Regional block-maxima modelling of precipitation extremes in climate model simulations

AND OBSERVED DATA

Martin Hanel, Adri Buishand

Technical University in Liberec (TUL)
Royal Netherlands Meteorological Institute (KNMI), De Bilt

ROBUST 2012
2012/09/09 Němčičky

evropský sociální fond v ČR
EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

OUTLINE

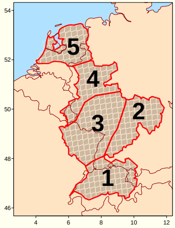
1. Introduction, study area & data
2. Statistical model
3. Assessment of daily and multi-day precipitation extremes in present and future climate for the Czech Republic
4. Model diagnostics
5. Sub-daily precipitation extremes in the Czech Republic
6. Conclusions

INTRODUCTION

- Work iniciated within the ENSEMBLES project - WP5.4 Evaluation of extreme events in observational and RCM data
- Questions:
  - are the precipitation extremes properly represented in the RCM simulations?
  - what are the projected changes in precipitation extremes?
  - how large is the uncertainty associated with the estimation of precipitation extremes and their changes in the climate model data?
- statistical model developed to answer these questions

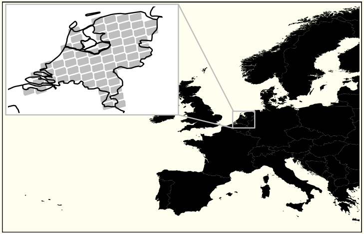
Statistical model applied to assess

- 1-day summer and 5-day winter precipitation extremes in the Rhine basin



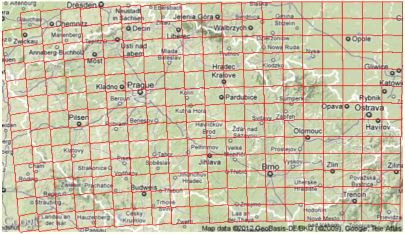
Statistical model applied to assess

- 1-day summer and 5-day winter precipitation extremes in the Rhine basin
- 1-day and 1-hour annual precipitation extremes in the Netherlands



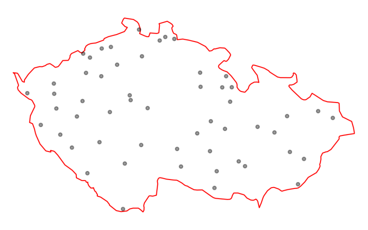
Statistical model applied to assess

- 1-day summer and 5-day winter precipitation extremes in the Rhine basin
- 1-day and 1-hour annual precipitation extremes in the Netherlands
- 1-, 3-, 5-, 7-, 10-, 15-, 20- and 30-day precipitation extremes for all seasons in the Czech Republic



Statistical model applied to assess

- 1-day summer and 5-day winter precipitation extremes in the Rhine basin
- 1-day and 1-hour annual precipitation extremes in the Netherlands
- 1-, 3-, 5-, 7-, 10-, 15-, 20- and 30-day precipitation extremes for all seasons in the Czech Republic
- *Observed sub-daily precipitation extremes (summer half year) in the Czech Republic*



DATA

RCM DATA

model	acronym	source	period
	— ECHAM5 driven —		
RACMO	**RACMO_EH5**	KNMI	1950-2100
REMO	**REMO_EH5**	MPI	1951-2100
RCA	**RCA_EH5**	SMHI	1951-2100
RegCM	**RegCM_EH5**	ICTP	1951-2100
HIRHAM	**HIR_EH5**	DMI	1951-2100
	— HadCM3Q0, HadCM3Q3, HadCM3Q16 driven —		
HadRM	**HadRM_Q0**	Hadley Centre	1951-2099
CLM	**CLM_Q0**	ETHZ	1951-2099
HadRM	**HadRM_Q3**	Hadley Centre	1951-2099
RCA	**RCA_Q3**	SMHI	1951-2099
HadRM	**HadRM_Q16**	Hadley Centre	1951-2099
RCA	**RCA_Q16**	C4I	1951-2099
	— ARPEGE driven —		
HIRHAM	**HIR_ARP**	DMI	1951-2100
CNRM-RM	**CNRM_ARP**	CNRM	1951-2100
ALADIN-CLIMATE/CZ	**ALA_ARP**	CHMI	1961-2100
	— BCM driven —		
RCA	**RCA_BCM**	SMHI	1961-2100

