data and smoothness
- Measure *roughness* of z using (second order) differences
- Use penalized regression (Δ is differencing operator), minimize

$$Q = S + \lambda R = \sum_i (y_i - z_i)^2 + \lambda \sum_i (\Delta^2 z_i)^2$$

- Tune smoothness with λ
- This goes back at least to Whittaker (1923)

Implementing the Whittaker smoother

- Matrix D, such that $Dz = \Delta^2 z$

$$Q = ||y - z||^2 + \lambda ||Dz||^2$$

- Explicit solution:

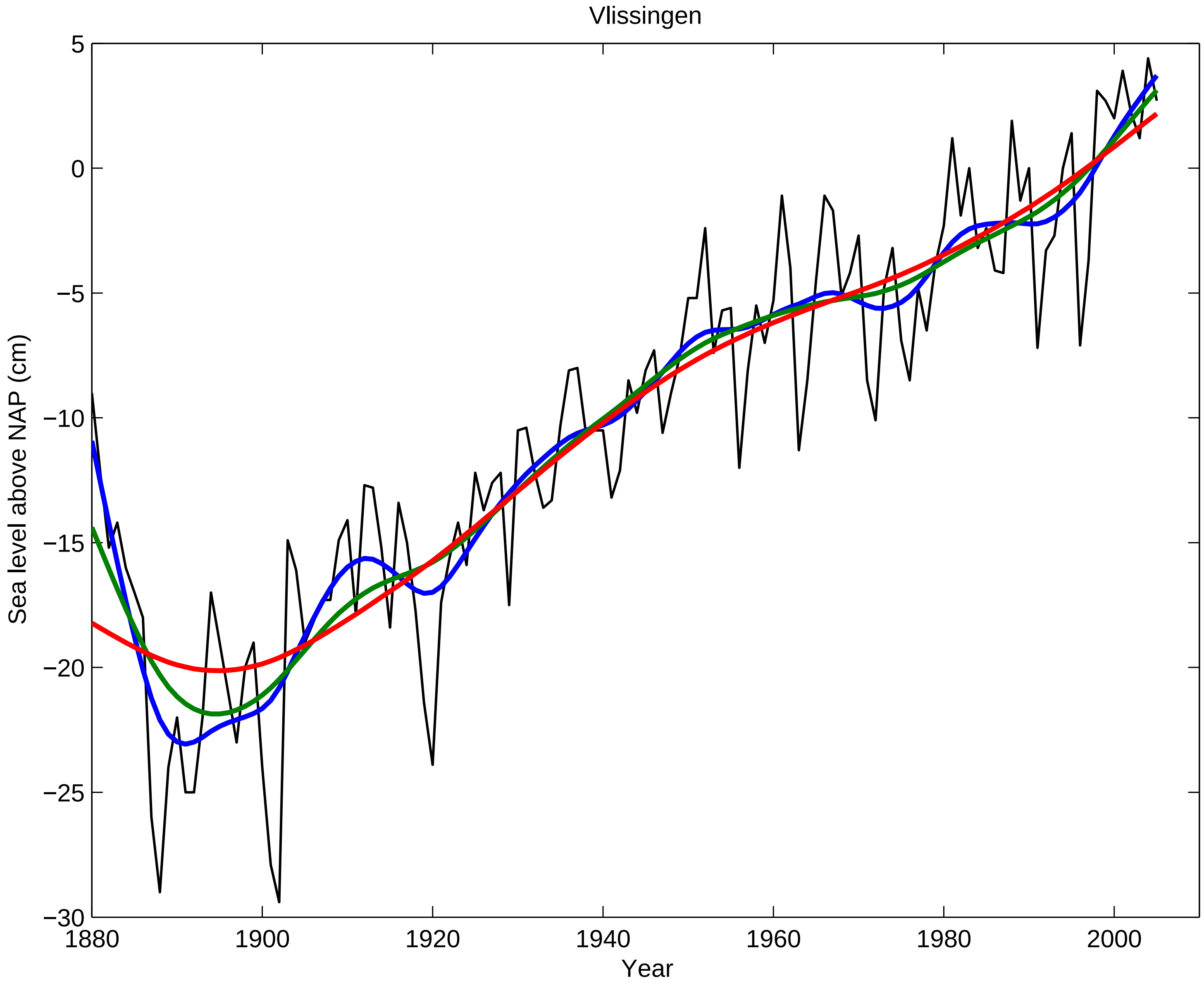
$$(I + \lambda D'D)\hat{z} = y$$

- Large linear system (m equations), but very sparse

- Five lines of Matlab (or R) code

- `D = diff(diag(m), diff = 2)`

Results for $\lambda = 100, 1000, 10000$



Automatic linear extrapolation

- What if data are missing?
- Introduce a 0/1 weight vector w, and matrix $W = \text{diag}(w)$
- New objective function and solution

$$Q = (y - z)'W(y - z) + \lambda||Dz||^2; \quad (W + \lambda D'D)\hat{z} = Wy$$

- Extend observations for $i > m$ with "missing" data
- Automatic linear extrapolation will occur

Effective model dimension

- Hat matrix H defined by $\hat{y} = Hy$
- Here we have
$$H = (W + \lambda D'D)^{-1} W$$
- Effective model dimension: $\mathrm{ED} = \mathrm{trace}(H) = \sum_i h_{ii}$
- Follows directly from Ye's (1998) definition
$$\mathrm{ED} = \sum_i \frac{\partial \hat{y}_i}{\partial y_i}$$

