Rainfall trends in the USA: data quality, prediction skill, and presentation of results

Holger Rootzén, *Mathematical Sciences, Chalmers, Gothenburg*



Joint with: Helga Olafsdottir | David Bolin | Rick Katz

Aim: Inform infrastructure planning, both for protection against high-impact catastrophes and for local planning of roads and sewers.

Use the best data and the best prediction methods for this

1. Data quality: A crucial issue for water resources modeling

- *River systems modeling*: Translation of water speed to water flow
- *Temperature and rainfall*: Station measurements available for much longer time periods and are more precise, but satellite measurements have much better coverage
- *Rainfall*: Often annual maxima series are much better checked than partial duration series, but partial duration series contain more information (partial duration series = PoT data)

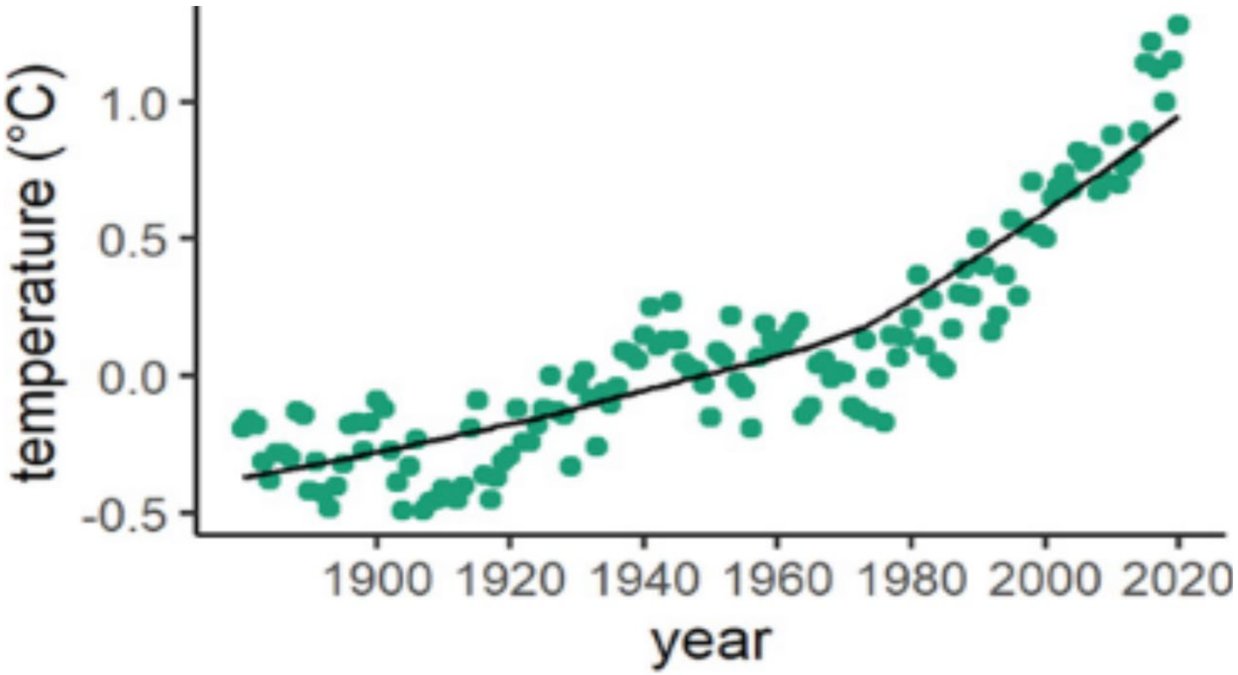
Is climate change making individual extreme rainfall events in the United States more frequent, more intense, or both?

The GHCN data contains daily measurements but is less checked

NOAA Atlas 14, vol. 10 contains carefully checked annual maxima of daily rainfall for 333 stations in northeastern USA which have at least 60 years of data and cover a time period up to at least 2010

Applying PoT to the GHCN data answers the questions directly – but how does the lower data quality affect these answers? We have instead developed a method to use the annual maxima data to answer the questions

Olafsdottir et al. J. Climate 2022



Loess smoothing of yearly average temperature in Northern Hemispere

Temperature as covariate:

- Catches broken trend
- Physically motivated
- Can be extrapolated to future climate scenarios

Standard approach in a non-stationary climate

(t = time or yearly average temperature or …)

