Local estimation of extremal trends in North Sea surge elevations

at the *IMA meeting on Flood Risk Assessment, 7-8 September 2004, University of Bath*

Adam Butler
Lancaster University
a.butler@lancaster.ac.uk

Janet E. Heffernan
Lancaster University
j.heffernan@lancaster.ac.uk

Roger A. Flather
Proudman Oceanographic Laboratory
raf@pol.ac.uk

Jonathan A. Tawn
Lancaster University
j.tawn@lancaster.ac.uk

Introduction

- Storm surges are a source of coastal flood risk
- Defense structures are typically overtopped under conditions of storm surge and high tide
- Return level: coastal engineering design criterion

- Return levels can be estimated from data or from model output using statistical methods e.g. extreme value (EV) methods
- Standard EV approaches assume stationarity and fail to exploit spatial structure

Extreme sea levels

- Sea level = Mean sea level + Tide + Surge + Tide-surge interaction + Waves
- Surges are generated by wind stress and by differences in air pressure (Flather, 2001)
- Interactions mean that storm surges tend to occur on rising tides (Prandle and Wolf, 1978)

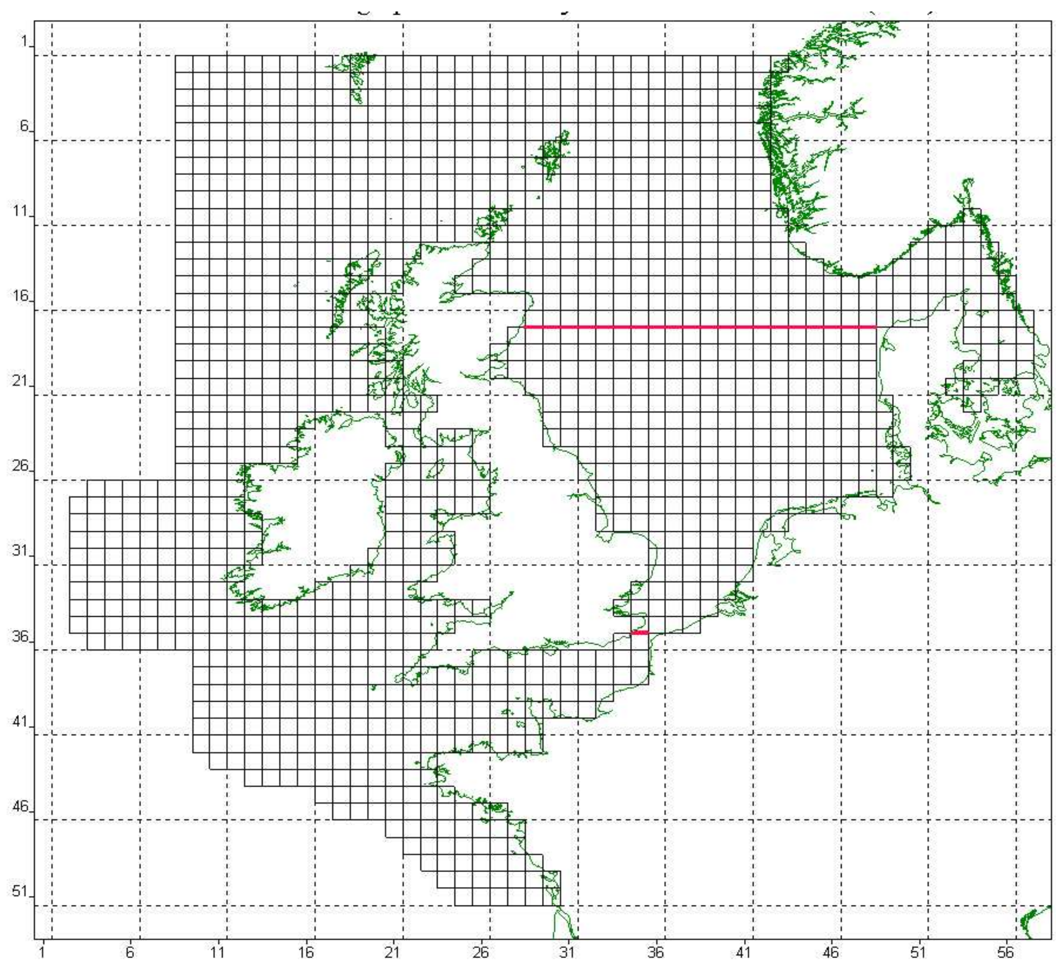
- Observational data derived from tide gauges
- Hydrodynamical models can be used to study surge characteristics (Bode and Hardy, 1997)

The CSX-DNMI hindcast

- CSX is a 2d numerical storm surge model for the European Continental Shelf (Flather et al., 1991)
- Developed by POL as an operational surge model.
- CSX has a resolution of 35km × 35km × 15 min.

- DNMI have reconstructed a high-quality air pressure dataset for 1955- (Reistad and Iden, 1995)
- Flather et al. (1998) analyse surges generated by forcing CSX with interpolated DNMI data
- We reanalyse this CSX-DNMI hindcast

CSX model grid



Identification of storm surges

- Use an extreme value approach (Coles, 2001)
- Proceed with a site-by-site analysis
- At each site: hourly data for n years $t_1, \ldots, t_n$

- Extract the r-largest storm peaks per year, using a declustering algorithm (Tawn, 1988)
- Let $x_j^{(k)}$ = k-th largest storm peak in year t_j

The r-largest model for extremes

- Assume $\mathbf{X}_j = \left(X_j^{(1)}, \ldots, X_j^{(r)}\right)$ are jointly distributed according to the r-largest model (Weissman, 1978; Smith, 1986; Tawn, 1988).

- r-largest model has parameters $\theta_j = (\mu_j, \sigma_j, \xi_j)$; (location, scale and shape parameters)
- A design parameter is the level $z_{j,N}$ such that

$$\mathbb{P}\left[X_j^{(1)} > z_{j,N}\right] = 1/N;$$

  the "N-year return level"
- Flather et al. (1998) assume that $\theta_1 = \ldots = \theta_n$ & use maximum likelihood methods for inference

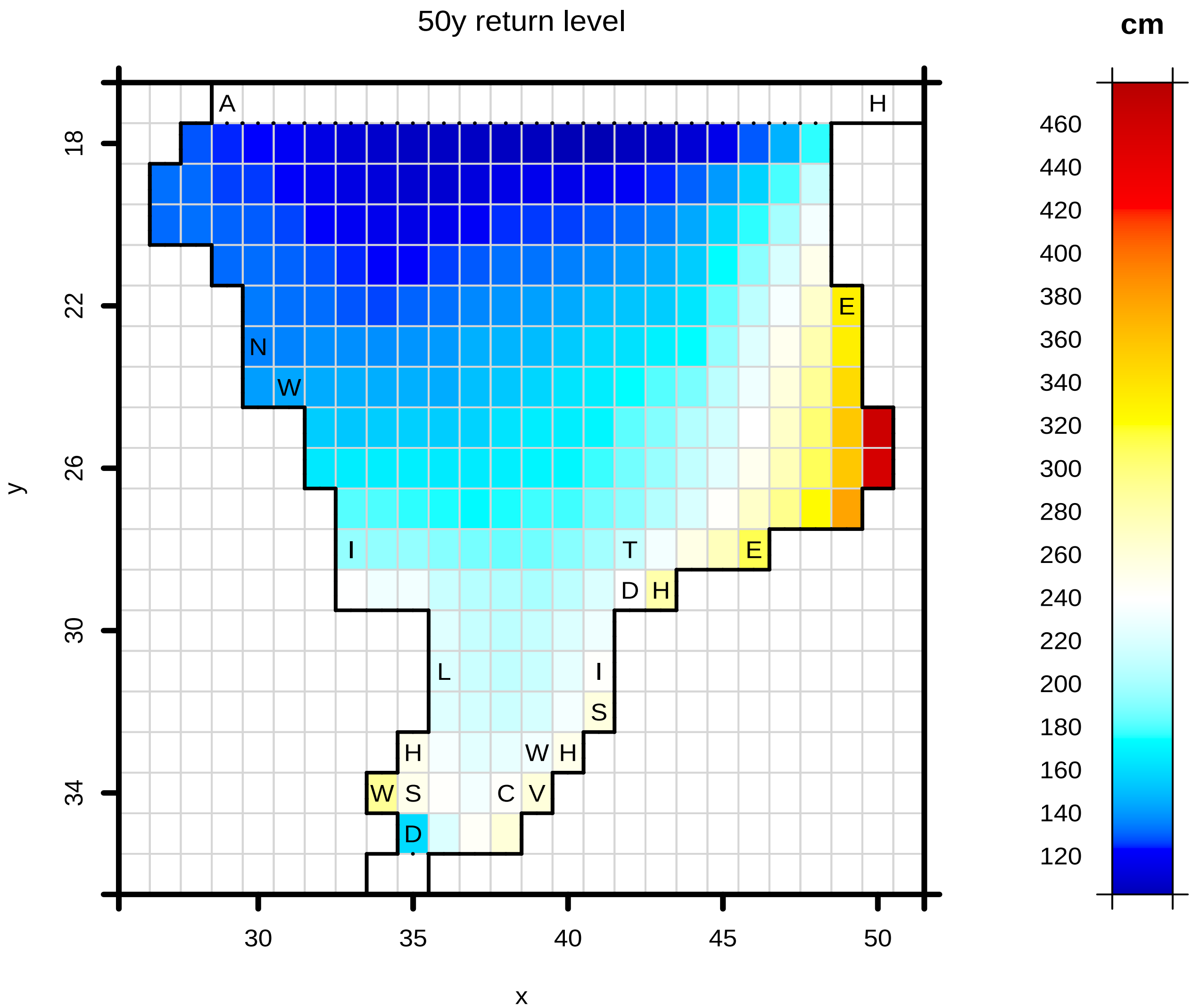
Choice of r

- r-largest model motivated by asymptotic theory: the application relies on r being sufficiently small
- If $r = 1$ the r-largest model reduces to the GEV model for annual maxima

- Can choose r based on parameter stability plots
- Flather et al. (1998) use $r = 7$;
  we suggest using more data, e.g. $r = 20$.

