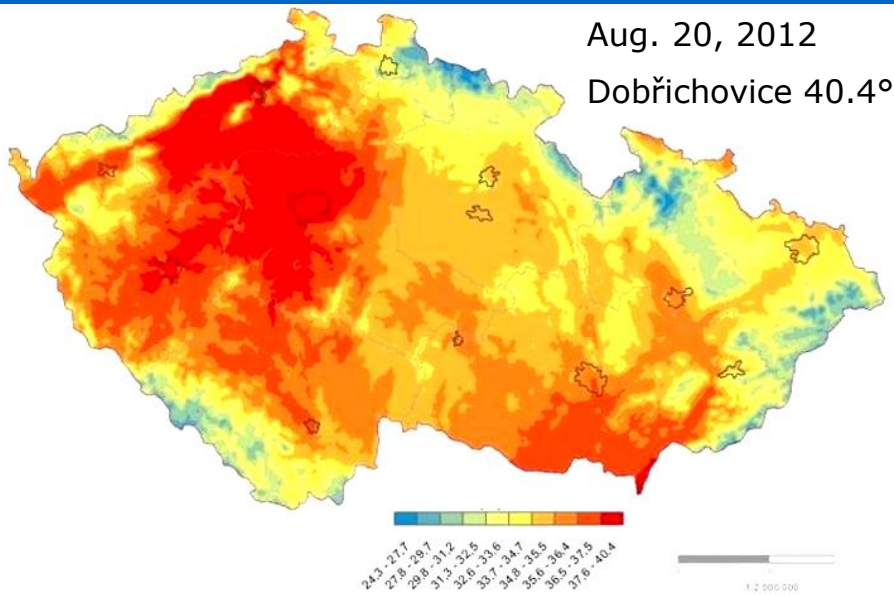
Extreme value analysis in climatology

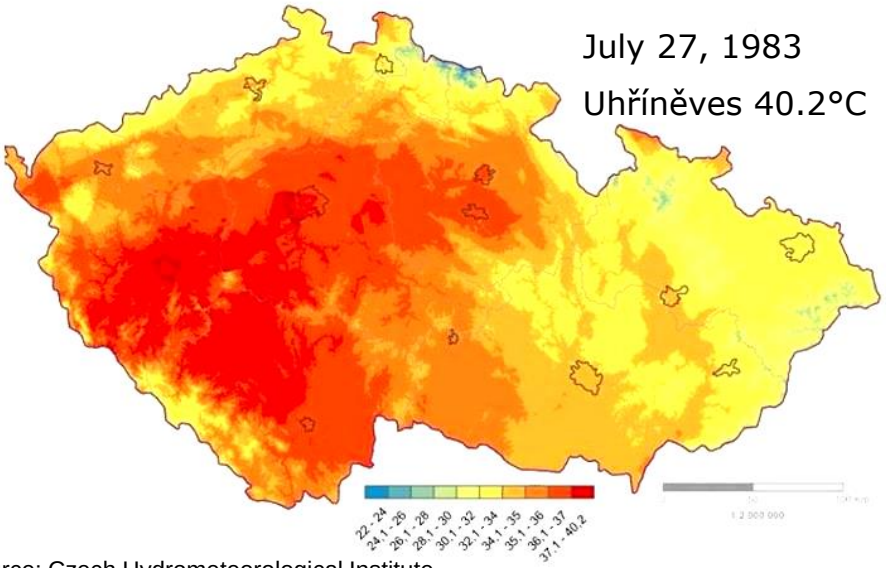
Jan Kyselý (1,2), Jan Picek (2)

(1) Institute of Atmospheric Physics AS CR, Prague
(2) Technical University of Liberec

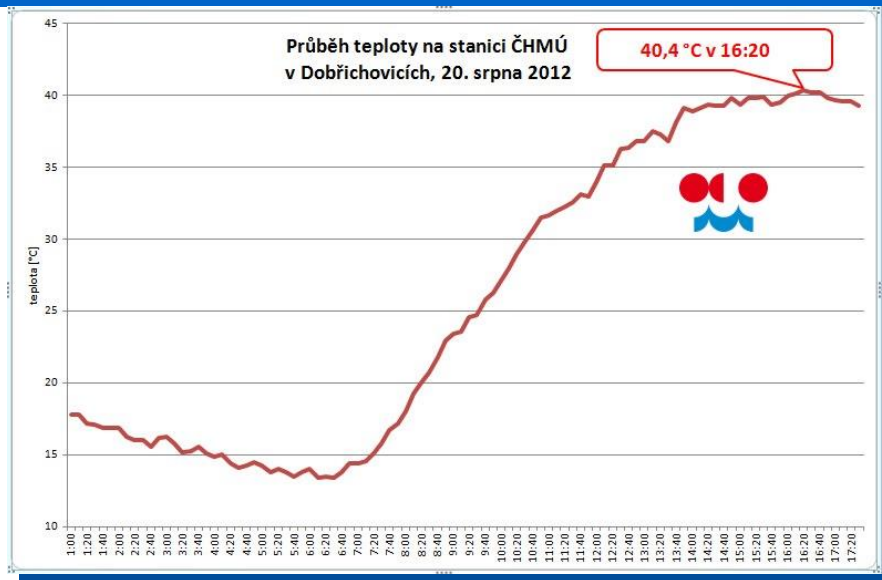
KLIMATEXT session: Advanced statistical models of extremes and their applications