Annual Maxima

$\mathrm{GEV}(\mu(t), \sigma(t), \gamma)$

location $\mu(t) = \mu_0 + \mu_1 t$

scale $\log(\sigma(t)) = \sigma_0 + \sigma_1 \mathrm{t}$

shape γ

Peaks over Threshold

excesses of threshold u are $\mathrm{GP}(\sigma(t), \gamma)$

Poisson intensity $\log(\lambda(t)) = \lambda_0 + \lambda_1 t$

scale $\log(\sigma(t)) = \sigma_u + \sigma_{u,1} \mathrm{t}$

shape γ

$\boldsymbol{\lambda_1 > 0}$***: rains become more frequent***

$\boldsymbol{\sigma_{u,1} > 0}$***: rains become more extreme***

Maximum likelihood estimation

A variant of Langbein's 1949 formula

Assume excesses of threshold u are GP(σ_u, γ), exceedance times are Poisson process with intensity λ. Then, for $x > u$, yearly maxima have cdf

$$\exp\{-\left(1+\gamma\frac{x-\mu}{\sigma}\right)_+^{-\frac{1}{\gamma}}\} = \exp(-\left\{1+\gamma\frac{x-\left(u+\sigma_u({\lambda_u}^{\gamma}-1)/\gamma\right)}{\lambda_u^{\gamma}\sigma_u}\right\}^{-1/\gamma})$$

The PGEV model: A reparametrized GEV distribution with

location $\mu = u + \sigma_u(\lambda_u^{\gamma} - 1)/\gamma$

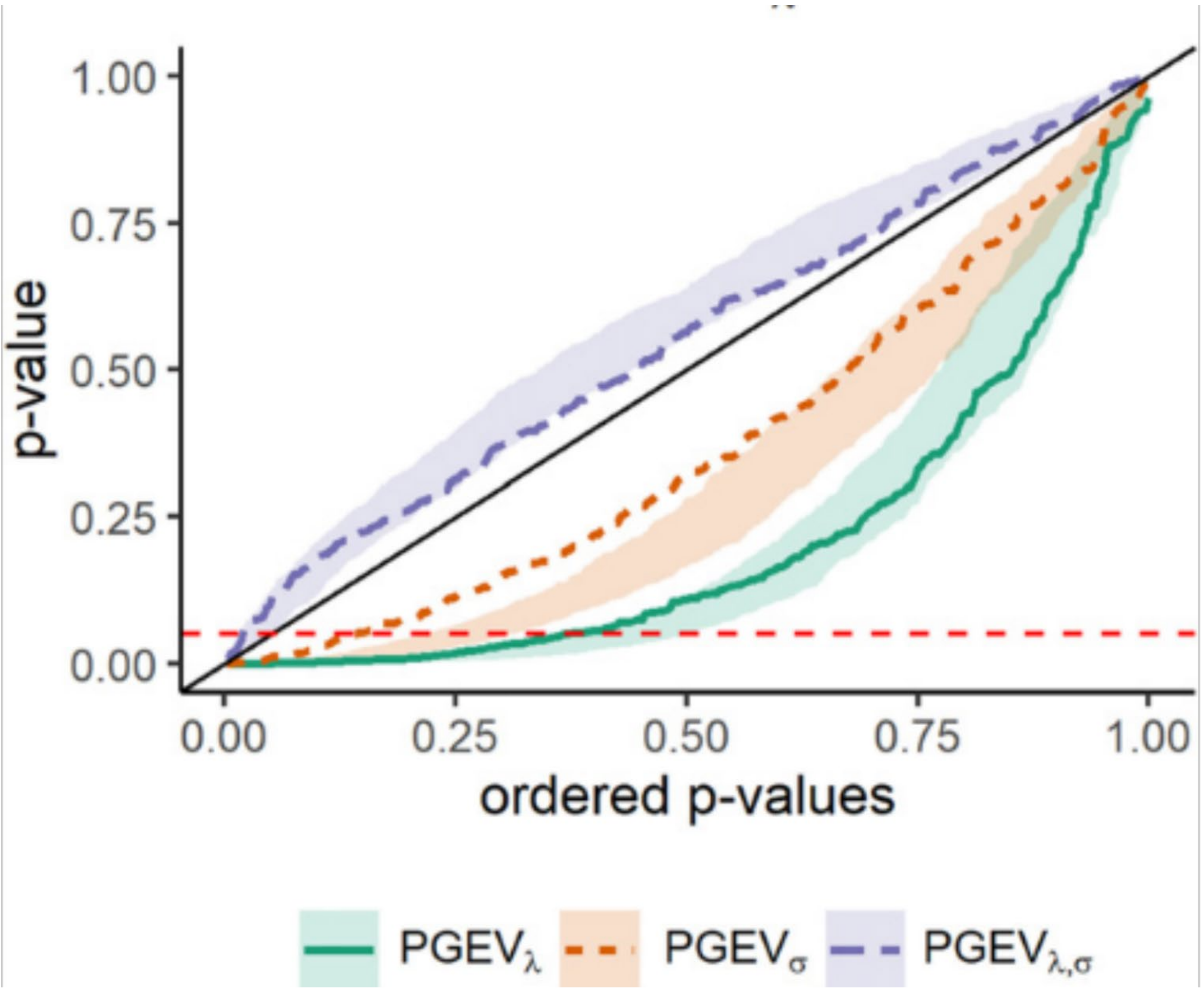
scale $\sigma = \lambda^{\gamma}\sigma_u$