Estimated **50y surge return levels** in the North Sea

Time-constant r-largest model ($r = 20$) fitted to CSX-DNMI output



Temporal trends in extremes

- Nonparametric modelling of extremal trends via local likelihood (Tibshirani and Hastie, 1987)
- Likelihood-based inference from local (nonparametric) regression models
- Developed in the context of extremes by Hall and Tajvidi (2000), Davison and Ramesh (2000).

- Alternative penalized likelihood (smoothing spline) approach - developed in the EV context by: Chavez-Demoulin (1999), Pauli and Coles (2001)

Local likelihood estimators

- Let $f(\mathbf{x}_j; \theta_j)$ denote the density of $\mathbf{X}_j$
- Local constant estimator for θ_j: maximise

$$\sum_{J=1}^{n} K(t_J - t_j; h) \log \left\{ f \left(\mathbf{x}_J; \theta_j \right) \right\},$$

- K is a kernel function with bandwidth h
- Local regression estimator for θ_j: model θ_j as a function of time using a local regression model.

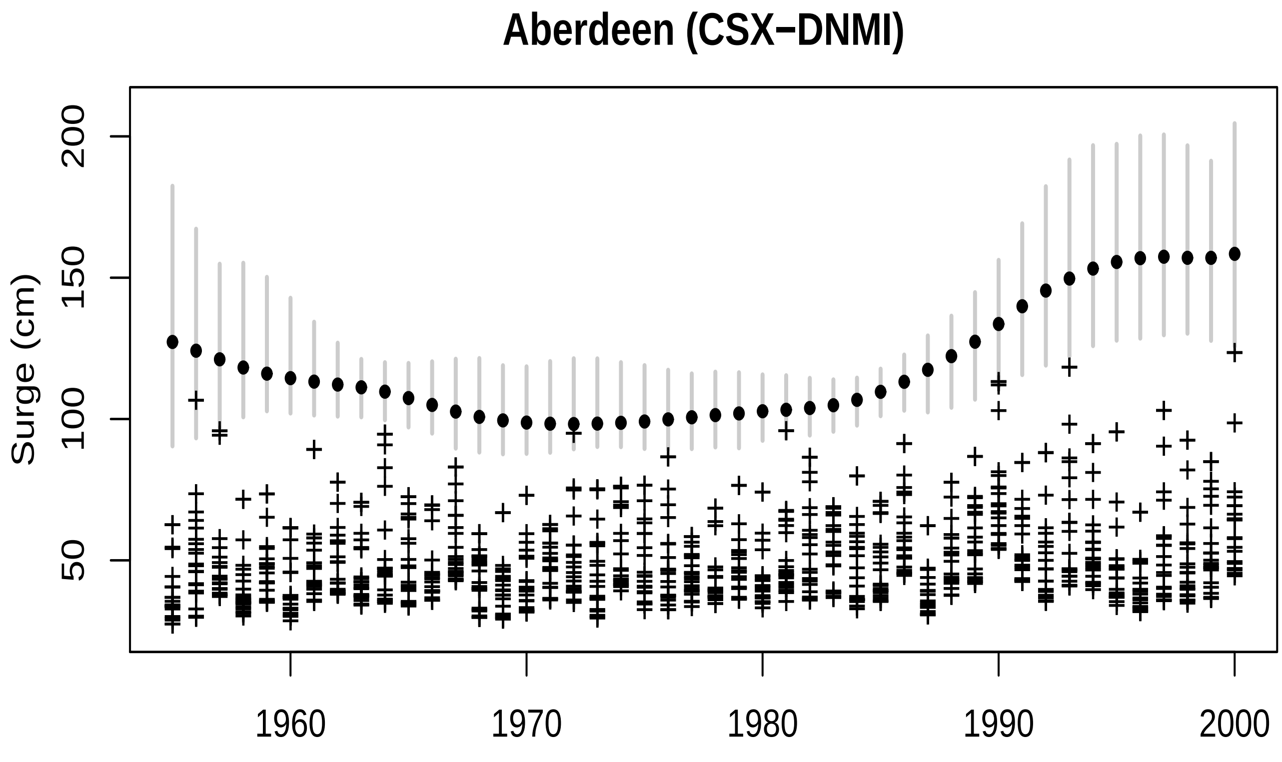
Local likelihood: statistical issues

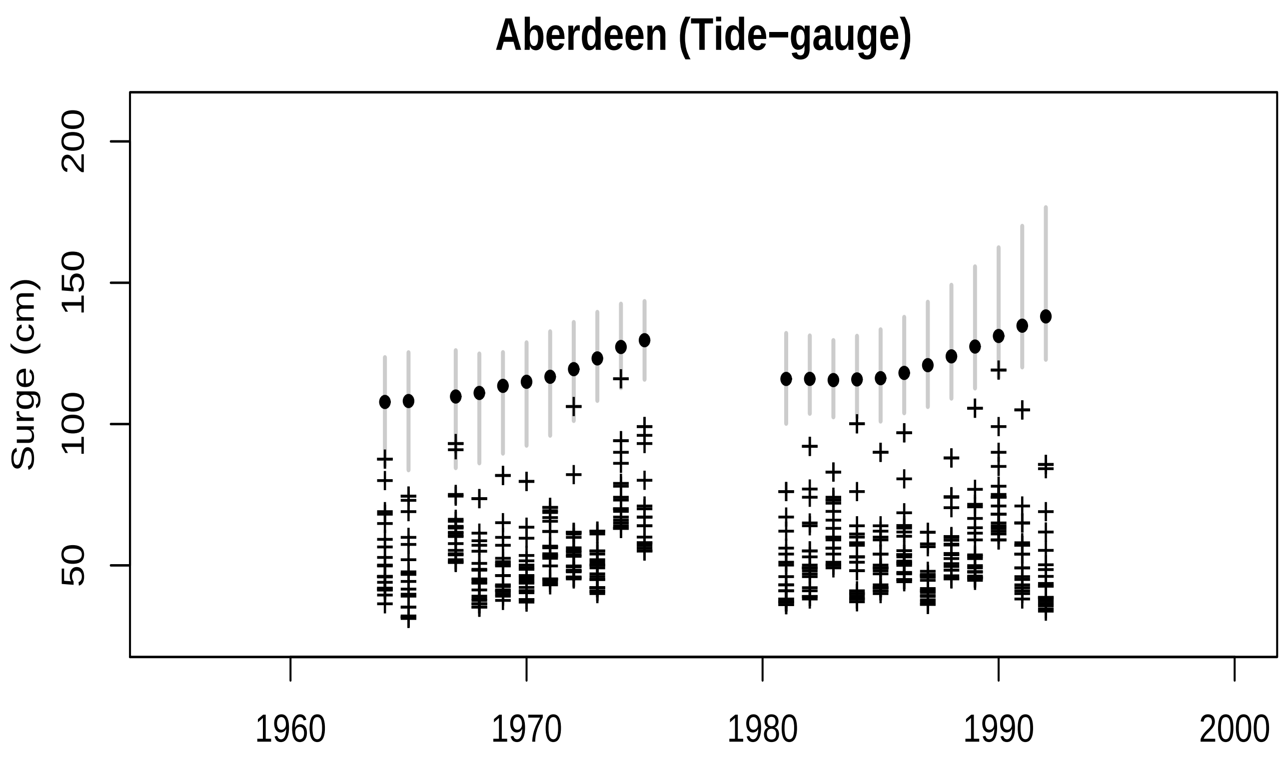
- Selection of the kernel bandwidth h
- Variability bands: use the semiparametric bootstrap (Davison and Ramesh, 2000)
- Model diagnostics: quantile-quantile plots, for each $k = 1, \ldots, r$

