Introduction to Earth System

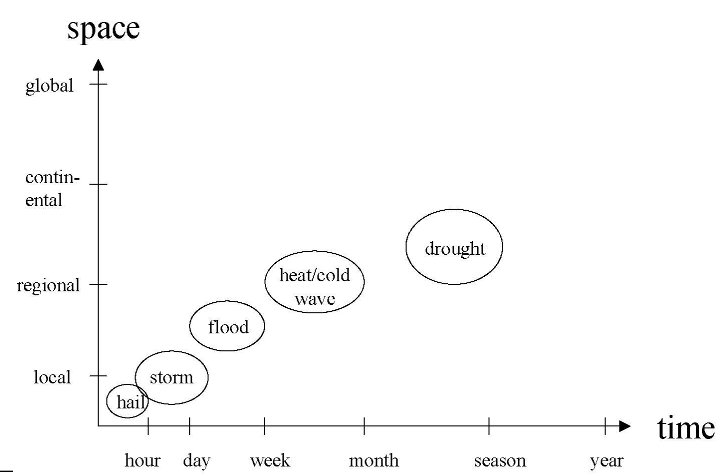
Working with data

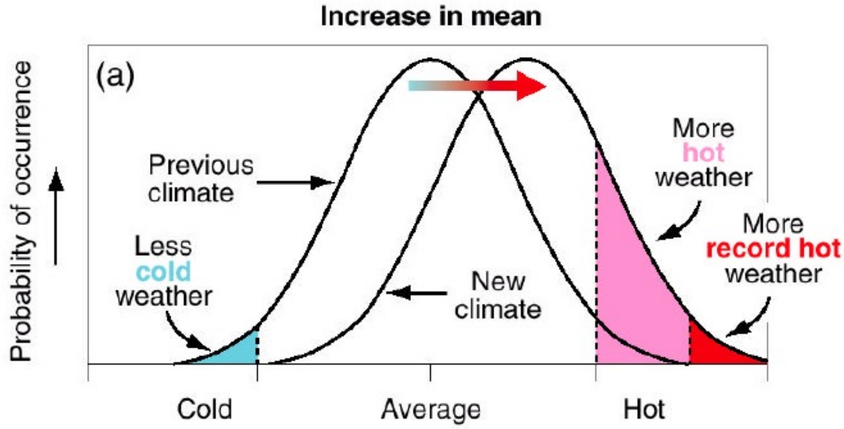
Extreme events

Extreme events

Definition:

Extreme event … very **rare** and very **intense** event with **severe** impacts on society and biophysical systems.



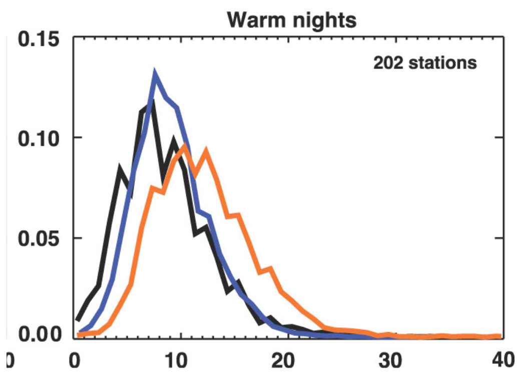


IPCC - TAR

3

E.g. Changes in extreme temperature

- Obs. showed a decrease in the number of frost day in the mid-latitude regions; a decrease in the number of cold nights, an increase in the number of hot days

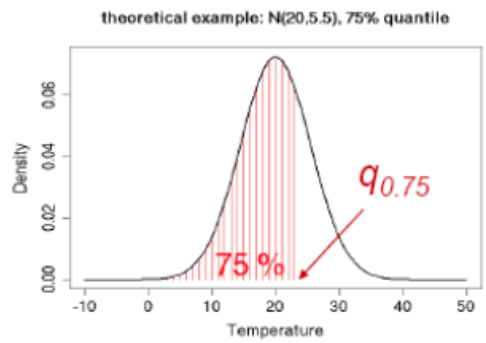


3 periods: **1901-1950 (black)**, **1951-1978 (blue)** & **1979-2003 (orange)**

4

Percentile

The ***P*^th^ percentile** of a list of *N* ordered values (sorted from least to greatest) is the smallest value in the list such that *P* percent of the data is less than or equal to that value.



Exercise:
{5,10,25,30,40}
→ What are the 30th, 40th, 50th, and 100th of the list?