shape γ

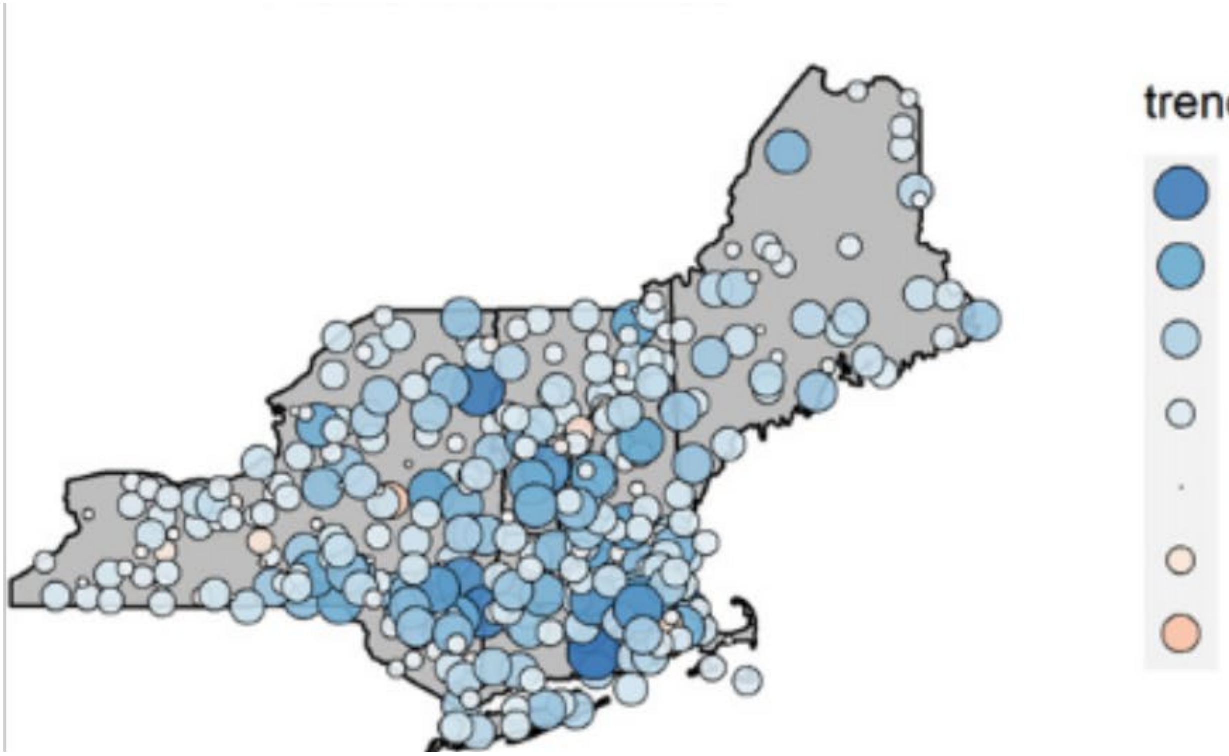
Ordered p-values from LR tests against best submodel for the NOAA annual maxima of daily rainfall for the 333 stations in northeastern USA

