Data Rich Hydrology, Perugia – 28 January, 1 February 2019

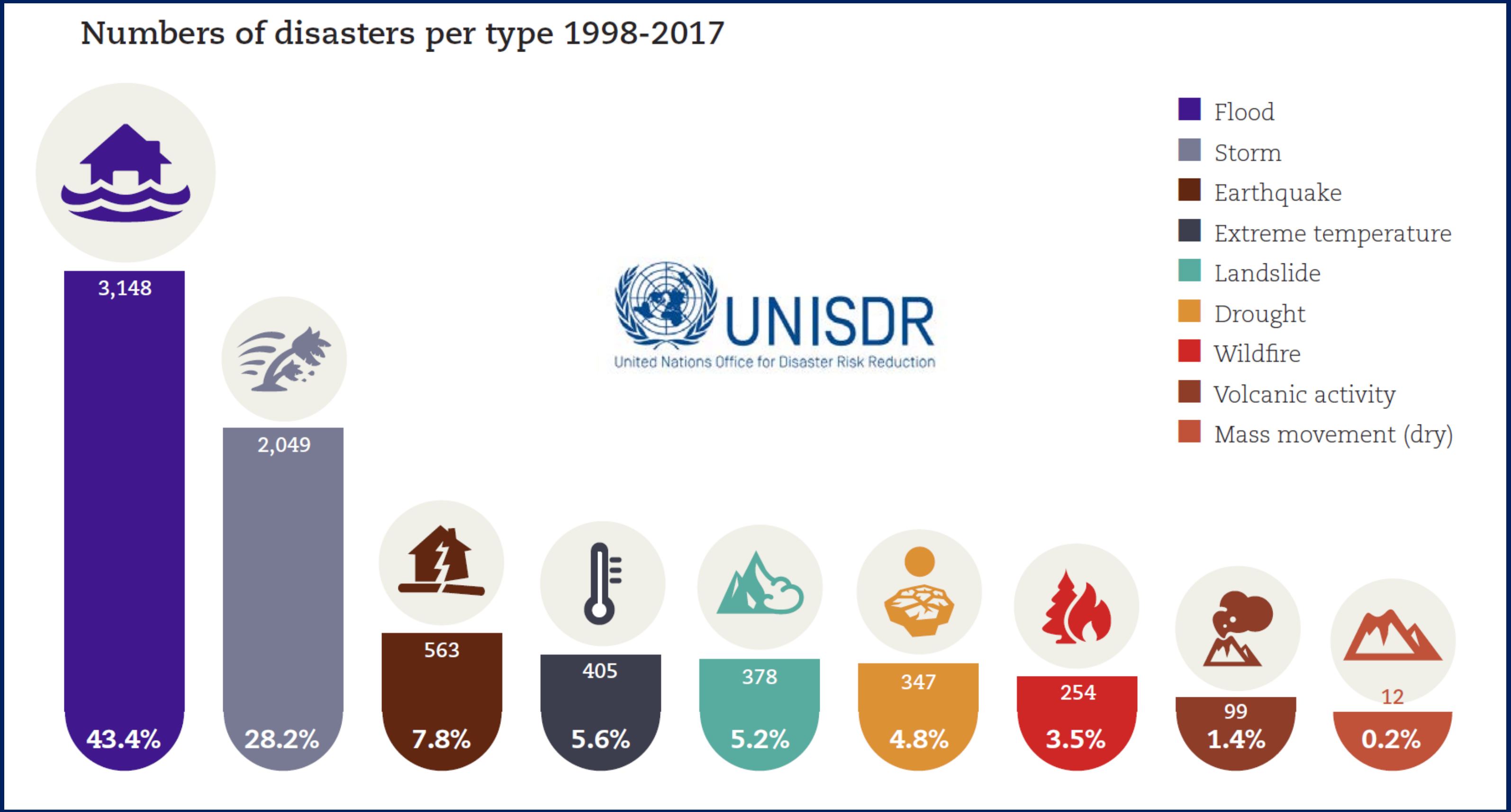
Beyond traditional extreme value theory: lessons learned from rainfall and hurricane intensity

Marco Marani[1,2], Enrico Zorzetto[2], Arianna Miniussi[1],
Seyed Reza Hosseini[3], Marco Scaioni[3], Gabriele Villarini[4]
[1]Universita' di Padova, [2]Duke University, [3]Polytechnic Milan, [4]University of Iowa

CINID CONSORZIO INTERUNIVERSITARIO PER L'IDROLOGIA · UNESCO United Nations Educational, Scientific and Cultural Organization · World Water Assessment Programme · WAREDOC · Università per Stranieri di Perugia

A little motivation. Extremes of all types…

Extreme events globally



- 91% of disasters worldwide caused by floods, storms, droughts, heatwaves, and other extreme weather events;
- Extreme storms/floods affected ~3 billion people in 1998-2017, causing 69% of disaster-related economic losses (~ $2 \cdot 10^{12}$ US$).

United Nations Office for Disaster Risk Reduction, Economic 1998-2017 Losses, Poverty & Disasters, 2018

Some motivation

Given a stochastic quantity x,

Expected lifetime (V_N)
Level of Risk accepted (R)

↓

Return Time (T_r)

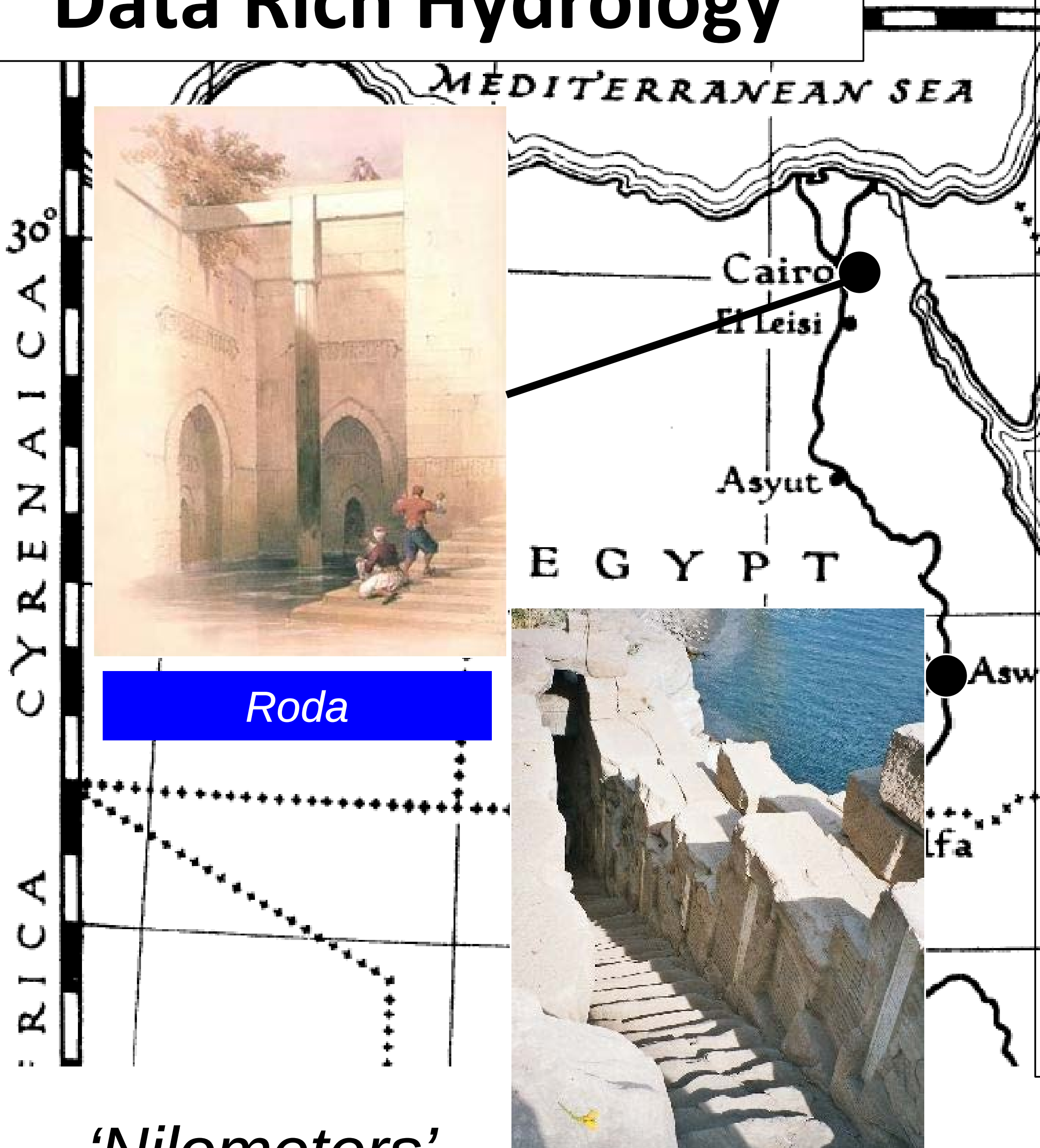
↓

Estimated quantile $\hat{x}(T_r)$

↓

-Design of related structures
-Optimal water planning

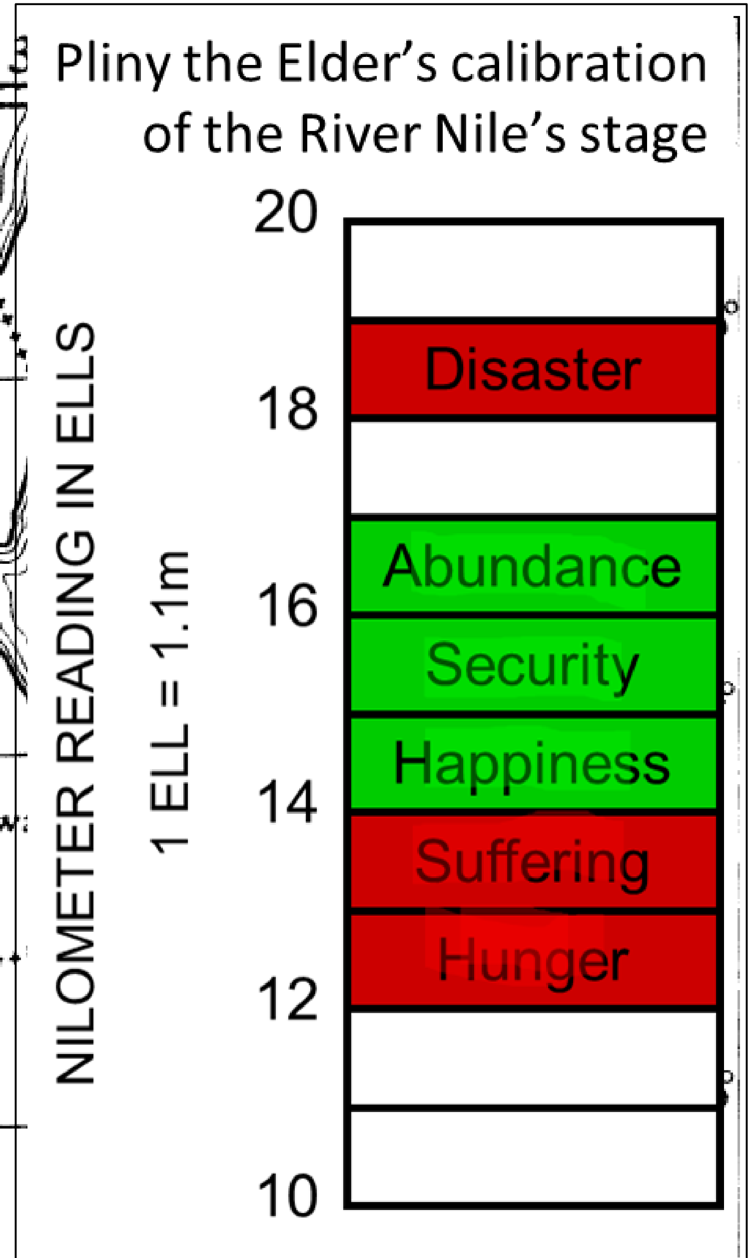
Data Rich Hydrology



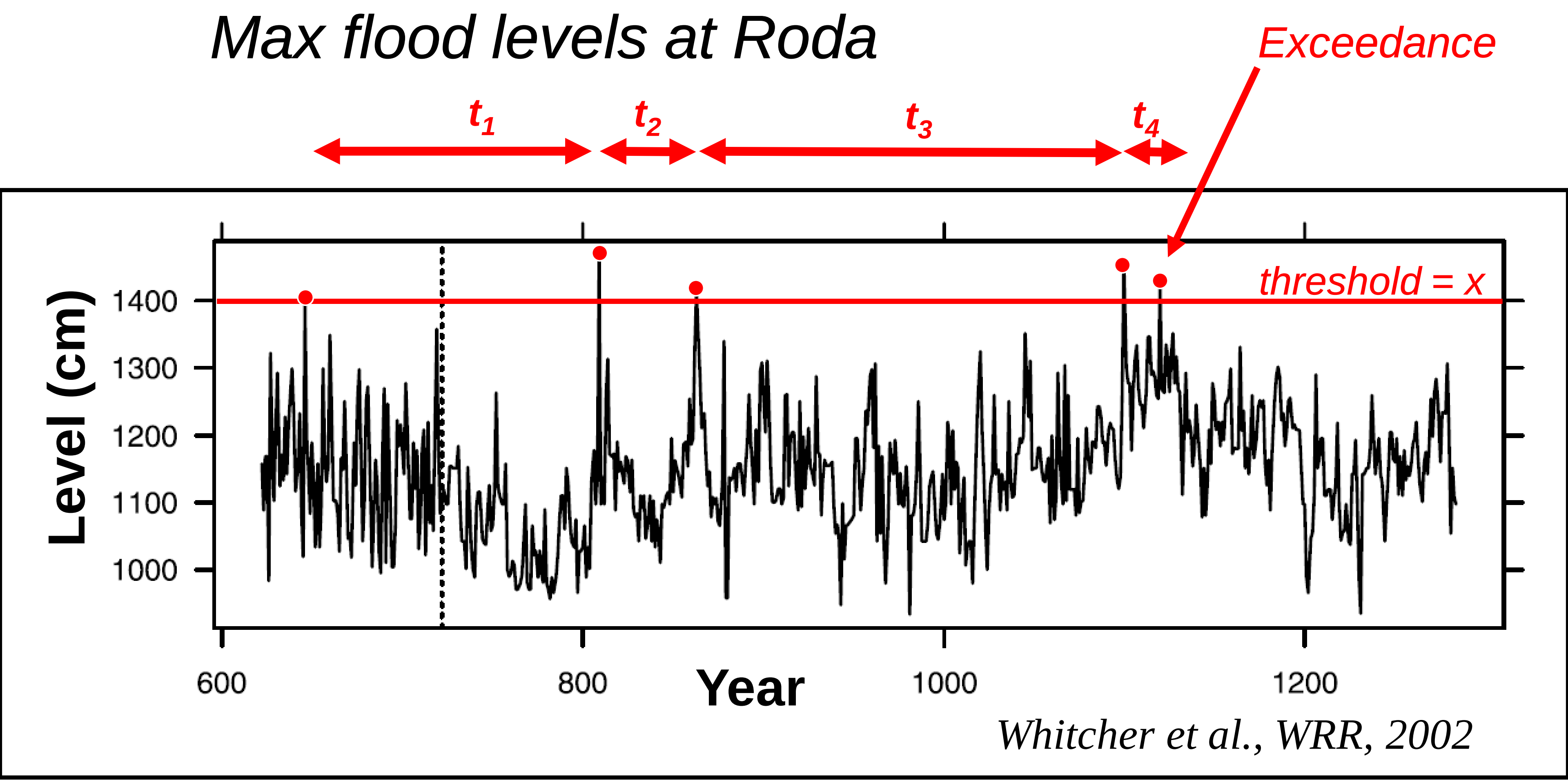
Roda

Elephantine island

‘Nilometers’

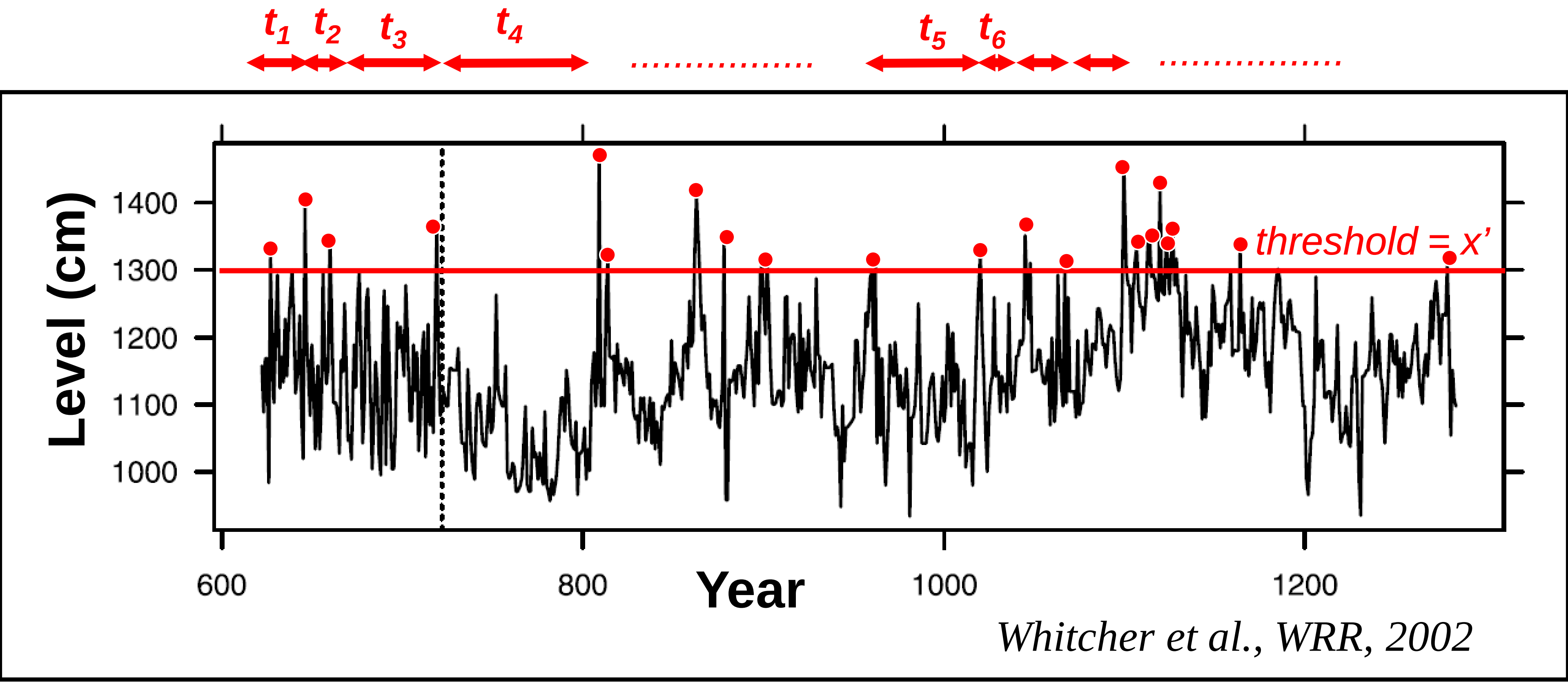


Toussoun, 1922
Hurst, 1927



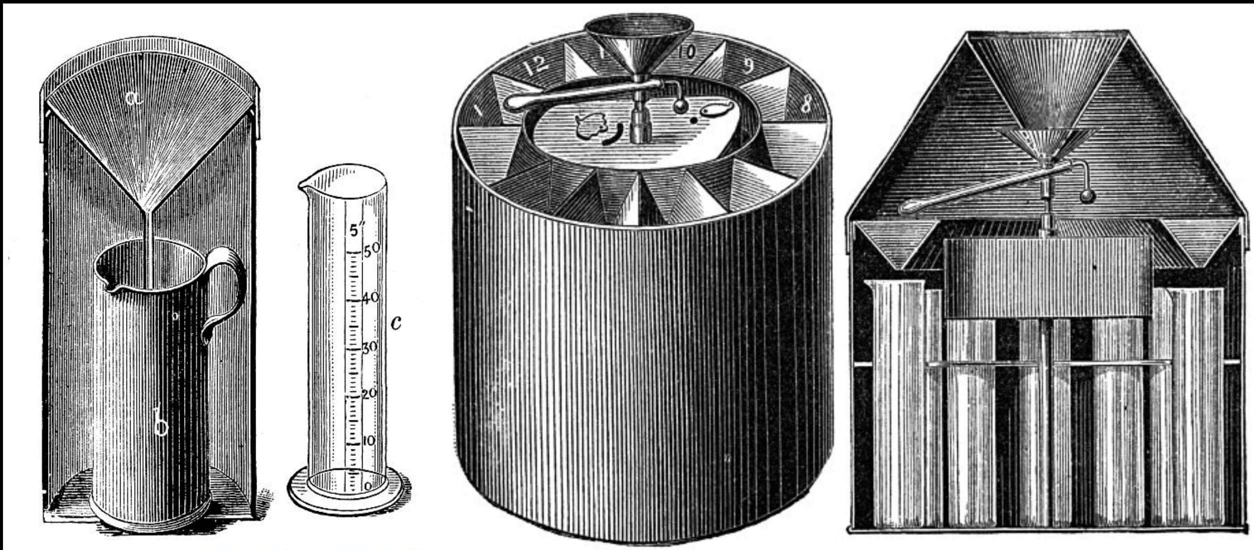
Average recurrence interval or return time: Tr(x)= ***average*** *time intervals between two successive exceedances of x*

Max floods levels at Roda



Because x'<x, then Tr(x') < Tr(x)

An important application: precipitation.