- The Nearest Rank Method

n=[P*N/100+0.5]

Exercise #1

- The daily rainfall values (mm/d) for 20 days are: 62, 66, 71, 75, 75, 78, 81, 83, 84, 85, 85, 87, 89, 89, 91, 92, 93, 94, 95, 99

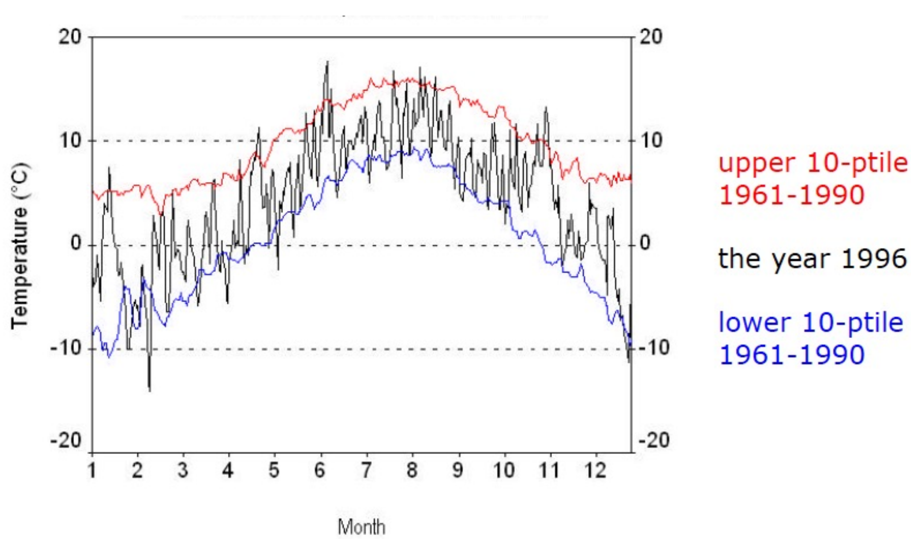
1. Find the value in the list corresponding to the 30th percentile
2. Find the percentile rank for a rainfall value of 85

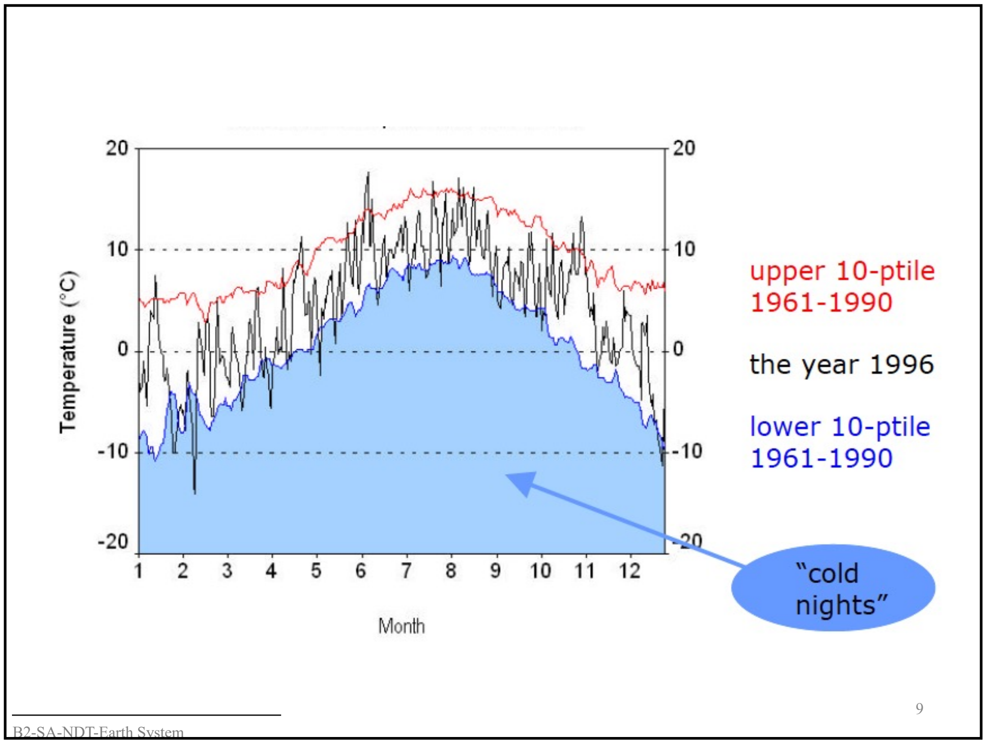
Exercise #2

The daily temperature values (ºC) for 14 consecutive days are: 19, 15, 20, 19, 21, 24, 23, 21, 28, 25, 28, 32, 29, 34