Confidence intervals from simulation

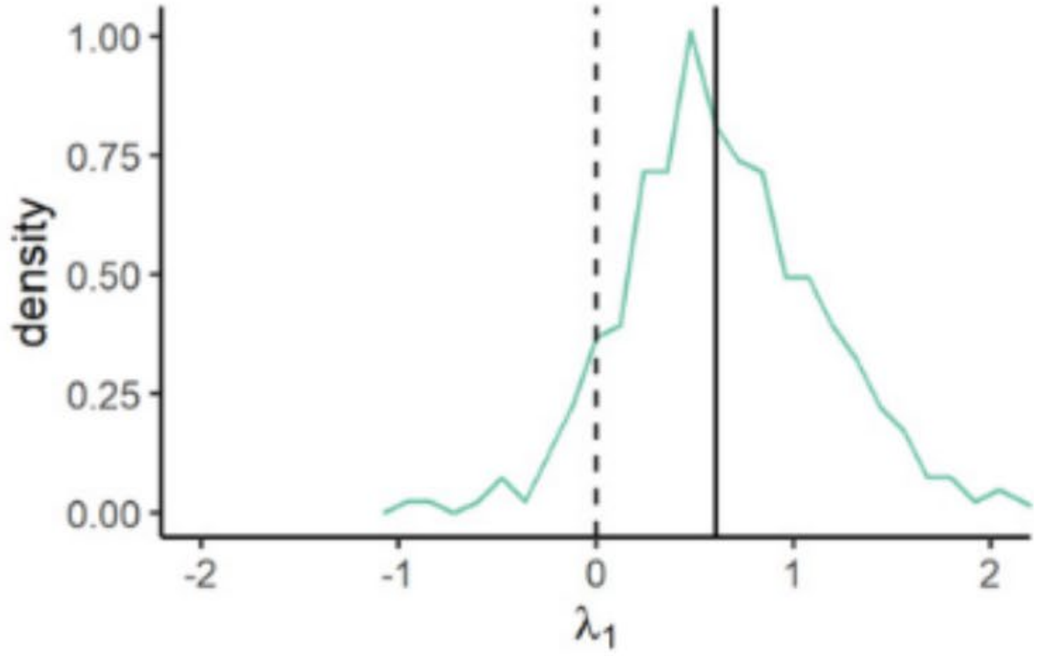
Indicates trends in frequency, no trends in scale (even if no trend in sigma the PGEV$_\sigma$ p-values are likely to be small)



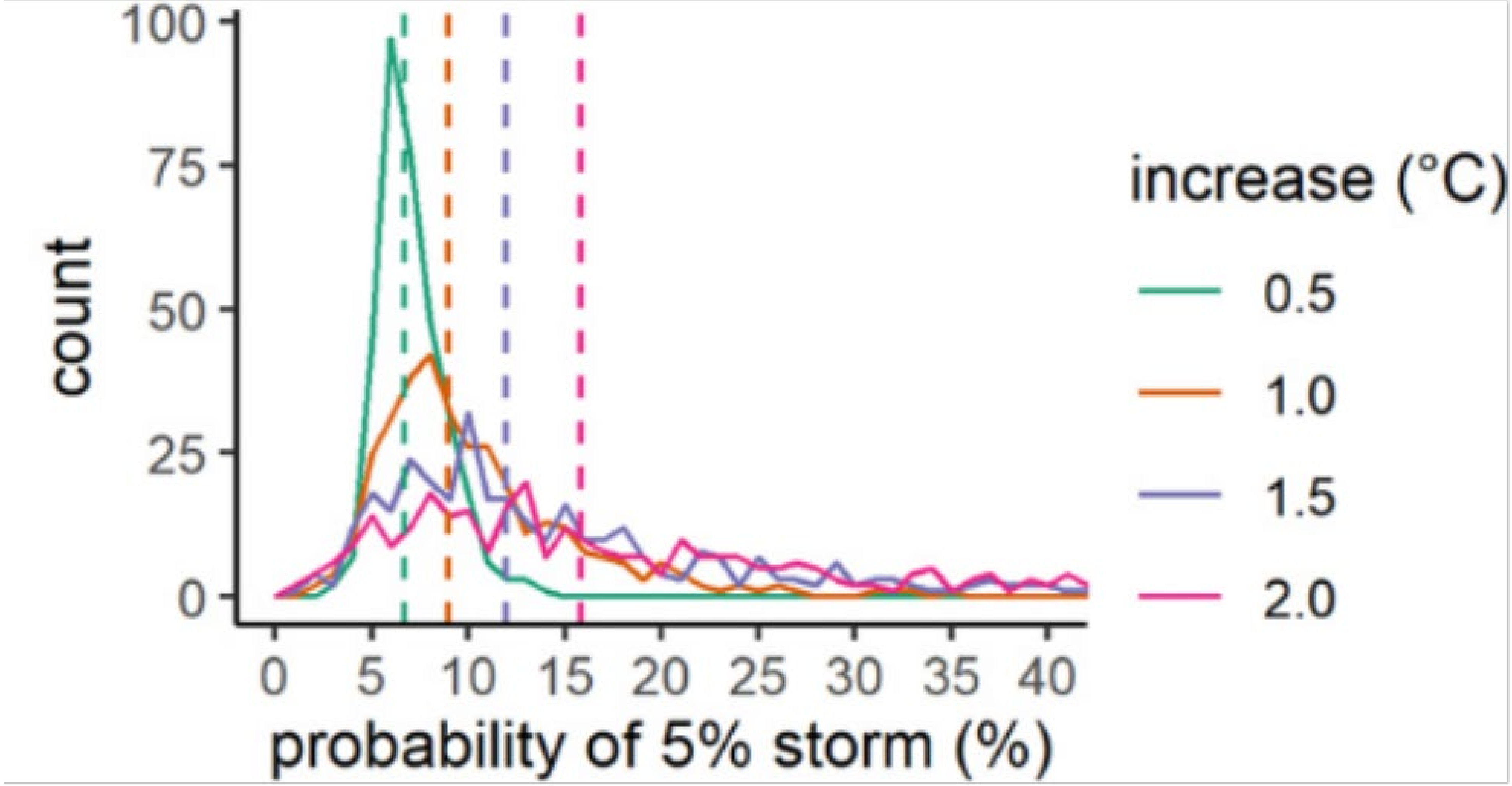
Results from $\mathbf{PGEV}_{\lambda}$ for the annual maxima data