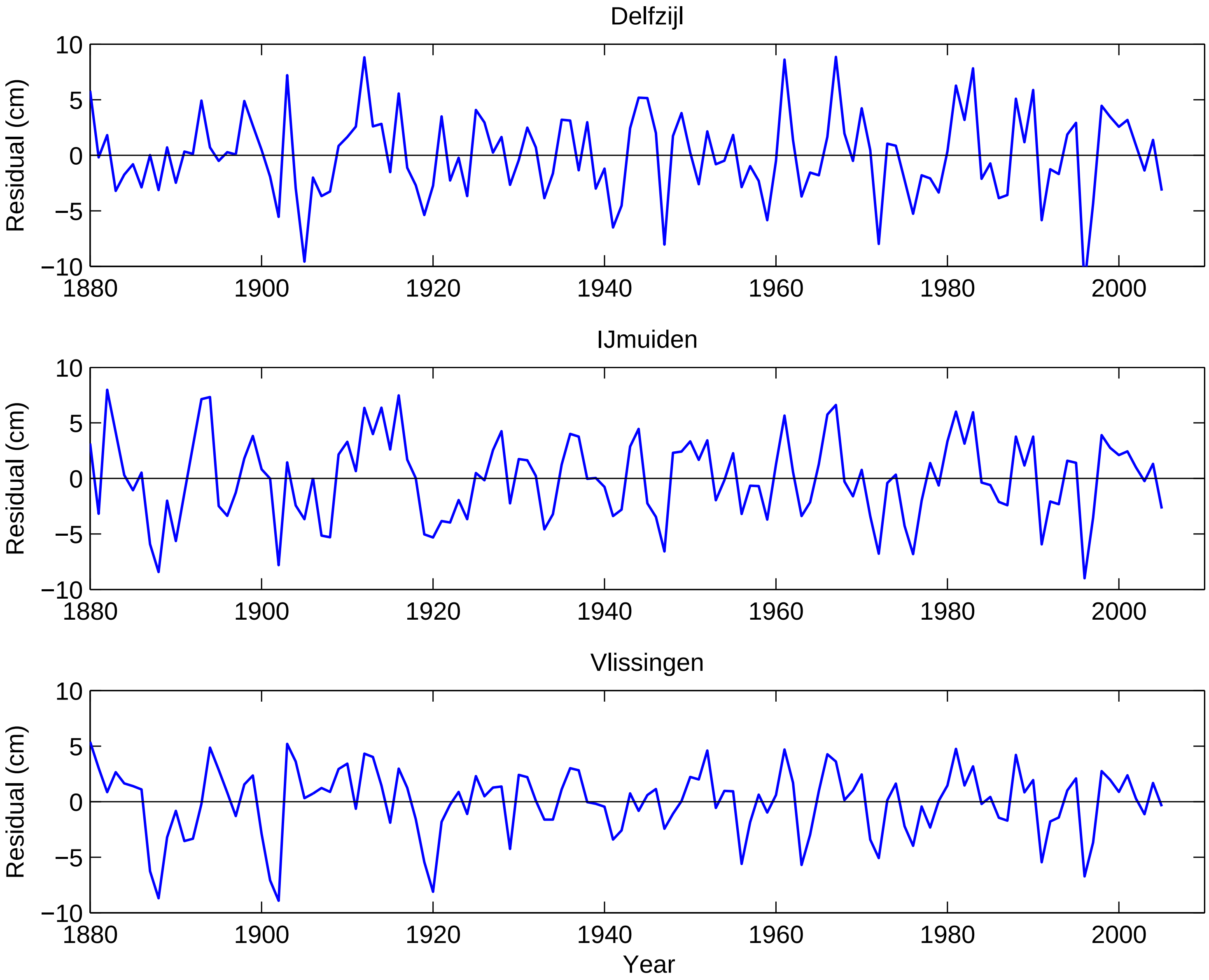
Optimal smoothing?

- It would be nice to optimize smoothing automatically
- Using cross-validation or AIC
- Unfortunately the errors are not independent at all
- A very light amount of smoothing is indicated; not useful
- Thus we use our carpenter's eye
- Standard error estimates are too optimistic

The crucial question

- Is it wrong? Perhaps.
- Is it useful? Yes!
- Is it useful enough? No!
- We can do better
- The residuals show special patterns
- Combining stations gives extra power

Residuals look very similar



Modeling the structure in residuals

- The residuals look very similar
- In pattern and in size
- New model for multiple stations ($j = 1, \ldots, n$):

$$y_{ij} = z_{ij} + u_i + e_{ij}$$

- Smooth trend z_{ij} per station
- Shared disturbances u_i
- Errors e_{ij}

Is this new?

- No: a special case of Seemingly Unrelated Regressions (SUR)
- A popular model in econometrics
- Correlations between "errors" connect regression models
- This has been extended to non-parametric models
- Here we do penalized regression
- Hence: Seemingly Unrelated Penalized Regressions
- Nickname: SUPR