Data Rich Hydrology:

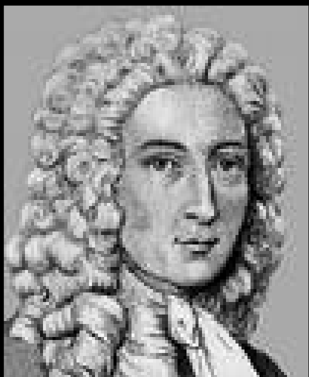
First obs.: India IV century BC

In Europe: Benedetto Castelli (1578–1643)

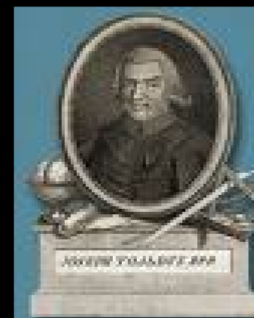
Sir Christopher Wren (1632–1723)

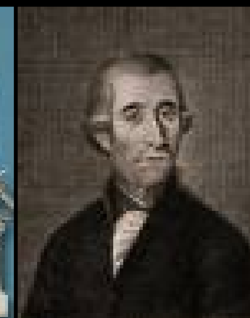
Giovanni Poleni (Venice 1683- Padova 1761)

Data Rich Hydrology: precipitation in Padova (1725- today)

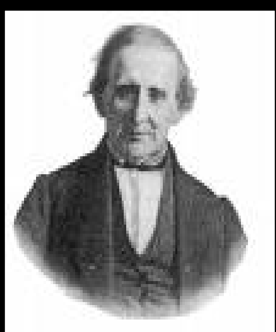


Giovanni e Francesco Poleni
via Beato Pellegrino
1725-1764





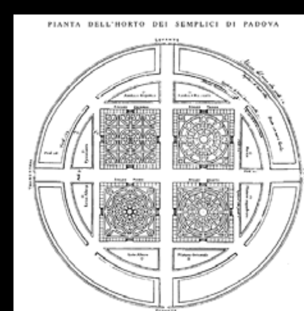
Giuseppe Toaldo
e
Vincenzo Chiminello:
Specola 1768-1813



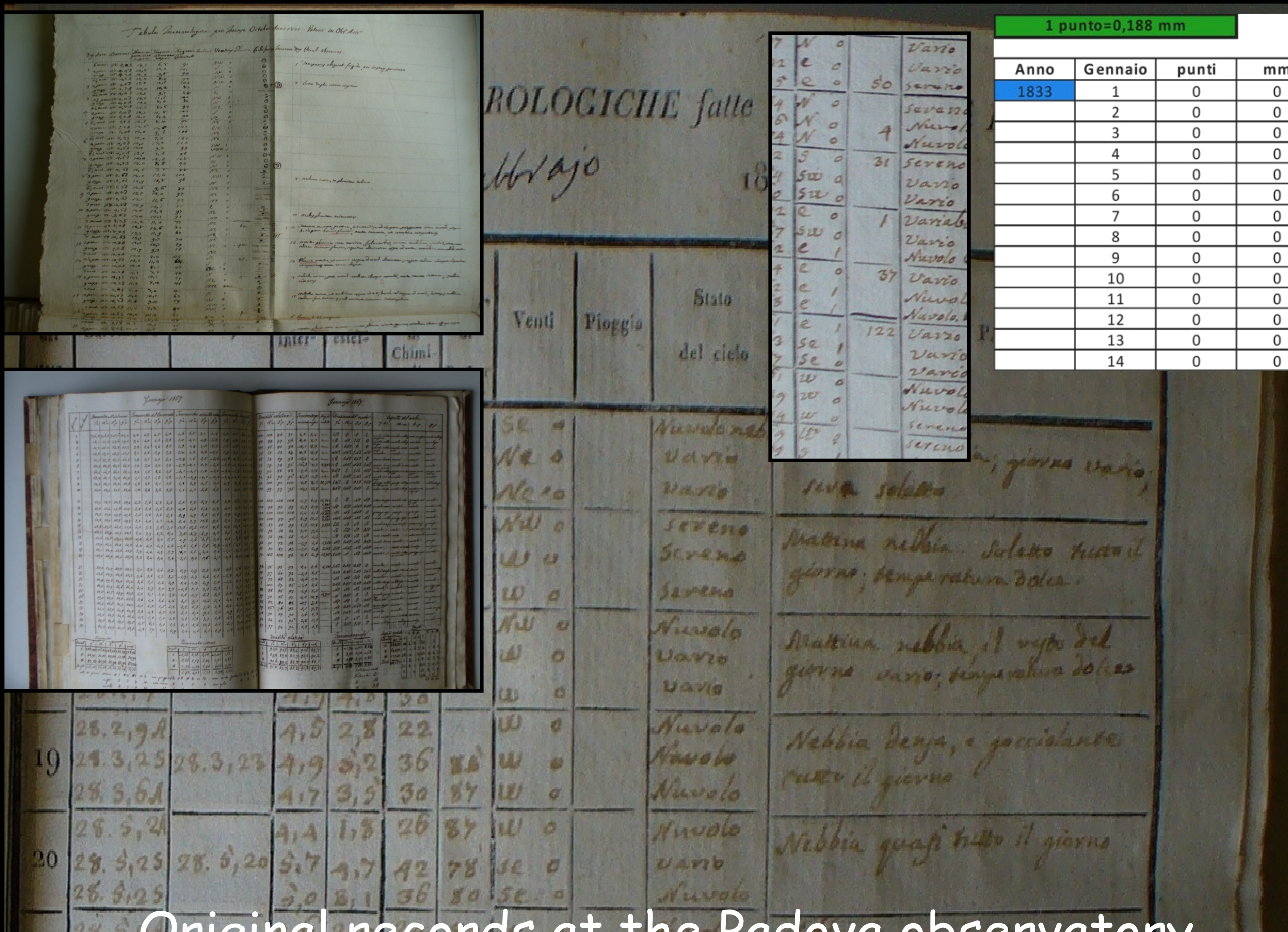
Giovanni Santini:
Specola 1823-1877

Giuseppe Lorenzoni:
Specola 1878-1934

Magistrato alle Acque di
Venezia: Osservatorio
G. Magrini
1936-1996



ARPAV
Orto Botanico:
1997-presente

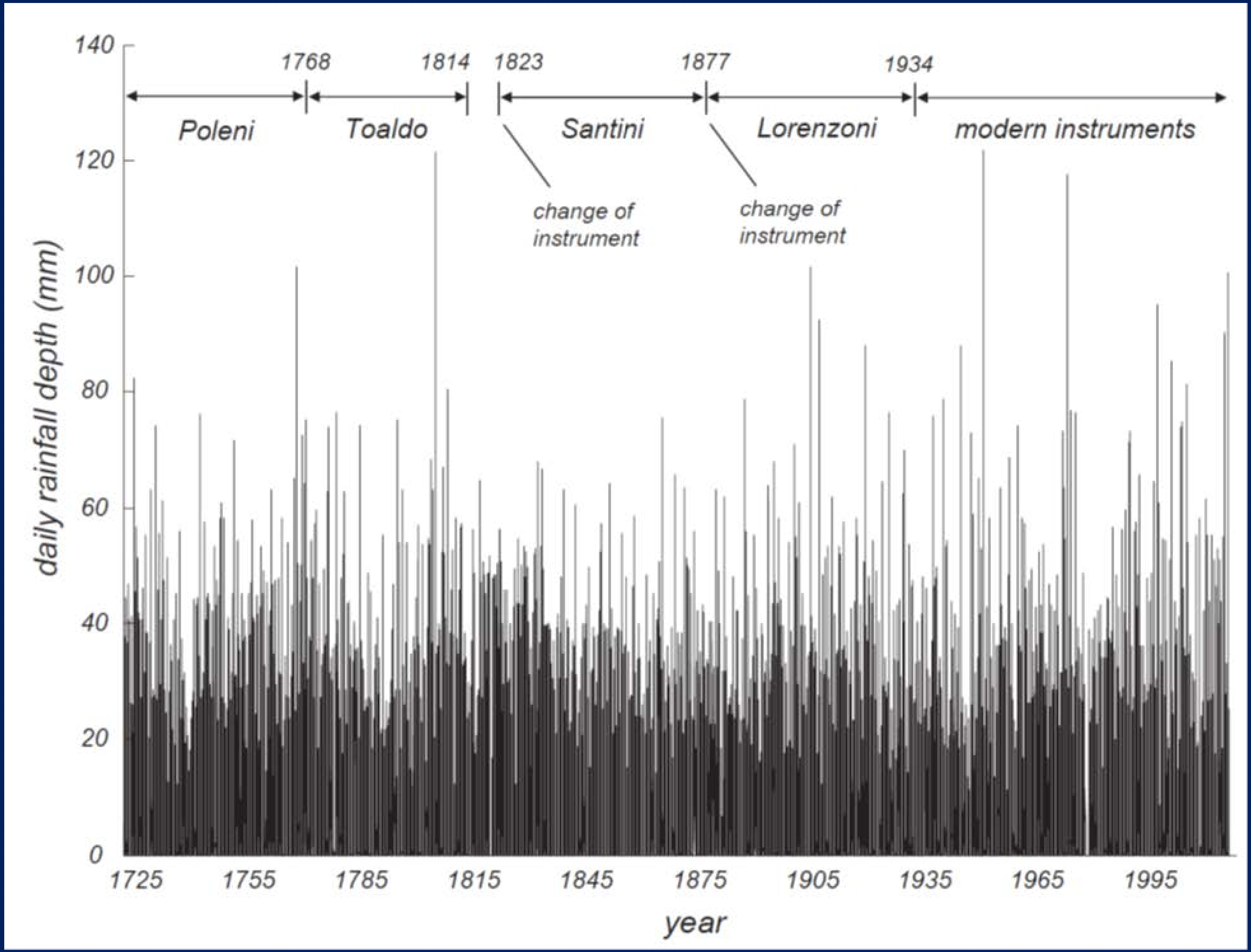


1 punto=0,188 mm

Anno	Gennaio	punti	mm
1833	1	0	0
	2	0	0
	3	0	0
	4	0	0
	5	0	0
	6	0	0
	7	0	0
	8	0	0
	9	0	0
	10	0	0
	11	0	0
	12	0	0
	13	0	0
	14	0	0

Original records at the Padova observatory

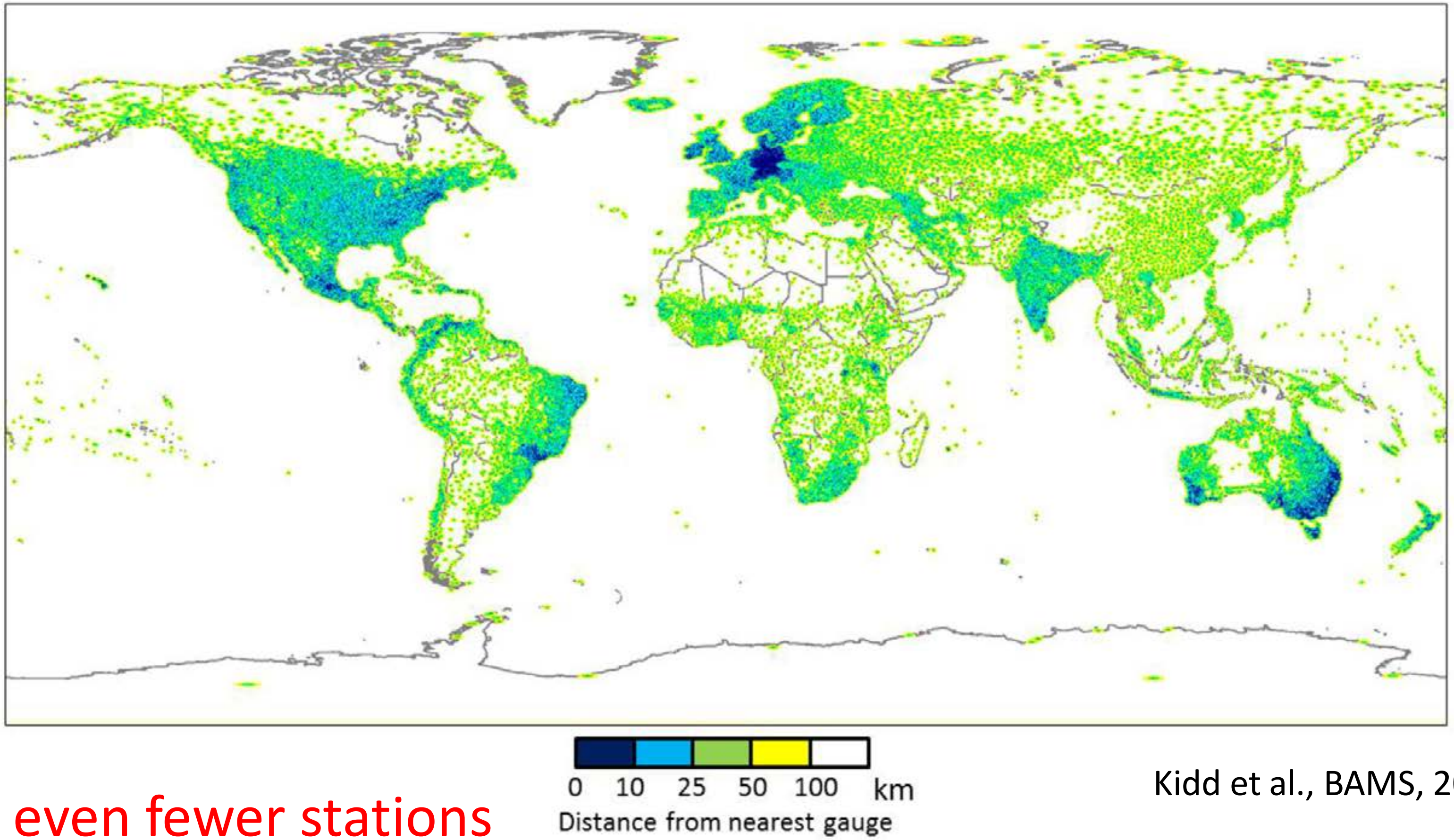
The Padova daily precipitation time series 1725-2006



(Marani and Zanetti, 2015)

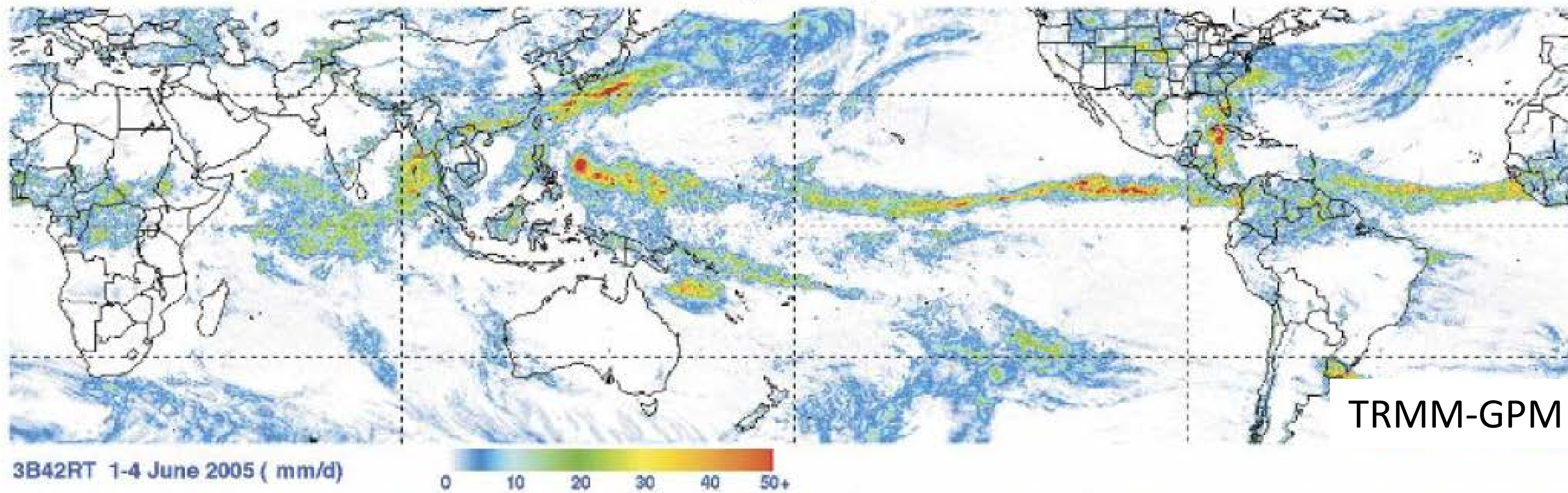
Data Rich Hydrology?

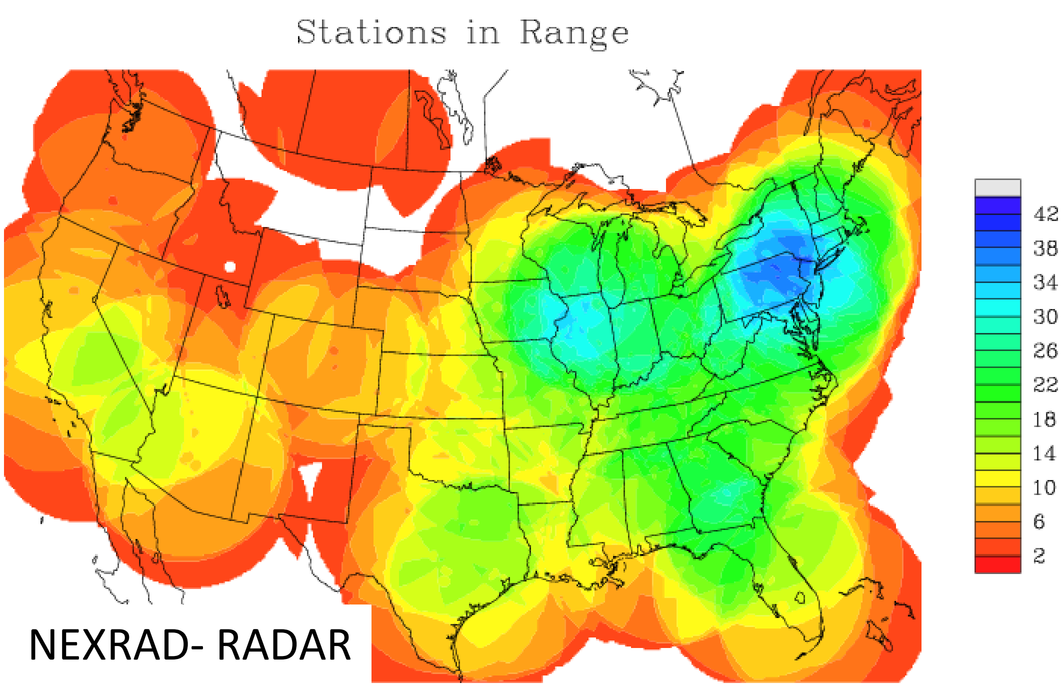
Global Rainfall Extremes ? Large gaps in observational networks



Kidd et al., BAMS, 2016

And even fewer stations are long enough for extreme rainfall analysis





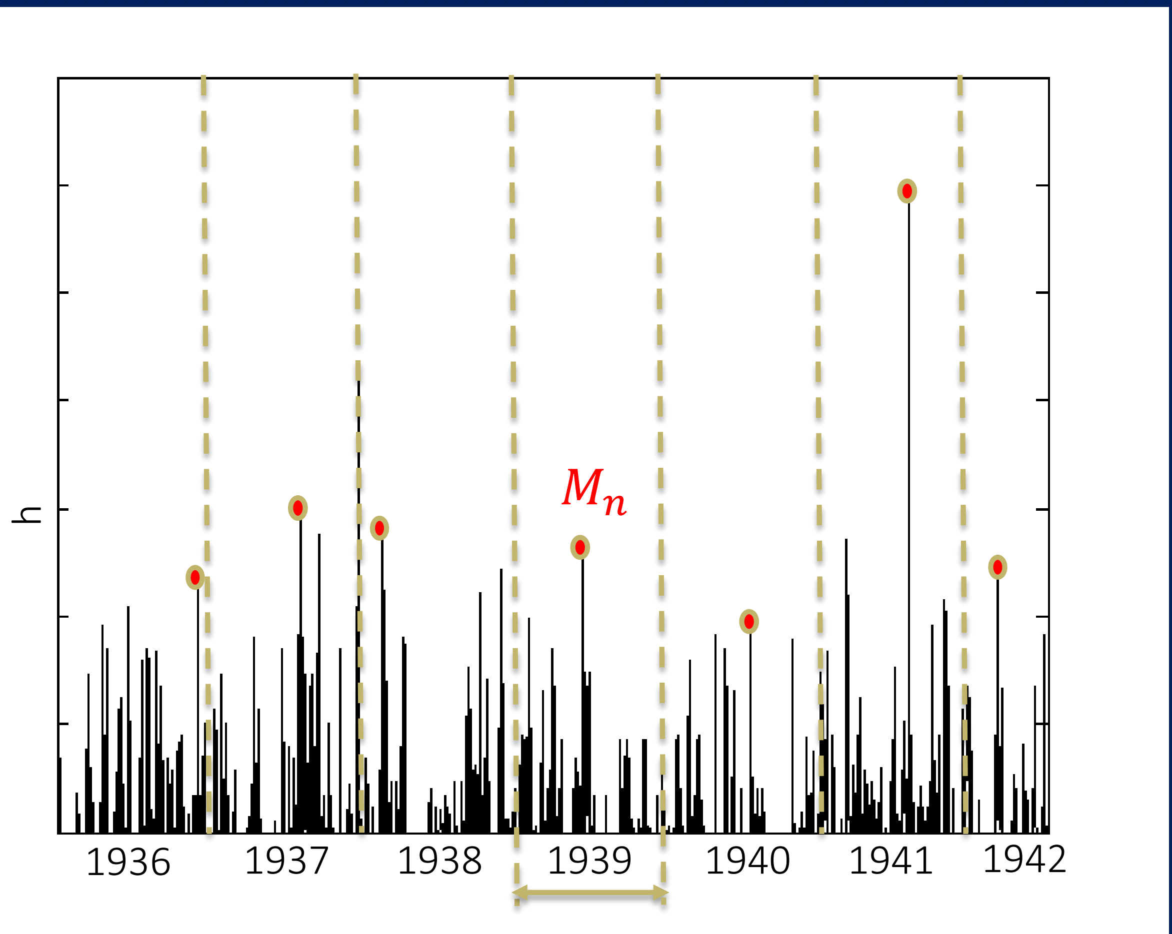
Remote sensing estimates of rainfall

- Continental to quasi-global coverage;
- Short observational period (max < 20 yrs);
- Testing vs point obs ?
- Downscaling ?