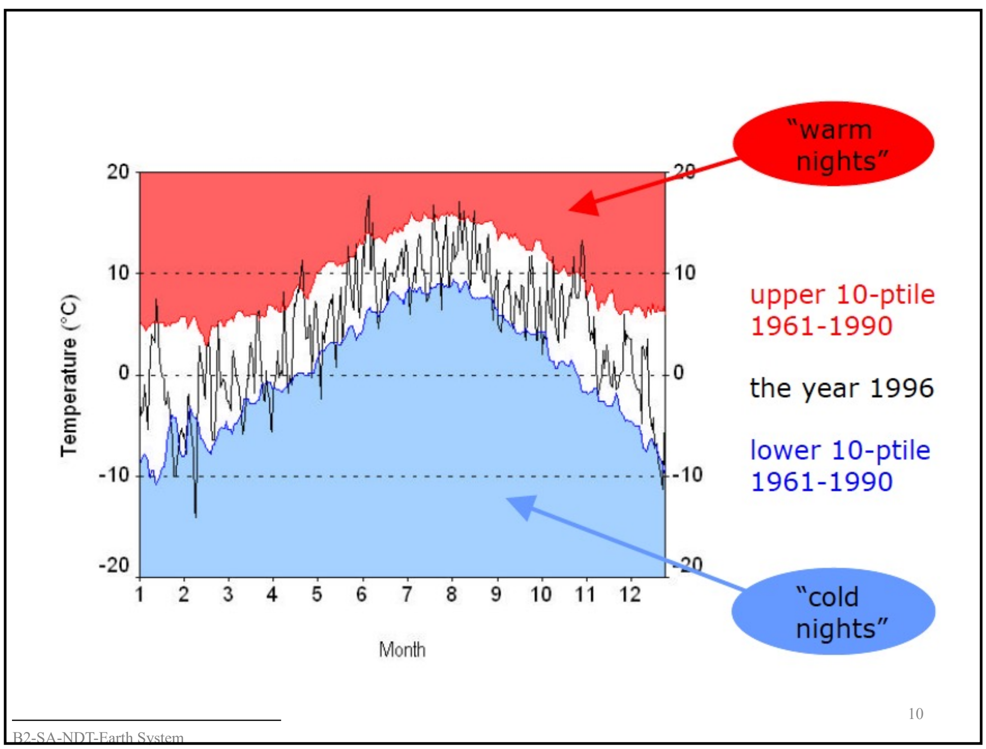
- Find the percentile rank for a temperature value of 21

Python: np.percentile(a, p)
a: array of values
p: percentile, eg. 10




20
10
0
-10
-20
Temperature (°C)
1 2 3 4 5 6 7 8 9 10 11 12
Month
upper 10-ptile
1961-1990
the year 1996
lower 10-ptile
1961-1990
"cold
nights"
B2-SA-NDT-Earth System
9

9


"warm
nights"
20
10
0
-10
-20
Temperature (°C)
1 2 3 4 5 6 7 8 9 10 11 12
Month
upper 10-ptile
1961-1990
the year 1996
lower 10-ptile
1961-1990
"cold
nights"
B2-SA-NDT-Earth System
10

10

The Climate Extremes Index (CEI)

- The CEI was first introduced in early 1996 (Karl et al. 1996)
- General lack of observational data on extremes
- Provided a means to communicate with policy makers regarding our understanding of changes in climate
- Originally looked at historical data (1+ years old) on an annual basis

Extreme Temperature Indices

- **Warm days** no of days with maximum temperature > 90th percentile
- **Warm nights** no of nights with minimum temperature > 90th percentile
- **Cool days** no of days with maximum T < 10th percentile
- **Cool nights** no of nights with minimum T < 10th percentile
- **Warm spell duration** Annual count of days with at least 4 consecutive days when daily maximum temperature > 90th percentile
- **Cold spell duration** Annual count of nights with at least 4 consecutive nights when daily minimum temperature < 10th percentile