20-largest surges per year, & estimated **50y return levels**

Estimates from a local linear r-largest model with $h = 3.5$y and $r = 20$

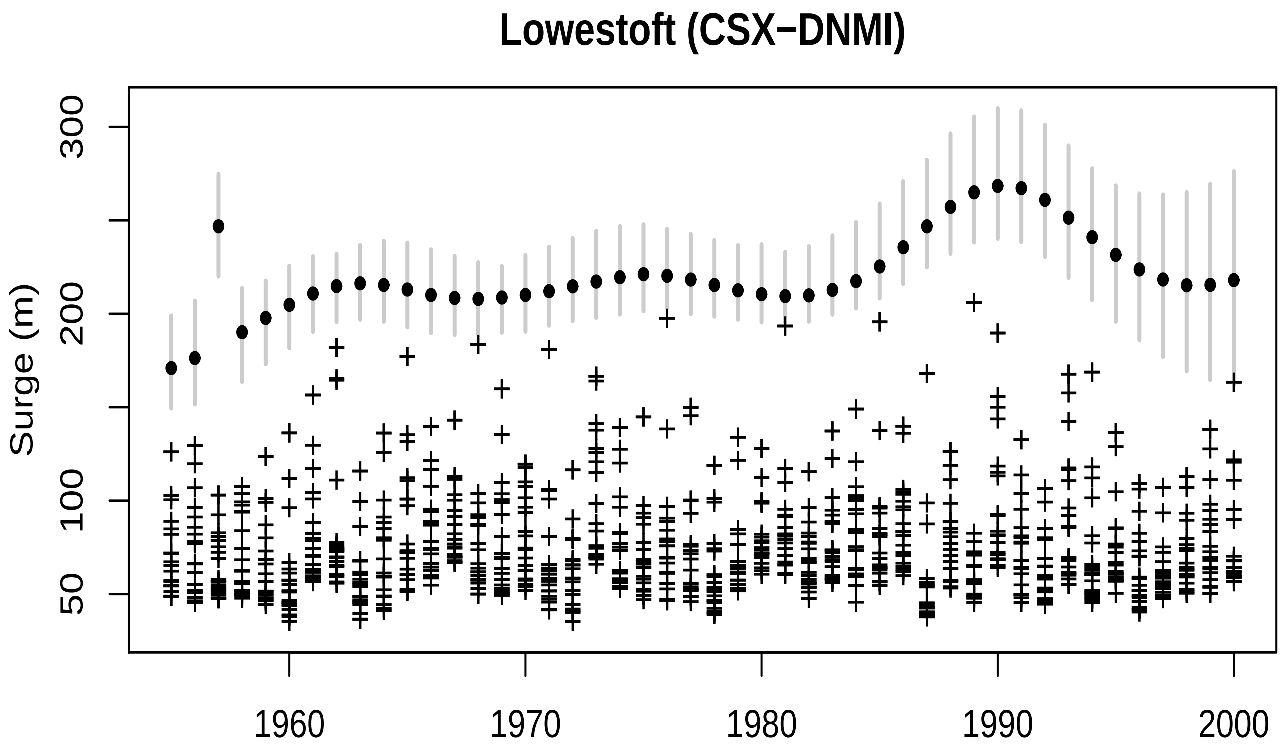
95% pointwise variability bands shown (grey)

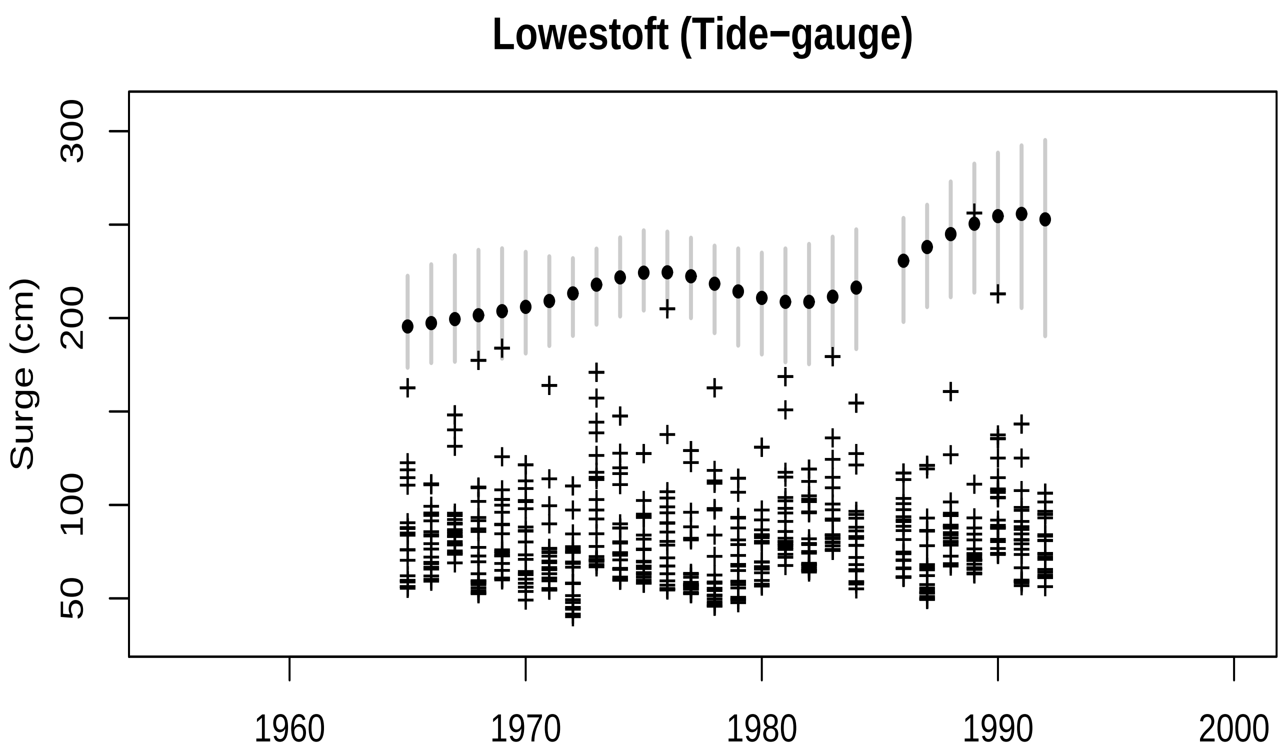




20-largest surges per year, & estimated **50y return levels**

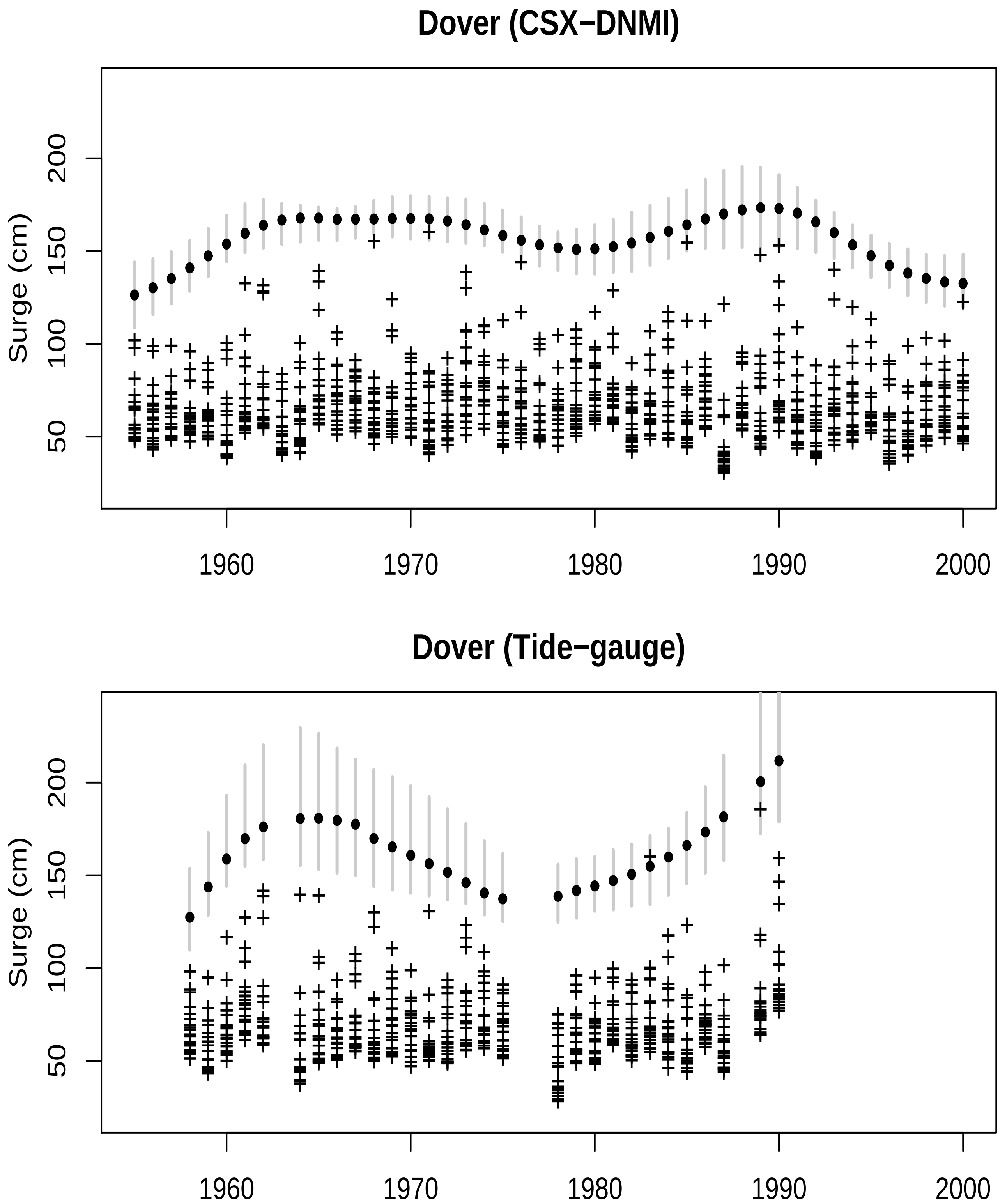
Estimates from a local linear r-largest model with $h = 3.5$y and $r = 20$
95% pointwise variability bands shown (grey)