Probability of record-breaking daily temperature…





Source: Czech Hydrometeorological Institute

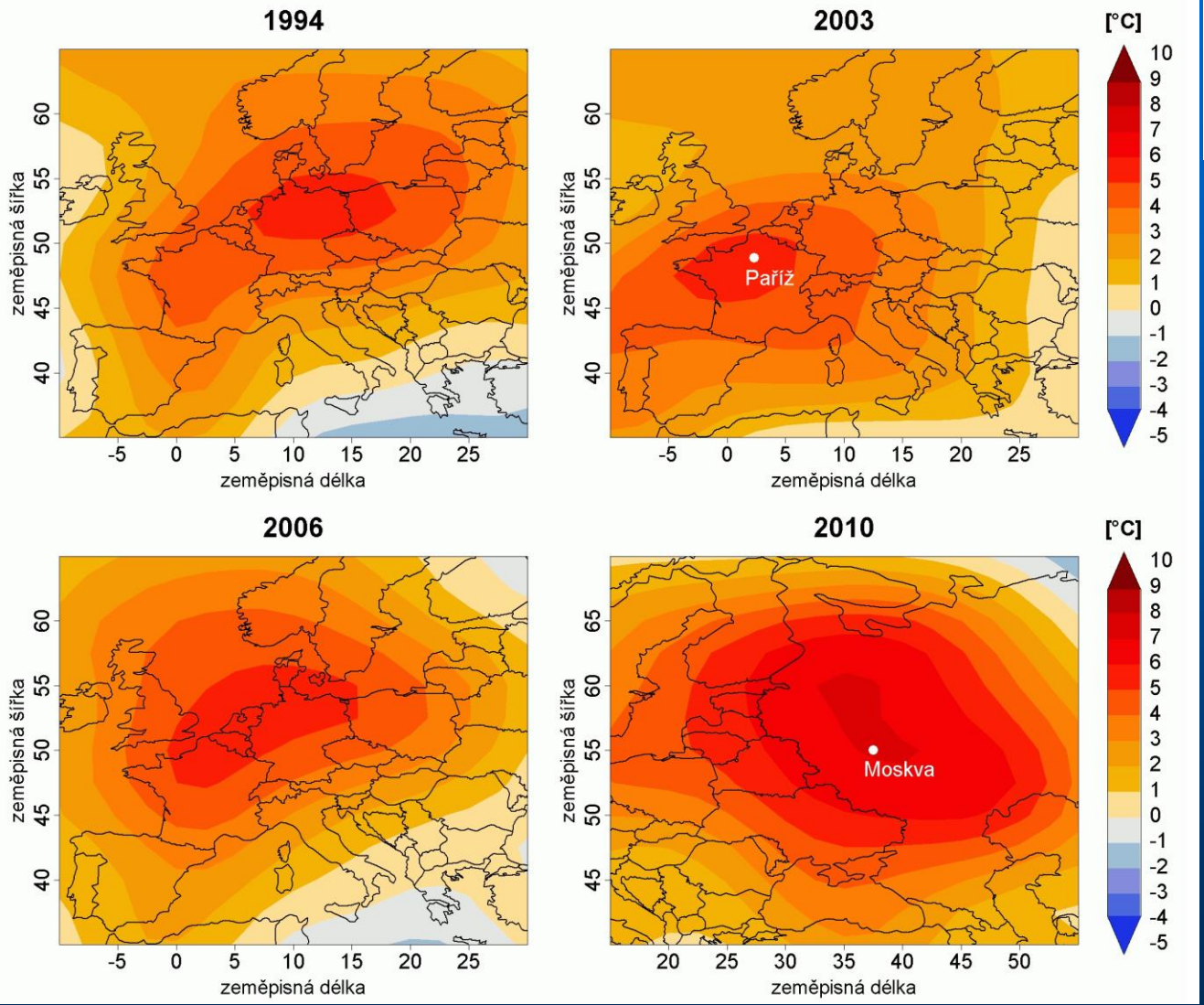


- probability at the given station?
- probability anywhere in the Czech Rep.?
- probability in that part of year? (annual cycle…)
- how does the probability change due to warming trend?
- why did the previous record hold for nearly 30 yrs?