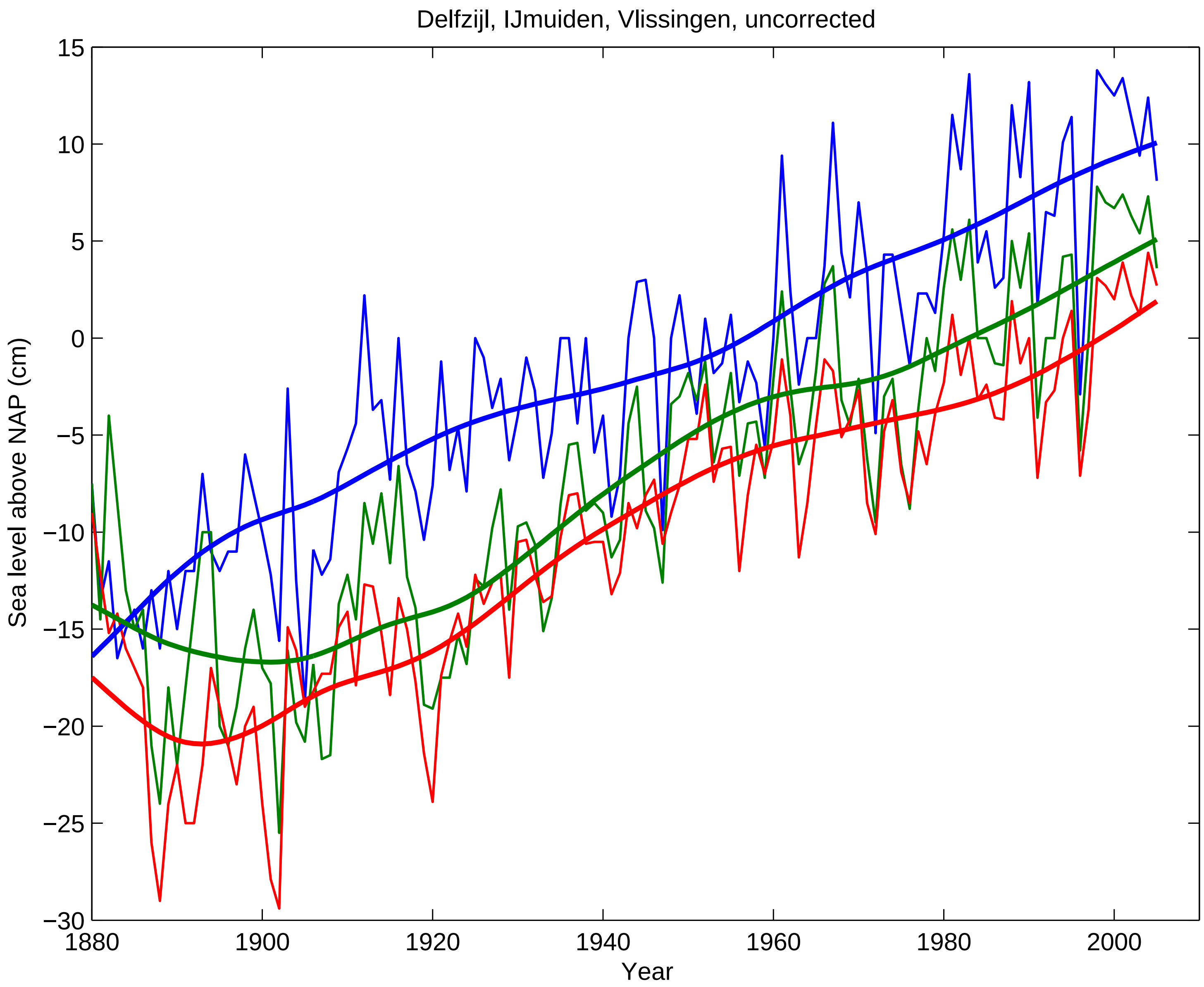
Estimation

- Equation system shown for $n = 3$
- Notation: y_j and z_j are the vectors for stations j

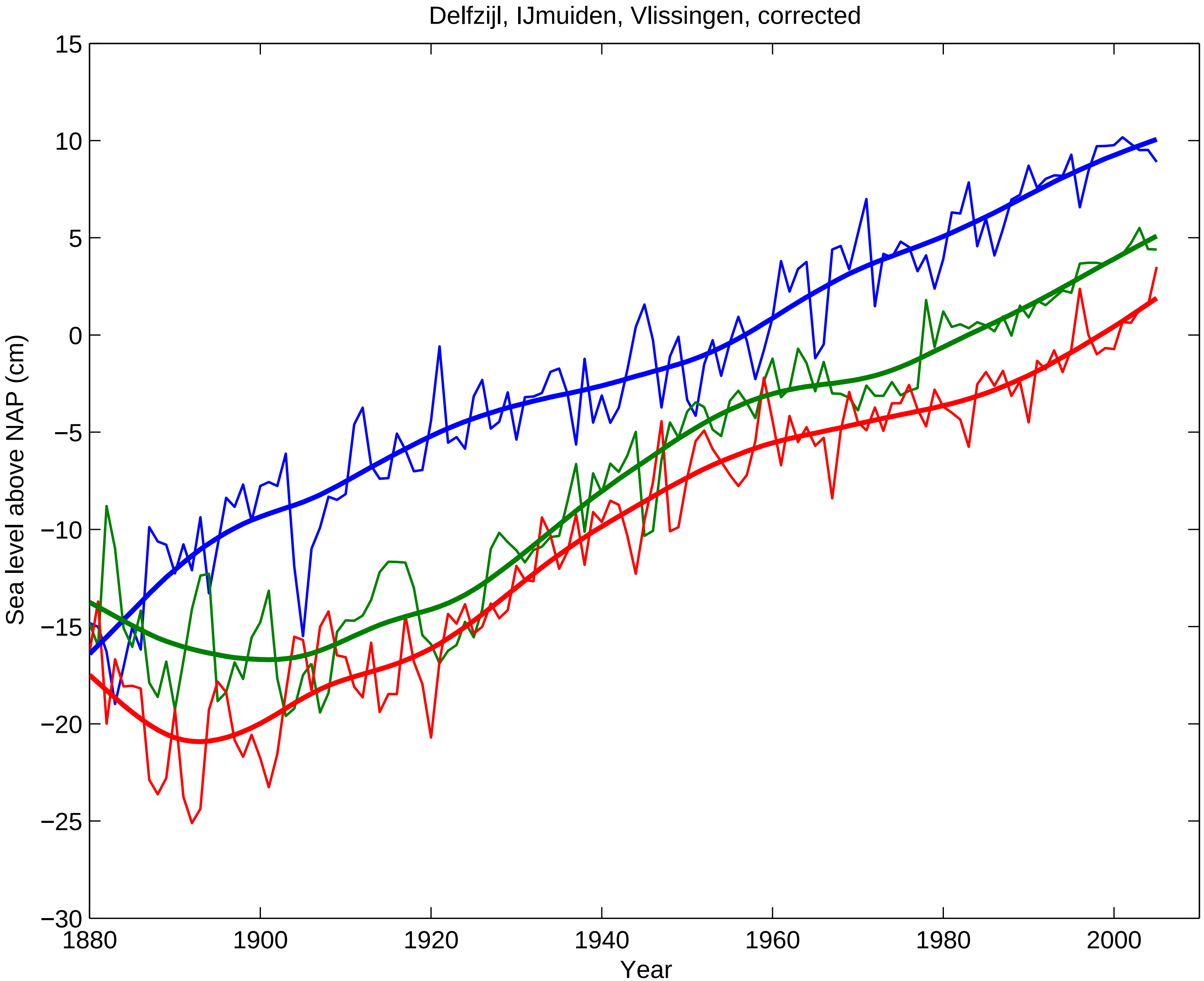
$$\begin{bmatrix} I + \lambda D'D & \mathbf{0} & \mathbf{0} & I \\ \mathbf{0} & I + \lambda D'D & \mathbf{0} & I \\ \mathbf{0} & \mathbf{0} & I + \lambda D'D & I \\ I & I & I & (3+\kappa)I \end{bmatrix} \begin{bmatrix} \hat{z}_i \\ \hat{z}_2 \\ \hat{z}_3 \\ \hat{u} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_1 + y_2 + y_3 \end{bmatrix}$$