Classical Extreme Value Theory (EVT)

[Fischer-Tippett-Gnedenko, 1928-1943]

$x_n = \max_n(x_i)$ $H_n(x) = F(x)^n$ for independent, identically distributed x_i



Often block size = one year

'Block Maxima':

M_n= Maximum value occurred within a n-event block

What is the distribution of M_n?

Three-Type Theorem:

-As $n \longrightarrow \infty$

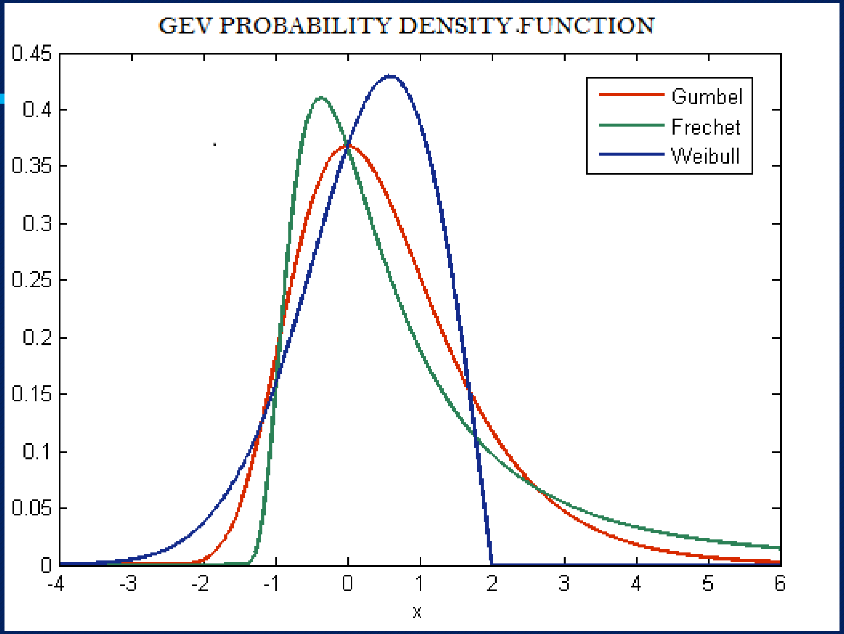
-After proper renormalization, are only three 'types' of asymptotic distribution for M_n:

- *Gumbel*
- *Frechet*
- *Weibull*

The GEV distribution

[Von Mises, 1936]

$$H(x) = \exp\left\{-\left[1 + \frac{\xi}{\psi}(x - \mu)\right]_+^{-\frac{1}{\xi}}\right\}$$



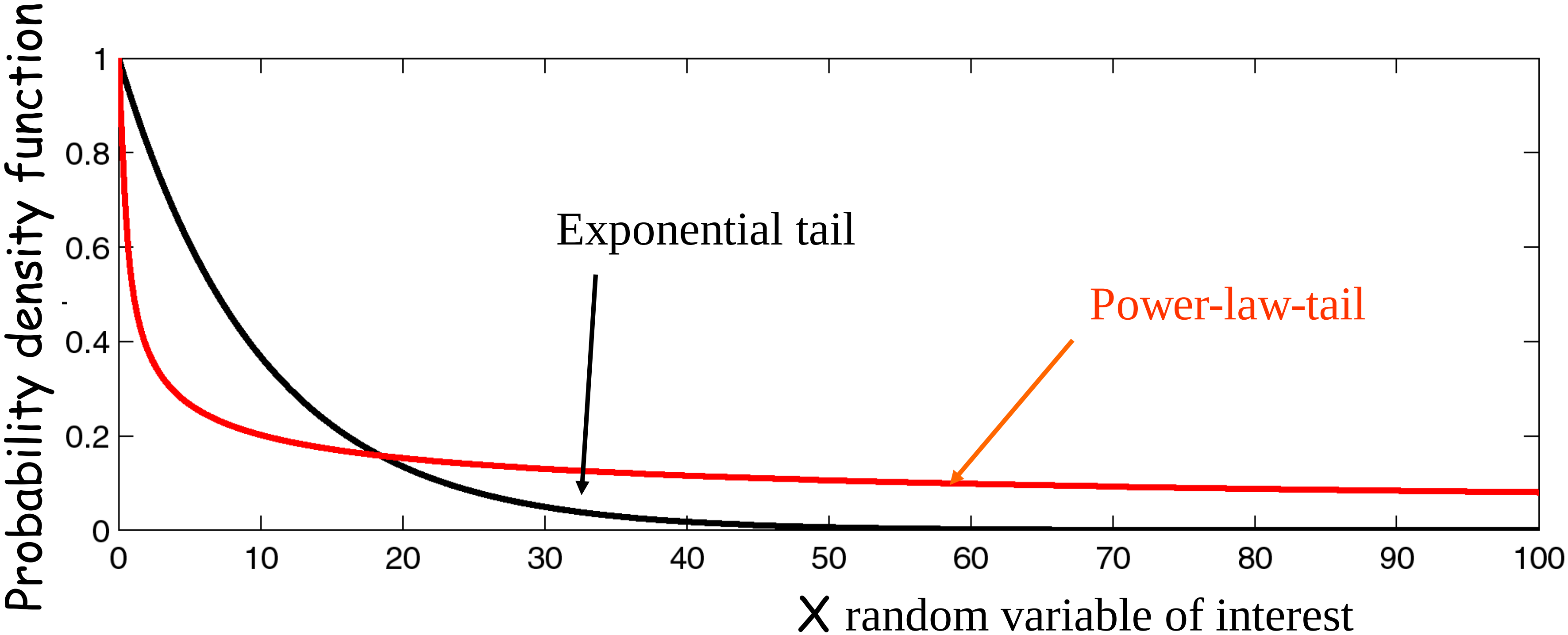
Shape Parameter:

- *ξ>0 Frechet, 'heavy' tailed*
- *$\xi \rightarrow 0$ Gumbel, exponential tail;*
- *$\xi < 0$ Inverse Weibull, upper bounded.*

Quantiles corresponding to return periods of interest can be retrieved:

$$F_r = 1 - \frac{1}{T_r} \qquad \hat{h} = H^{-1}(F_r)$$

Exponential and Power-law decays



Two radically different types of extremes!

Peak Over Threshold Method (POT)

[Balkema, De Haan & Pickand, 1975; Davison and Smith, 1990]

For a fixed threshold q $\rightarrow$ Exceedances $Y_i = H_i - q \qquad i.i.d. \quad r.v.$

$$P(Y_{max} < \mathrm{x}) = \sum_{n=1}^{\infty} p(n) \cdot F(x)^n = \sum_{n=1}^{\infty} \frac{\lambda^n e^{-n}}{n!} \cdot \left\{ 1 - \left[1 + \frac{\xi}{\psi} \cdot (x - q) \right]^{-1/\xi} \right\}^n$$

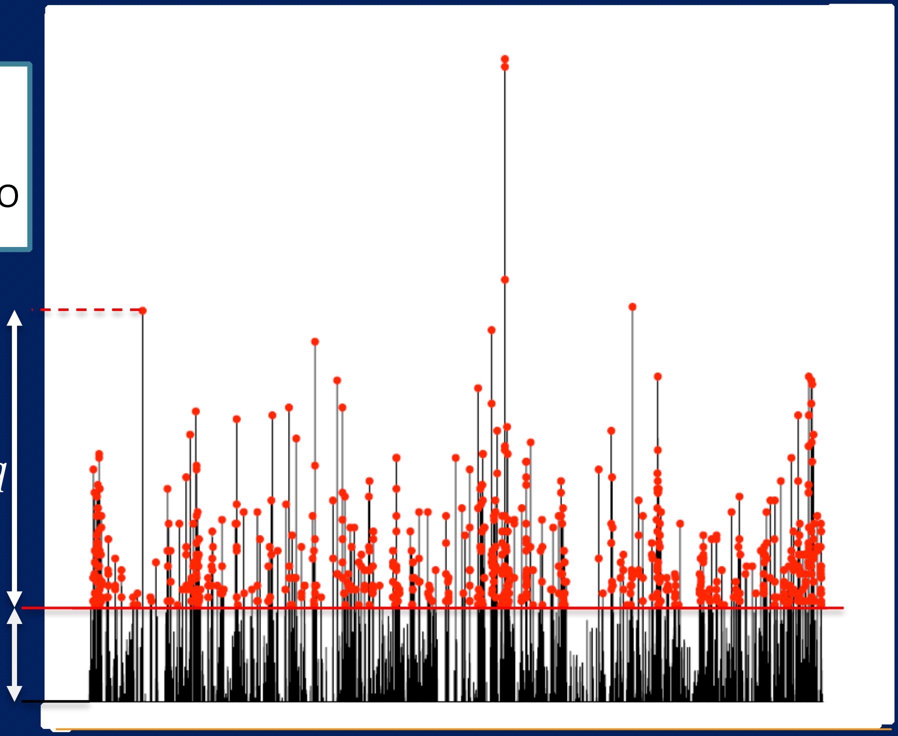
- Exceedances arrivals → Poisson
- Distribution of excesses → Generalized Pareto

Advantages:

1. Better description of the 'tail'
2. Consistent with GEV

$y_i = h_i - q$

q



Considerations on the application of the classical EVT

- Uncertain convergence of actual distribution to limiting one (some suggest observed Frechet is a result of incomplete convergence to Gumbel, e.g. Koutsoyiannis, 2013; Serinaldi and Kilsby, 2014).

- Fitting of GEV using Maximum Likelihood only uses yearly maxima and neglects most of the data. Important in the presence of short observations.

- Use of POT requires identification of threshold, uses more data but still a fraction of all available information.

- When number of events is small (dry climates, hurricanes, ...), yearly maxima also co me from bulk of distribution, not just the tail (we are far from a limiting form)

A Metastatistical Extreme Value Distribution (MEVD)

The Block-maxima distribution

$$x_n = \max_n(x_i)$$

$$H_n(x) = F(x)^n$$

for independent, identically distributed $X_i's$

The compound block-maxima distribution accounting for stochastic n and parameters of "ordinary" events pdf

$$\zeta(x) = \sum_{n=1}^{\infty} \int_{\Omega_{\vec{\Theta}}} F(x;\vec{\theta})^n \; g(n,\vec{\theta}) d\vec{\theta}$$

G(n,θ) = joint prob distrib. of the parameters.

Marani and Ignaccolo, 2015; Zorzetto et al., 2016

A Metastatistical Extreme Value distribution (MEV)

The MEV block-maxima distribution:

$$\zeta(x) = \sum_{n=1}^{\infty} \int_{\Omega_{\vec{\Theta}}} F(x;\vec{\theta})^n \, g(n,\vec{\theta}) d\vec{\theta}$$

G(n,θ) = joint prob distrib. of the parameters.

Can be approximated using sample averages:

$$\zeta(x) \cong \frac{1}{T} \sum_{j=1}^{T} F(x;\overrightarrow{\theta_j})^{n_j}$$

T = # sub-periods over which n and θ are estimated

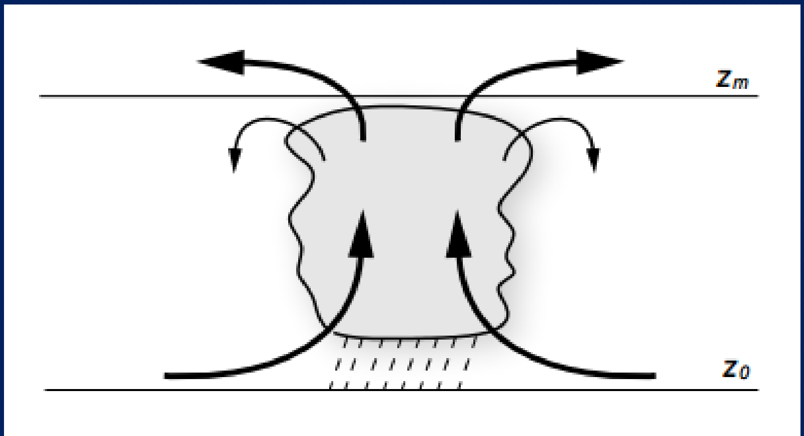
Marani and Ignaccolo, 2015; Zorzetto et al., 2016

Rainfall: what F(x) - the pdf of daily «ordinary» rainfall?

[Wilson e Tuomi, 2005]

Assumptions

- Heavy precipitation events dominated by moisture advection
- Negligible contribution from local evaporation



-Simple two-layers atmospheric model
-Temporal average

$\bar{k}$=*precipitation efficiency*

$\bar{q}$=*specific humidity*

m=advection mass

$$R_{acc} = \bar{k}\bar{q}m$$

$$F(x) = 1 - e^{\left(\frac{x}{c}\right)^{w}}$$

Weibull Parent distribution

Weibull ∈ domain of attraction of Gumbel, but convergence is very slow! →

It would need a disproportionally large number of events N for Gumbel to hold. Frechet is improperly used instead.

MEV-Weibull distribution

The MEV expression:

$$\zeta(x) \cong \frac{1}{T}\sum_{j=1}^{T} F(x;\overrightarrow{\theta_j})^{n_j}$$

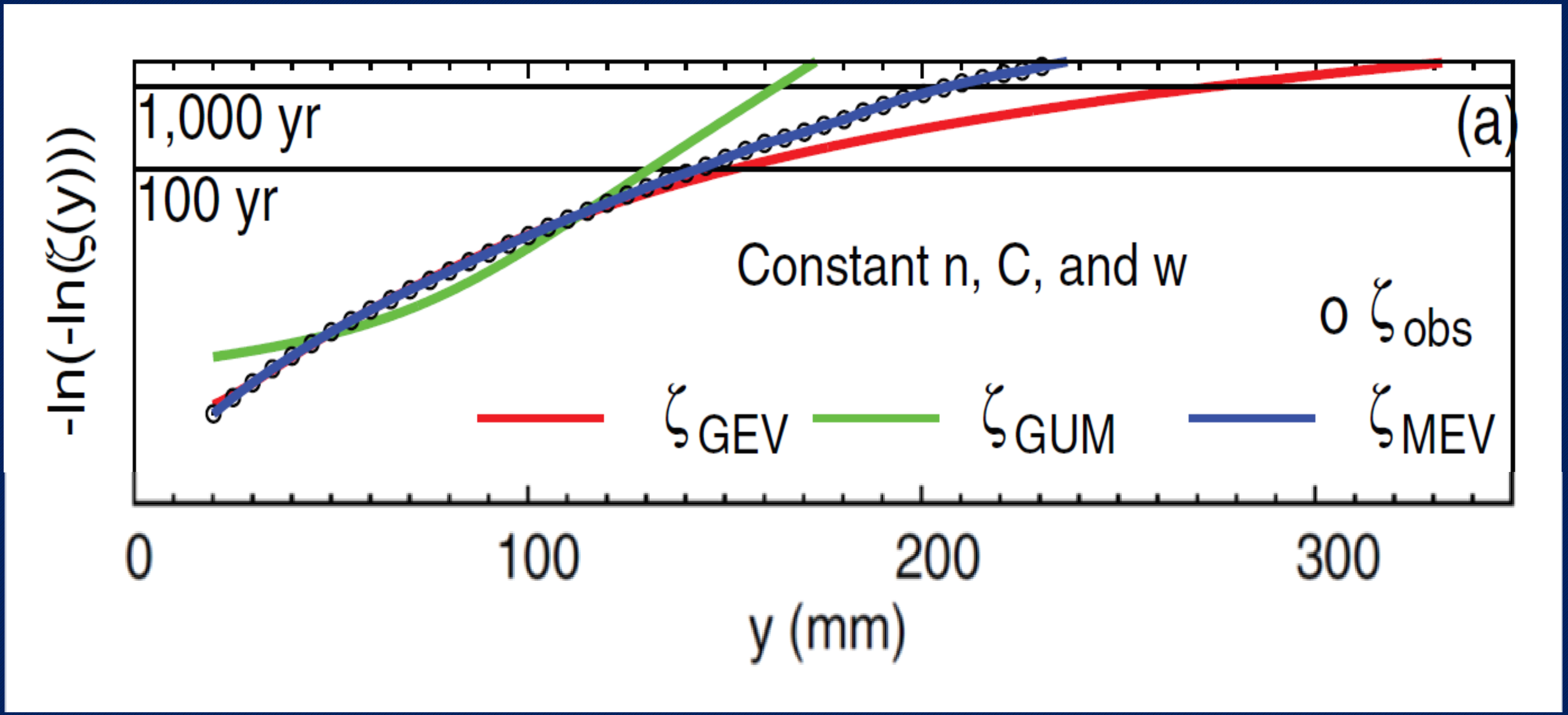
T = # sub-periods over which n and θ are estimated

In the Weibull case becomes:

$$\zeta(x) \cong \frac{1}{T}\sum_{j=1}^{T}\left[1 - e^{\left(\frac{x}{C_j}\right)^{w_j}}\right]^{n_j}$$