RCM simulations

- $\approx$ 25 km $\times$ 25 km
- SRES A1B
- transient simulations

Definition of maxima

- block maxima, i.e. the largest precipitation amount in a year/season
- various aggregations of data (e.g. 1-day, 5-day ...)

OBSERVATIONAL DATA ($\approx$ 25 km x 25 km)

acronym	source	period
CHMI_OBS	CHMI	1950-2007
E_OBS	KNMI	1950-present

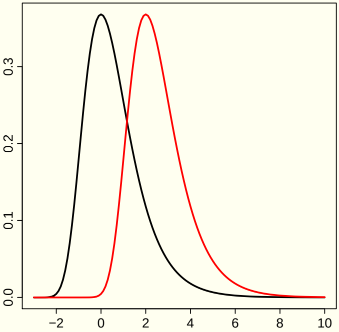
GEV DISTRIBUTION FUNCTION

$$F(x) = \exp\left\{-\left[1+\kappa\left(\frac{x-\xi}{\alpha}\right)\right]^{-\frac{1}{\kappa}}\right\}, \qquad \kappa \neq 0$$

$$F(x) = \exp\left\{-\exp\left[-\left(\frac{x-\xi}{\alpha}\right)\right]\right\}, \qquad \kappa = 0$$

- ξ ... location parameter

PROBABILITY DENSITY FUNCTION



The GEV parameters can

- vary over the region (spatial heterogeneity)
- vary with time (climate change)

GEV DISTRIBUTION FUNCTION

$$F(x) = \exp\left\{-\left[1+\kappa\left(\frac{x-\xi}{\alpha}\right)\right]^{-\frac{1}{\kappa}}\right\}, \qquad \kappa \neq 0$$

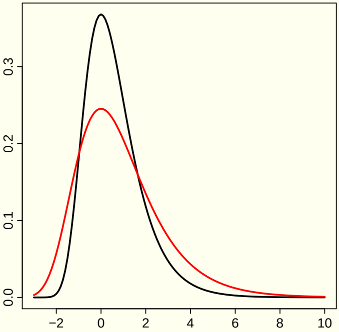
$$F(x) = \exp\left\{-\exp\left[-\left(\frac{x-\xi}{\alpha}\right)\right]\right\}, \qquad \kappa = 0$$

- ξ ... location parameter
- α ... scale parameter

The GEV parameters can

- vary over the region (spatial heterogeneity)
- vary with time (climate change)

PROBABILITY DENSITY FUNCTION



GEV DISTRIBUTION FUNCTION

$$F(x) = \exp\left\{-\left[1+\kappa\left(\frac{x-\xi}{\alpha}\right)\right]^{-\frac{1}{\kappa}}\right\}, \qquad \kappa \neq 0$$

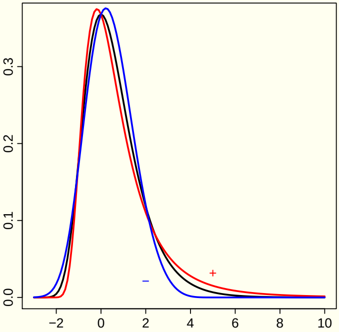
$$F(x) = \exp\left\{-\exp\left[-\left(\frac{x-\xi}{\alpha}\right)\right]\right\}, \qquad \kappa = 0$$

- ξ ... location parameter
- α ... scale parameter
- κ ... shape parameter

The GEV parameters can

- vary over the region (spatial heterogeneity)
- vary with time (climate change)