20-largest surges per year, & estimated **50y return levels**

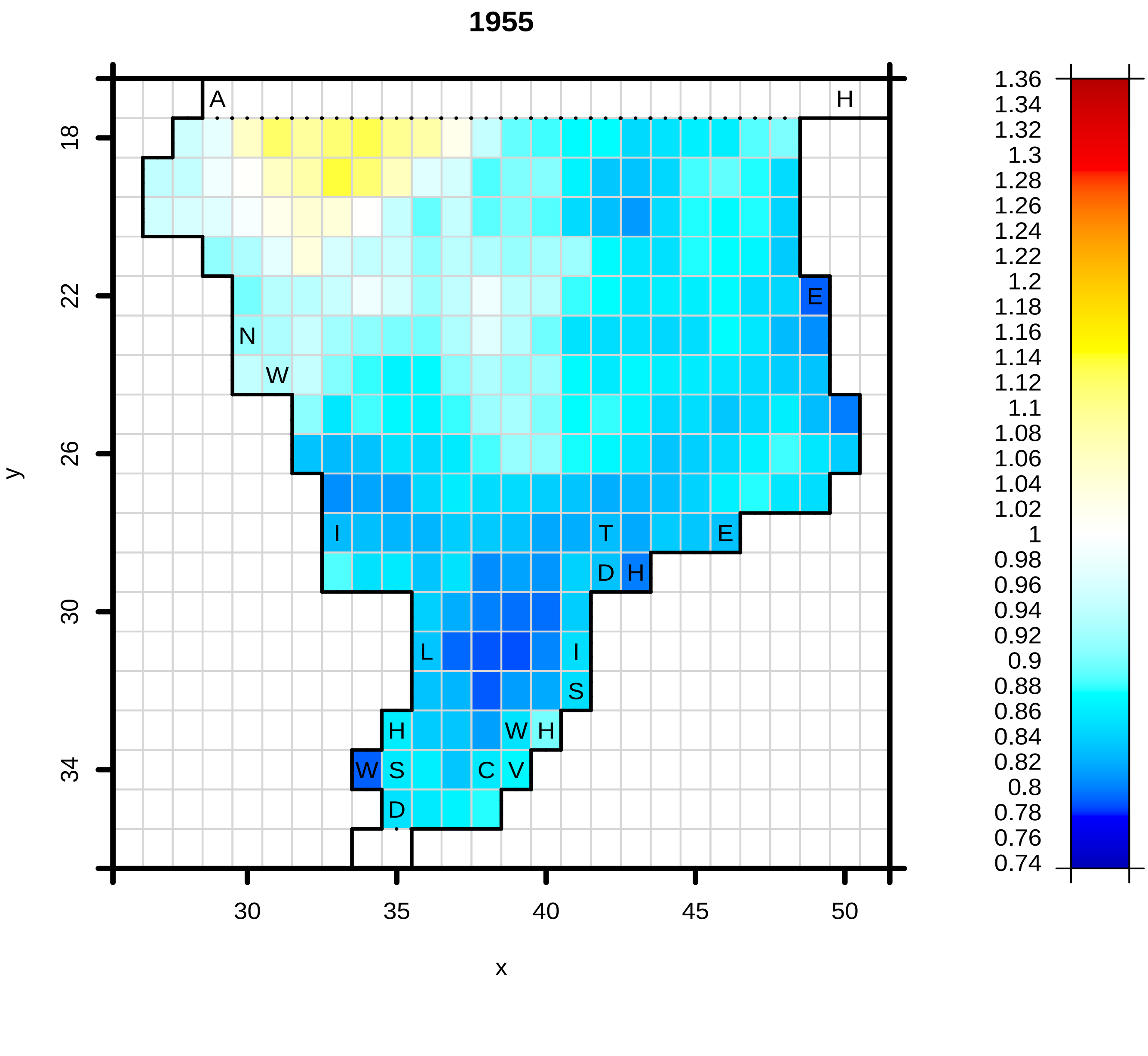
Estimates from a local linear r-largest model with $h = 3.5$y and $r = 20$
95% pointwise variability bands shown (grey)



Spatial distribution of **scaled** 50y return level estimates

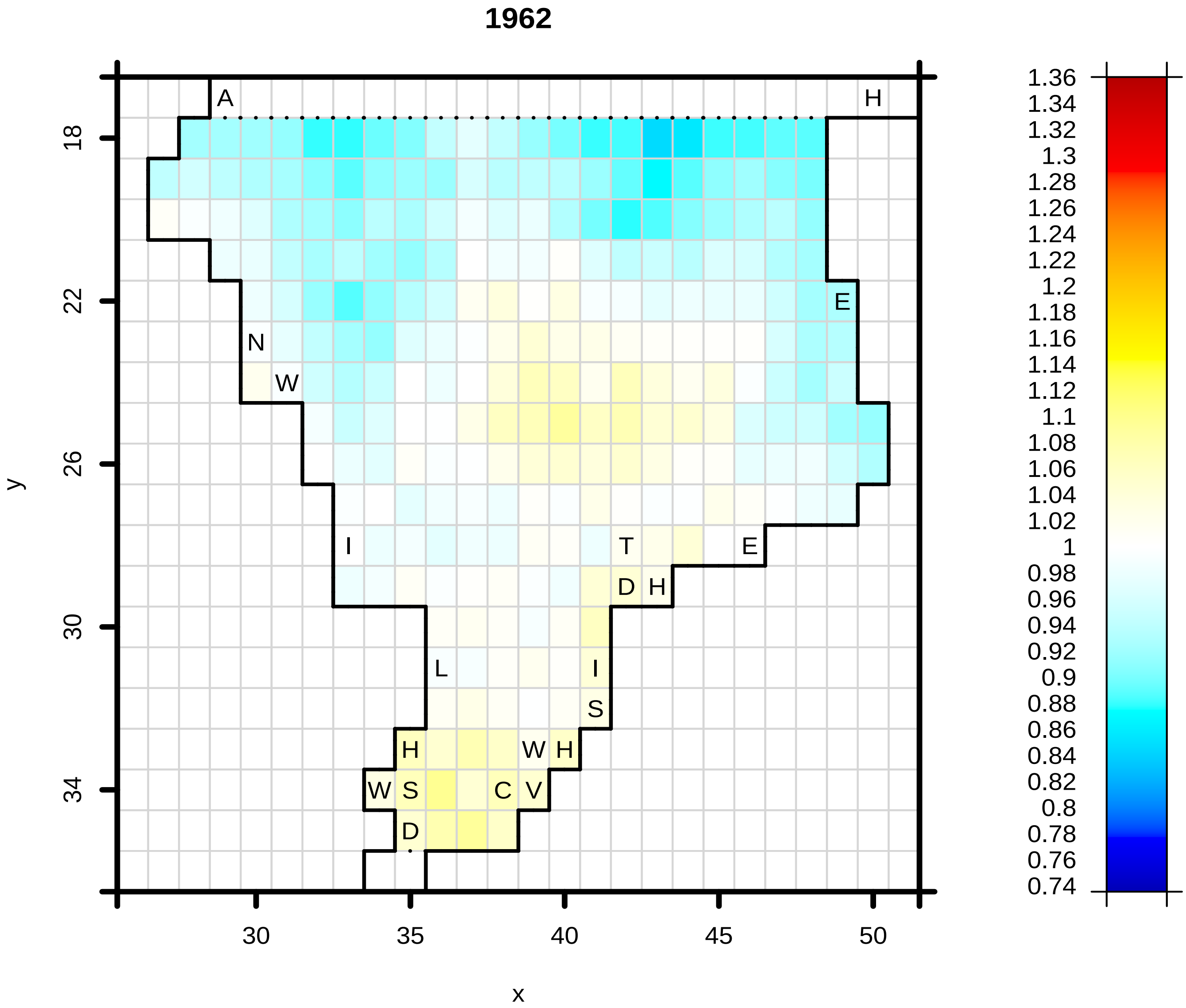
Estimates from a local constant r-largest model with $h = 3.5$y and $r = 20$

Estimates scaled by dividing by the time-constant return level estimate



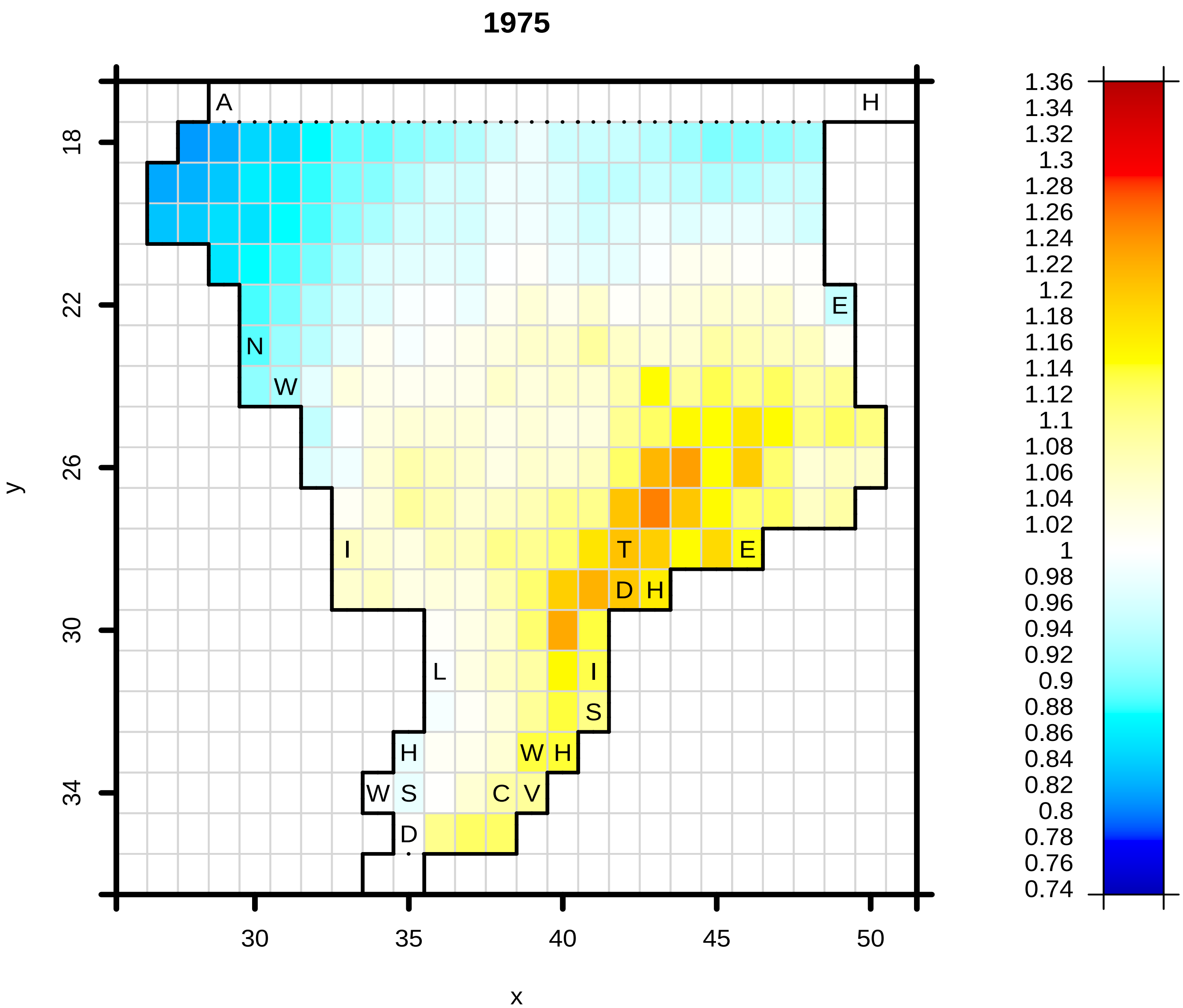
Spatial distribution of **scaled** 50y return level estimates