Extreme Precipitation Indices

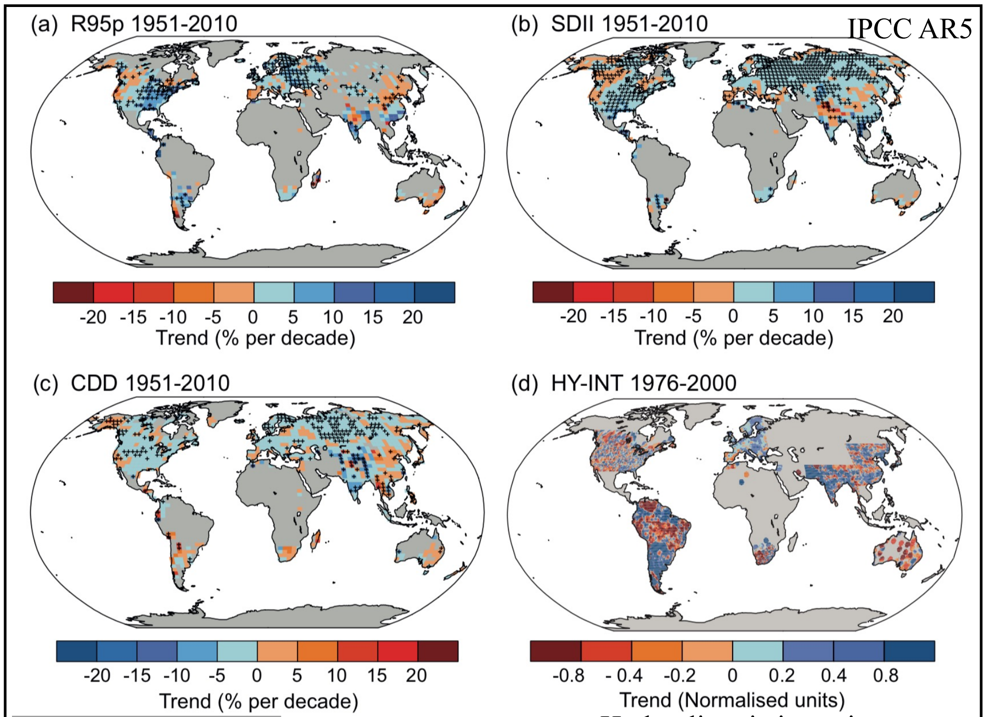
- **Very wet day precipitation** Annual total precipitation when daily precipitation > 95th percentile
- **Extremely wet day precipitation** Annual total precipitation when daily precipitation > 99th percentile

Other Extreme Temperature Indices

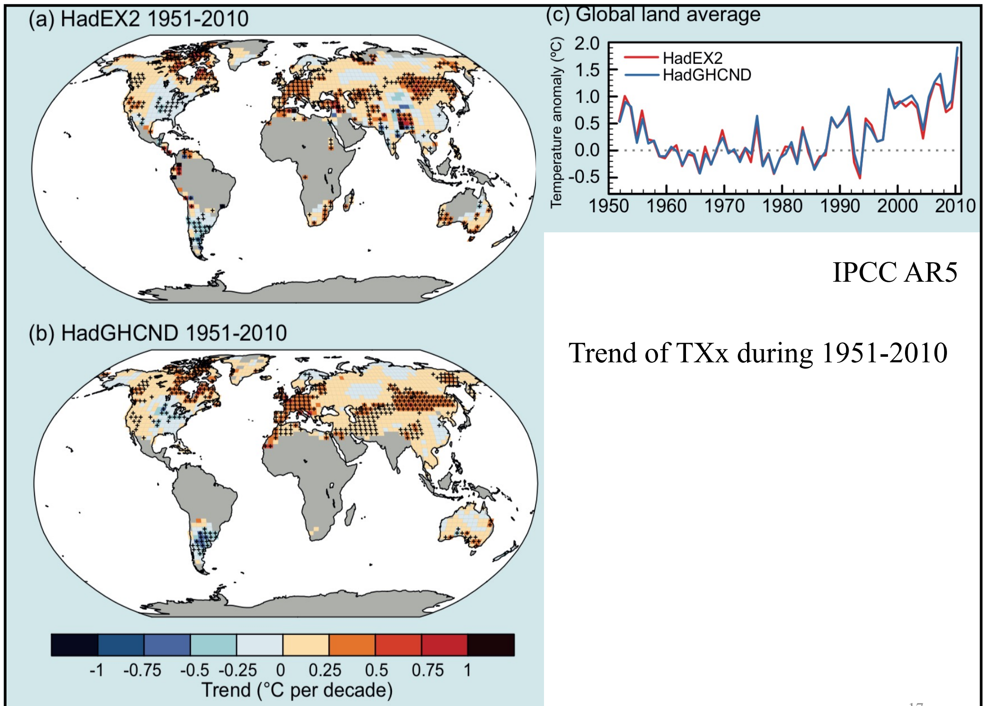
- **Very hot days** Annual count of days with maximum temperature > 40°
- **Hot days** Annual count of days with maximum temperature > 35° C
- **Very hot nights** Annual count of nights with minimum temperature > 25° C
- **Hot nights** Annual count of nights with minimum temperature > 20° C
- **Cold days** Annual count of days with maximum temperature < 15° C
- **Very cold days** Annual count of days with maximum temperature < 10° C
- **Cold nights** Annual count of nights with minimum temperature < 5° C
- **Frost nights** Annual count of nights with minimum temperature < 0° C
- **Highest maximum temperature** Annual maximum value of daily maximum temperature
- **Highest minimum temperature** Annual maximum value of daily minimum temperature
- **Lowest maximum temperature** Annual minimum value of daily maximum temperature
- **Lowest minimum temperature** Annual minimum value of daily minimum temperature