PROBABILITY DENSITY FUNCTION



STATISTICAL MODEL

- Index flood method assumes that precipitation maxima over a region are identically distributed after scaling with a site-specific factor

SPATIAL HETEROGENEITY

- ξ varies over the region
- κ, $\gamma = \frac{\alpha}{\xi}$ are constant over the region
  γ is the dispersion coefficient analogous to the coefficient of variation
- T-year quantile at any site s can be represented as

$$Q_T(s) = \mu(s) \cdot q_T, \qquad q_T = 1 - \gamma \cdot \frac{1 - \left[-\log\left(1 - \frac{1}{T}\right)\right]^{-\kappa}}{\kappa}$$

  where q_T is a common dimensionless quantile function (growth curve) and $\mu(s)$ is a site-specific scaling factor ("index flood")

- it is convenient to set $\mu(s) = \xi(s)$

STATISTICAL MODEL

NON-STATIONARITY

- location parameter varies over the region, but with common trend
$$\xi(s,t) = \xi_0(s) \cdot \exp\left[\xi_1 \cdot I(t)\right]$$
where $I(t)$ is the time indicator
- dispersion coefficient and shape parameter are constant over the region, but vary with time
$$\gamma(t) = \exp\left[\gamma_0 + \gamma_1 \cdot I(t)\right]$$
$$\kappa(t) = \kappa_0 + \kappa_1 \cdot I(t)$$

UNCERTAINTY

- assessed by a bootstrap procedure (1000 samples for each RCM simulation, season and duration)

RELATIVE CHANGES

- T-year quantile in time t and location s:

$$Q_T(s,t) = \xi(s,t) \cdot q_T(t)$$

- Relative change between t_2 and t_1 ($t_2 > t_1$) is:

$$\frac{Q_T(s,t_2)}{Q_T(s,t_1)} = \frac{\xi(s,t_2)}{\xi(s,t_1)} \cdot \frac{q_T(t_2)}{q_T(t_1)}$$

$$\xi(s,t) = \xi_0(s) \cdot \exp\left[\xi_1 \cdot I(t)\right] \Rightarrow \frac{\xi(s,t_2)}{\xi(s,t_1)} = \exp\left\{\xi_1 \cdot [I(t_2) - I(t_1)]\right\}$$

$\Rightarrow$ the relative change in quantiles can be written as

$$\frac{Q_T(s,t_2)}{Q_T(s,t_1)} = \exp\left\{\xi_1 \cdot [I(t_2) - I(t_1)]\right\} \cdot \frac{q_T(t_2)}{q_T(t_1)}$$

which does not depend on s.

TIME INDICATOR

Various choices for $I(t)$ are possible:

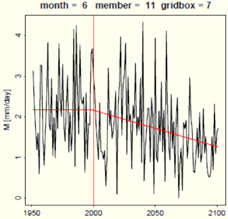
- year t

  e.g., $I(t) = t$, but more complicated functions are needed
- temperature/ temperature anomaly

We use:

- seasonal global temperature anomaly of the driving GCM

CHANGE IN PRECIPITATION

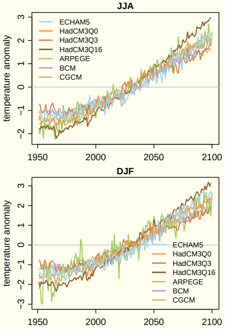


TIME INDICATOR

Various choices for $I(t)$ are possible:

- year t
  e.g., $I(t) = t$, but more complicated functions are needed
- temperature/ temperature anomaly

SEASONAL GLOBAL TEMPERATURE ANOMALY



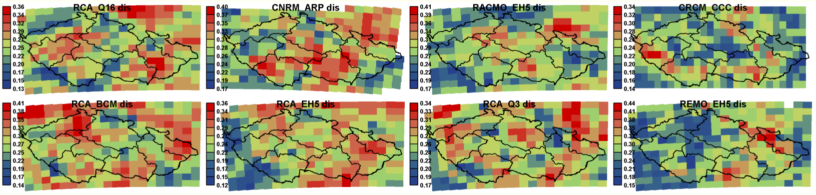
We use:

- seasonal global temperature anomaly of the driving GCM

IDENTIFICATION OF HOMOGENEOUS REGIONS

For each RCM simulation and season we assessed

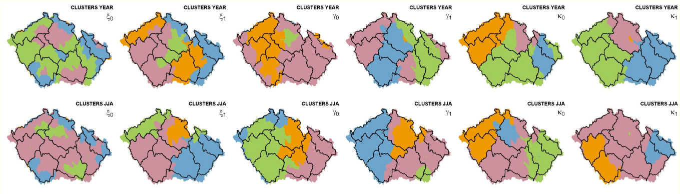
- the grid box estimates of the GEV parameters for the period 1961-1990 and 2070-2099
- their changes between these two periods
- focus mainly on the dispersion coefficient and changes in location parameter and dispersion coefficient
- formation of homogeneous regions common for all RCM simulations and seasons is challenging



IDENTIFICATION OF HOMOGENEOUS REGIONS

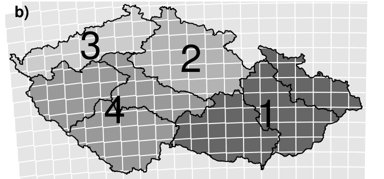
For each RCM simulation and season we assessed

- the grid box estimates of the GEV parameters for the period 1961-1990 and 2070-2099
- their changes between these two periods
- focus mainly on the dispersion coefficient and changes in location parameter and dispersion coefficient
- formation of homogeneous regions common for all RCM simulations and seasons is challenging



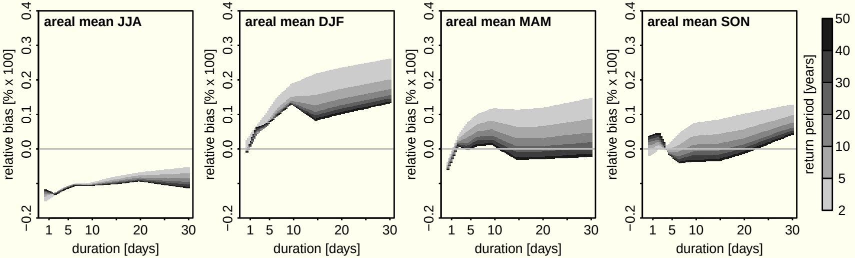
IDENTIFICATION OF HOMOGENEOUS REGIONS

- obviously it is not possible to identify strictly homogeneous regions common for all RCM simulations and seasons
  - however, regional frequency analysis is more accurate than the at-site analysis even in not strictly homogeneous regions
  - allowing more heterogeneity of precipitation maxima in the statistical model improves goodness-of-fit, however, the estimated changes are not much different when compared to the standard model
- it turned out that the regions could be acceptably based on the areas of the eight river basin districts in the Czech Republic
- different groupings of the river basin districts were examined with respect to the lack-of-fit of the statistical model