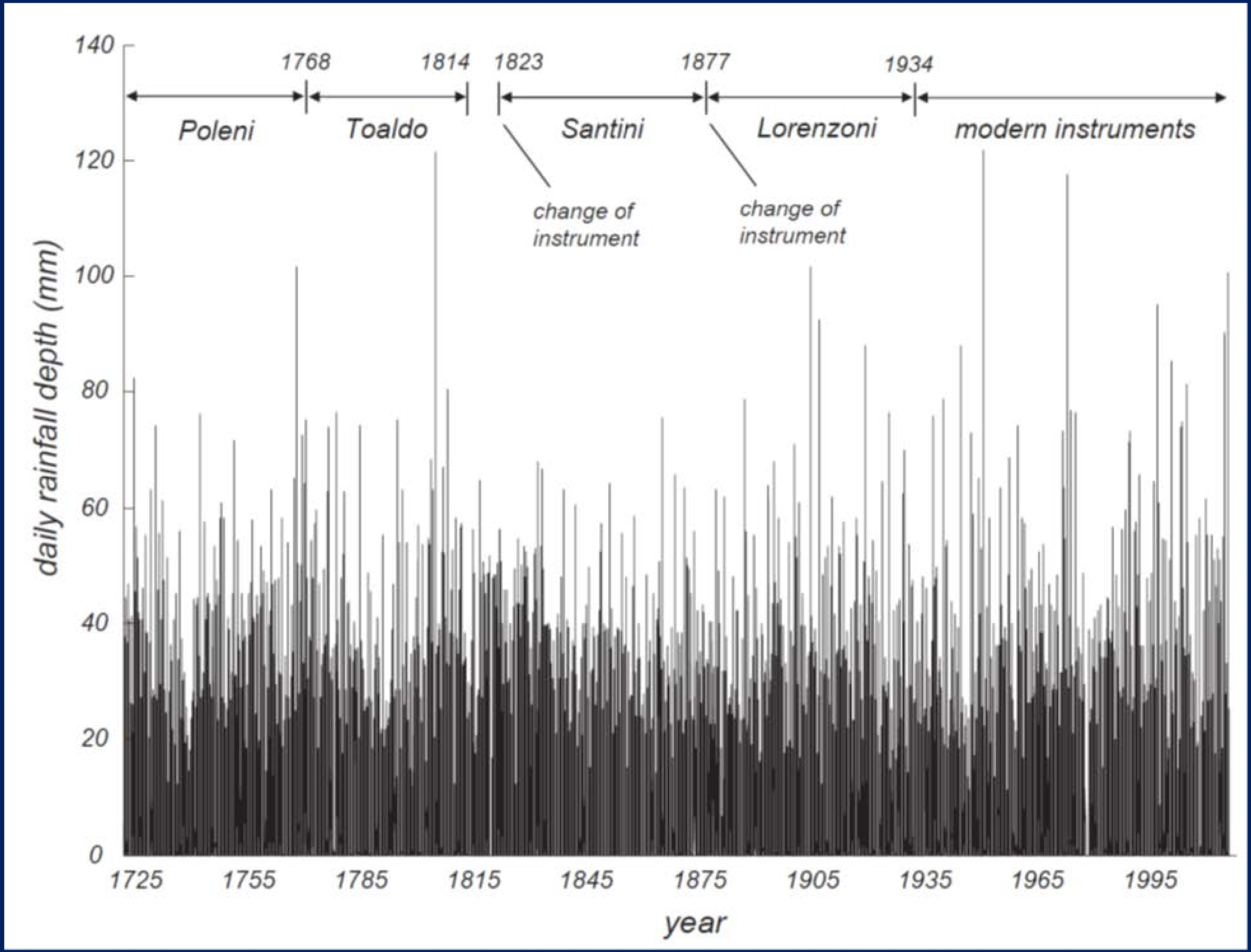
Marani and Ignaccolo, 2015; Zorzetto et al., 2016

Weibull-distributed synthetic data

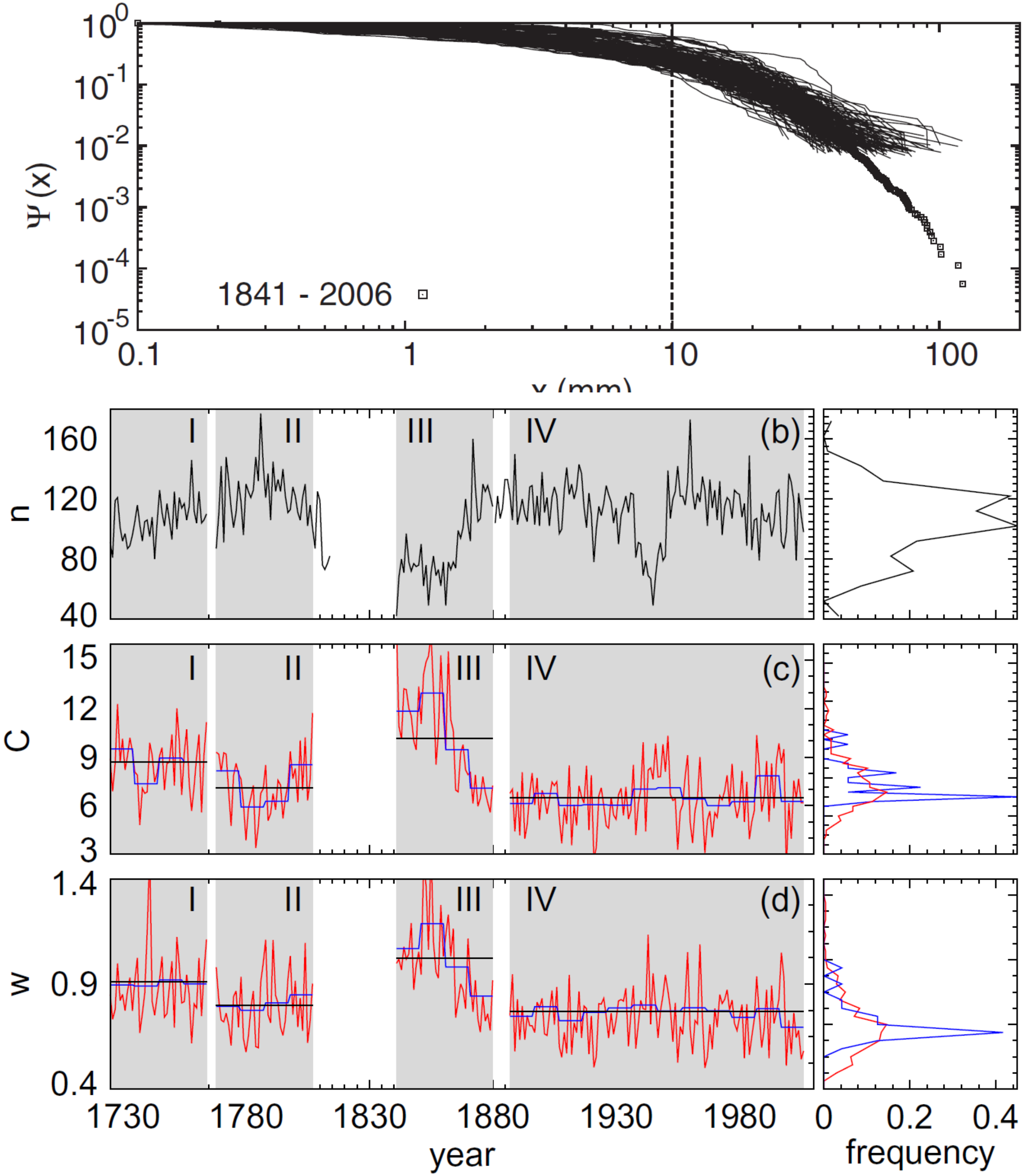


Marani and Ignaccolo, 2015

The Padova daily precipitation time series 1725-2006



(Marani and Zanetti, 2015)



Padova series:

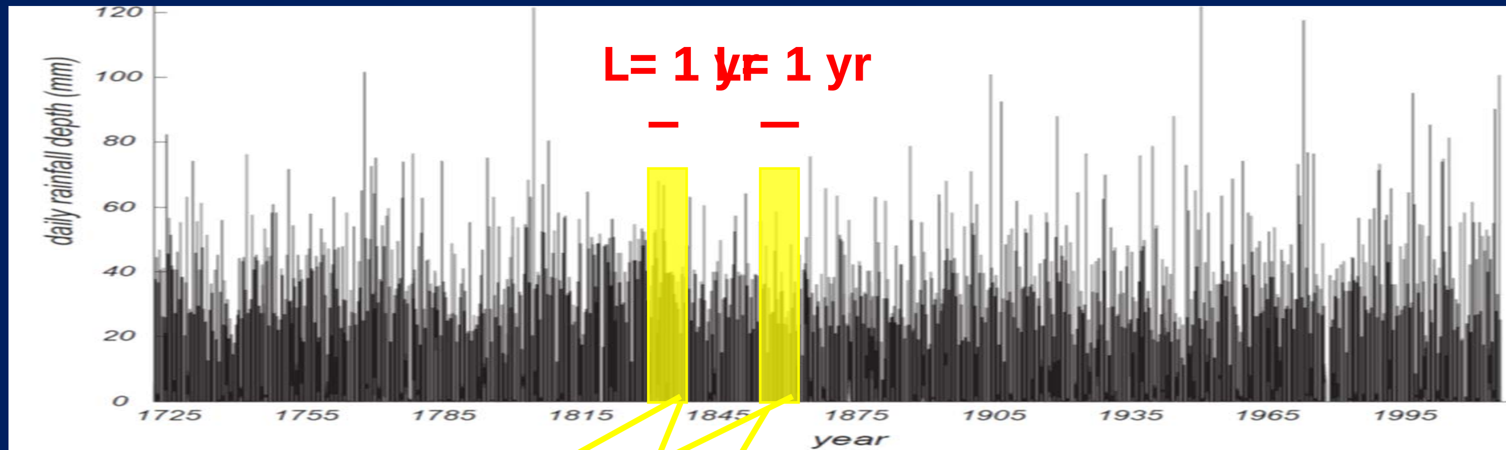
Wide fluctuations in pdf parameters and in number of events.

(Marani and Ignaccolo, 2015)

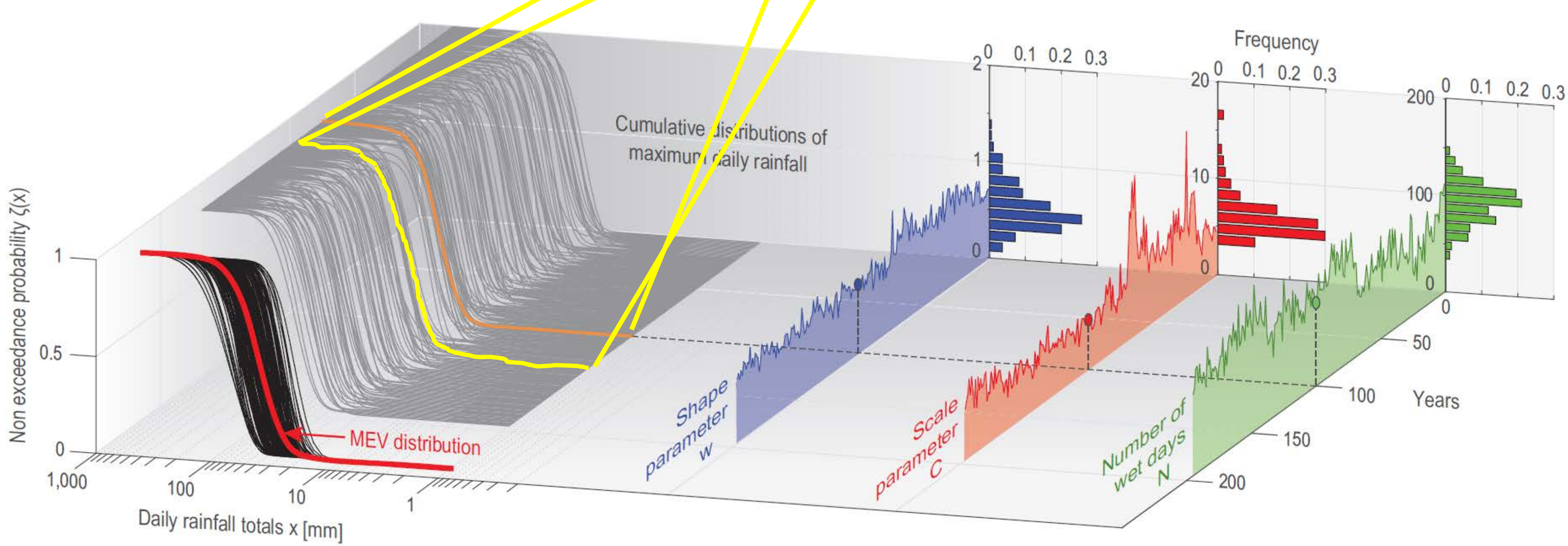
The Metastatistical Extreme Value distribution (MEV)

(Marani and Ignaccolo, AWR, 2015; Zorzetto et al., GRL, 2016)

Padova:
Daily rainfall
1725- today



Marani and Zanetti, WRR, 2015



Reality Check

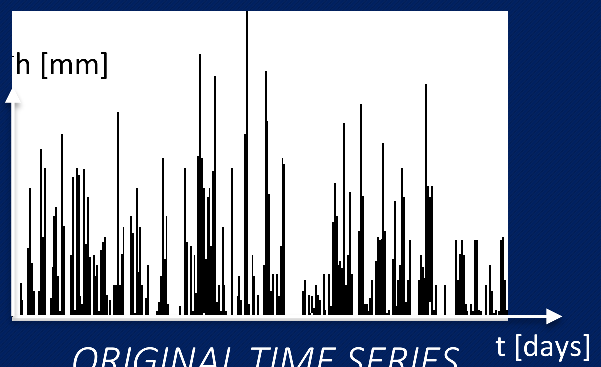
36 datasets, 106 -275 years of daily observations, (<L> =135 yrs)

Less than 5% of missing data

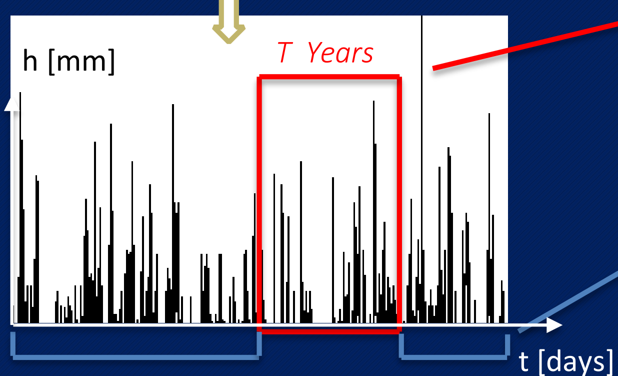
Method of analysis

Reshuffling of daily data preserving
(1) yearly number of events, and
(2) observed values (i.e. Pdf's)

- To eliminate correlation and non-stationarity
- Preserving the true (unknown) distribution of the parameters and numbers of wet days.



ORIGINAL TIME SERIES



RANDOMLY RESHUFFLED TIME SERIES

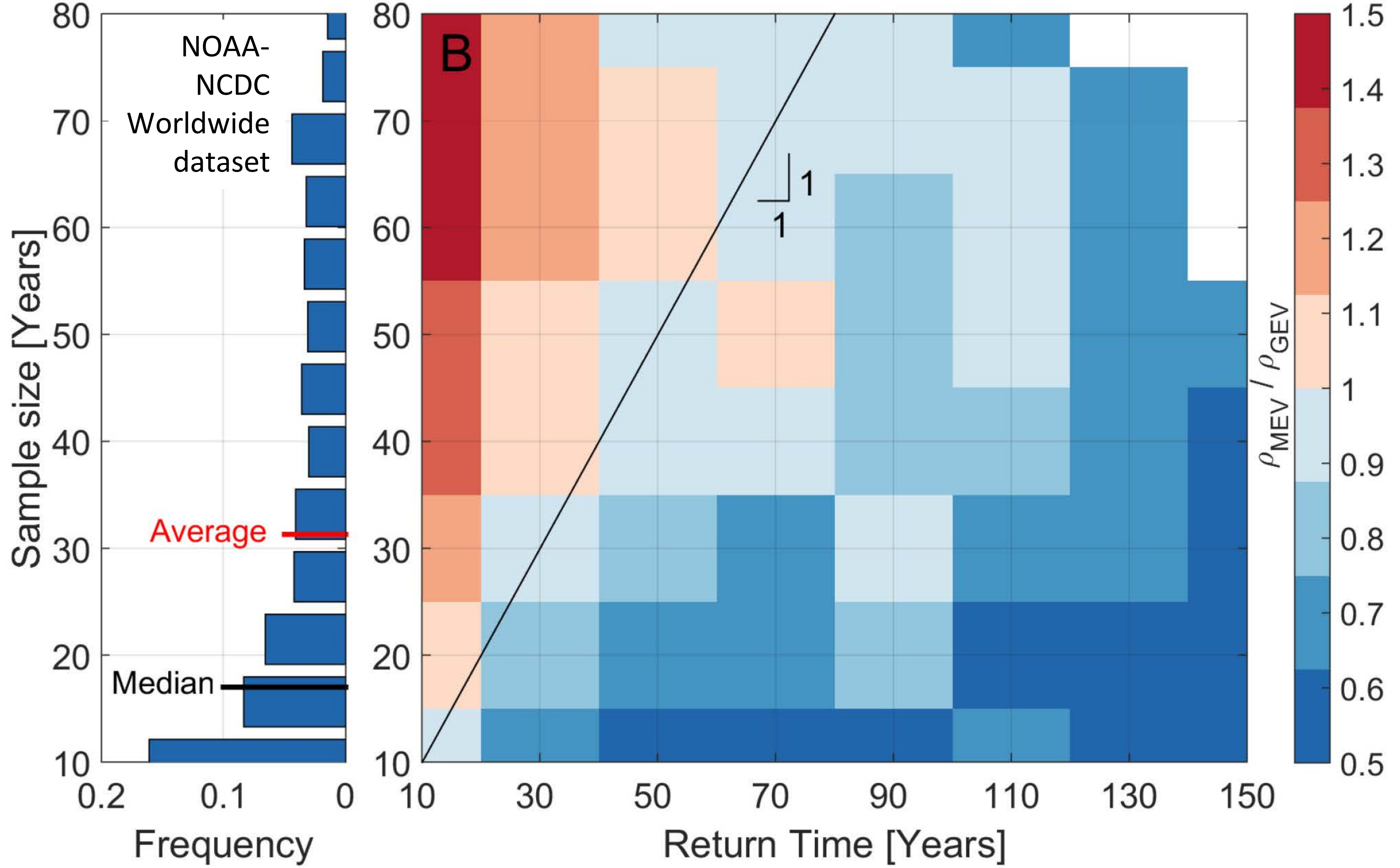
Selection of a T-year sample from which compute $\hat{x}\,(T_r, T)$ with GEV, POT and MEV.

The remainder of the dataset is used to observe the 'true value' $x_{obs}(T_r)$ to assess the performance of EV methods

Non dimensional estimation error

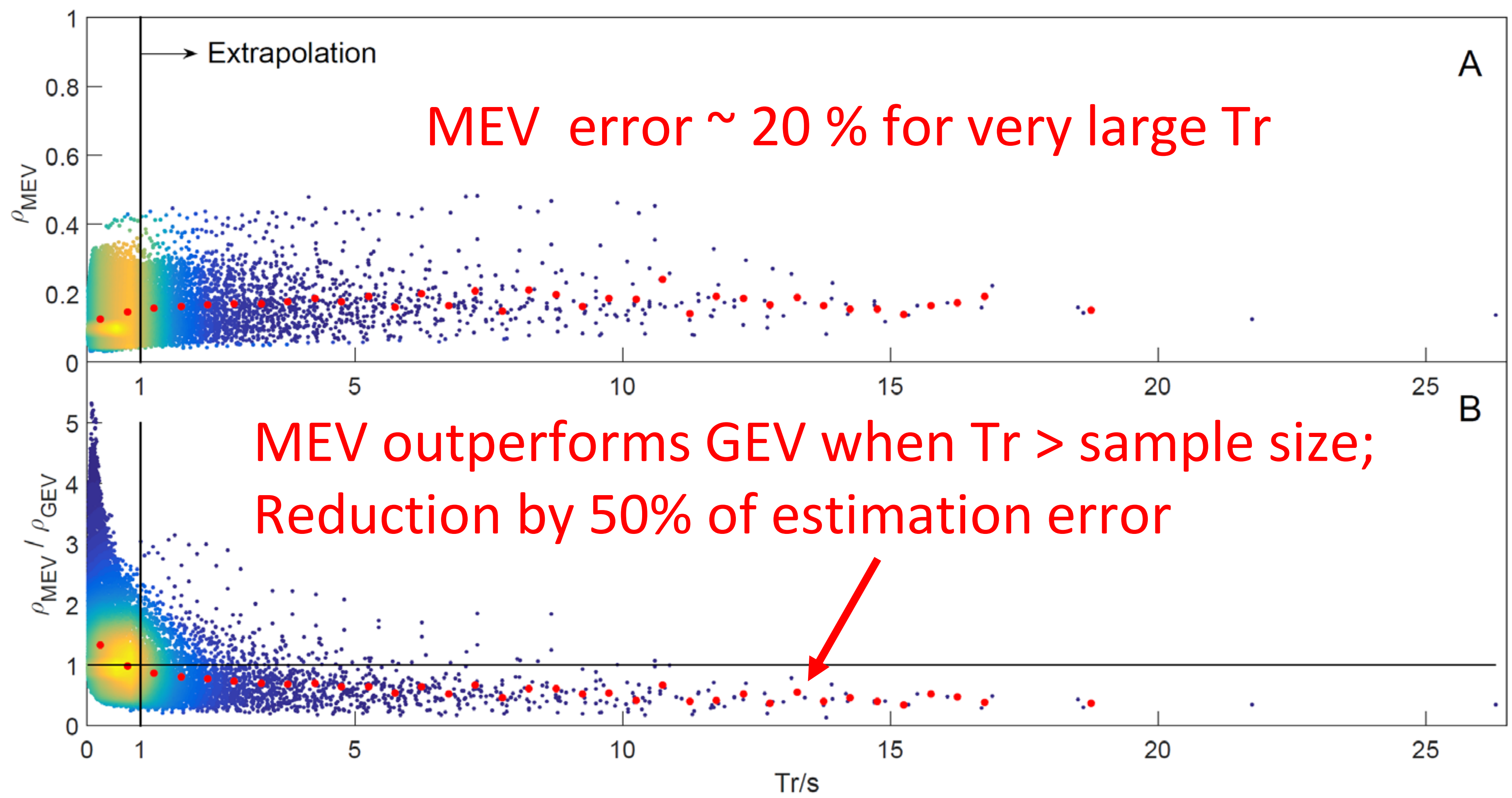
$$\epsilon = \sqrt{\frac{1}{N}\sum \left(\frac{\hat{x} - x_{obs}}{x_{obs}}\right)^2}$$