Estimates from a local constant r-largest model with $h = 3.5$y and $r = 20$
Estimates scaled by dividing by the time-constant return level estimate



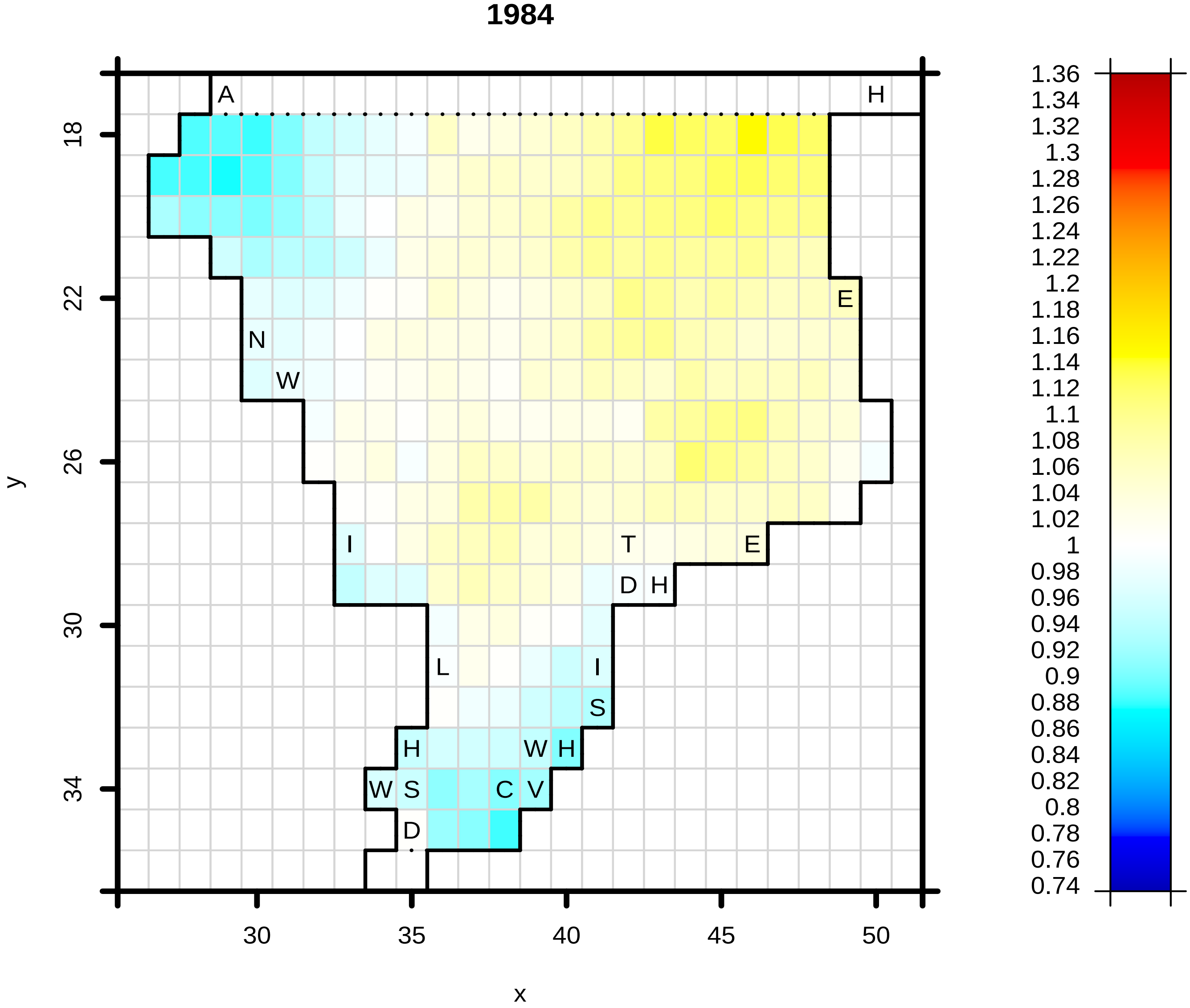
Spatial distribution of **scaled** 50y return level estimates

Estimates from a local constant r-largest model with $h = 3.5$y and $r = 20$
Estimates scaled by dividing by the time-constant return level estimate



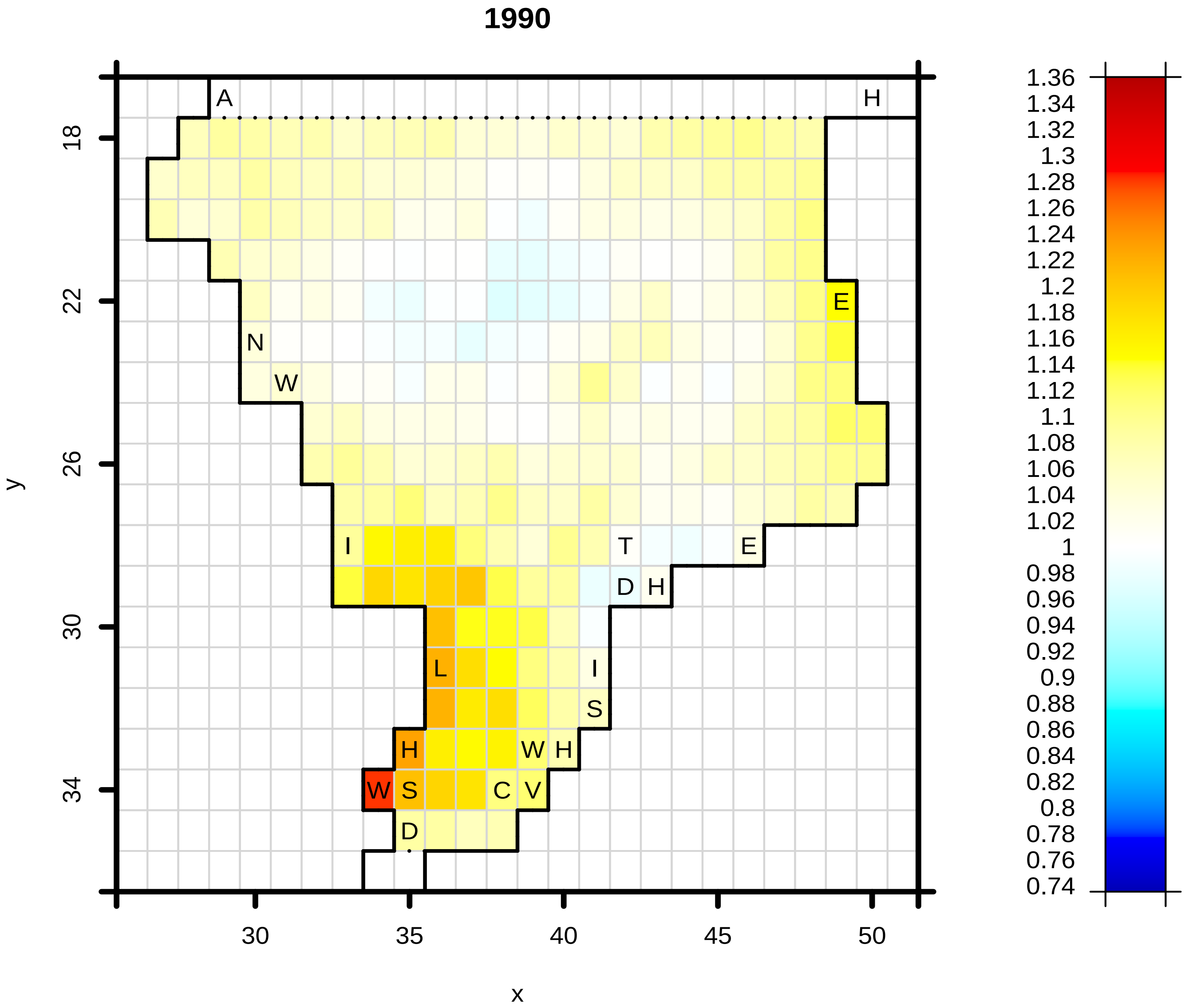
Spatial distribution of **scaled** 50y return level estimates

Estimates from a local constant r-largest model with $h = 3.5$y and $r = 20$
Estimates scaled by dividing by the time-constant return level estimate