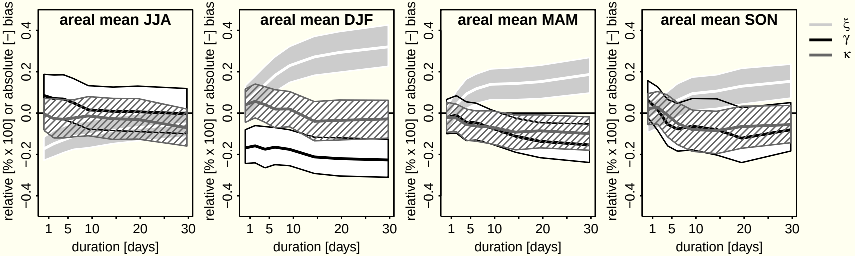


EVALUATION FOR CONTROL CLIMATE

Relative bias in quantiles

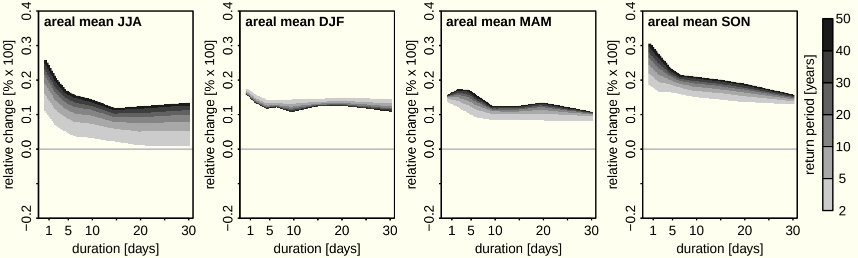


Relative (ξ, γ) and absolute (κ) bias in parameters

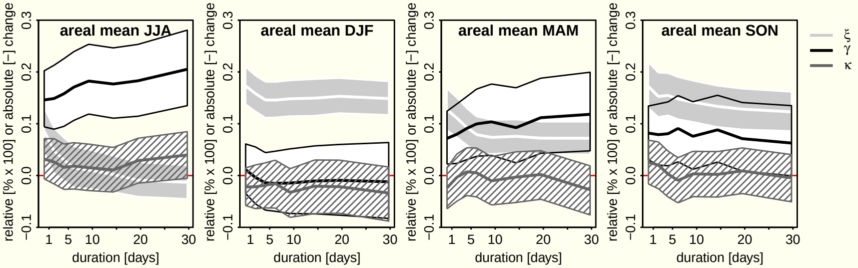


CHANGES BETWEEN 1961-1990 AND 2070-2099

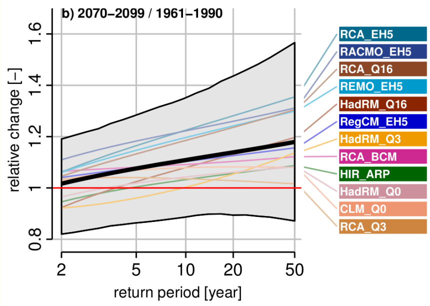
Relative changes in quantiles

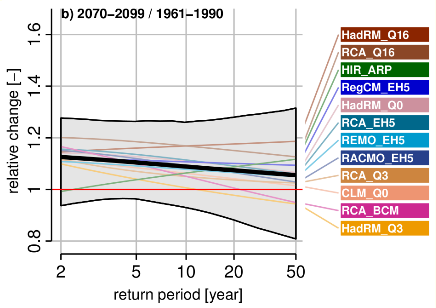


Relative (ξ, γ) and absolute (κ) changes in parameters



CHANGES IN 1-DAY SUMMER AND 5-DAY WINTER PRECIP EXTREMES, RHINE BASIN





MODEL DIAGNOSTICS

For each grid box we calculate the residuals:

- we transform the seasonal maxima X_t to:

  $$\widetilde{X}_t = \frac{1}{\hat{\xi}(t)} \cdot \log\left[1 + \frac{\hat{\xi}(t)}{\hat{\gamma}(t)} \cdot \left(\frac{X_t}{\hat{\mu}(t)} - 1\right)\right],$$

  which should have a standard Gumbel distribution;