Extreme Precipitation Indices

- **Wet days**: Annual count of days with daily R ≥ 1 mm
- **Heavy precipitation days:**Annual count of days with daily R ≥ 10 mm
- **Very heavy precipitation days:** Annual count of days with daily R ≥ 30 mm
- **Maximum 1-day precipitation**: Annual maximum 1-day precipitation total
- **Maximum 5-day precipitation**: Annual maximum consecutive 5-day precipitation total
- **Annual total wet day precipitation**: Annual total precipitation on wet days (daily R ≥ 1 mm)
- **Simple daily intensity:** Annual total precipitation divided by the number of wet days (daily precipitation ≥ 1 mm)
- **Consecutive dry days**: Maximum number of consecutive days with daily R < 1 mm
- **Consecutive wet days**: Maximum number of consecutive days with daily R ≥ 1 mm



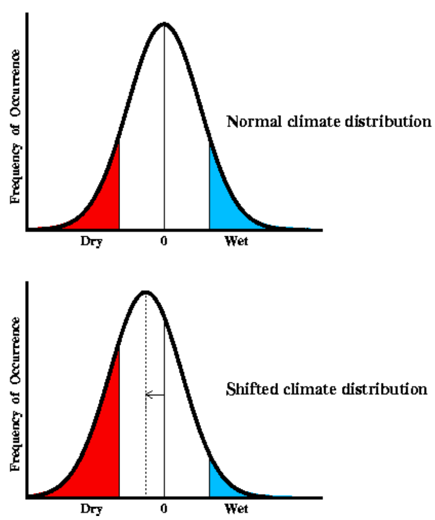
Hydroclimatic intensity



IPCC AR5

Trend of TXx during 1951-2010

Extreme Value Distributions



A "little" change of 1ºC in mean temperature→ a large change in extremes

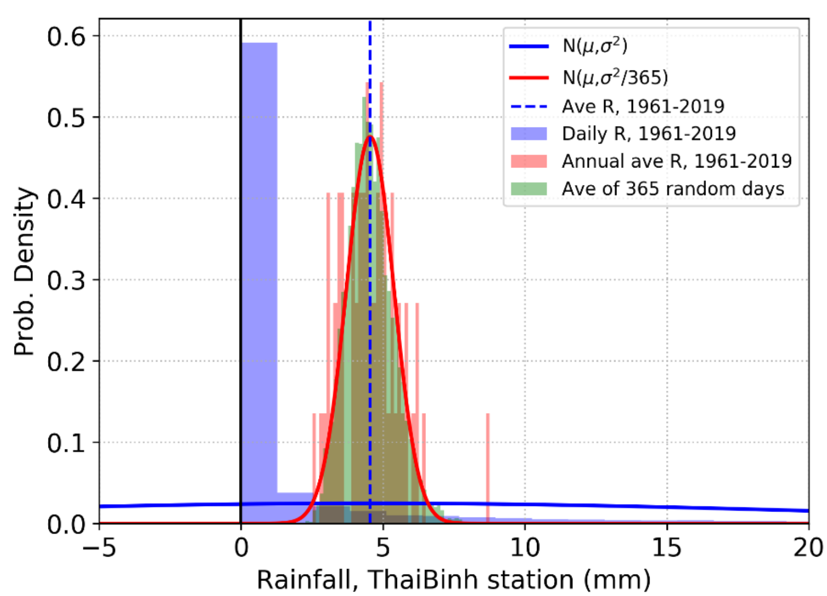
Central Limit Theorem (CLT)

- The distribution of sample means approximates a normal distribution as the sample size gets larger
- x_i are independent with mean μ and variance σ^2

$$a_n = \sum_{i=1}^{n} x_i$$

$a_n/n \rightarrow N(\mu, \sigma^2/n)$ with n large enough

Central Limit Theorem (CLT)



Extreme Value Distributions

- $x_1, x_2, \ldots, x_n$ random, independent variables with the same distribution.
- $x_1, x_2, \ldots, x_n$ → mean μ & variance σ^2
- Central Limit Theorem →

$$\left(\sum_{i=1}^{n} x_i - n\mu\right)/\sqrt{n}\,\sigma \quad \rightarrow \text{N}(0,1) \text{ when } n \rightarrow \infty$$

- **Question:** The extreme values of $x_1, x_2, \ldots, x_n$ has which distribution?