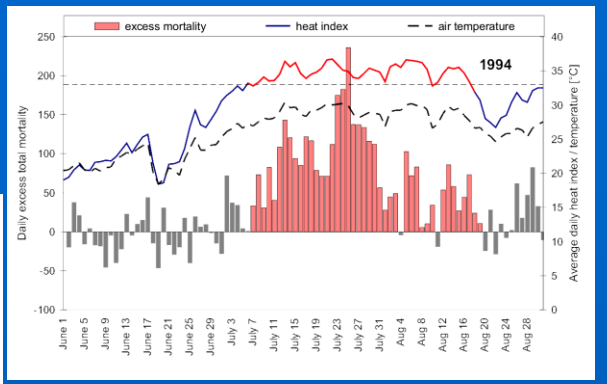
→ special and unexpected event!

… vs. severe heat waves? (multi-day events)

(& how does it change under climate change?)



Monthly temperature anomalies at 850 hPa during heat wave periods in Europe, with respect to 1960–1990.
(Data source: NCEP/NCAR reanalysis, http://www.esrl.noaa.gov/psd/data/reanalysis/reanalysis.shtml)



Magnitude & duration

Major impacts on environment & society

- often associated with droughts
- 2003: ~70,000 excess deaths in Europe
- 2010: ~55,000 excess deaths in Russia

Extreme value analysis

- widely used in climatological, hydrological and other environmental applications

- estimates of

  return levels of (observed, simulated) extremes

  design values (expected to occur with a given probability)

  uncertainty

  changes/trends in observed extremes

  changes projected by climate models for a perturbed climate

(e.g. late 21st century – not a "future climate" but "projected climate" under given assumptions, e.g. increased radiative forcing due to greenhouse gases)

- estimates (and their uncertainty/changes) important because of impacts of extremes and their role in practical applications

Extreme value analysis

- "routine methods":

  block maxima, GEV distribution

- "less routine methods" (need additional "initialization"/setting):

  peaks-over-threshold (POT), GP distribution, Poisson process

  (choice of threshold, checking for independence of exceedances)

- "advanced methods":

  models including non-stationarity/covariates

  spatial/regional analysis

  multivariate models

  compound models (two-component distributions, …)

  time series modelling

  (choice of covariates and models for dependence on covariates,
  checking for regional homogeneity, etc.)

Jeseník nad Odrou, June 25, 2009

Strunkovice nad Blanicí, June 28, 2009

Road between Černětice and Malenice, June 28, 2009

Jeseník nad Odrou, June 25, 2009

- flash flood in the Odra river basin (the Nový Jičín district) in late evening of June 24, 2009, leaving 10 people dead and causing extensive damage to buildings and infrastructure

(daily precipitation amounts: Bělotín 123.8 mm, Hodslavice 120.2 mm)