Estimated trend in Poisson intensity



”Histogram” of estimated trend



5% storm is the daily rainfall that has the probability 5% of being exceeded during a year, in the current climate

Details and model checking

- Declustering
- Choice of threshold
- Does the method work for the real rainfall data?

Investigated using NOAA data on carefully checked daily rainfall measurements from 16 stations in northeastern USA

- Graphics and fitting a Matern random field to the estimated stationwise trends suggested little difference between the trends

Conclusions

- Extreme daily rainfall events in the USA are becoming more frequent
- Little evidence of increasing trends in the distribution of sizes of individual extreme daily rainfall events.
- Trends strongest in the northeastern US where for many measuring stations the frequency increase exceeds 150% for 1°C of average Northern hemisphere temperature increase
- And, of course the maximum of many rainstorms is larger than the maximum of fewer storms, so the worst events in the future still will be more extreme than those in the past

2. Prediction skill

- Scoring rule: takes a predicted distribution and the predicted observation and returns a number. Choice of prediction model often based on average scores from many predictions.
- Log likelihood score and CRPS examples of scoring rules. (CRPS $= \int (F(x) - 1\{x > y\})^2 \, dx$, for F the predicted distribution and y the predicted observation)
- Special scoring rules needed for extremes. One issue is local scale invariance – scale invariance often desirable for modeling of extreme rainfalls

- Olafsdottir et al. J. Forecasting 2024

Annual maxima of daily rainfall for 333 stations in northeastern USA.

Our new scoring rule: tailored to extremes, locally scale invaraiant

	LS	CRPS	SCRPS	LS_q		wCRPS		swCRPS	
				90%	99%	90%	99%	90%	99%
Gumbel	-1.148	-0.483	-1.048	-0.427	-0.0903	-0.0865	-0.010178	-0.1556	-1.531
GEV	-1.129	-0.481	-0.966	-0.419	-0.0856	-0.0860	-0.010201	-0.0795	0.893
GEV_μ	-1.108	-0.473	-0.958	-0.415	-0.0850	-0.0857	-0.010199	-0.0644	0.906
$PGEV_\lambda$	**-1.107**	**-0.472**	**-0.956**	**-0.414**	**-0.0847**	**-0.0854**	-0.010195	**-0.0603**	0.901

Bold means p-value of t-test between $\mathbf{PGEV}_\lambda$ and $\mathbf{GEV}_\mu$ less than 0.05

Here all scoring rules point to the same prediction method ($PGEV_\lambda$). This is not the case for some of the other examples we have looked at.