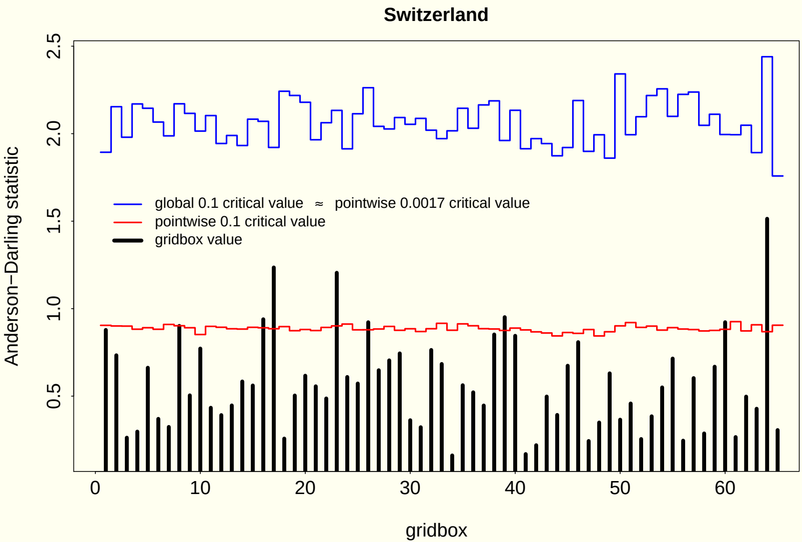
  $$\Pr\{\widetilde{X}_t \leq x\} = \exp\left[-\exp(-x)\right]$$

  if the model is true.
- these residuals can be inspected visually and/or e.g. by the Anderson-Darling statistics:

  $$A^2 = N \int_{-\infty}^{\infty} \frac{[F_N(x) - F(x)]^2}{F(x)[1 - F(x)]} dF(x),$$

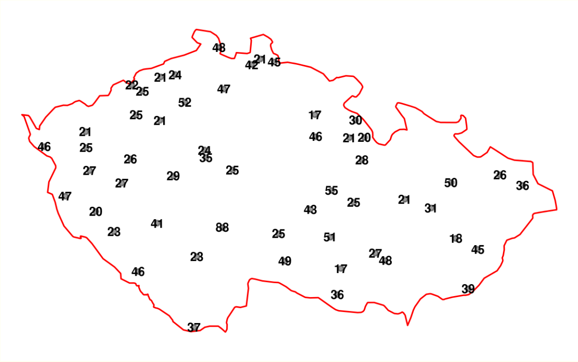
  where $F_N(x)$ is the empirical distribution function of $\widetilde{X}_t$ and $F(x)$ standard Gumbel distribution function.
- critical values can be derived using simulation
- bootstrap procedure based on resampling of standard Gumbel residuals

MODEL DIAGNOSTICS



DATA

- 54 stations, combination of pluviographic (10-min) and automatic (30-min) rain gages, max 7 years overlap
- from 1921-1991 to 2011, May-September
- 33 years on average, in total 1807 station-years

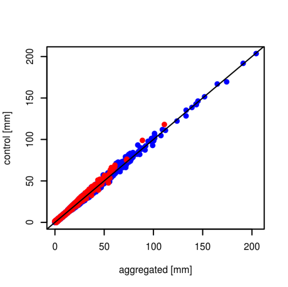


QUALITY CHECK

for both data sources available control daily total - checked against aggregated amount from pluviograph/automatic station

- records for days with difference > 1.5 mm (for totals < 15 mm) or > 10 % marked unreliable
- only years with < 10 % of unreliable data considered

- sources combined
- automatic rain gages on average 5 % underestimation, pluviograph 1 % underestimation
- no correction considered (yet?)



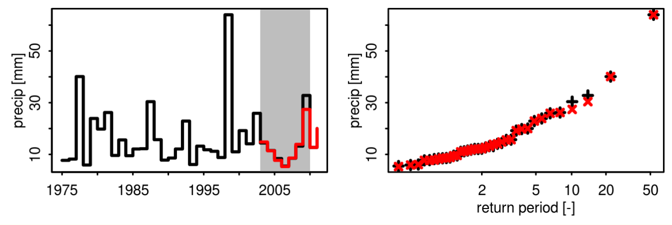
SUB-DAILY PRECIPITATION EXTREMES FOR THE CZECH REPUBLIC

QUALITY CHECK

for both data sources available control daily total - checked against aggregated amount from pluviograph/automatic station

- records for days with difference > 1.5 mm (for totals < 15 mm) or > 10 % marked unreliable
- only years with < 10 % of unreliable data considered

- sources combined
- automatic rain gages on average 5 % underestimation, pluviograph 1 % underestimation
- no correction considered (yet?)