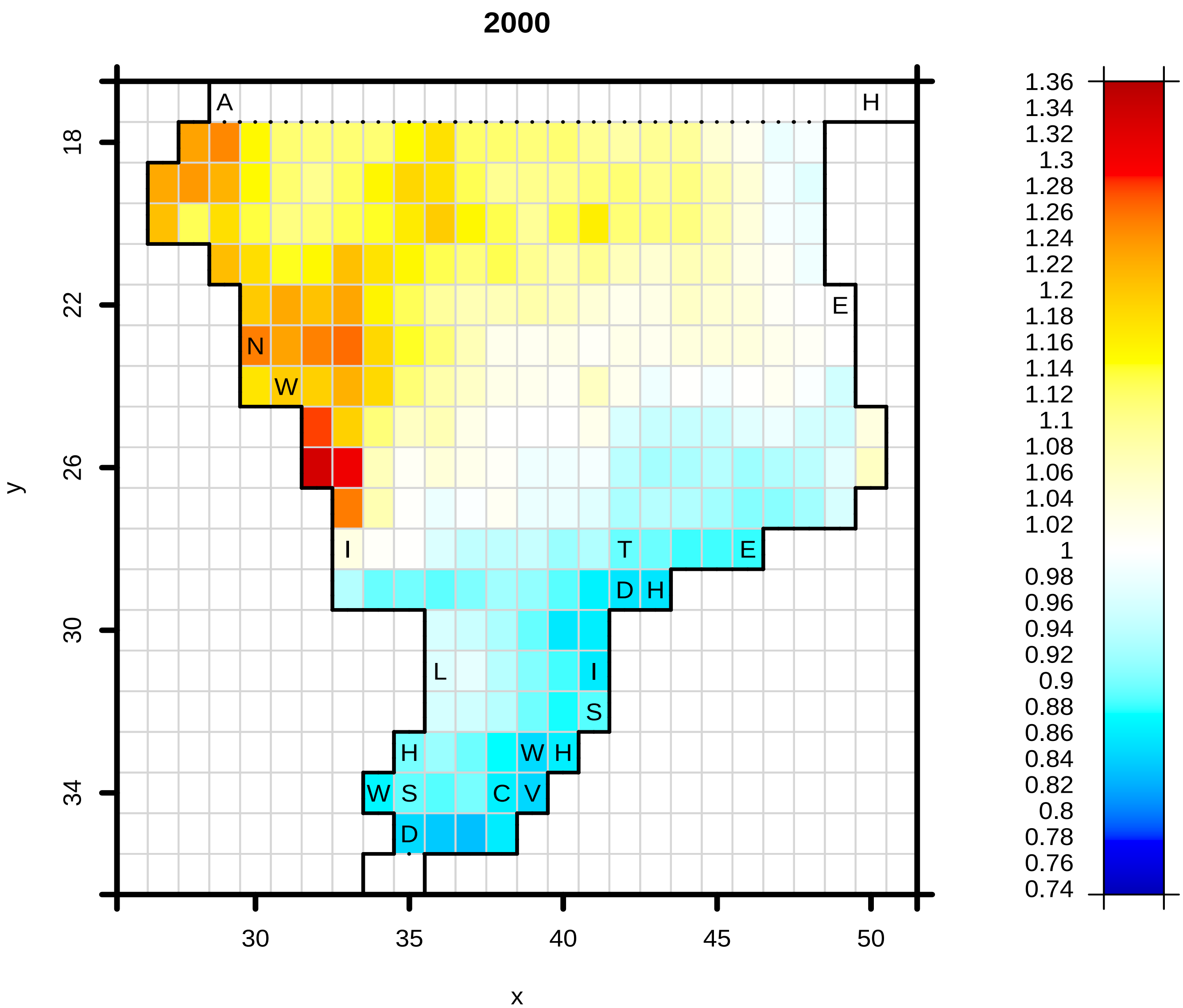
Spatial distribution of **scaled** 50y return level estimates

Estimates from a local constant r-largest model with $h = 3.5$y and $r = 20$
Estimates scaled by dividing by the time-constant return level estimate



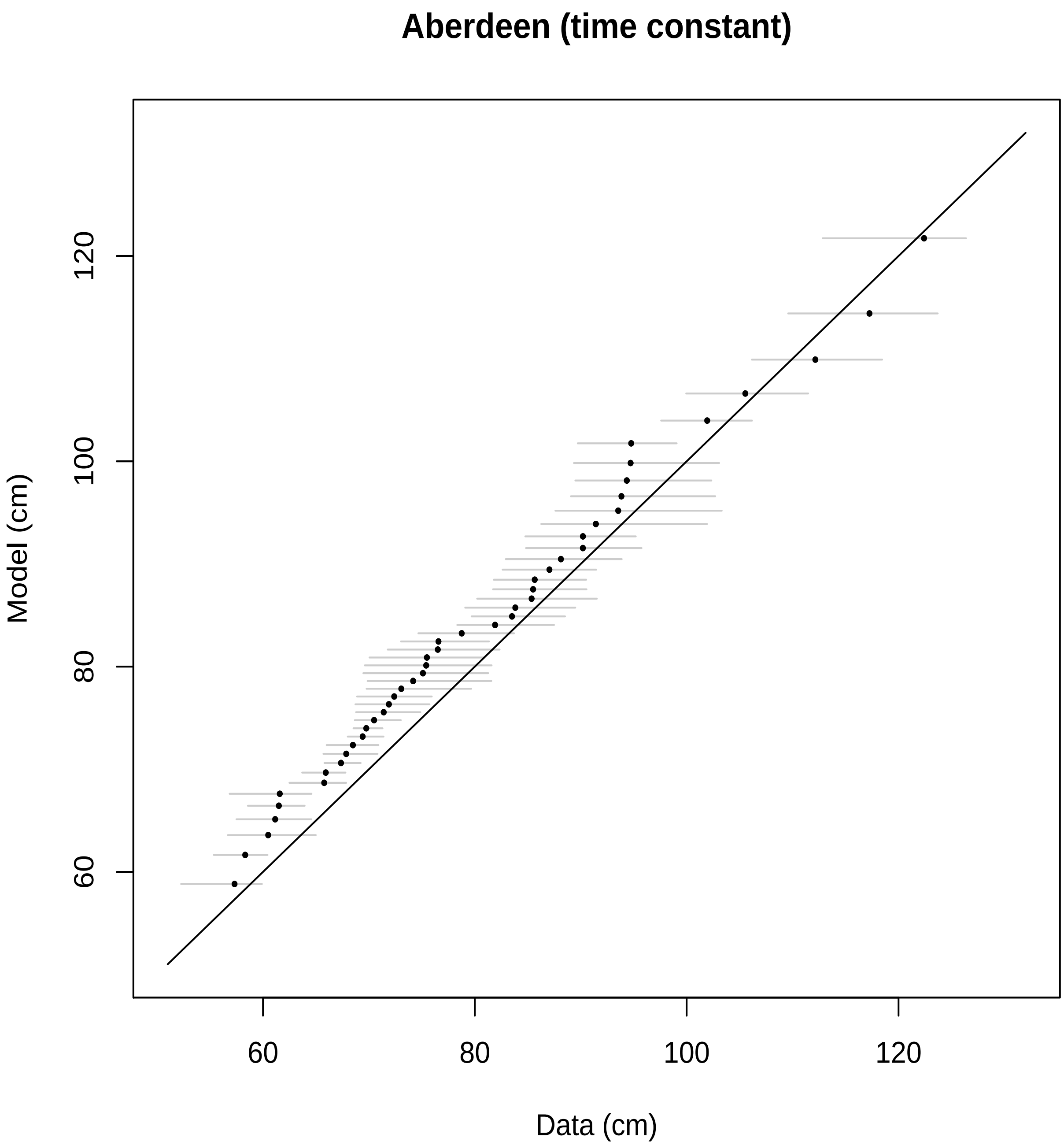
Spatial distribution of **scaled** 50y return level estimates

Estimates from a local constant r-largest model with $h = 3.5$y and $r = 20$
Estimates scaled by dividing by the time-constant return level estimate