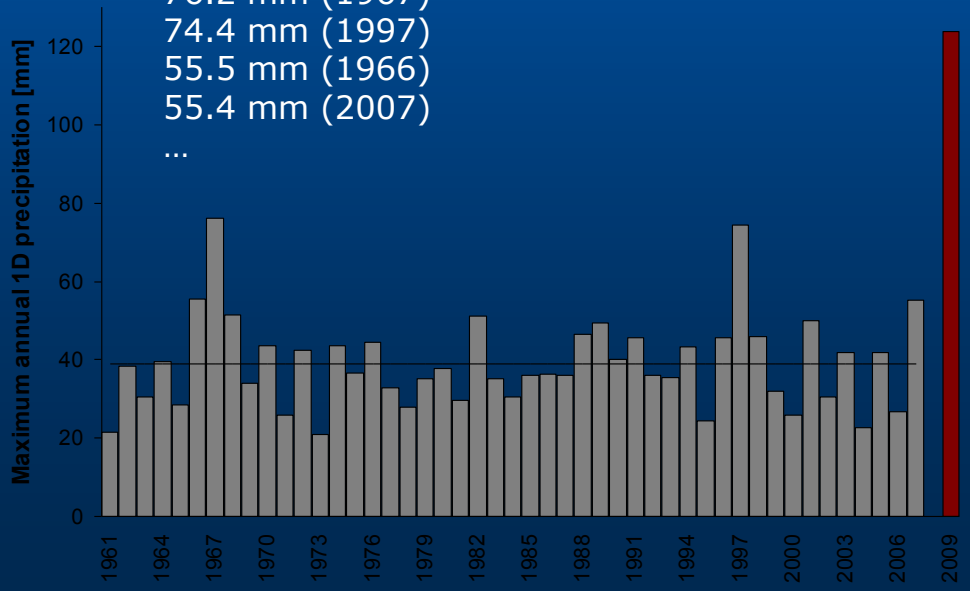
Jeseník nad Odrou, June 25, 2009

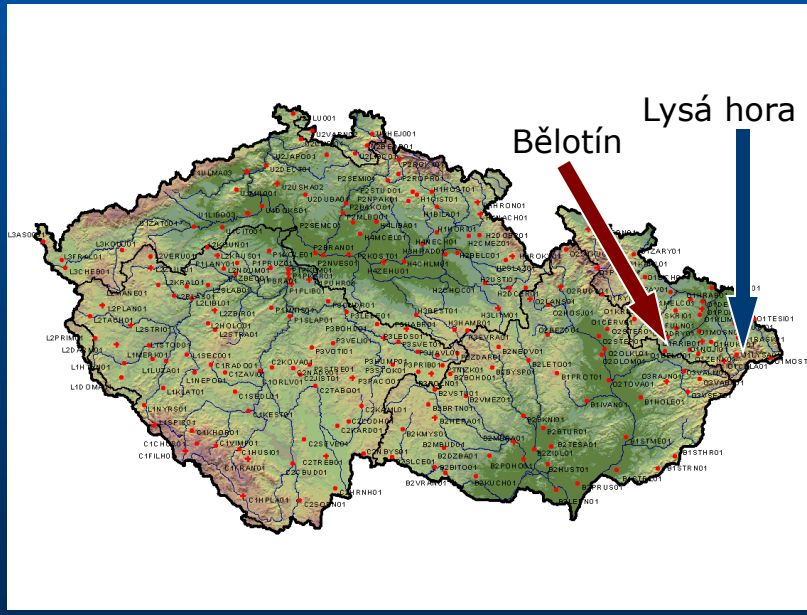
Bělotín, 298 m a.s.l.: 24-hour precipitation amount 123.8 mm (114 mm during 3 hours, 19-22)

'local' event – Lysá hora 4.5 mm

previous daily maxima at Bělotín (1961-2007):

76.2 mm (1967)
74.4 mm (1997)
55.5 mm (1966)
55.4 mm (2007)
...





Block maxima, GEV distribution

At-site estimation does not provide meaningful answers

Regional methods tend to give similar answers in spite of different concepts

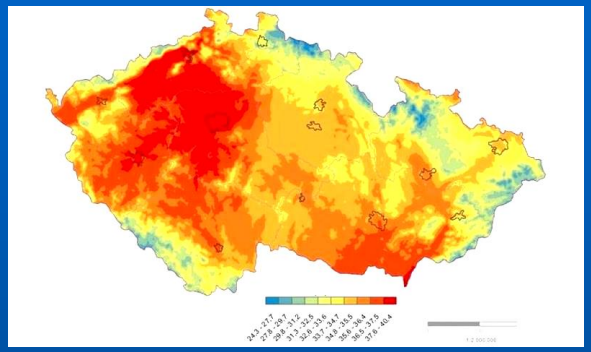
Table IV. Return periods and their 90% CIs associated with the 1D precipitation amount observed at the Bělotín station on 24 June 2009 (123.8 mm), estimated by different methods.

	At-site	*ROIgeo2*	*ROIhyb*	HW
Return period (90% CI) estimated from the 1961–2007 data (years)	45110 (605–2.123 × 10^9)	657 (458–1340)	695 (392–1568)	483 (359–920)
Return period (90% CI) estimated from the data supplemented with the single observation on 24 June 2009 (years)	283 (85–69 730)	419 (333–864)	415 (279–944)	353 (275–641)