- Note κ: (small) ridge penalty for identifiability
- Again a sparse system

Before removal of shared disturbances



After removal of shared disturbances



Standard errors

- Estimate of error variance (no weights):

$$\hat{\sigma}^2 = ||y - z||^2/(m - \text{ED})$$

- With weights:

$$\hat{\sigma}^2 = (y - z)'W(y - z)/[\text{tr}(W) - \text{ED})]$$

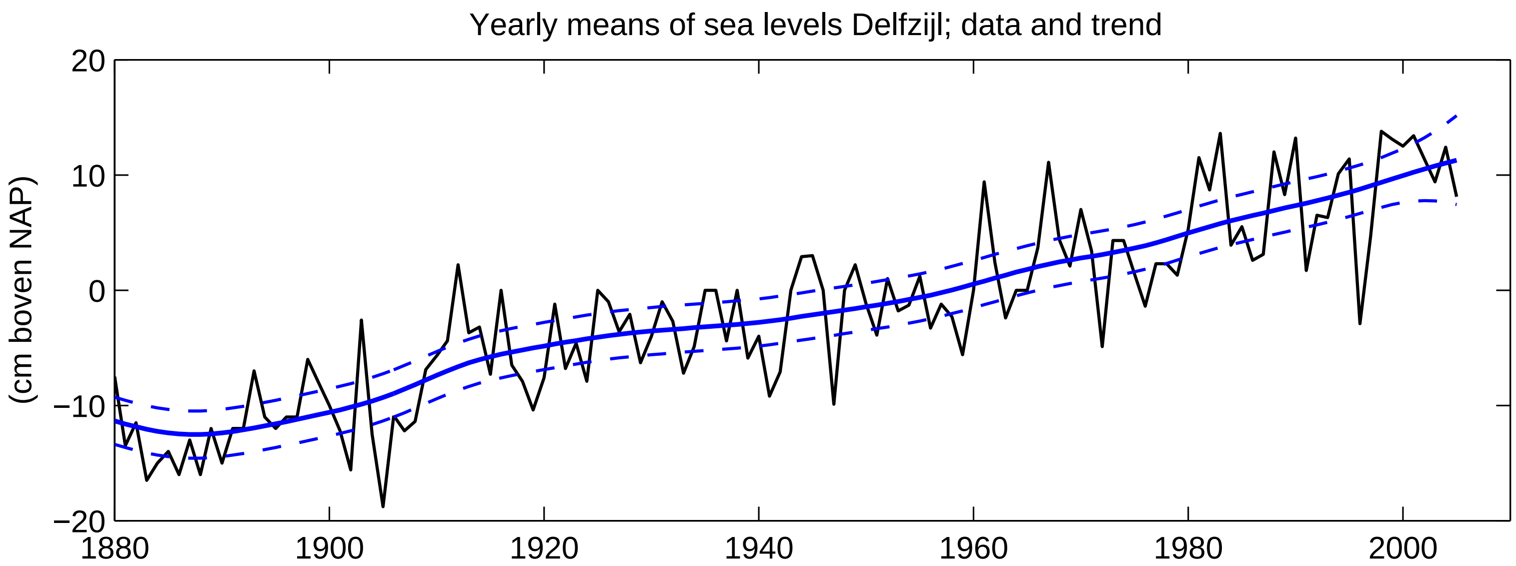
- Covariance matrix of $\hat{y}$:

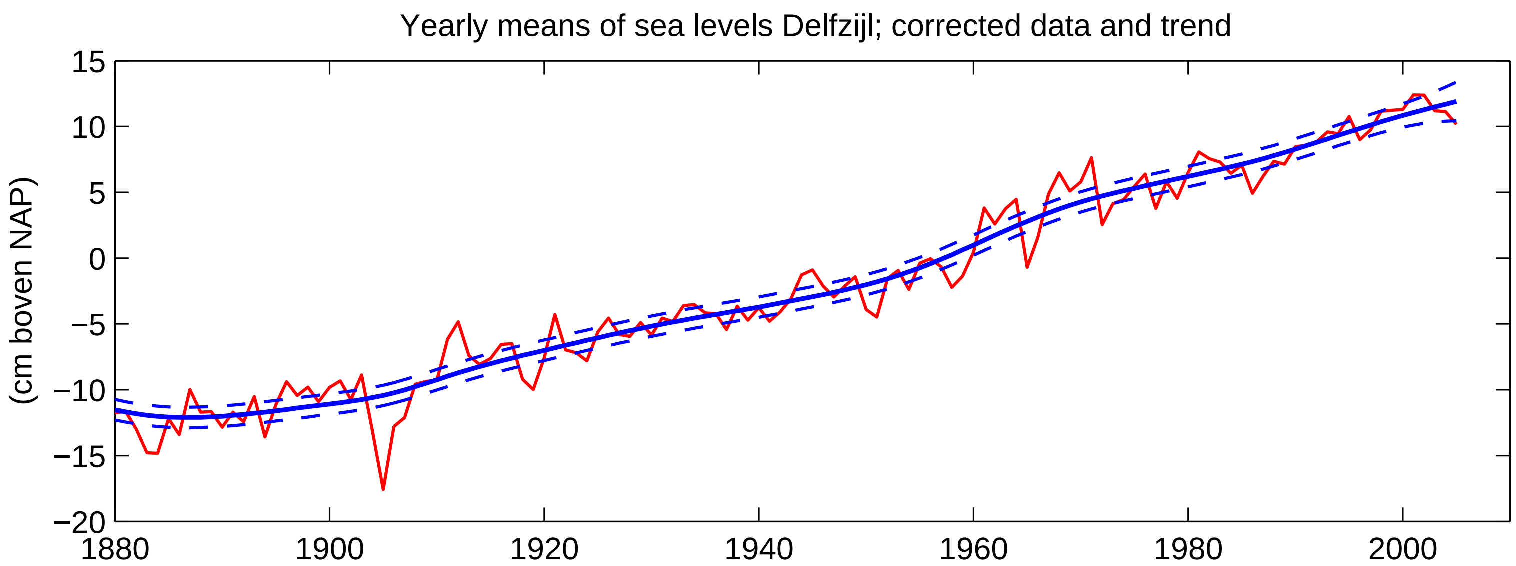
$$C = \hat{\sigma}^2(I + \lambda D'D)^{-1}$$