3. In the past the concepts of return levels and return periods have been standard and important tools for engineering design – but they do not apply to a changing climate, whether local or global

Rootzen & Katz Water Resources Research 2013

And anyway, nobody except hydrologists and statisticians can understand return levels and return periods – and it is hard for them too.
Hydrologists in the USA have started to understand this

- But, hydrologists in Europe and statisticians from all around the world have not ☺

In a nonstationary climate the basic information needed for engineering design is

(i) the design life period (e.g. the next 50 years, say 2017-2066) and
(ii) the probability (e.g., 5% chance) of a hazardous event (typically in the form of the hydrological variable exceeding a high level) during this period

The *Design Life Level* is defined as an upper quantile (e.g. 5%) of the distribution of the maximum value of the hydrological variable (e.g. water level) over the design life period.

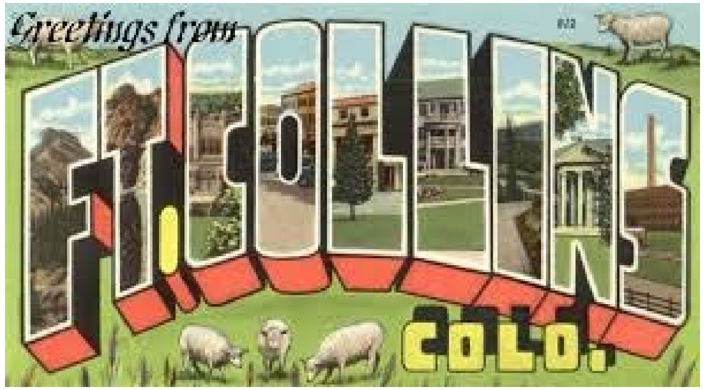
This concept – or related ones -- should be used in hydrology and climate science

- Technical quantification/communication:

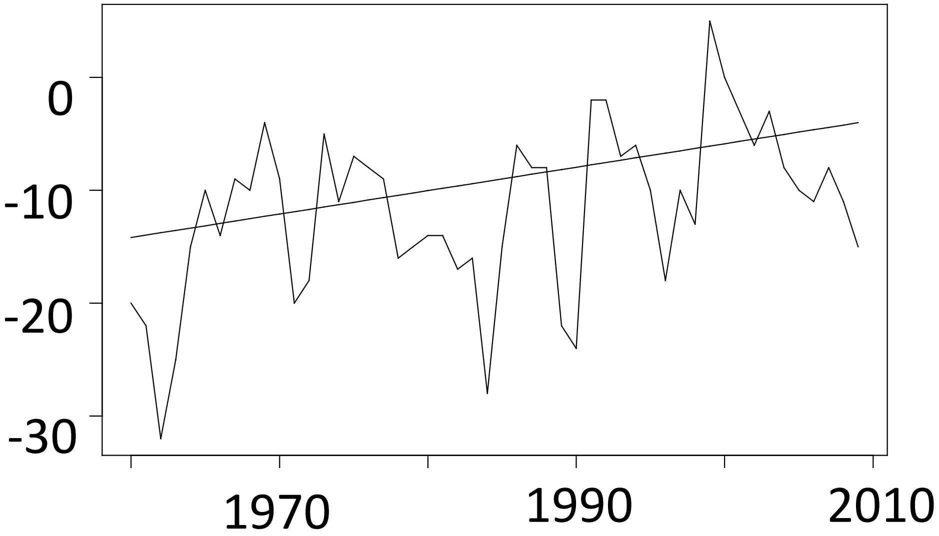
  “the 2015-2064 5% highest water level is 11.5 m”

- Communication with the public:

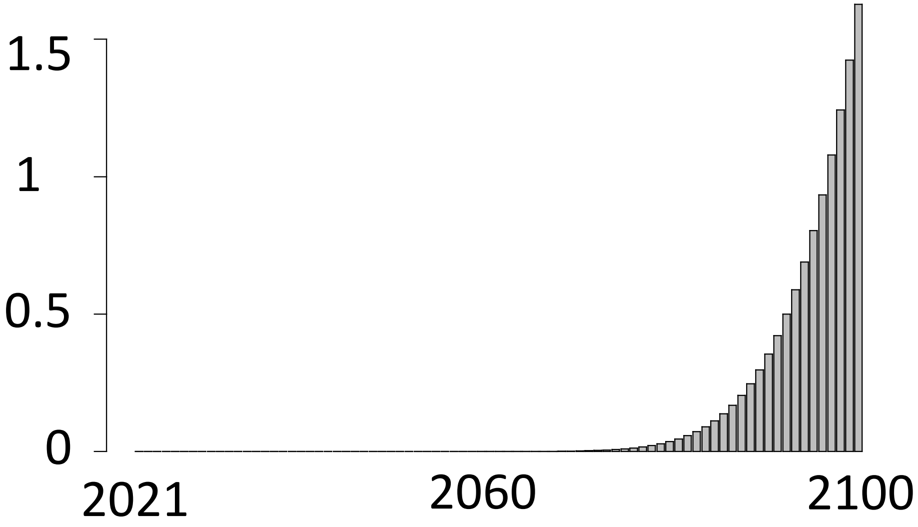
  ”there is a 5% risk that the biggest flood during 2015-2064 will be higher than 11.5 m”



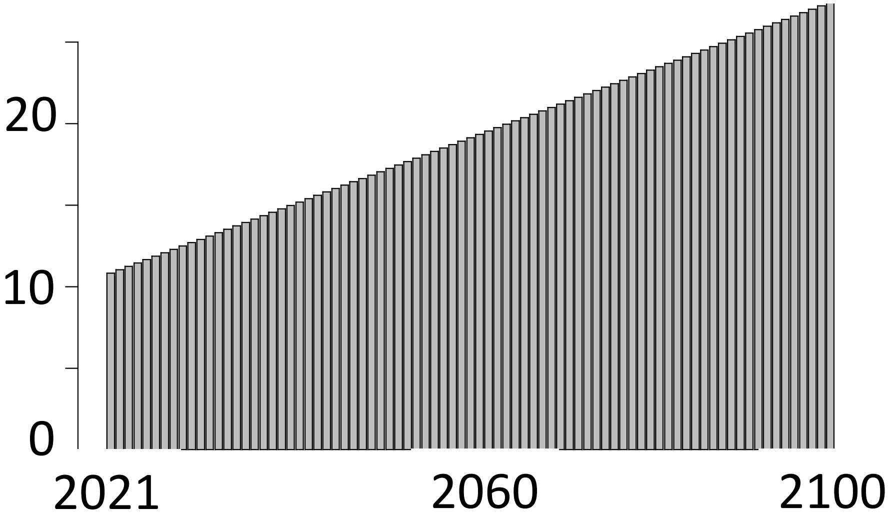
the 2021–2060 10 % highest minimum winter temperature in Fort Collins is 24 degrees