Ratio of MEV estimation error to GEV-LMOM error

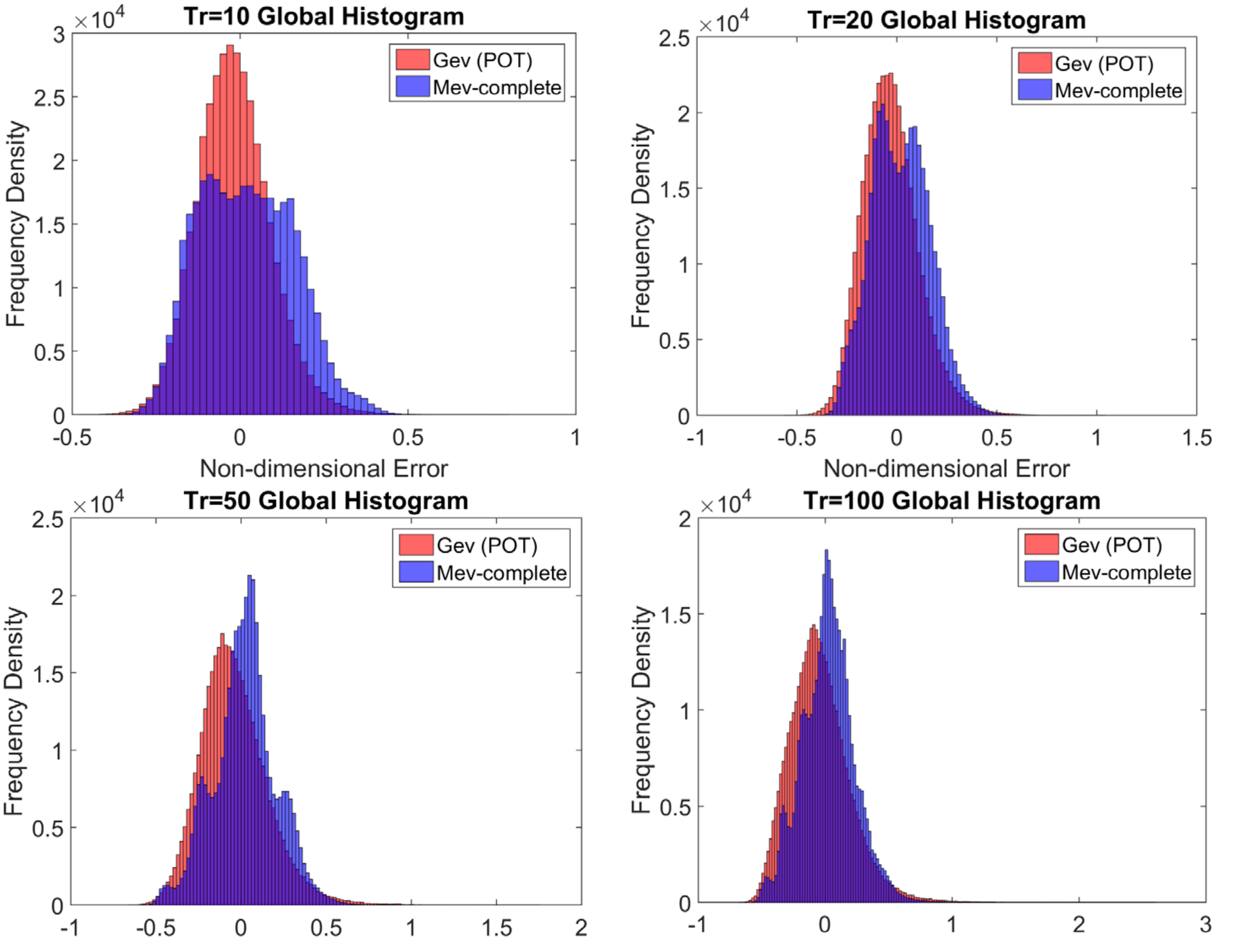


Relative estimation error as a function of Tr/(sample size) when training and testing data are separate

MEV vs. GEV-LMOM

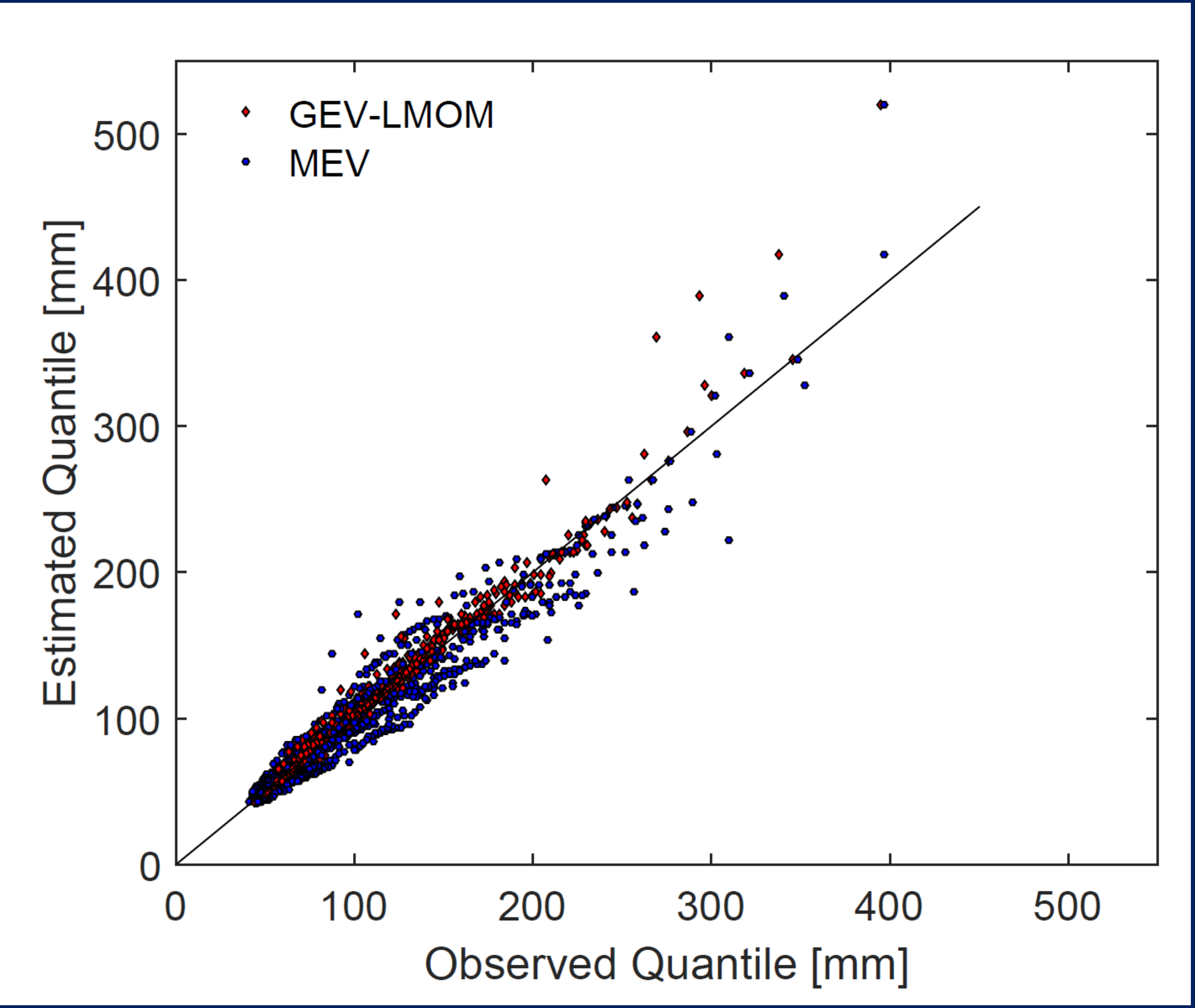


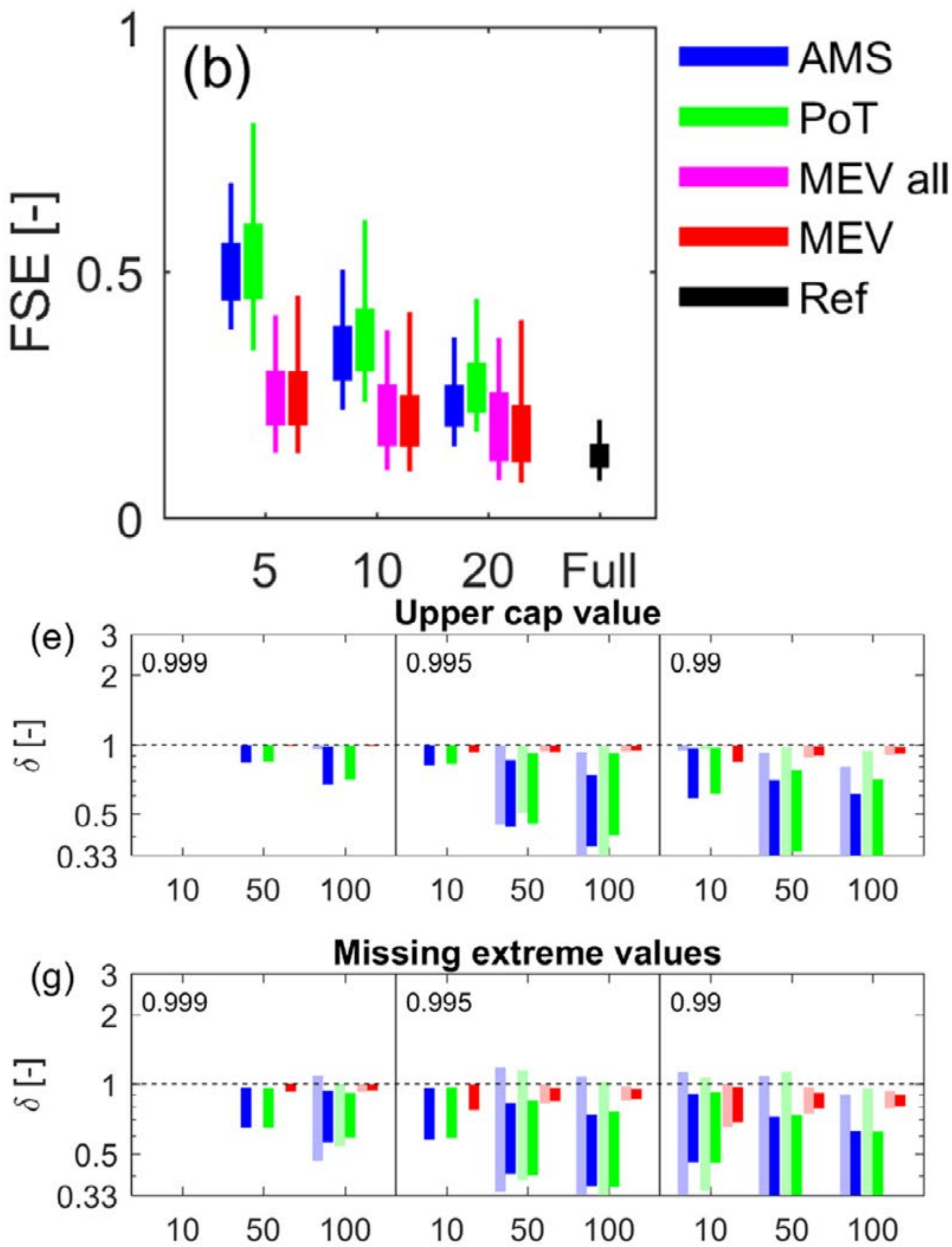
Zorzetto et al., GRL 2016



Tr=10 Global Histogram
Tr=20 Global Histogram
Tr=50 Global Histogram
Tr=100 Global Histogram
Gev (POT)
Mev-complete
Frequency Density
Non-dimensional Error
×10^4

An interesting observation: GEV performs better if calibration data=testing data





Hourly rainfall (Marra et al., AWR 2018):

- MEV uncertainty ~50%< GEV

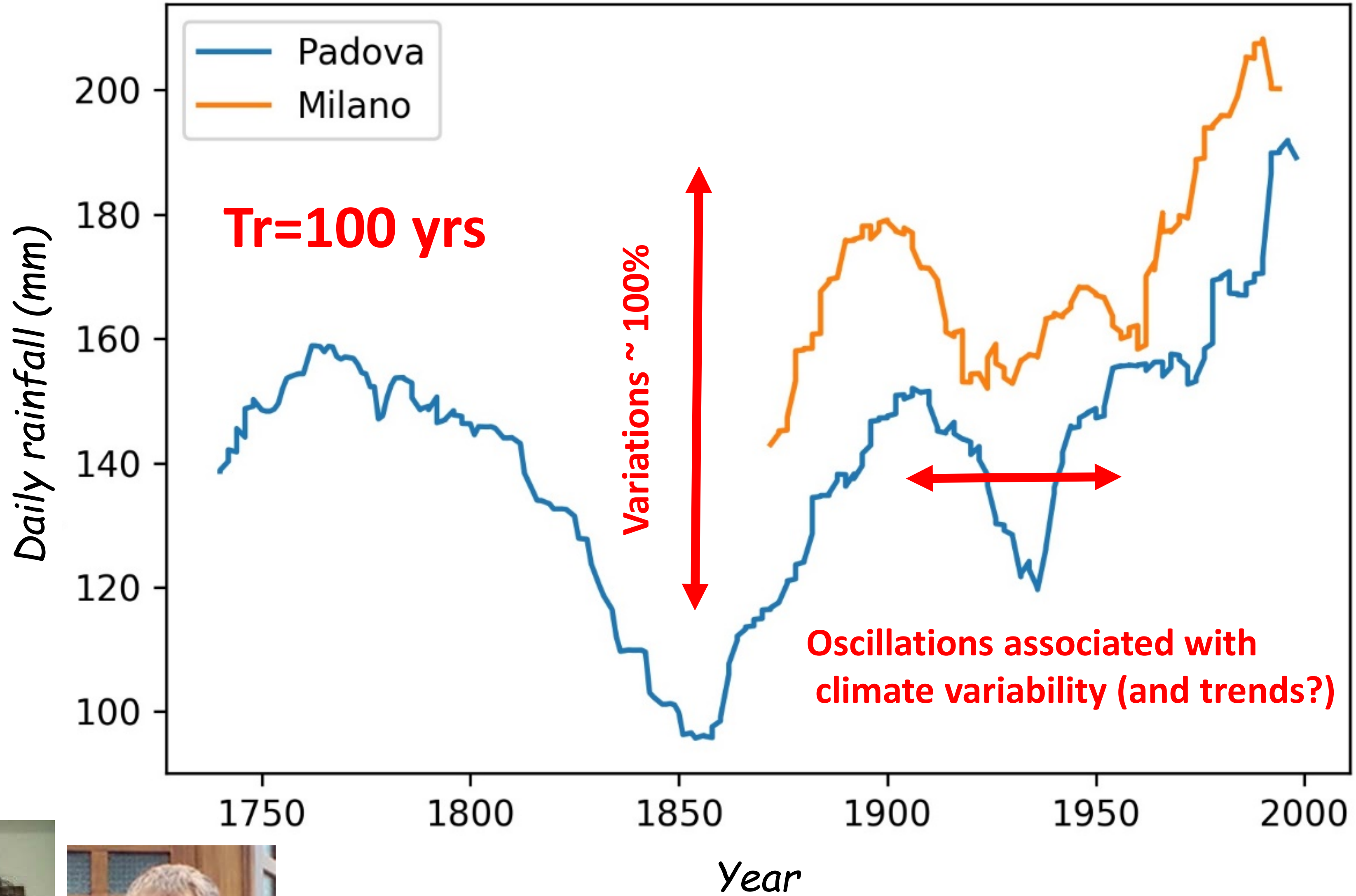
- MEV less sensitive than GEV to errors in data (saturation at high values, missing largest values, etc.

Conclusions (1/3)

MEV ouperforms classical EV distributions providing:

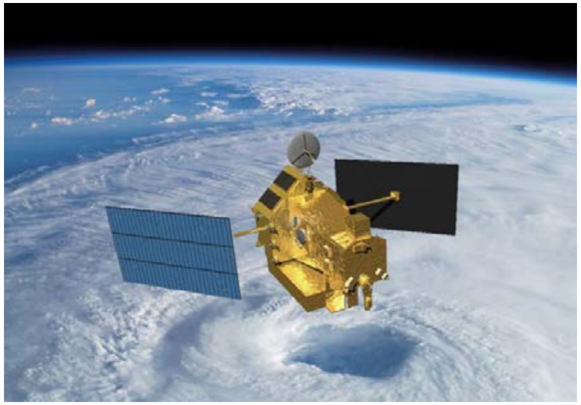
1. Reduced uncertainty (50%) for Tr > sample size
2. Better use of the available daily data
3. Removal of asymptotic/Poisson hypotheses

Slight detour: return time under climate fluctuations

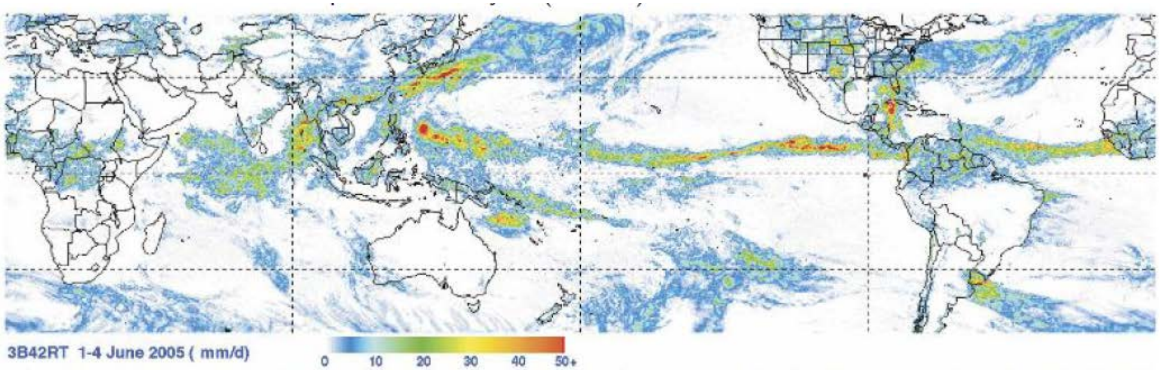



(Marani and Zanetti, WRR 2015; Zorzetto et al., in preparation)

TRMM Multi satellite Precipitation Analysis (TMPA)



(source: NASA)



(Huffman et al, 2007)

3B42 V7 TMPA: research quality Combination of multiple sensors, chiefly microwave and infrared.

Fairly coarse 0.25°x0.25° spatial resolution

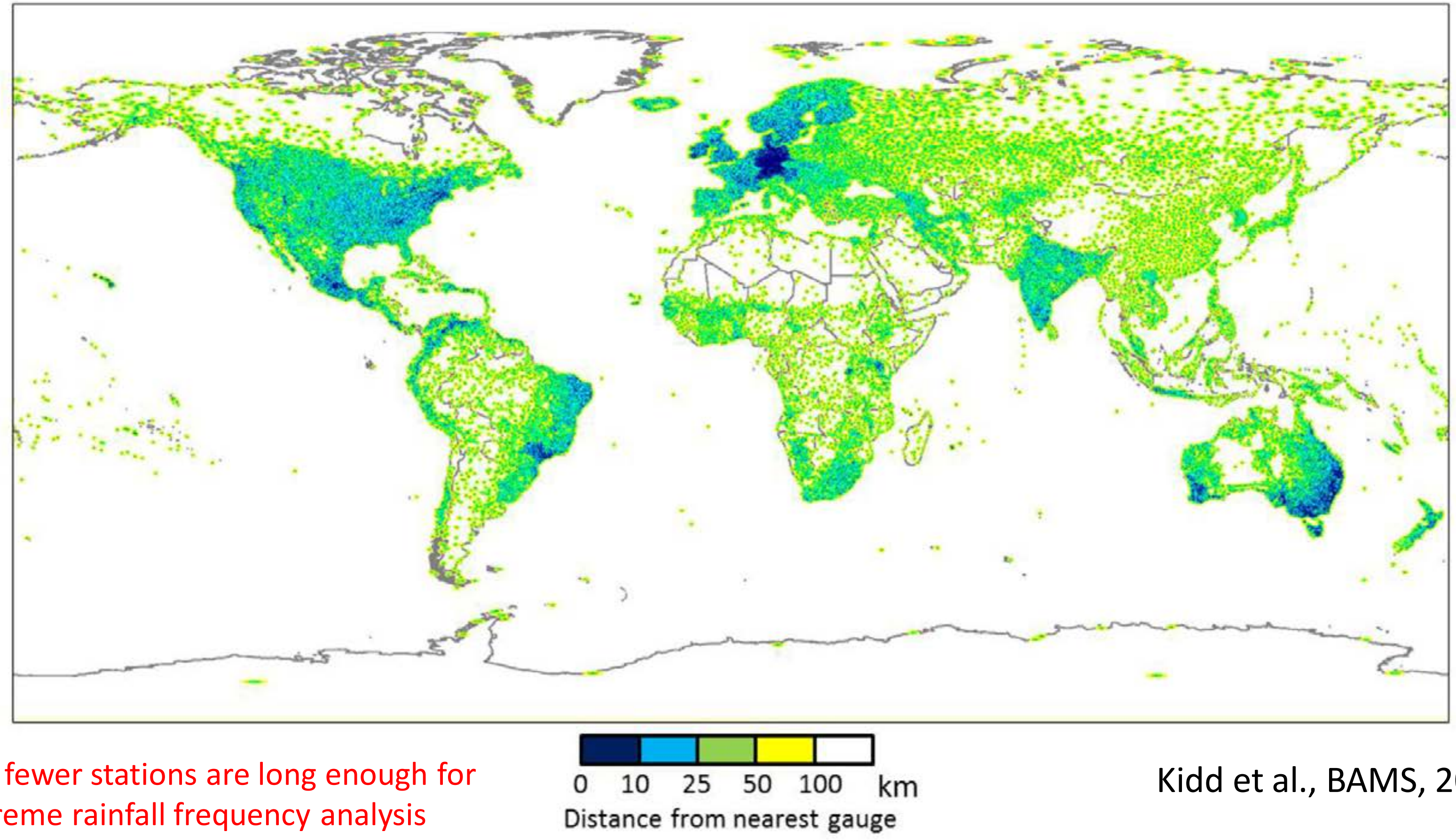
Up to 3 hours temporal resolution

Useful for Inference on rainfall pdfs and extremes?