- Diagonal of C gives us error bands

- Similar computations for the extended model

Standard error bands for the two models

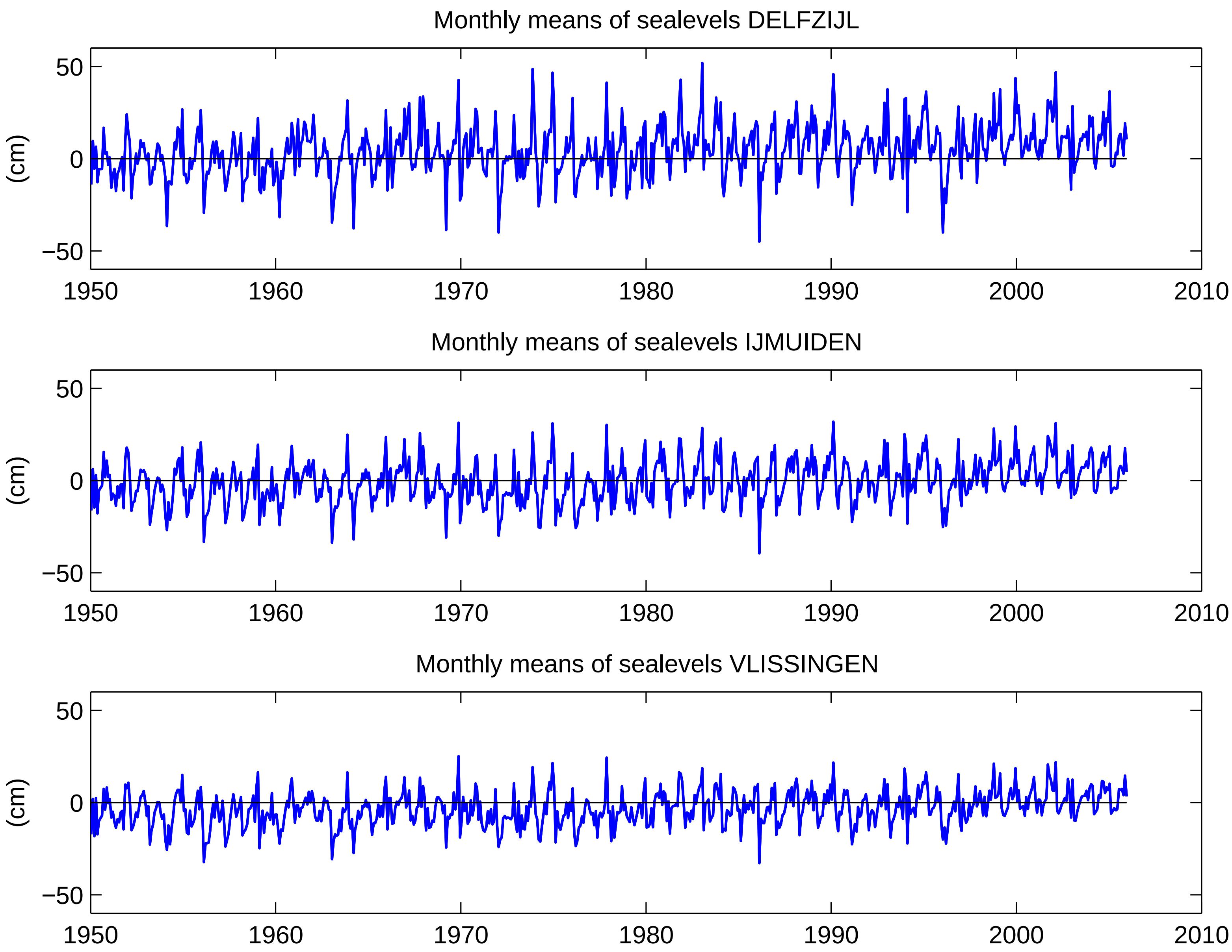




Monthly data

- Let's go to a higher frequency: monthly averages
- Does the same model hold?
- Yes, but with a twist
- Patterns of disturbances are very similar
- But their size varies between stations

Monthly data for three stations



An extended model

- Assign a scale parameter, s, to each station:

$$y_{ij} = z_{ij} + s_j u_i + e_{ij}$$

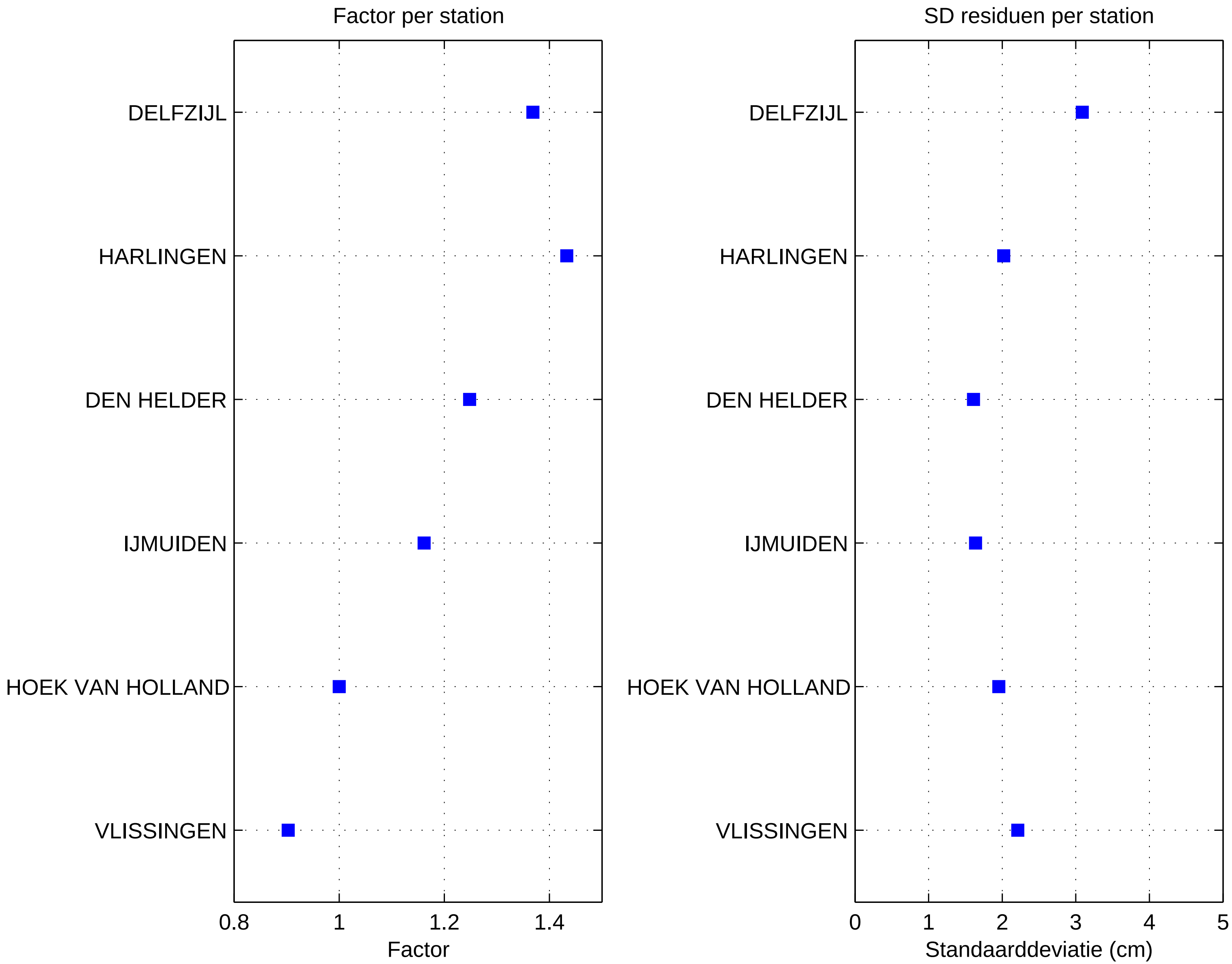
- Normalizing condition: $\sum_j s_j^2 = n$