→ Extreme value theory
→ Block maxima and threshold maxima

Extreme Value Distributions

Block maxima & Threshold maxima:

- If X_n= max($x_1, x_2, \ldots, x_n$). Eg. X_n is annual TX_x or Rx1day → X_n: block maximum.
- If Z_n are rainfall values of days in a year of which the daily rainfall amount ≥ 20mm → Z_n is a series of threshold maxima
- What are the distributions of X_n and Z_n?

Extreme Value Distributions

Generalized Extreme Value (GEV) distribution

- The CDF of a block maxima z :

$$G(z) = exp\left[-\left\{1 + \xi(\frac{z-\mu}{\sigma})\right\}^{-1/\xi}\right]$$

G(z) has 3 parameters, including:

– location μ

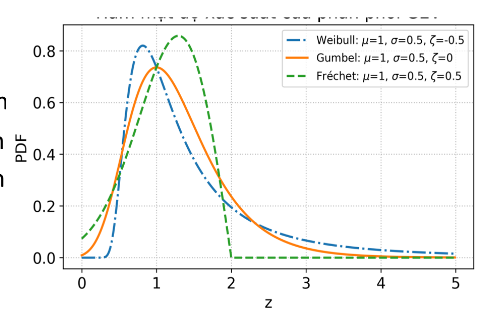
– scale σ

– shape ξ;

ξ=0 → Gumbel distribution

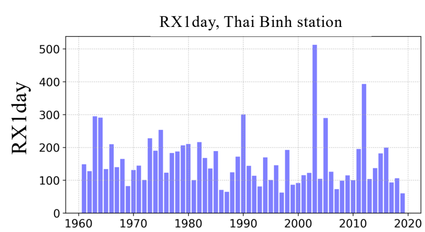
ξ>0 → Fréchet distribution

ξ<0 → Weibull distribution



Example

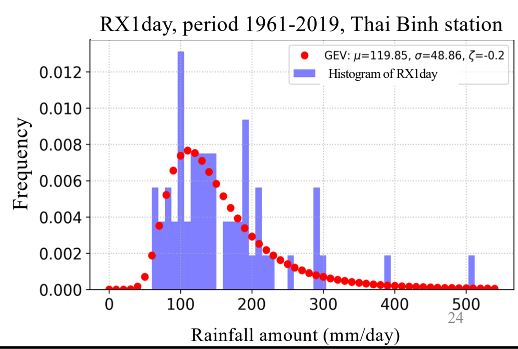
- Rx1day is the annual maximum daily rainfall, ThaiBinh station, period 1961–2019 → a series of 59 values



From 59 RX1day values → the GEV distribution with 3 parameters:

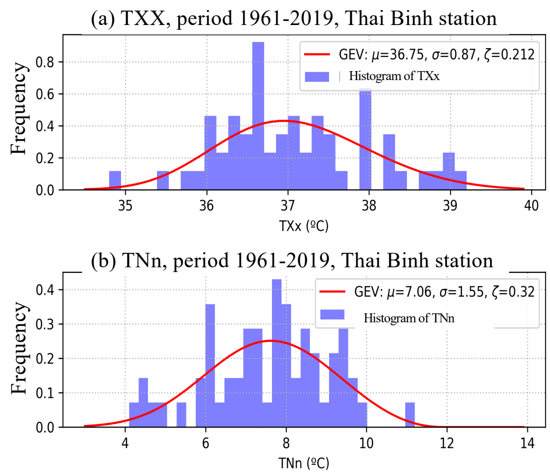
μ=119.85, σ=48.86, & ξ=-0.2

→ Rx1day → approximated by the Weibull distribution.