- QPE errors
- Short record length

(Enrico Zorzetto, under NASA Fellowship NESSF 17-EARTH17F-270)

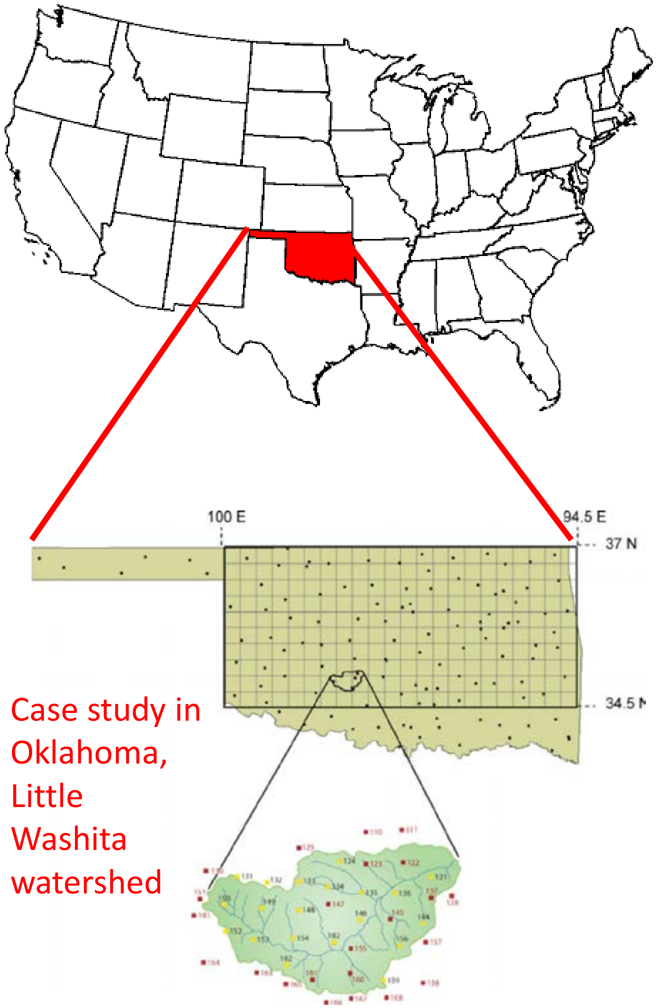
Global Rainfall Extremes: Gaps in global observational networks



But fewer stations are long enough for
Extreme rainfall frequency analysis

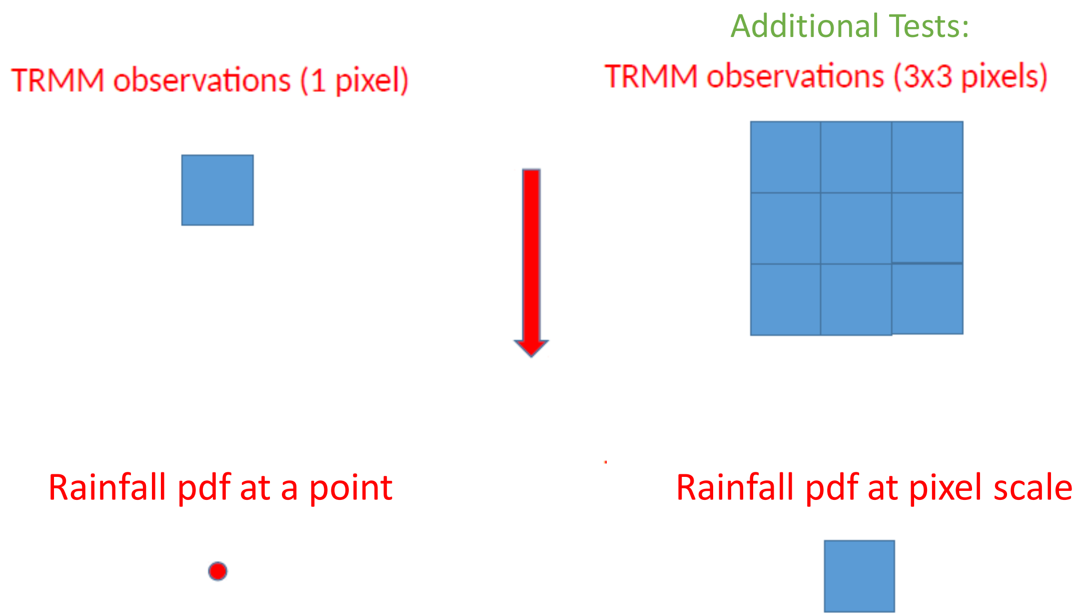
Kidd et al., BAMS, 2016

How to validate QPE rainfall pdfs using point measurements?



Case study in Oklahoma, Little Washita watershed

Parameters of rainfall pdf downscaled using closed-form relations from space-time stochastic rainfall models



A Metastatistical Extreme Value distribution (MEV)

The MEV block-maxima distribution:

$$\zeta(x) = \sum_{n=1}^{\infty} \int_{\Omega_{\vec{\Theta}}} F(x;\vec{\theta})^n \, g(n,\vec{\theta}) d\vec{\theta}$$

G(n,θ) = joint prob distrib. of the parameters.

Can be approximated using sample averages:

$$\zeta(x) \cong \frac{1}{T} \sum_{j=1}^{T} F(x;\overrightarrow{\theta_j})^{n_j}$$

T = # sub-periods over which n and θ are estimated

Marani and Ignaccolo, 2015; Zorzetto et al., 2016

A downscaling approach for TMPA gridded data

We transform across scales the following quantities

The mean is conserved across scales

$\mu_0 = \mu_L.$

The variance increase of a factor γ

Effect of spatial correlation

$$\sigma_0^2 = \frac{1}{\gamma}\sigma_L^2, \qquad \gamma = \frac{\sigma_{c_L}^2}{\sigma_{c_0}^2} = \frac{4}{L_x^2 L_y^2}\int_0^{L_x}\int_0^{L_y}(L_x - x)(L_y - y)\,\rho(x, y)\,dx\,dy$$

[Vanmarcke, 2010]

Variation of the yearly number of daily events/year

$\gamma_s(L) = \frac{N_L}{N_0}$

[Zorzetto and Marani, 2018]

Can we infer these quantity from TMPA alone?

Precipitation spatial correlation at the point scale

$$\triangle(L_{x,k}, L_{y,l}) = 4\int_0^{L_{x,k}}\int_0^{L_{y,l}} (L_{x,k} - s_1)(L_{y,l} - s_2)\,\rho(s_1, s_2)ds_1 ds_2$$

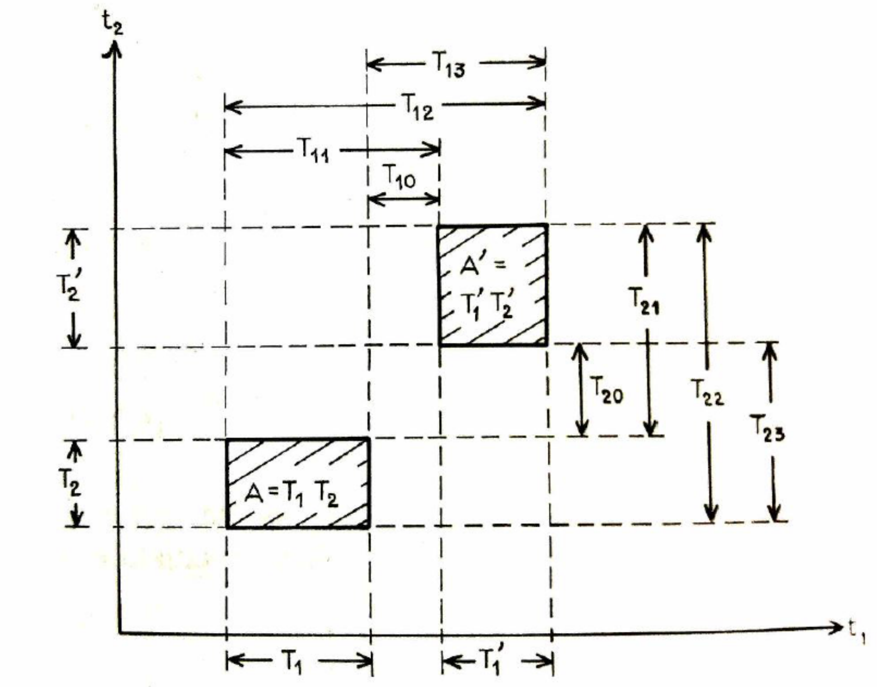
Correlation of local averages

$$Cov\left[h_L, h_{L'}\right] = \frac{\sigma_0^2}{4\left(L_x L_y\right)^2}\sum_{k=0}^{3}\sum_{l=0}^{3}(-1)^k(-1)^l\,\triangle\left(L_{x,k}, L_{y,l}\right)$$

Distances $L_{x,k}$ and $L_{y,k}$ with $k, l = 0, 1, 2, 3$ encode the relative position of the two pixels Assume correlation function:

$$\rho(s) = \begin{cases} e^{-\frac{\alpha s}{\epsilon}} & s < \epsilon \\ \left(\frac{\epsilon}{es}\right)^{\alpha} & s \geq \epsilon, \end{cases} \qquad s = \sqrt{x^2 + y^2}$$

From Marani, 2003



From Vanmarcke, 2010

(Zorzetto and Marani, WRR, 2018)

Precipitation spatial correlation at the point scale

$$\triangle(L_{x,k}, L_{y,l}) = 4 \int_0^{L_{x,k}} \int_0^{L_{y,l}} (L_{x,k} - s_1)(L_{y,l} - s_2)\,\rho(s_1, s_2)\,ds_1\,ds_2$$

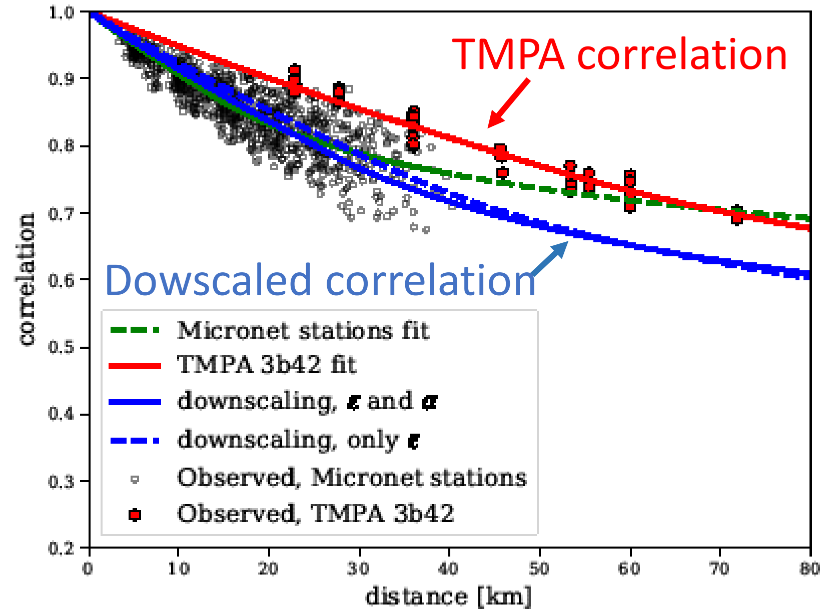
Correlation of local averages

$$Cov\left[h_L, h_{L'}\right] = \frac{\sigma_0^2}{4\left(L_x L_y\right)^2} \sum_{k=0}^{3} \sum_{l=0}^{3} (-1)^k (-1)^l \triangle\left(L_{x,k}, L_{y,l}\right)$$

Distances $L_{x,k}$ and $L_{y,k}$ with $k, l = 0, 1, 2, 3$ encode the relative position of the two pixels Assume correlation function:

$$\rho(s) = \begin{cases} e^{-\frac{\alpha s}{\epsilon}} & s < \epsilon \\ \left(\frac{\epsilon}{es}\right)^{\alpha} & s \geq \epsilon, \end{cases} \qquad s = \sqrt{x^2 + y^2}$$

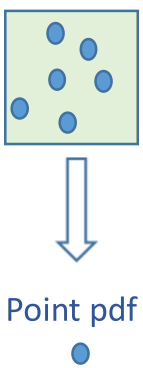
From Marani, 2003

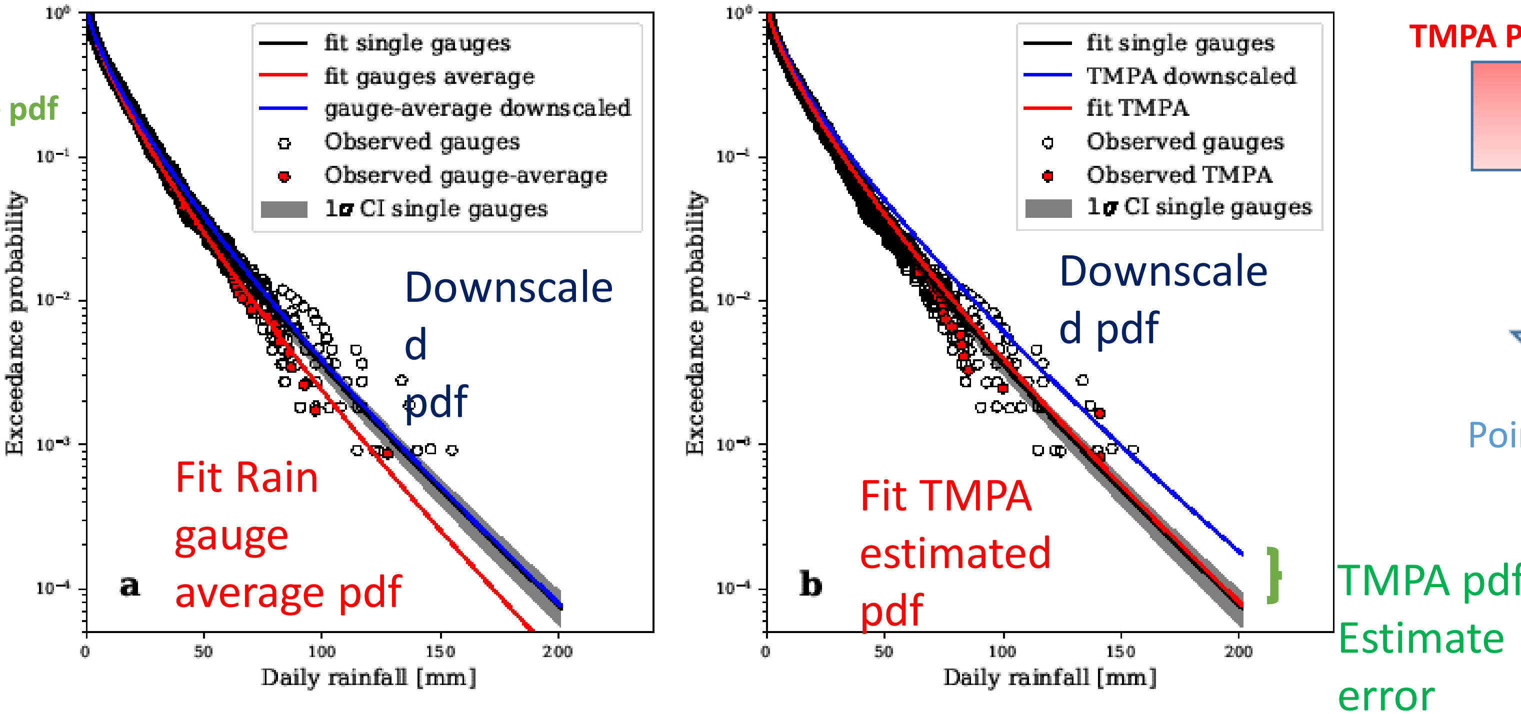


Assumed Correlation function shape

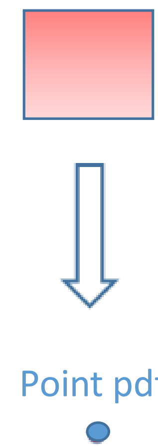
Downscaling of the pdf of ordinary daily rainfall

Station Average pdf





TMPA Pixel Average



(Zorzetto and Marani, WRR, 2018)

A downscaling approach for TMPA gridded data

We transform across scales the following quantities

The mean is conserved across scales

$\mu_0 = \mu_L$.

The variance increase of a factor γ

$$\sigma_0^2 = \frac{1}{\gamma}\sigma_L^2, \qquad \gamma = \frac{\sigma_{c_L}^2}{\sigma_{c_0}^2} = \frac{4}{L_x^2 L_y^2}\int_0^{L_x}\int_0^{L_y}(L_x - x)(L_y - y)\,\rho(x,y)\,dxdy$$

[Vanmarcke, 2010]

Variation of the yearly number of daily events/year

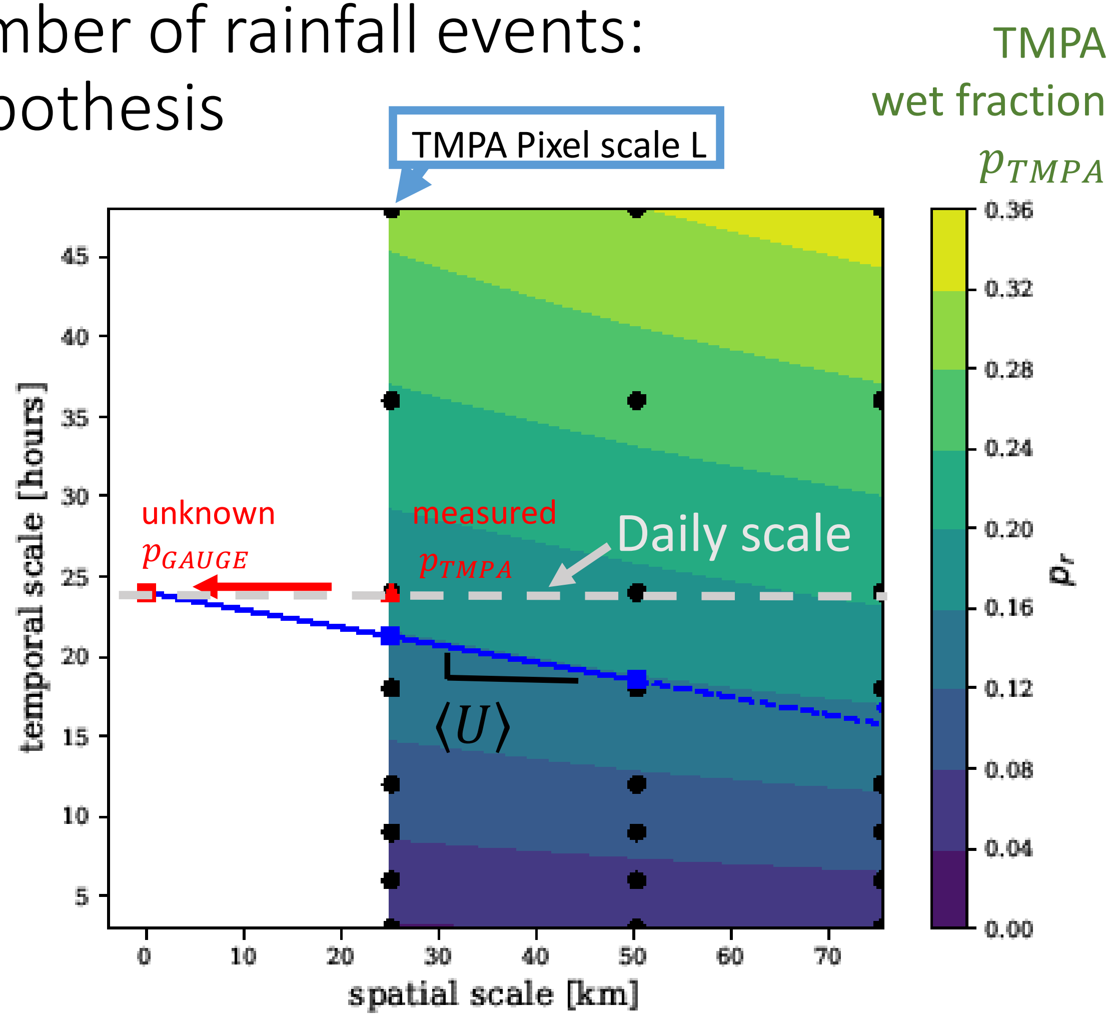
$\gamma_s(L) = \frac{N_L}{N_0}$ Effect of spatial intermittency

[Zorzetto and Marani, 2018] Can we infer these quantity from TMPA alone?