- Alternating least squares algorithm:
  - smooth $y_{ij} - \tilde{s}_j \tilde{u}_i$, to update z
  - compute singular value decomposition of $Y - \tilde{Z}$
  - "largest" singular vectors give new $\tilde{s}$ and $\tilde{u}$

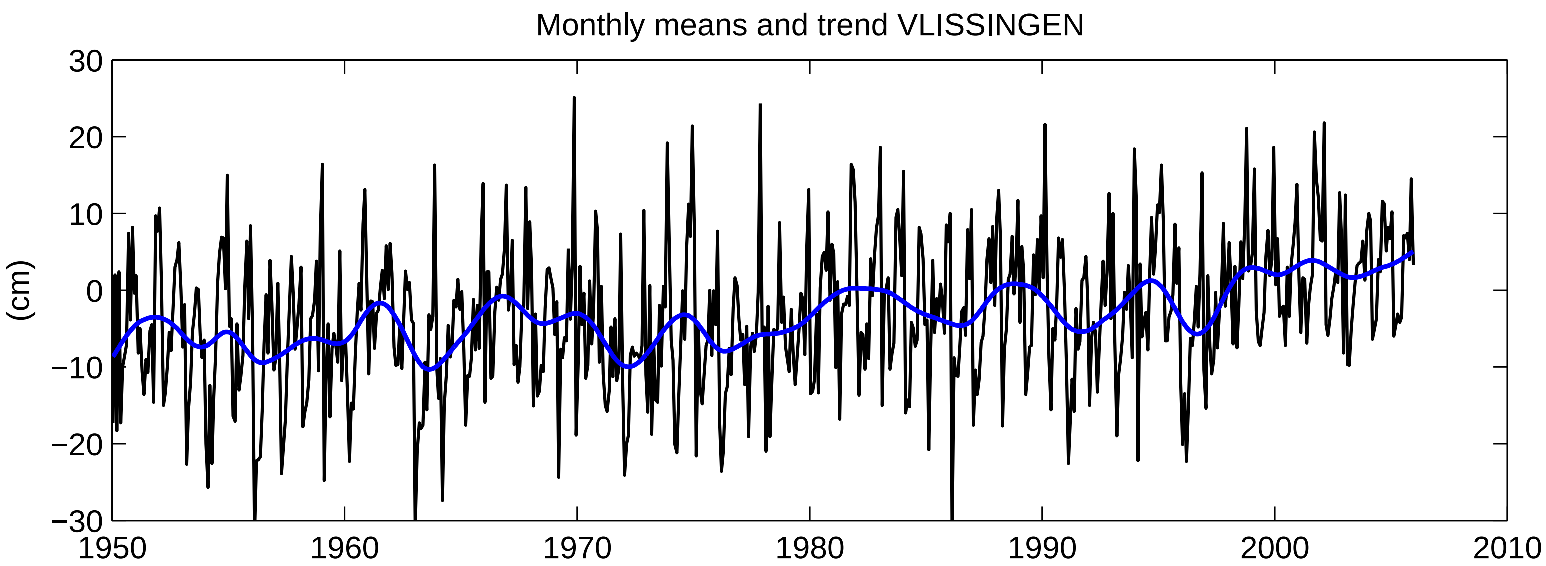
- Start with all zero $\tilde{s}$ and $\tilde{u}$