Example

- Distributions of TXx & TNn at ThaiBinh station



- TXx → GEV distribution with 3 parameters: $\mu=36.75, \sigma=0.87, \& \xi=0.212$
- TNn → GEV distribution with 3 parameters: $\mu=7.06, \sigma=1.55, \& \xi=0.32$

→ approximated by the Fréchet distribution

Extreme Value Distributions

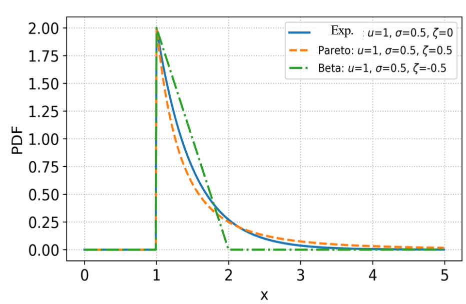
Generalized Pareto (GP) distribution

GP distribution is often applied for threshold maxima

$$H(x) = 1 - \left[1 + \xi\left(\frac{x-u}{\sigma}\right)\right]^{-1/\xi}$$

H(x) has 3 parameters:

- threshold μ
- scale σ
- shape ξ;
  - $\xi>0$ → Pareto distribution
  - $\xi<0$ → Beta distribution
  - $\xi \to 0$ → exponential distribution

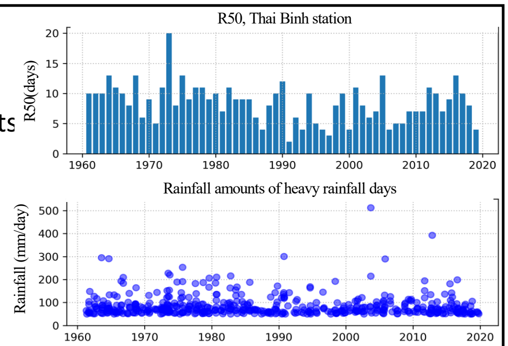


80/20 rule (or Pareto principle): in many events, 80% of outcomes result from 20% of all causes

Example

- ? Distribution of the rainfall amounts of heavy rainfall days, i.e. days with daily rainfall ≥ 50mm, ThaiBinh station, period 1961–2019

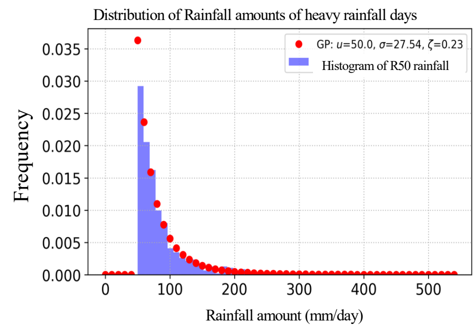
→ 500 days with heavy rainfall



- Rainfall values of 500 heavy rainfall days → GP distribution with 3 parameters:

  $\mu=50$, $\sigma=27.54$, & $\xi=0.23$

  $\xi>0$ → Pareto distribution.



Extreme Value Distributions

Return level

- From the distribution → can calculate the return level of an event
- x_T is the value that is expected to be repeated on average one time per one period T (with a probability $1/T$)
- T → return period

- If x is block maxima, x_T satisfies:

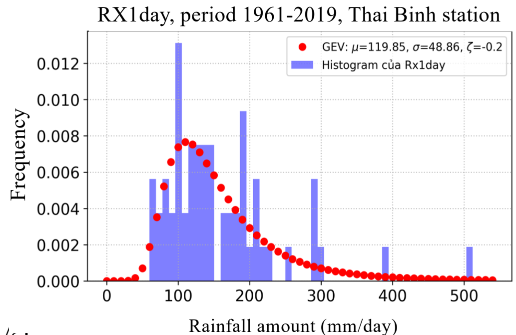
$$G(x_T) = 1 - 1/T$$

- If x is threshold maxima, x_T satisfies:

$$H(x_T) = 1 - 1/T$$

Example

- GEV Weibull distribution for Rx1day, ThaiBinh station, with 3 parameters: μ=119,85, σ=48,86, & ξ=-0,2.
- Calculate the return level of Rx1day fc each period 2 years, 5 years, 10 years, 20 years!?