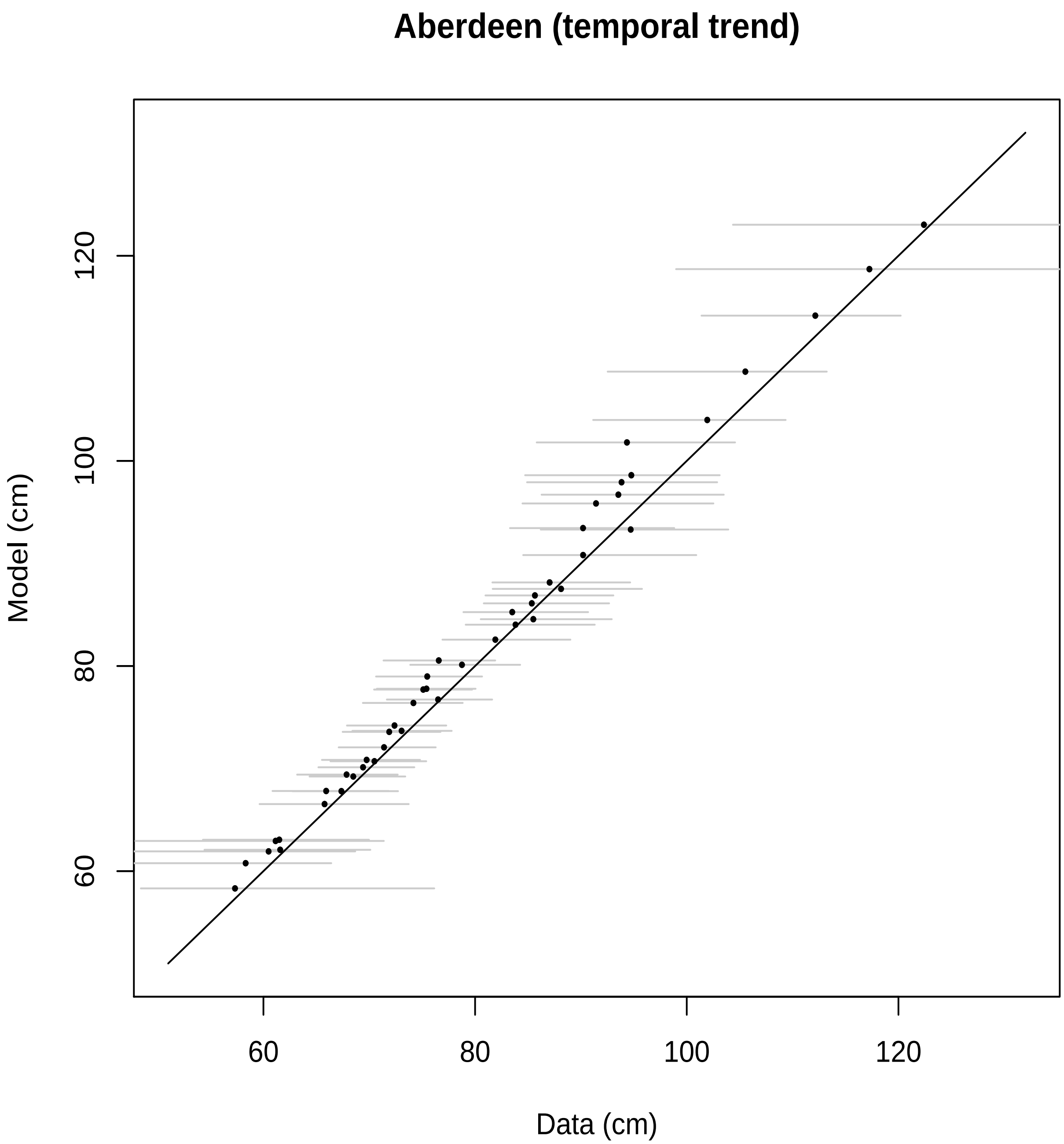
Q-Q plot for r-largest model fitted to CSX-DNMI surges

$k = 1$; $r = 20$; 95% pointwise variability bands shown (grey)



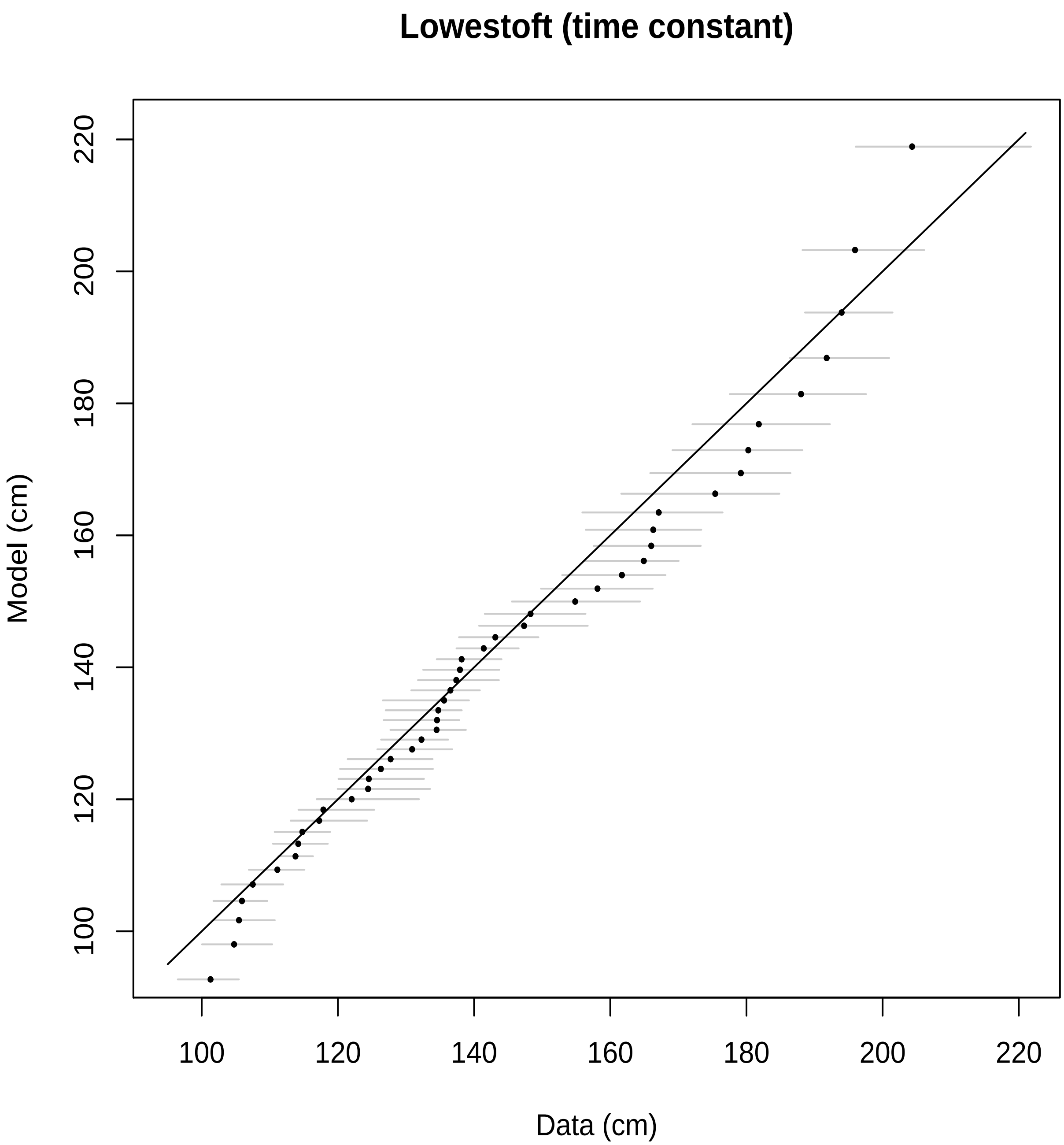
Q-Q plot for r-largest model fitted to CSX-DNMI surges

$h = 3.5$, $k = 1$; $r = 20$; 95% pointwise variability bands shown (grey)



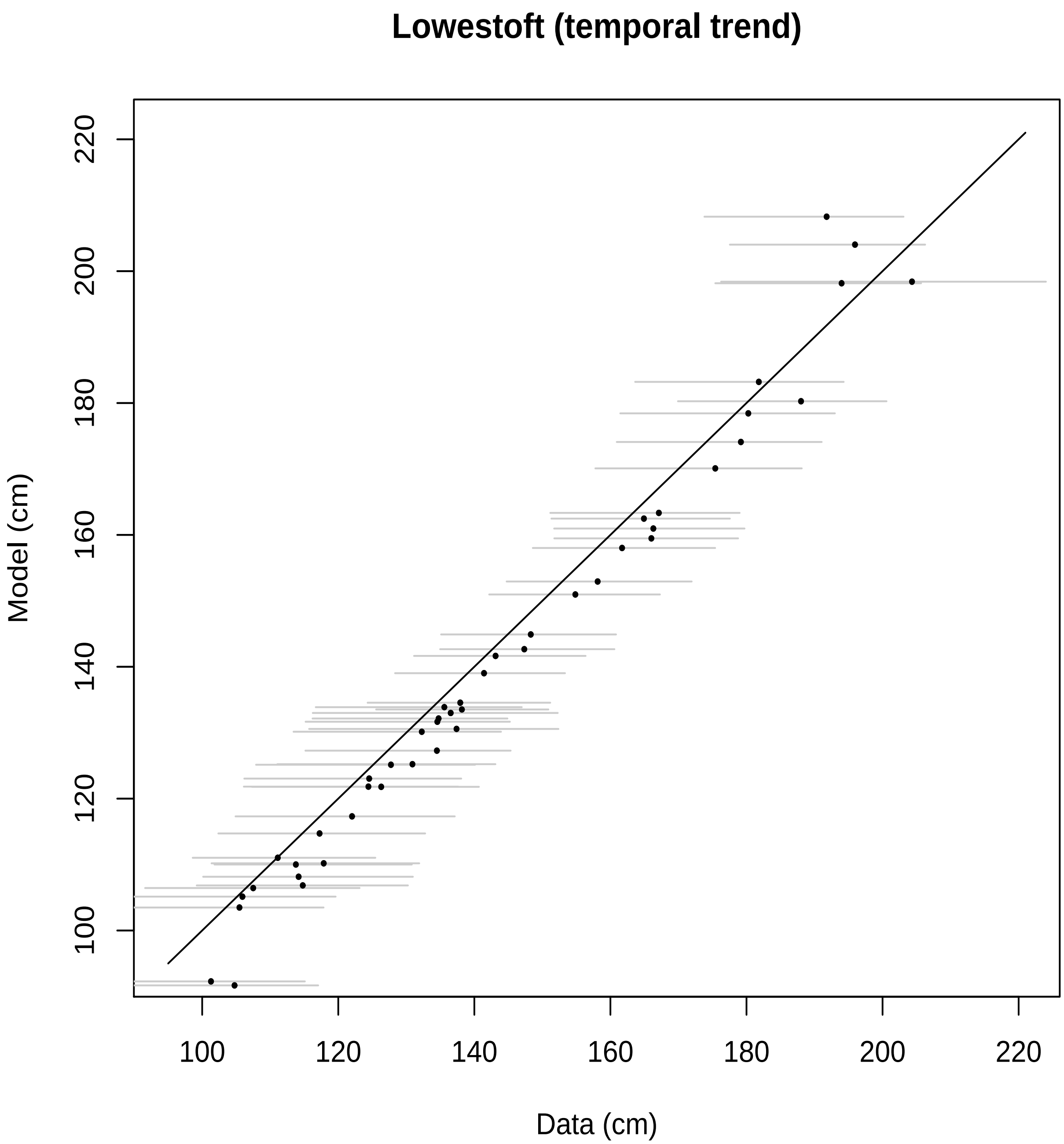
Q-Q plot for r-largest model fitted to CSX-DNMI surges