Source: Kyselý, Gaál & Picek, Int. J. Climatol. 2011

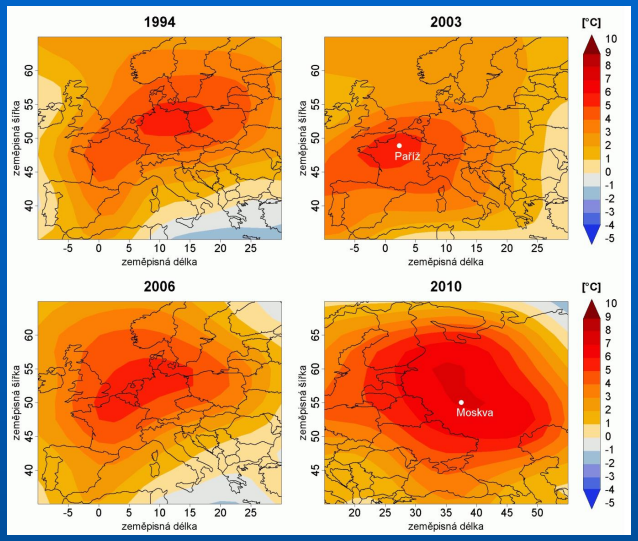
Different spatial and temporal scales of extremes → different methods needed

Air temperature:



One-day extremes, smaller areas

x



Multi-day extremes, large areas

Precipitation / floods:

Flash floods from localized short-term storms

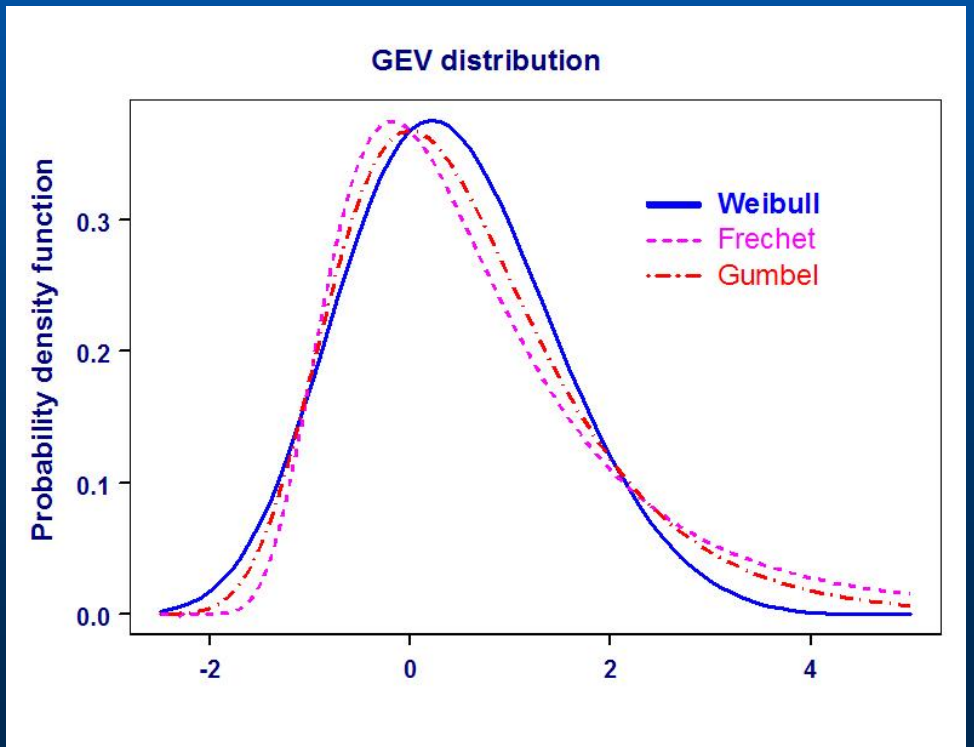
x

Large-scale floods from widespread rainfall (Prague 2002)

Extreme value analysis in climatology

variables most often examined:

- air temperature (2m, maximum/minimum daily values; close to Gumbel distribution)
- precipitation (strong evidence for heavy-tailed Frechet distribution)
- wind speed (light-tailed Weibull distribution)
- river discharges
- wave heights
- etc.