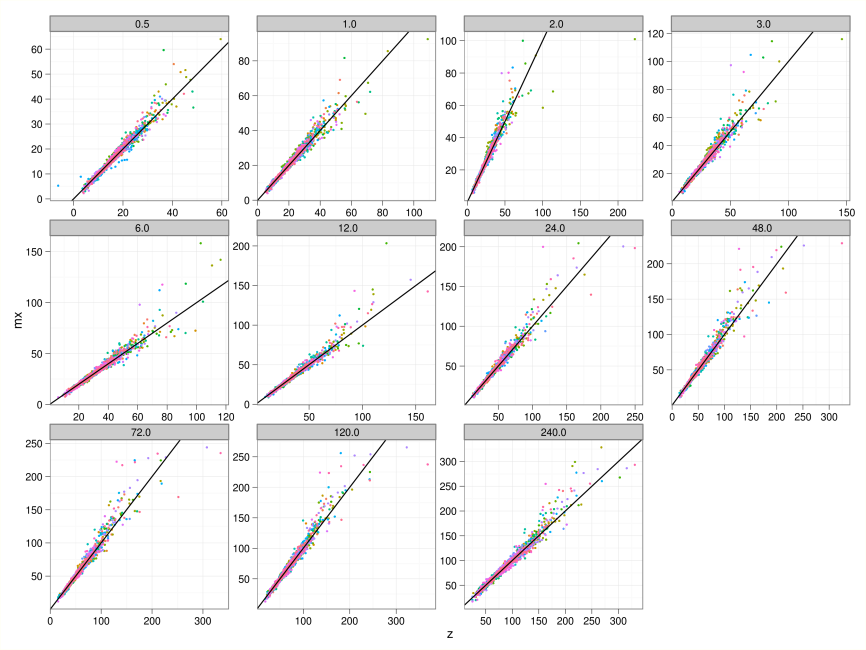


SUB-DAILY PRECIPITATION EXTREMES FOR THE CZECH REPUBLIC

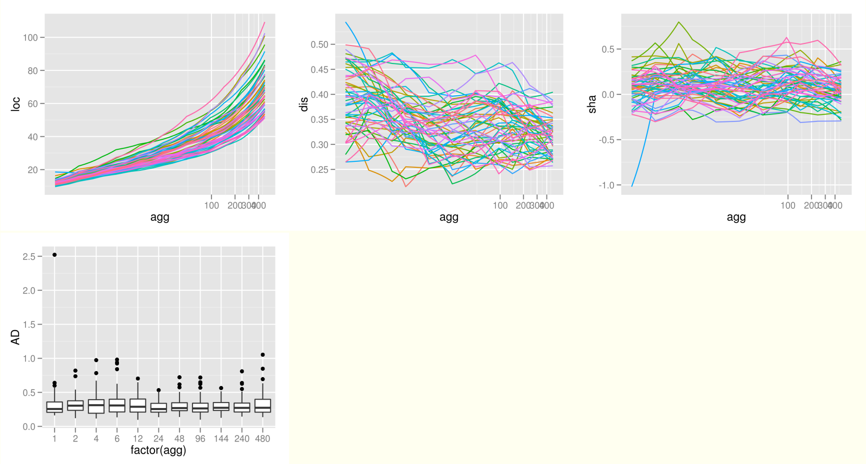
- assessment of 0.5, 1, 2, 3, 6, 12, (24, 48, 72, 120, 240) hour precipitation extremes

- stationary at-site model
- non-stationary index-flood model
- stationary at-site model across aggregations