$k = 1$; $r = 20$; 95% pointwise variability bands shown (grey)



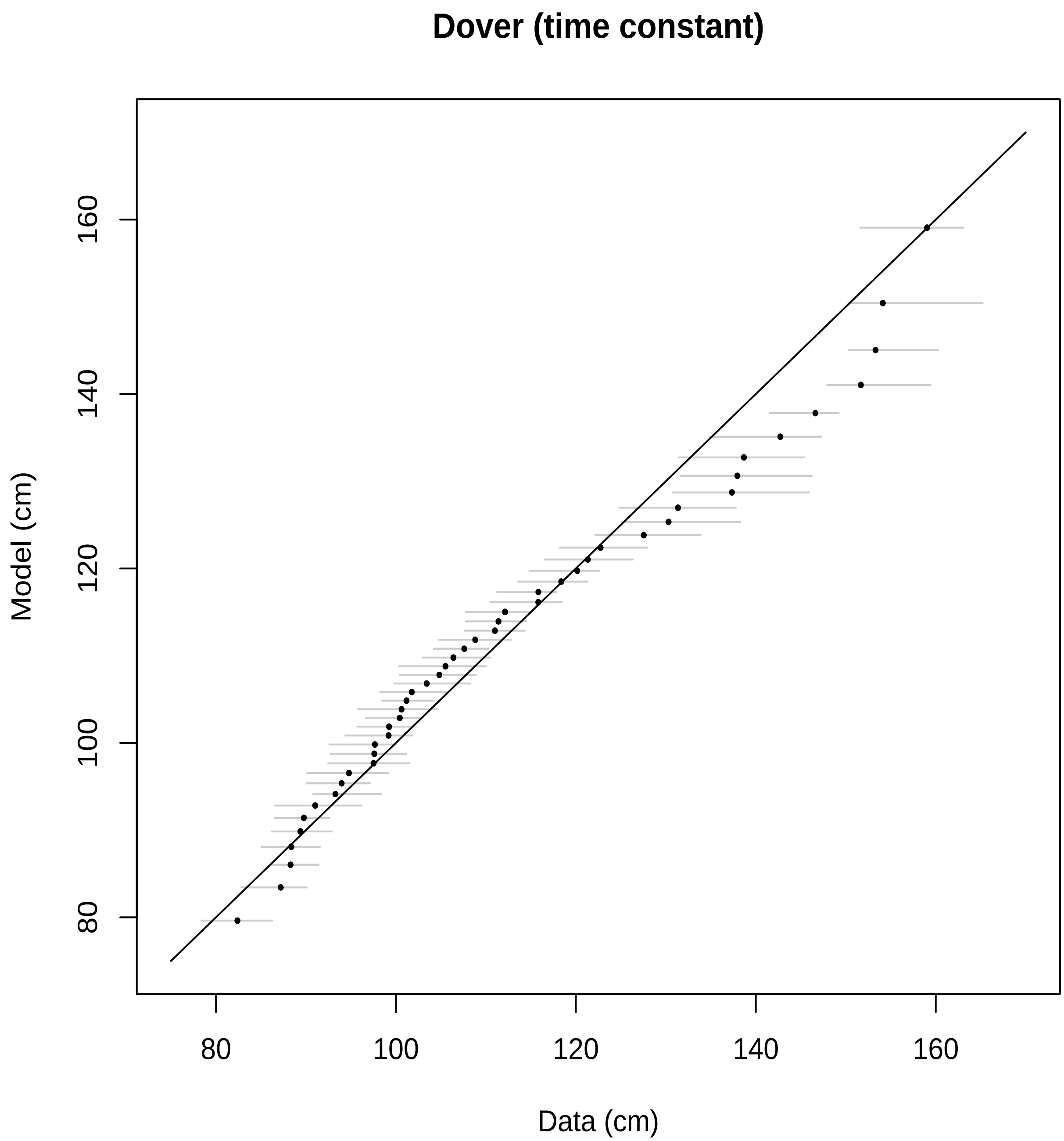
Q-Q plot for r-largest model fitted to CSX-DNMI surges

$h = 3.5$, $k = 1$; $r = 20$; 95% pointwise variability bands shown (grey)



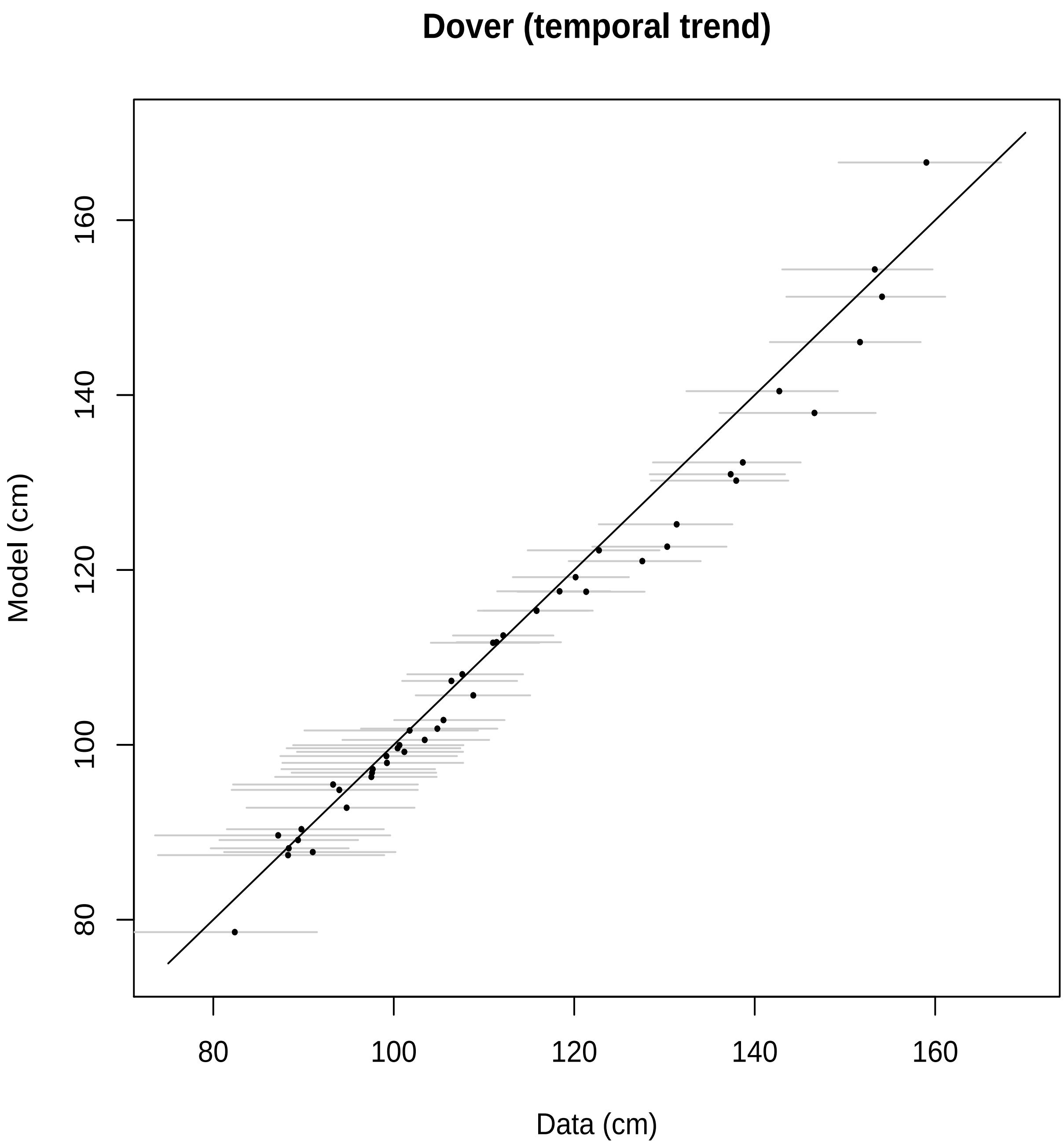
Q-Q plot for r-largest model fitted to CSX-DNMI surges

$k = 1$; $r = 20$; 95% pointwise variability bands shown (grey)



Q-Q plot for r-largest model fitted to CSX-DNMI surges

$h = 3.5$, $k = 1$; $r = 20$; 95% pointwise variability bands shown (grey)



Summary of temporal trends

- Sea areas Forth and Dogger -
  strong overall increase 1955-2000,
  dominated by rapid increase 1980-2000
- Sea area German Bight -
  moderate overall increase 1955-2000,
  but marked decrease 1990-2000
- Sea areas Thames and Humber -
  no clear long-term trend

Discussion

- Trends feed into the literature on:
  - Changing storminess (WASA, 1998)
  - Changing sea levels (Bijl et al., 1999)
- Interpretation of trends:
  - Climate change (natural or anthropogenic)
  - Creeping inhomogeneities in DNMI pressure data
- Interest in using GCM forcing to look at future climate change impacts (Lowe et al., 2001)
- Trends in storm surges - relevance to flood risk

- A novel application of modern statistical methods
- Use spatial information to reduce uncertainty in temporal trend estimates (Dixon and Tawn, 1992)