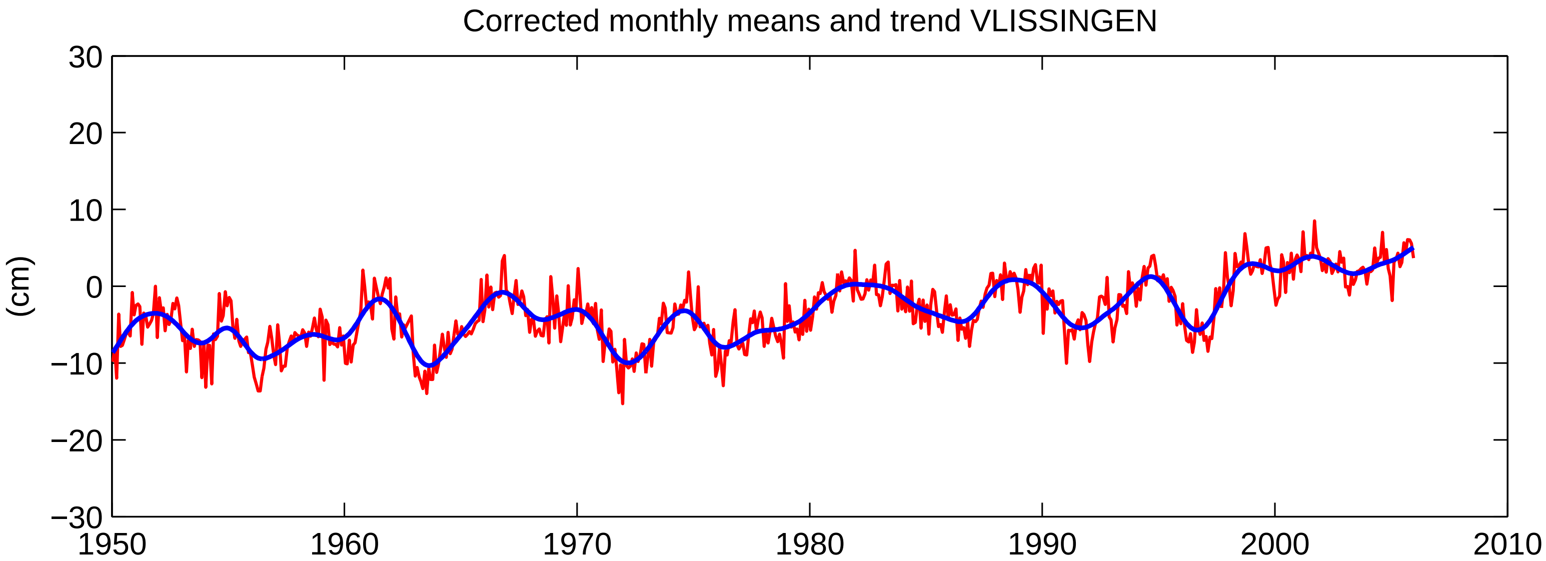
- This model seems to be new

Scale factors decrease from North to South



Correction works very well for monthly data





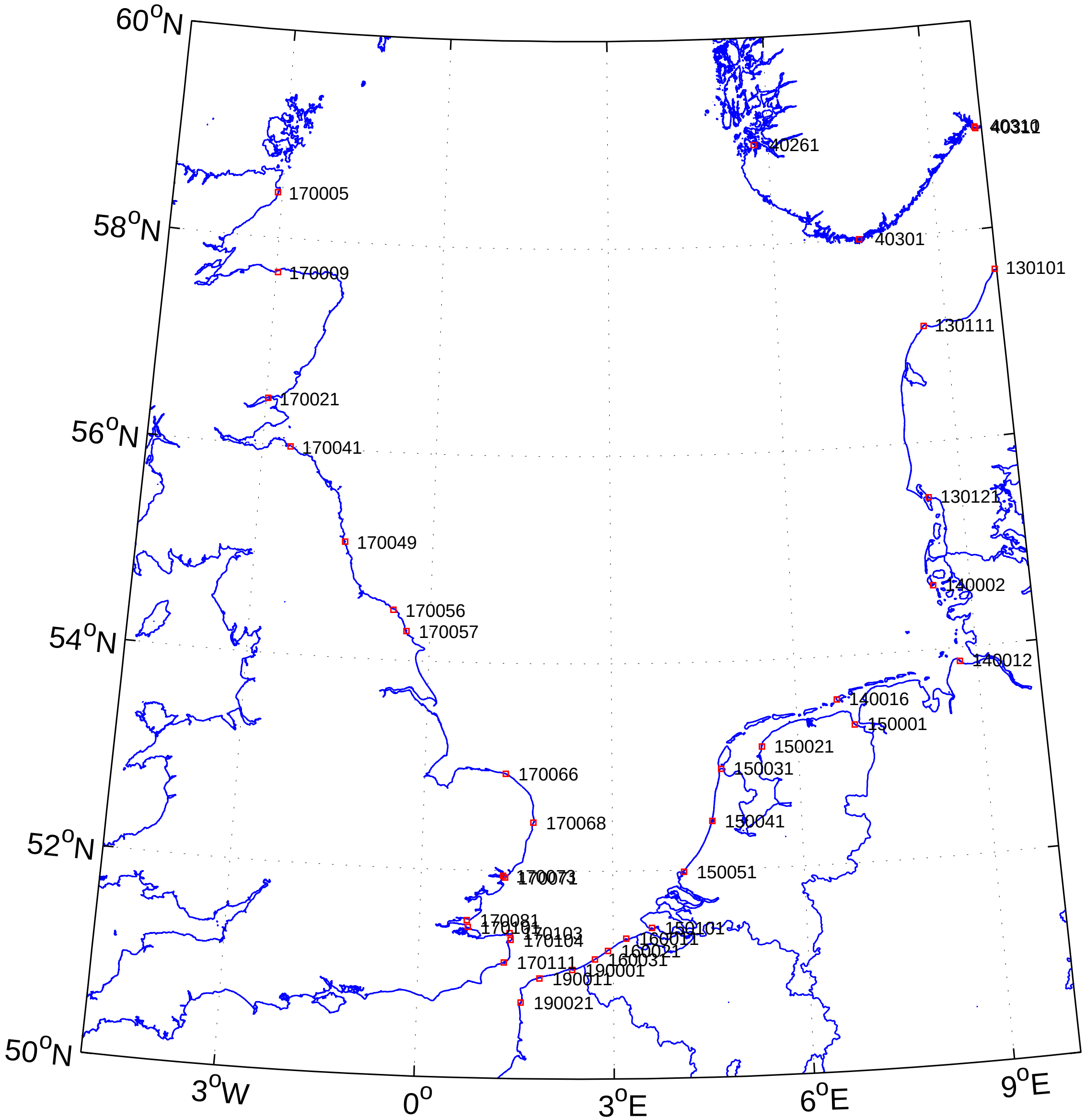
A complication and a solution

- Silent assumption: shared disturbances have no trend
- Only high frequencies
- Easy solution in the extended model
- Remove trends from $y_j - \tilde{z}_j$ in each iteration
- Before doing the SVD step

A larger region

- We also studied the Eastern North Sea Coast
- From Norway to France
- Similar patterns
- Largest disturbances near Cuxhaven (German Bight)
- Results not shown

The larger region on the map



To do: technical improvements

- Proper time series model for errors (AR?)
- Optimal smoothing
- Seasonal model for monthly disturbances
- Use small set (10 to 20) of penalize B-splines
- Smaller systems of equations
- Important for large sets of stations

To do: substantive improvements

- What explains the similar disturbances?
- Coupled large-scale water movements?
- Meteorological effects?
- What about the Western coast of the North Sea?
- What about other areas on the globe?
- The PSMSL contains over 1500 stations
- But many have only sparse data