STATIONARY AT-SITE MODEL

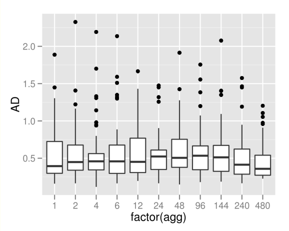


STATIONARY AT-SITE MODEL



NON-STATIONARY INDEX-FLOOD MODEL

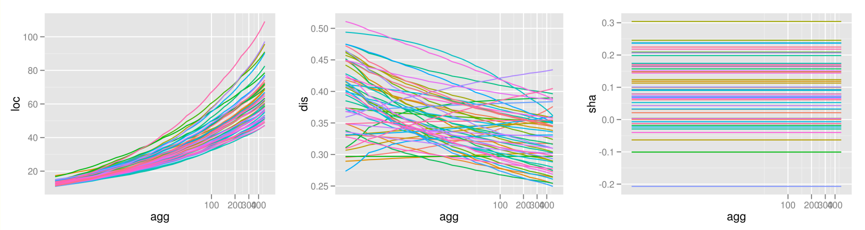
assumed constant shape parameter, location parameter and dispersion coefficient dependent on NH temperature anomaly

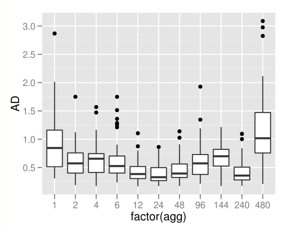


STATIONARY AT-SITE MODEL ACROSS AGGREGATIONS

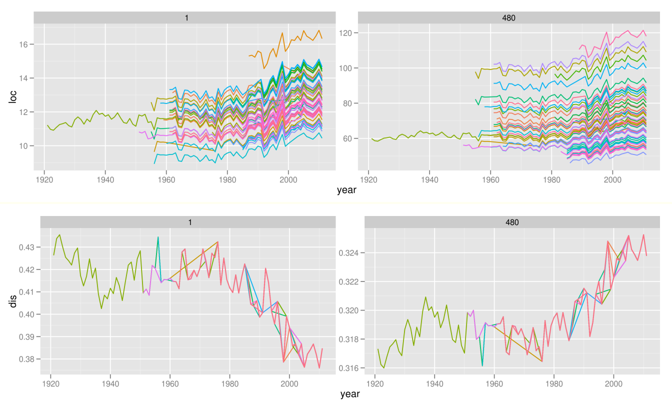
$$\theta = a + b \ln D^c,$$

with θ the GEV parameter, D aggregation and a, b, c parameters.
For the shape parameter assumed constant value for all aggregations.



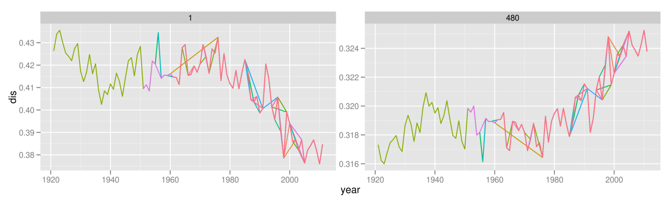


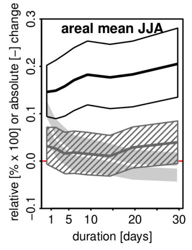
TRENDS



TRENDS


1
480
loc
year
16
14
12
10
120
100
80
60
1920
1940
1960
1980
2000


1
480
dis
year
0.43
0.42
0.41
0.40
0.39
0.38
0.324
0.322
0.320
0.318
0.316
1920
1940
1960
1980
2000


areal mean JJA
relative [% x 100] or absolute [–] change
duration [days]
0.3
0.2
0.1
0.0
−0.1
1
5
10
20
30

CONCLUSIONS

- Regional GEV modelling provides an informative summary of changes in parameters and various quantiles (rather than a single quantile only).
- Taking ξ and γ constant over a region strongly reduces standard errors.
- Consistent parameters across aggregations can be obtained by modification of the model.

Thank you for attention