Extreme value analysis in climatology

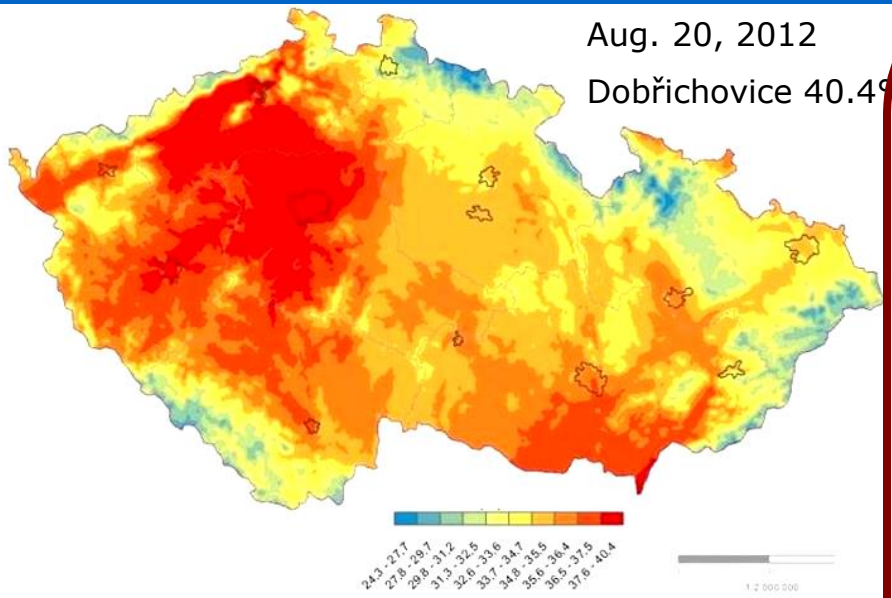
models in climatology:

- statistical (extreme value analysis, statistical downscaling)
- dynamical (Global Climate Models, Regional Climate Models)

interaction of both statistical models & dynamical models often needed when estimating recurrence probabilities of events:

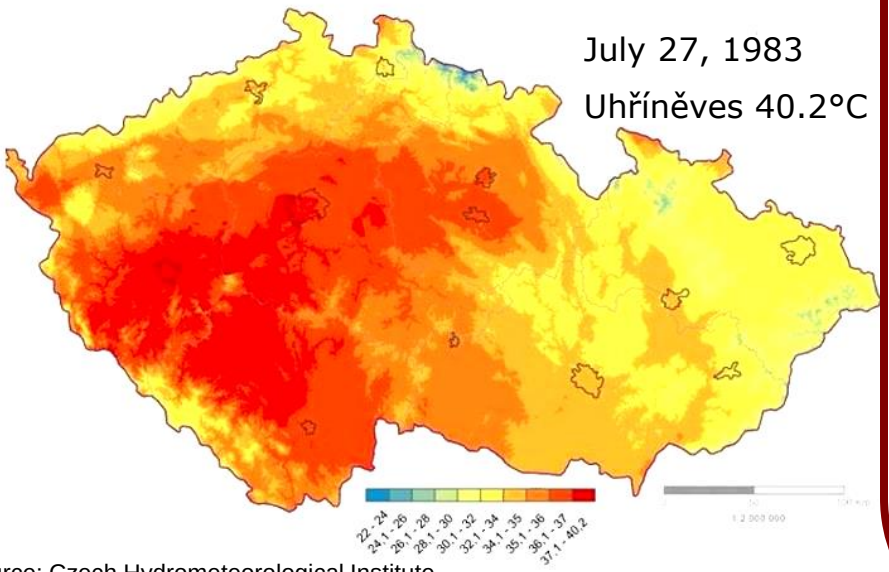
- to understand patterns in the data
- to understand physical processes behind the data