→ Solve the equation:

$$G(z) = exp\left[-\left\{1+\xi\left(\frac{z-\mu}{\sigma}\right)\right\}^{-1/\varsigma}\right] = 1 - 1/T$$

for T=2, 5, 10, 20 years

→ Results: the return levels are:

- ✓ 138.4 mm/day for 2-yr return period
- ✓ 205.3 mm/day for 5-yr return period
- ✓ 258.7 mm/day for 10-yr return period
- ✓ 318.0 mm/day for 20-yr return period

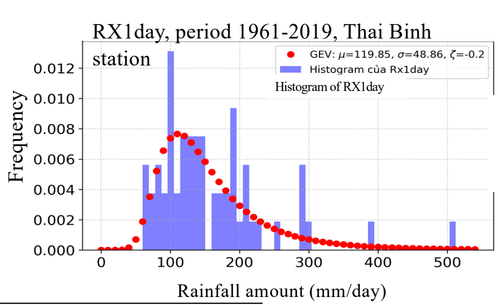
B2-SA-NDT-Earth System 29

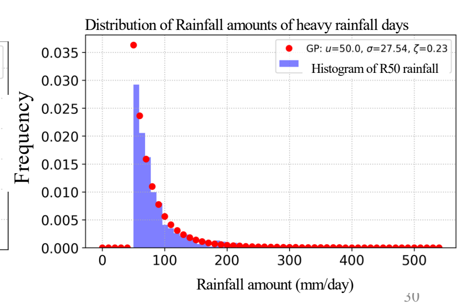
29

Practice Python #10

Given the VietTri rainfall data (see Practice Python #9)

1. Estimate RX1day for each year
2. Plot the distribution of RX1day
3. Calculate the return value of RX1day for the return period 5 yr, 10yr, 20 yr, 100 yr
4. **Do similarly for the rainfall amounts of the heavy rainfall days (days with rainfall ≥50mm)**





B2-SA-NDT-Earth System 30

30

GEV

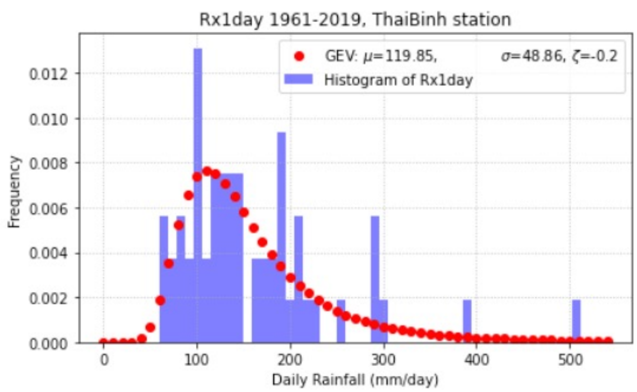
from scipy.stats import genextreme as gev

Determine the parameters of the GEV distribution
shape, loc, scale = gev.fit(Rx1day) # estimate the 3 parameters of the GEV
# Estimate the return values corresponding to the return periods 5 years
x5=gev.ppf((1-1/5.),shape,loc,scale)


```python
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import genextreme as gev
r2rain=np.loadtxt("R_THAIBINH_1961_2019.txt",comments="#")
nyr=int(len(r2rain[:,1])/31)     # number of years
# reformat to 3D: r3T2m[year,dayofmonth,month]
r3rain=np.reshape(r2rain[:,1:13],(nyr,31,12))
# Calculate Rx1day for each year
Rx1day=np.zeros(nyr)
iyr=0
while iyr < nyr:
    imn=0
    while imn <12:
        idy=0
        while idy<31:
            if r3rain[iyr,idy,imn]>= Rx1day[iyr]:
                Rx1day[iyr]=r3rain[iyr,idy,imn]
            idy=idy+1
        imn=imn+1
    iyr=iyr+1
# Determine the parameters of the GEV distribution
shape, loc, scale = gev.fit(Rx1day) # estimate the 3 parameters of the GEV
x = np.arange(0,550,10)
y = gev.pdf(x, shape, loc, scale)
# Estimate the return values corresponding to the return periods 5, 10, 20, 100 yrs
x5=gev.ppf((1-1/5.),shape,loc,scale)
x10=gev.ppf((1-1/10.),shape,loc,scale)
x20=gev.ppf((1-1/20.),shape,loc,scale)
x100=gev.ppf((1-1/100.),shape,loc,scale)
print(x5,x10,x20,x100)
```

```
## Plot
fig, ax1 = plt.subplots(figsize=(7,4))
ax1.hist(Rx1day,bins=50,density=True, histtype='stepfilled',color='blue',\
         alpha=0.5,label='Histogram of Rx1day')
plt.plot(x,y,'ro',label='GEV: $\mu$='+str(round(loc,2))+',\
         $\sigma$='+str(round(scale,2))+', $\zeta$='+str(round(shape,3)))
plt.legend(loc='upper right')
ax1.set(title='Rx1day 1961-2019, ThaiBinh station',\
        xlabel='Daily Rainfall (mm/day)',ylabel='Frequency')
```

```
205.2976342634028 258.67226113738667 317.96364801694955 488.3353055305761
```



33