Downscaling of the Number of rainfall events: "frozen turbulence" hypothesis

Effect of spatial intermittency

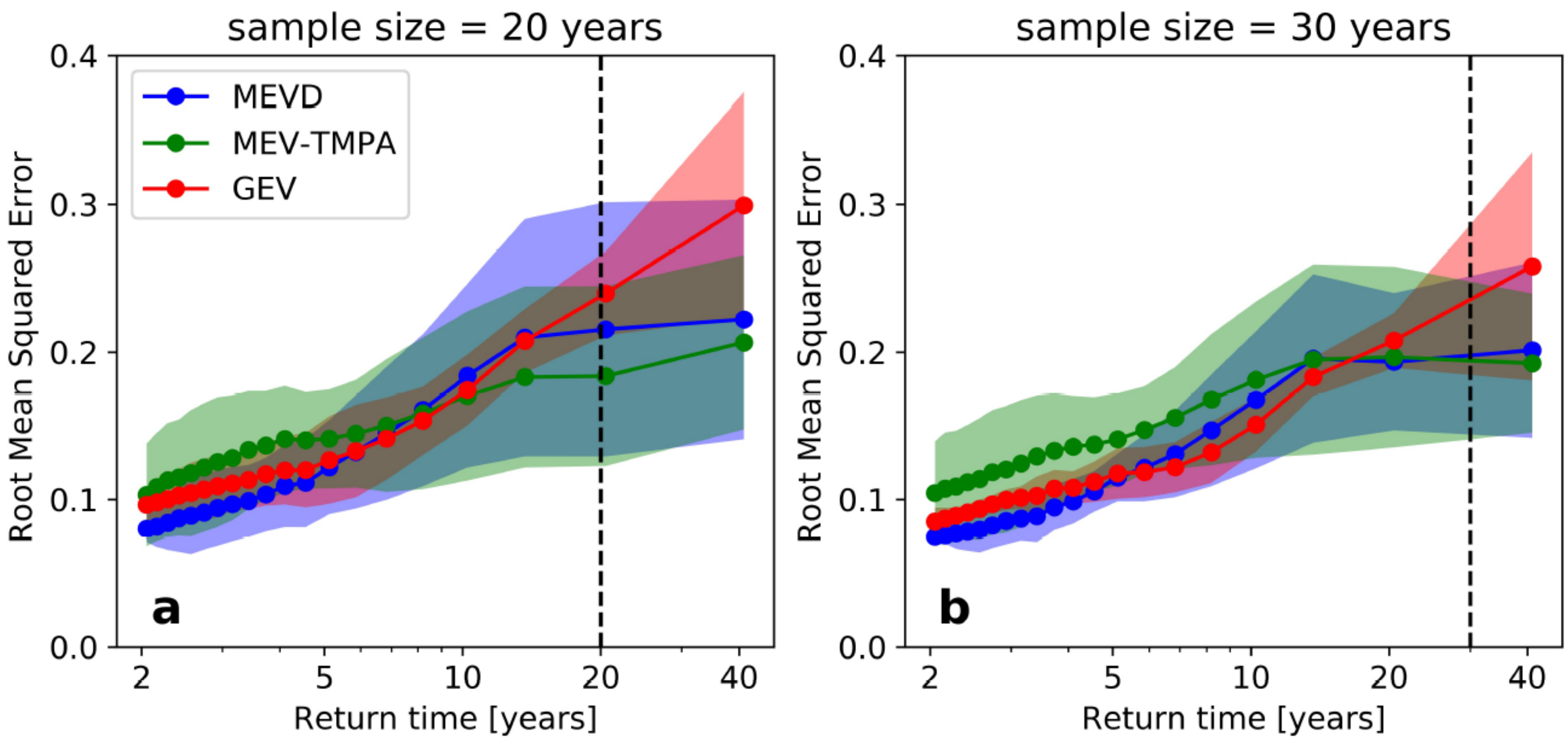
$$\gamma_s(L) = \frac{p_{TMPA}(L)}{p_{GAUGE}}$$



(Zorzetto and Marani, WRR, 2018)

Cross-Validation Analysis

Root Mean Square Error $R(s, Tr) = \left(\frac{1}{n_g} \sum_{i=1}^{n_g} \left[\frac{\hat{h}_i(Tr) - h_i^{obs}(s_s, T_r)}{h_i^{obs}(T_r)} \right]^2 \right)^{1/2}$

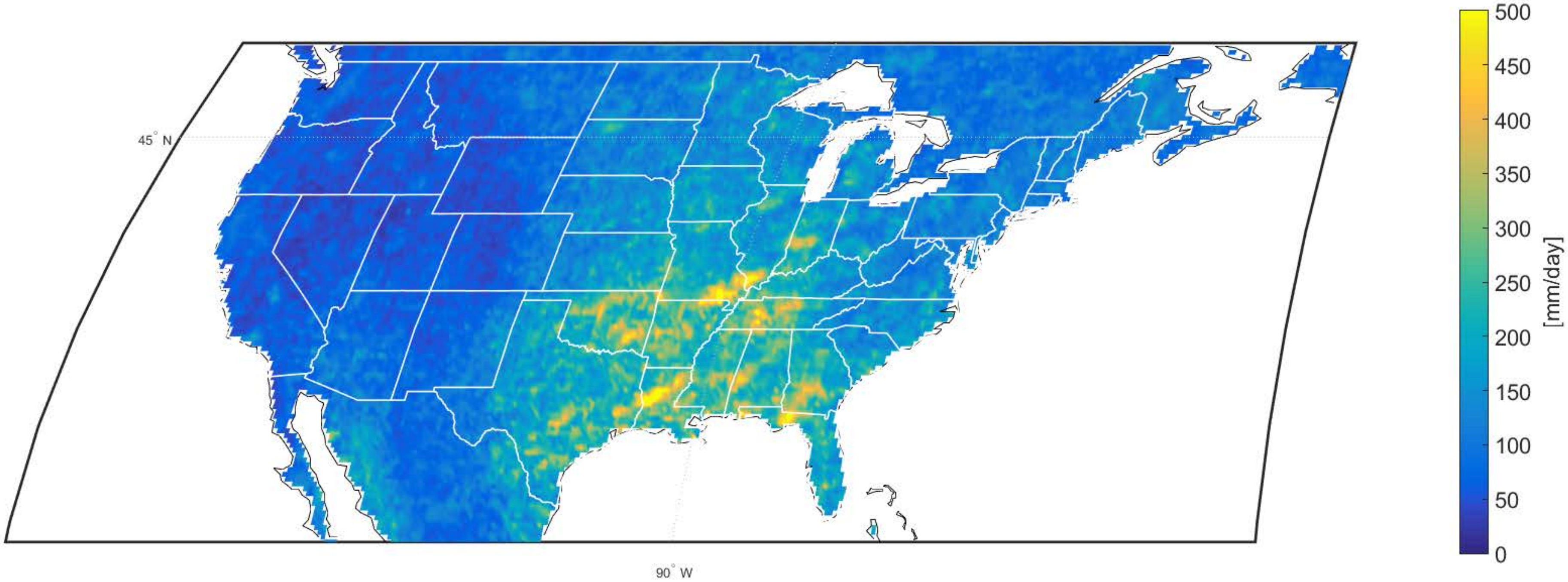


Error of the downscaling technique (green) vs ground observations (red, blue).

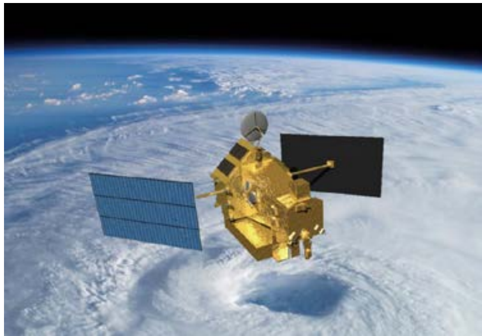
From TMPA satellite estimates alone

Zorzetto and Marani, WRR, 2018

Continental USA - GEV, 100-year daily rainfall event from all TRMM-TMPA observations (3B42 V7 - 17 yrs)

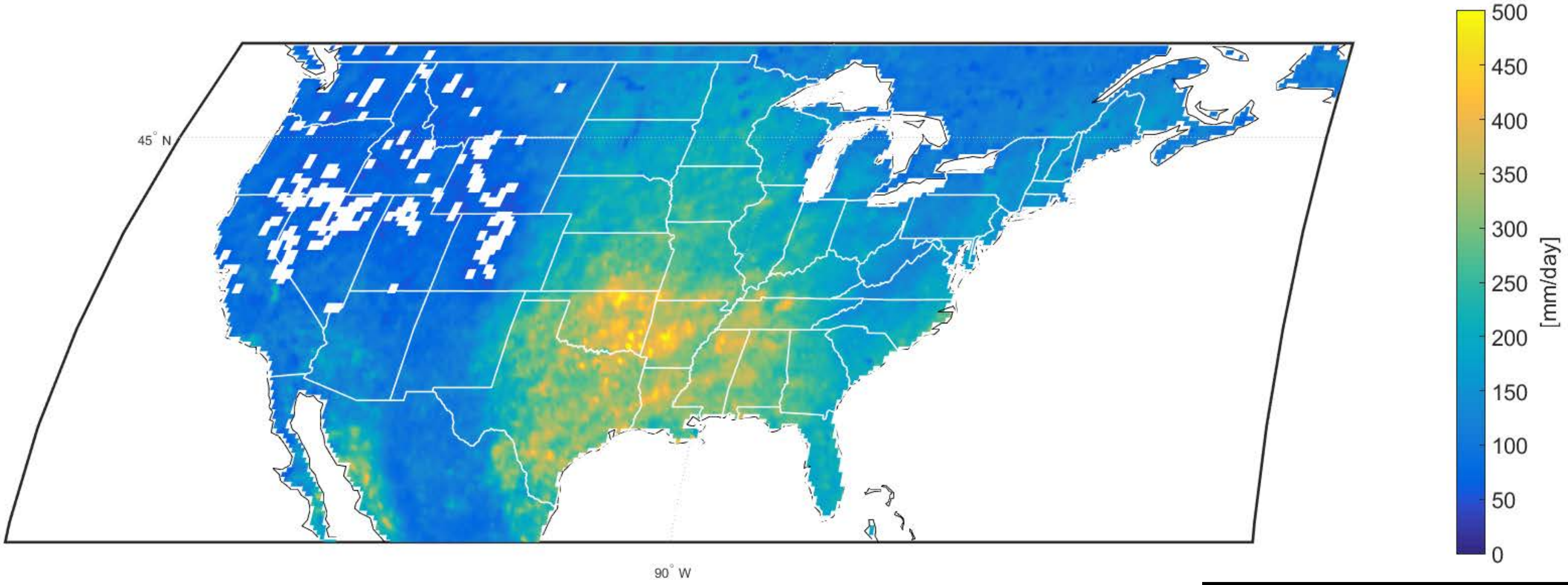


Daily TRMM 3B42 V7 data:
Quasi global (+50°N, -50°S);
0.25°x0.25° spatial resolution



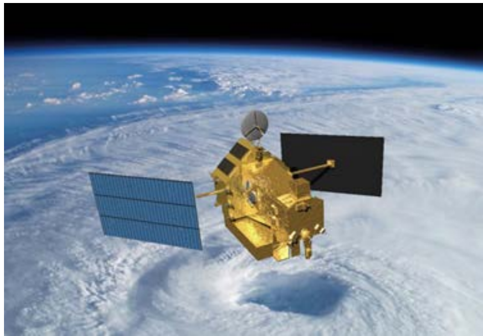
Zorzetto and Marani, in preparation, 2019

Continental USA – MEV, 100-yr daily rainfall event from all TRMM-TMPA observations (3B42 V7 - 17 yrs)



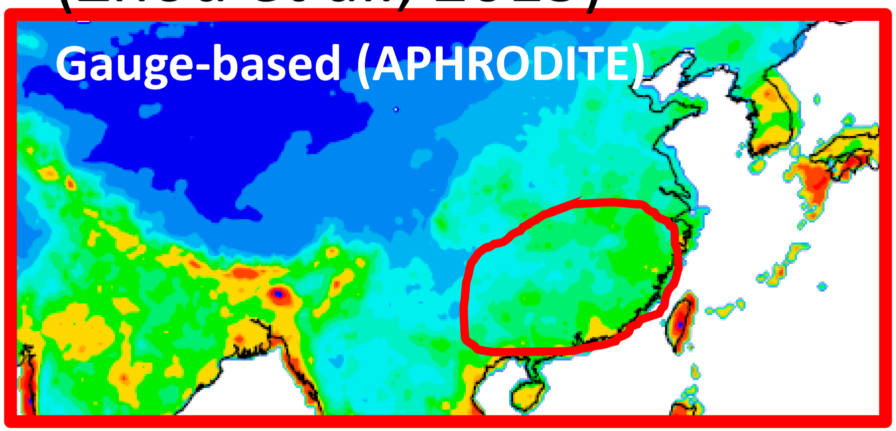
Zorzetto and Marani, In preparation, 2019

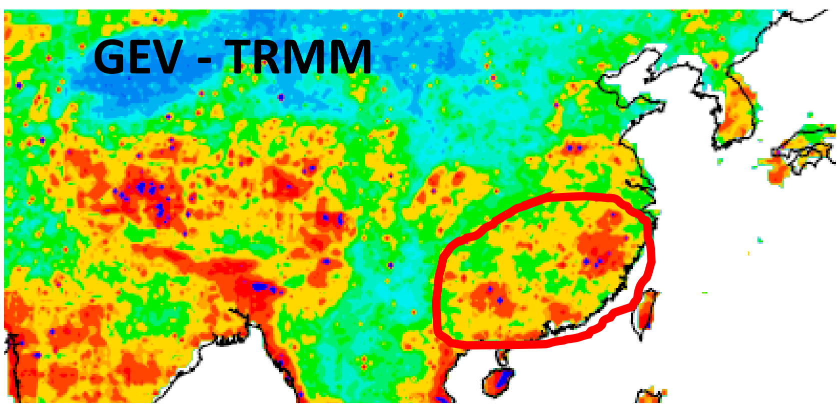
Daily TRMM 3B42 V7 data:
Quasi global (+50°N, -50°S);
0.25°x0.25° spatial resolution



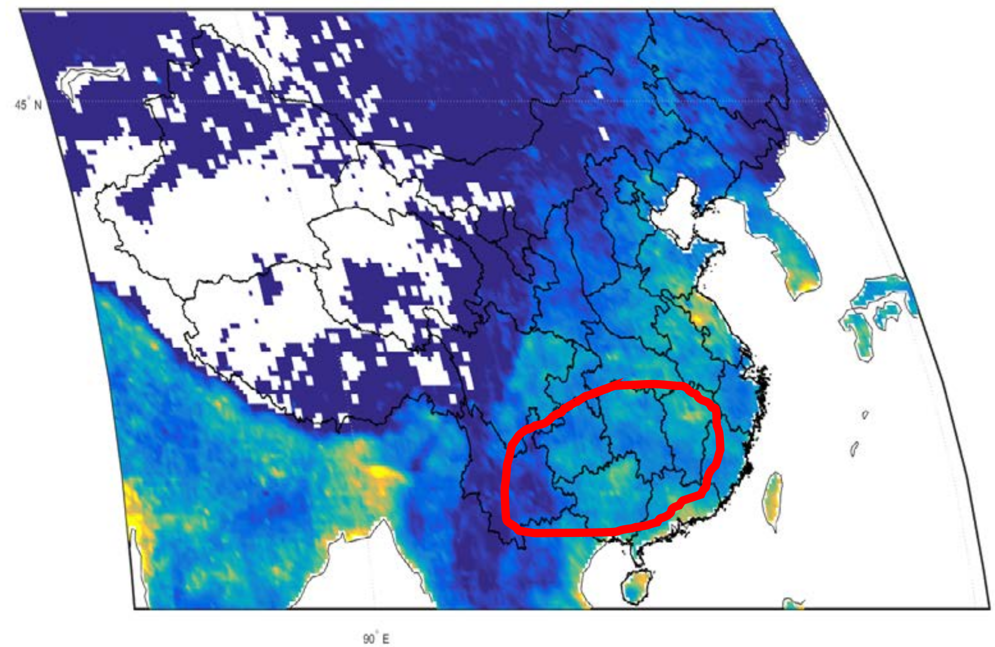
Comparisons of extreme event (Tr=100 yrs) distribution patterns

(Zhou et al., 2015)



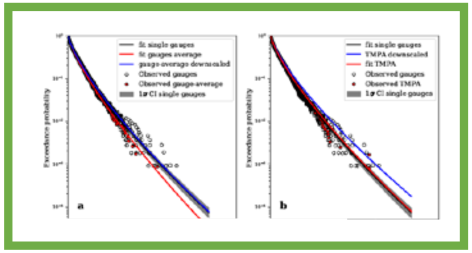


MEV-TRMM (this work)

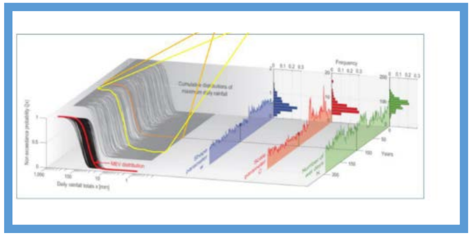


Zorzetto and Marani, in preparation, 2019

Conclusions (2/3):



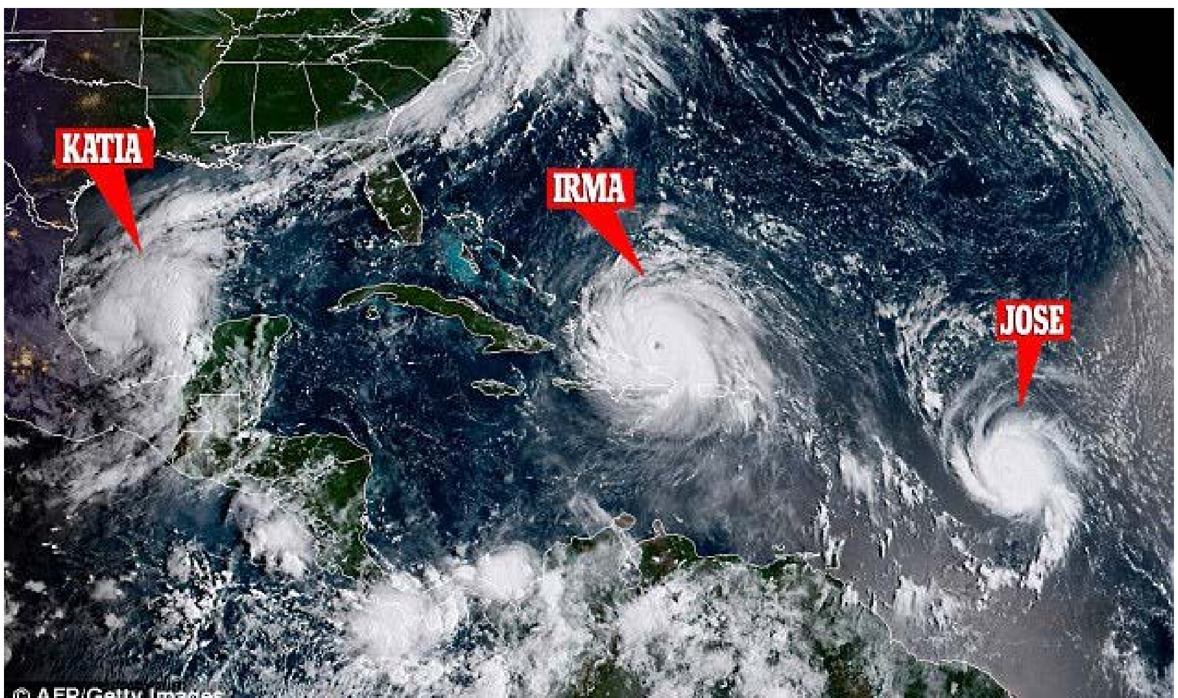
- Downscaling yields rainfall pdf at a point. **Validation of TMPA** feasible even with **sparse ground observations.**