Probability of record-breaking daily temperature…



Aug. 20, 2012

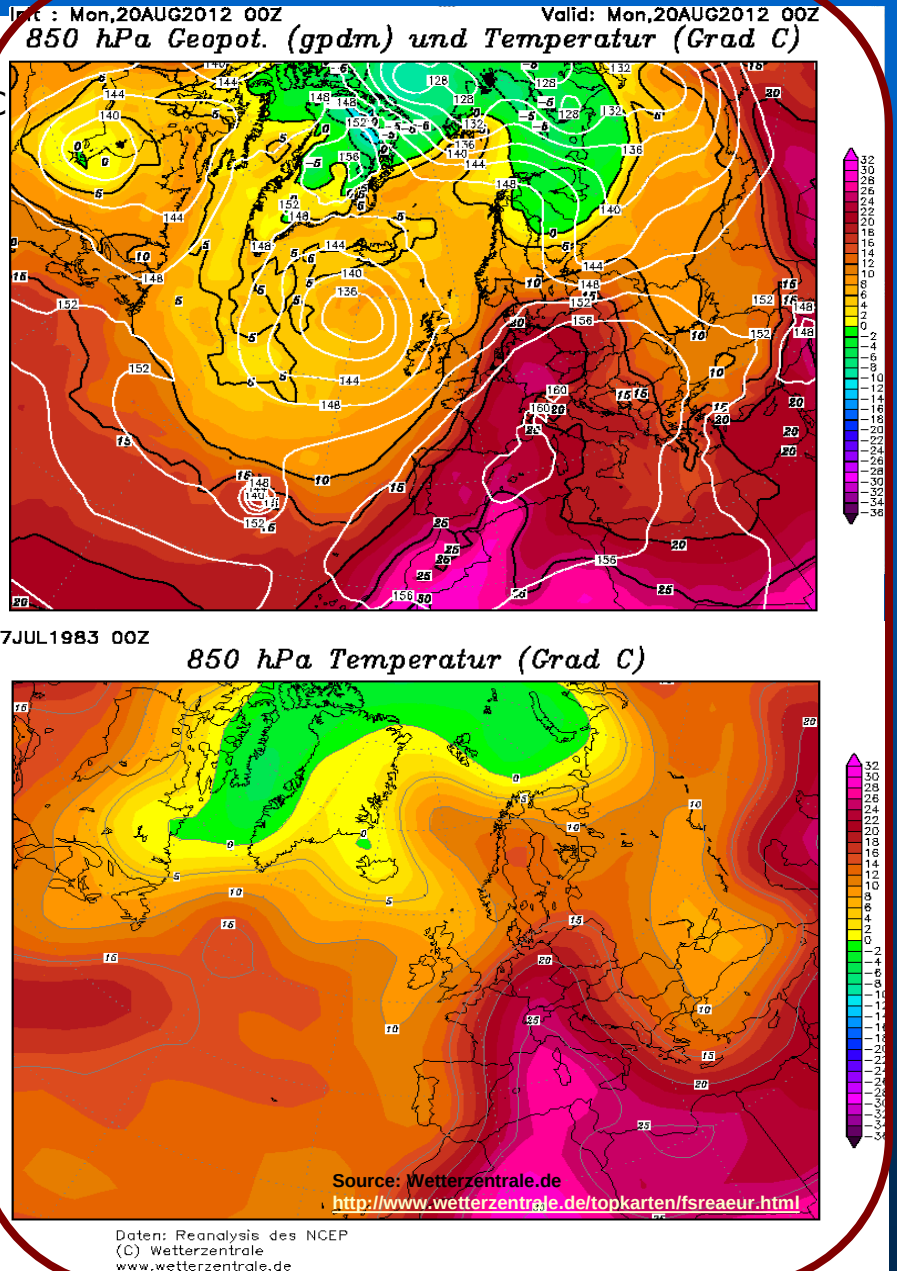
Dobřichovice 40.4°C

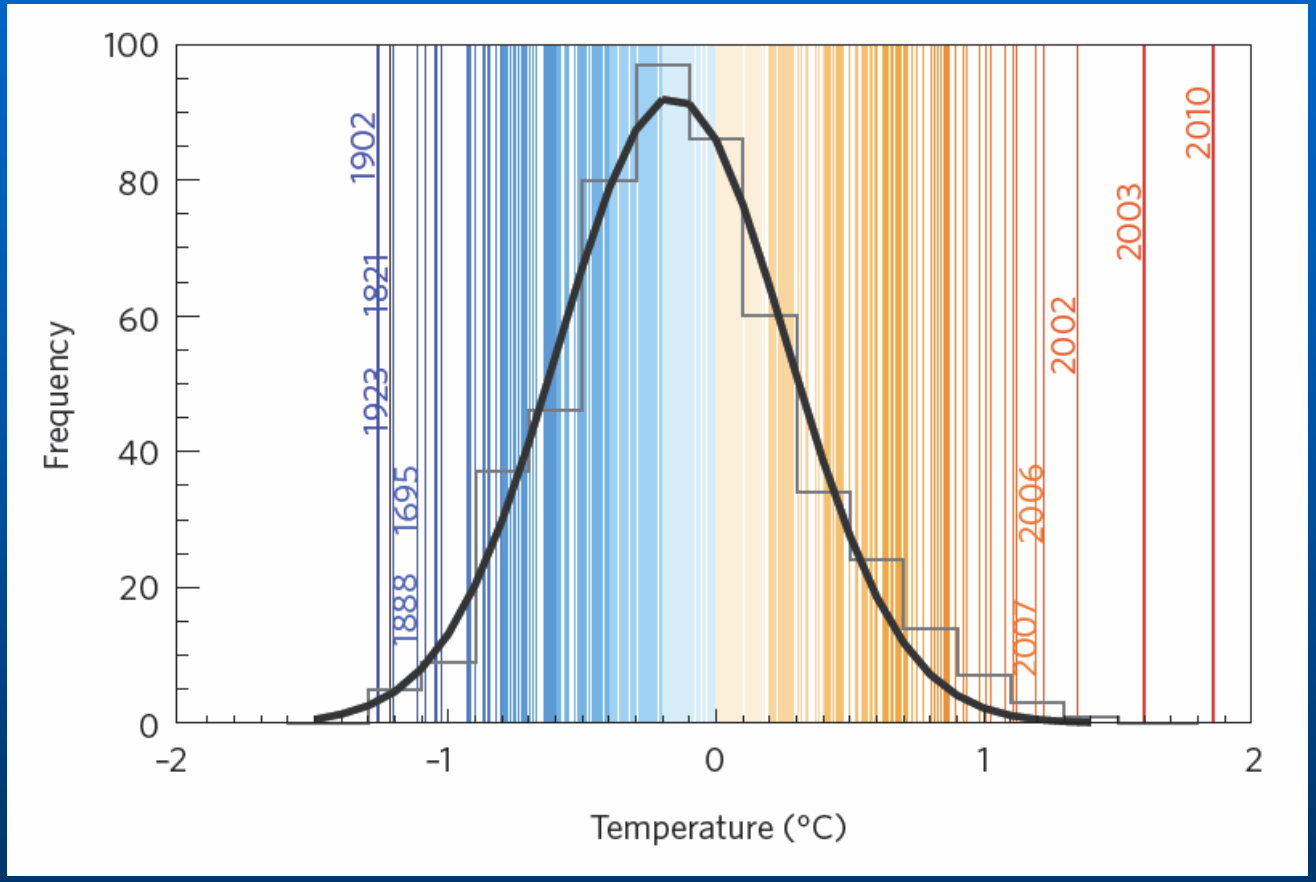


July 27, 1983

Uhříněves 40.2°C

Source: Czech Hydrometeorological Institute





European summer temperatures for 1500–2010. Statistical frequency distribution of best-guess reconstructed and instrument based European ([35°N, 70°N], [25°W, 40°E]) summer land temperature anomalies (°C, relative to the 1970–1999 period). The five warmest and coldest summers are highlighted. Gray bars represent the distribution for the 1500–2002 period, with a Gaussian fit in black.
(Source: Barriopedro et al., Science 2011)

Advanced statistical models of extremes

recent years – development of advanced statistical models for estimating probabilities of climate extremes:

1) non-stationarity/covariates

2) spatial/regional and multivariate models

attracting a lot of interest – "extreme progress" in development and implementation of these methodologies in the past 5-7 years in the international literature, due to availability of

1) statistical methodologies

2) computation/programming resources

3) "research need" – climate variability and change

- still many open and unresolved issues

Advanced statistical models of extremes

open issues – even "basic" ones:

- how to make most efficiently use of available data?
- how to choose threshold in a POT analysis? (M. Schindler)
- which spatial/regional models for extremes are most useful, what is the "added value" of more sophisticated models? (M. Hanel, M. Roth)
- which covariates (and forms of dependence of extremes on covariates) are most useful? (S. Begueria, P. Jonathan)