Yearly minimum temperature



The 24 degree risk plot



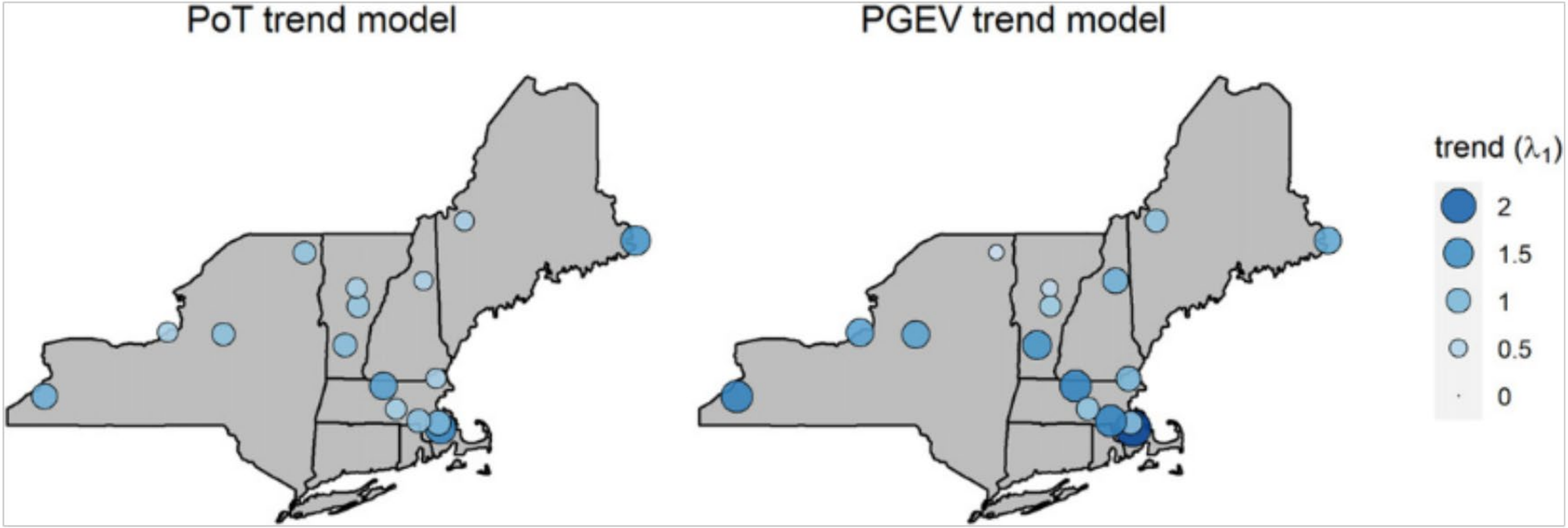
The 0.1% constant risk plot


SCANPIX

50
40
30
20

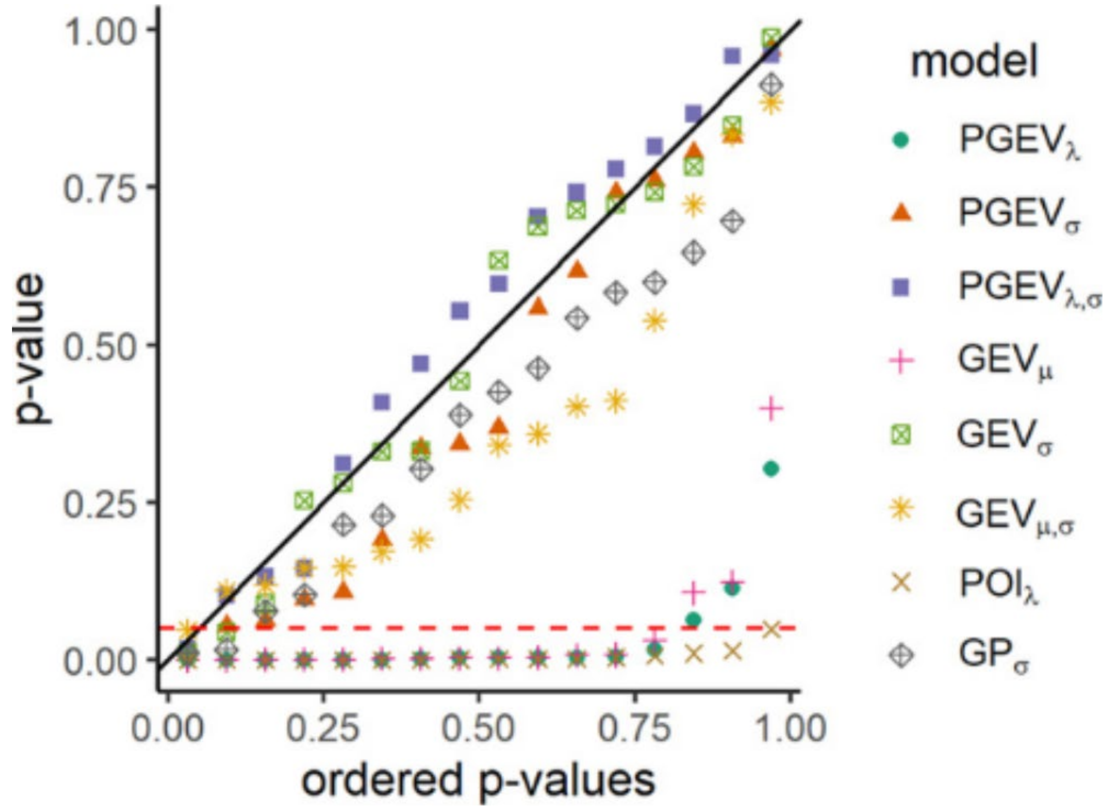
Thanks for listening

Validation of PGEV model on NE-PDS data



Trend λ_1 in Poisson intensity, no trend in shape parameter
Declustered

Model choice for NE-PDS data



Likelihood ratio tests against best submodel. PGEV_λ is PGEV model with trend in λ but not in σ, and so on. Benjamini-Hochberg, etc not suitable here