- how to deal with covariates from the statistical point of view? (J. Dienstbier)

- do results of individual methods and approaches meet? do we understand answers they give? were the questions well posed?
- to what extent may results of extreme value models be biased due to violated assumptions of the extreme value theory?

(almost always violated… – small samples, spatial and temporal dependence, various physical processes involved in generating extremes…)

KLIMATEXT session:

Advanced statistical models of extremes & their applications

Part I: **From regional analysis to the peaks-over-threshold method**

Part II: **Covariates**

- Key question: **How to improve estimates of extremes** (distributions, return levels, design values) **and make most efficiently use of available data?**

- Answer(s): may be useful in **other fields**, not just climate research…

Part I: From regional analysis to the peaks-over-threshold method

13:45-14:15 Martin Hanel: Regional block-maxima modelling of precipitation extremes in climate model simulations

14:15-15:00 Martin Roth: Regional peaks-over-threshold modelling with respect to climate change

15:00-15:30 Martin Schindler: How to choose threshold in a POT model?

15:30-15:50 Coffee/tea break

Part II: Covariates

15:50-16:35 Santiago Begueria: Covariate-dependent modelling of extreme events by nonstationary POT analysis

16:35-17:20 Philip Jonathan: Modelling covariate effects in extremes

17:20-17:50 Jan Dienstbier: Covariate effects in extemes - remarks and theory

17:50-18:20 Discussion

18:30-19:45 Dinner

Part III: Přednáška s diskuzí

20:00-21:30 Ladislav Metelka: Změna klimatu - mýty, fakta, statistika