- We can reconstruct **rainfall statistics at a point** in space, **uniquely based on areal-average satellite** precipitation estimates

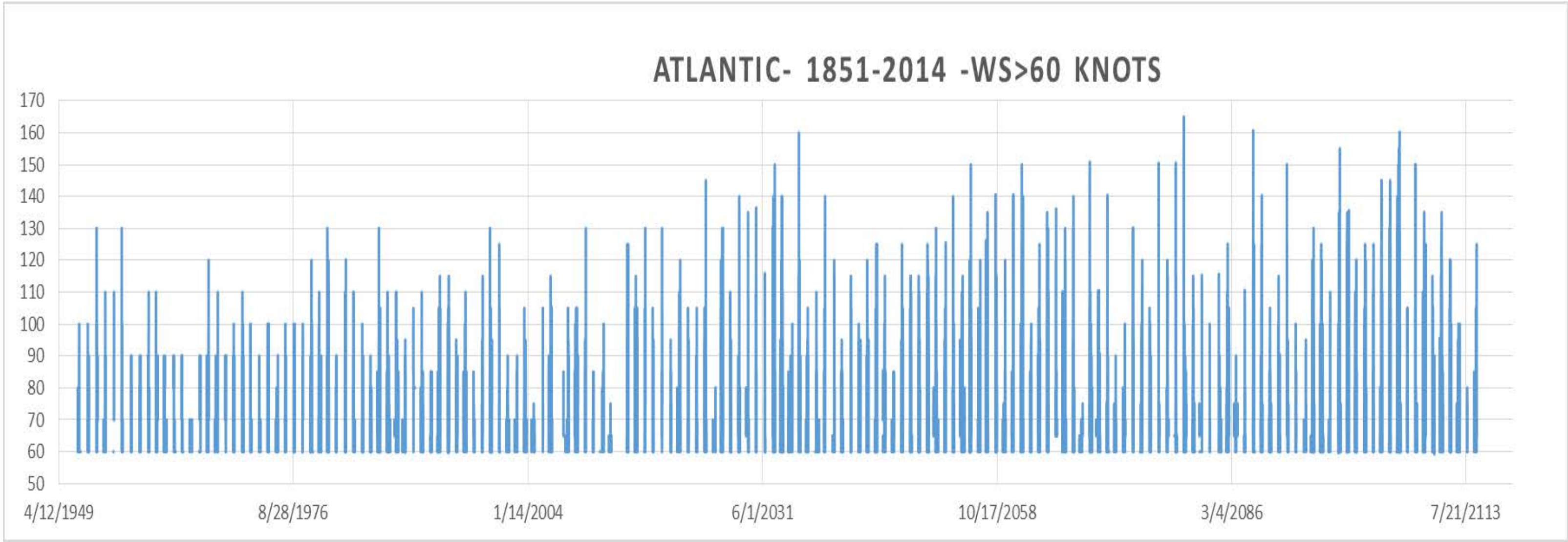
- ongoing work: investigation of method performance over **inhomogeneous and complex terrain**

Extreme Atlantic hurricanes



Reza Hosseini, Polytechnic Milan

Marco Scaioni, Polytechnic Milan



Category	Wind speeds
Five	≥70 m/s, ≥137 knots ≥157 mph, ≥252 km/h
Four	58–70 m/s, 113–136 knots 130–156 mph, 209–251 km/h
Three	50–58 m/s, 96–112 knots 111–129 mph, 178–208 km/h
Two	43–49 m/s, 83–95 knots 96–110 mph, 154–177 km/h
One	33–42 m/s, 64–82 knots 74–95 mph, 119–153 km/h

«Ordinary» Hurricanes: Generalize Pareto Distribution

The MEV expression is:

$$\zeta(x) = \frac{1}{T}\sum_{j=1}^{T} F(x; \overrightarrow{\theta_j})^{n_j}$$

T = # sub-periods over which n and θ are estimated

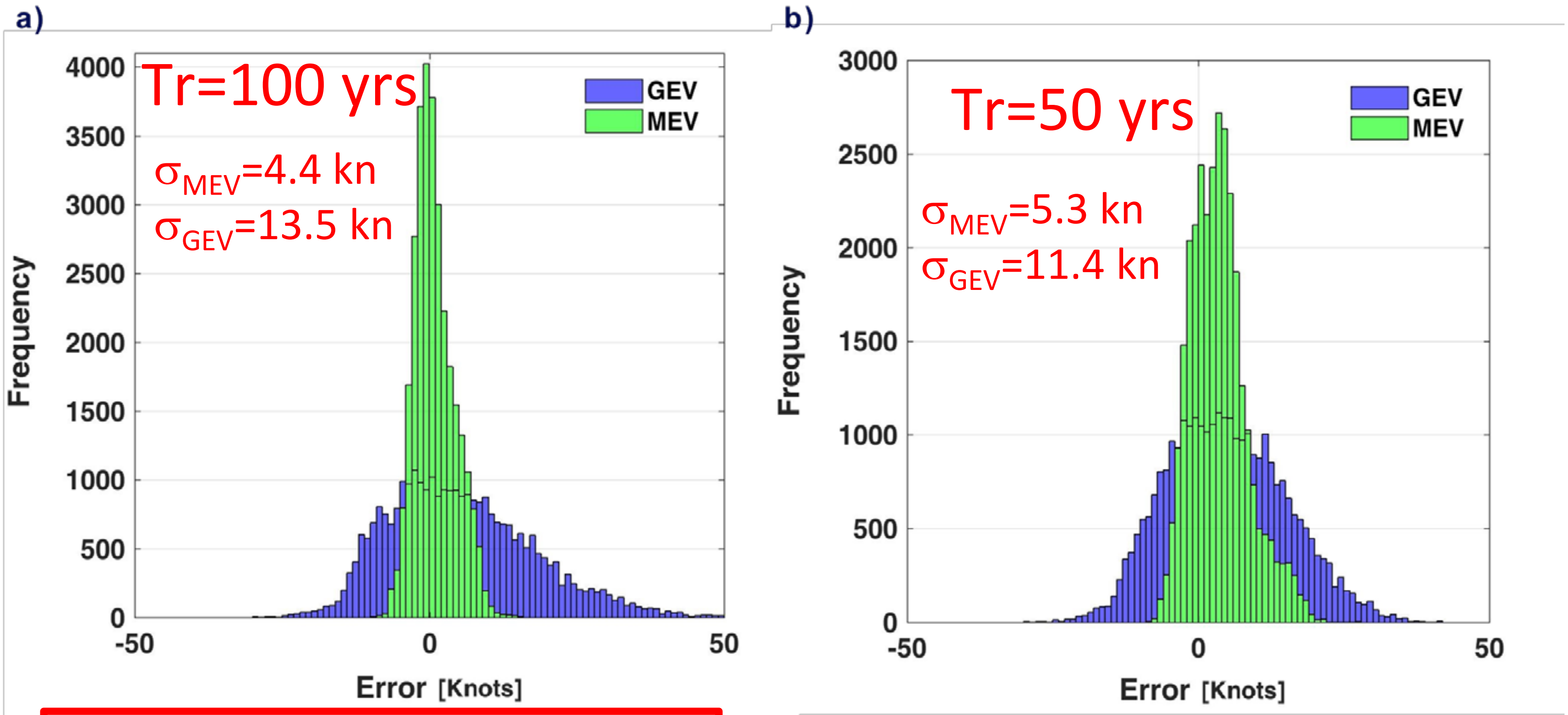
We find that the Generalized Pareto Distribution (GPD) best describes 'ordinary' Hurricanes:

$$\zeta(x) = \frac{1}{T}\sum_{j=1}^{T}\left[1 - \left(1 - k_j \frac{(x - \mathrm{u})}{\sigma_j}\right)^{\frac{1}{k}}\right]^{n_j}$$

Hosseini et al., submitted, 2019

Hurdat 2 analysis (cross validation with independent randomized calibration and test datasets).

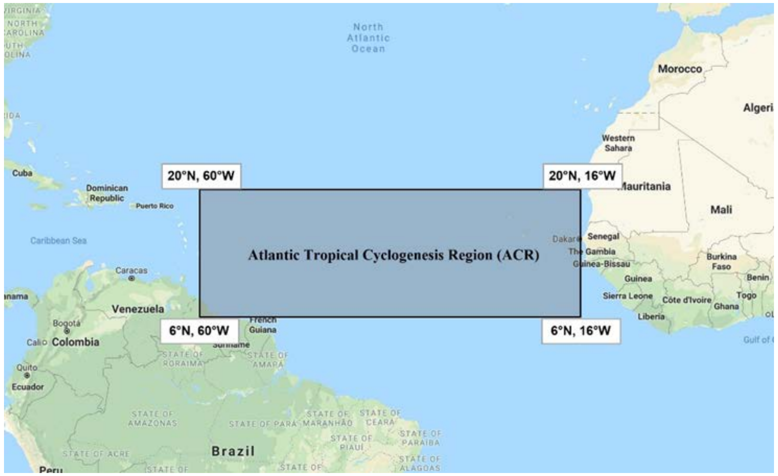
Extreme Hurricanes: MEVD (GPD) vs. traditional GEV

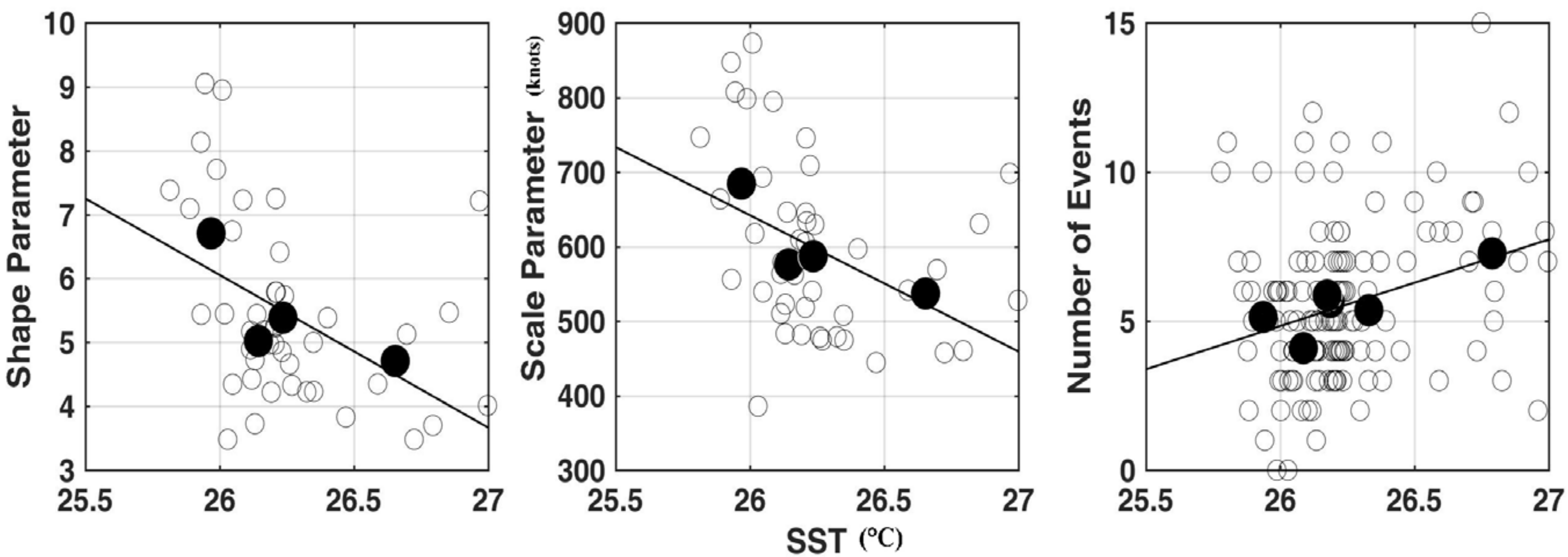


50-70% Improvement in estimation uncertainty

Hosseini et al, submitted, 2019

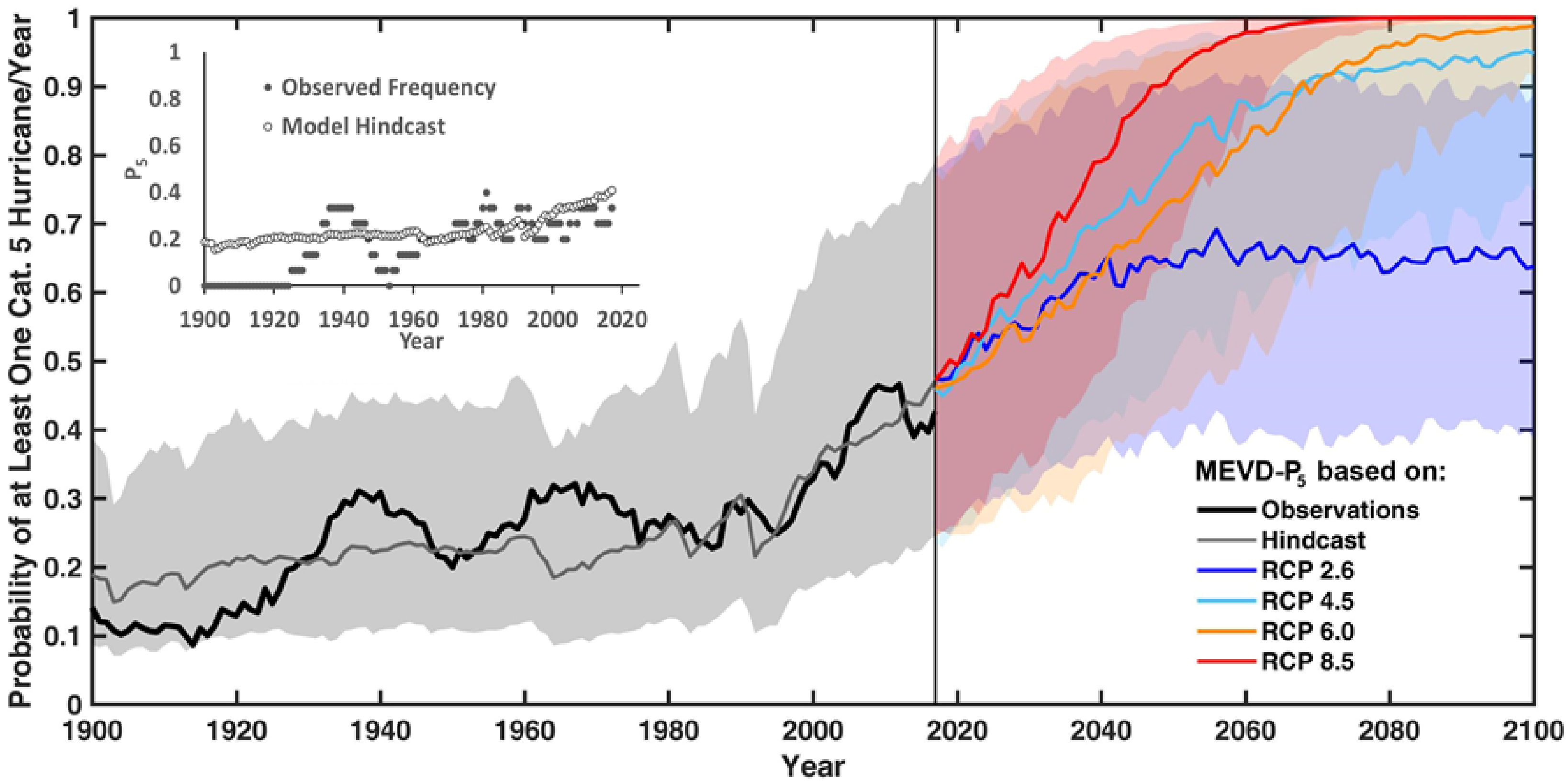
GPD parameters and # hurricanes/year as a function of Sea Surface Temperature (SST in the ACR)





Hosseini et al, submitted, 2019

Hindcast and projection of extreme-hurricane probability



Hosseini et al, submitted, 2019

Conclusions (3/3):

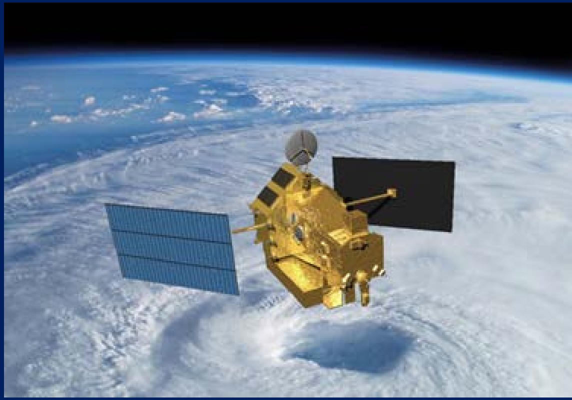
- Generalized Pareto Distribution of ordinary hurricane intensity;

- MEVD-GPD improves estimation uncertainty by 50%-70% in the case of atlantic hurricane intensity.

- MEVD can be successfully used to infer the effects on changes in ordinary evets on the distribution of large extremes

Summary:

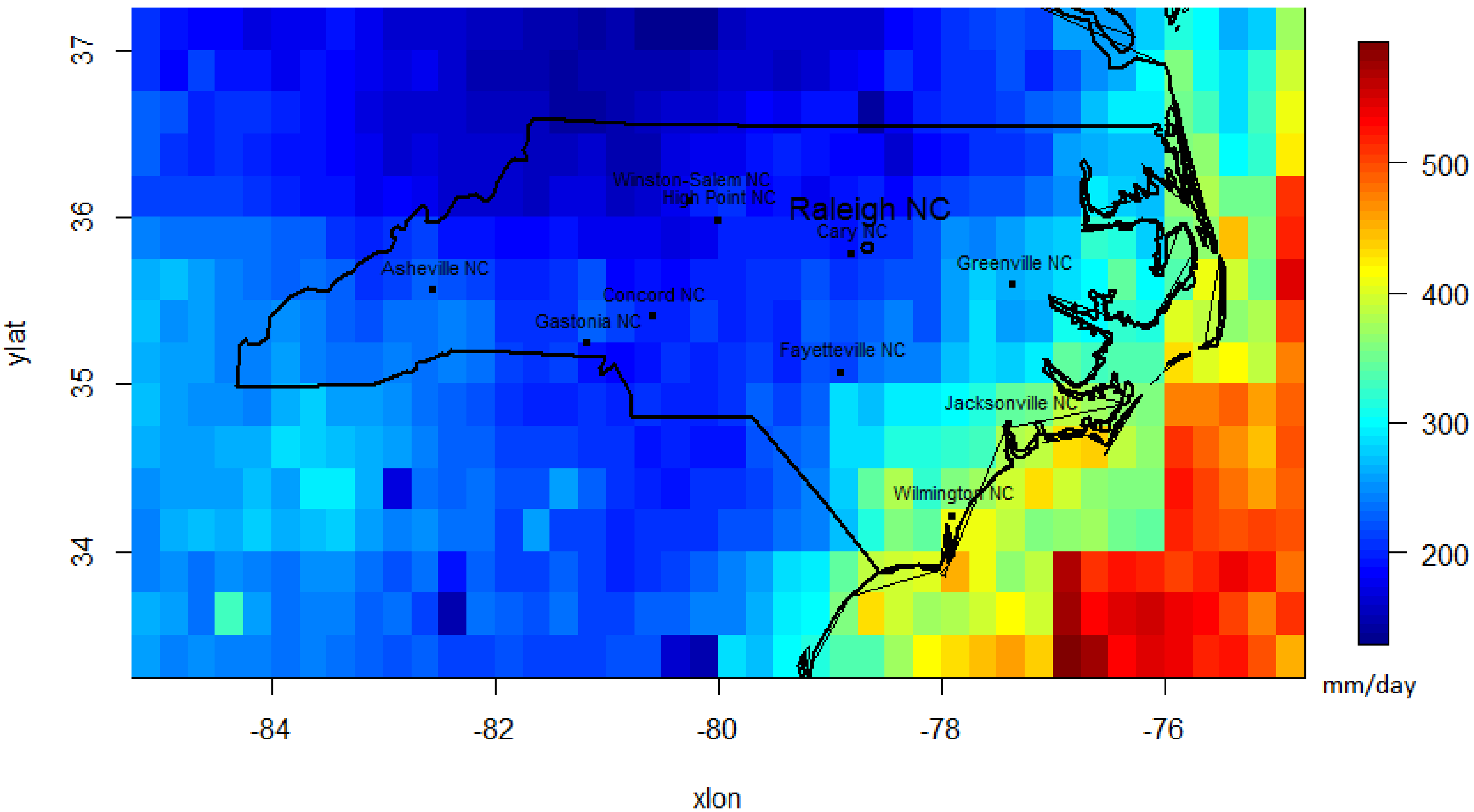
MEVD shows practical and conceptual advantages over traditional EVT (rainfall, floods, storm surges, wind)



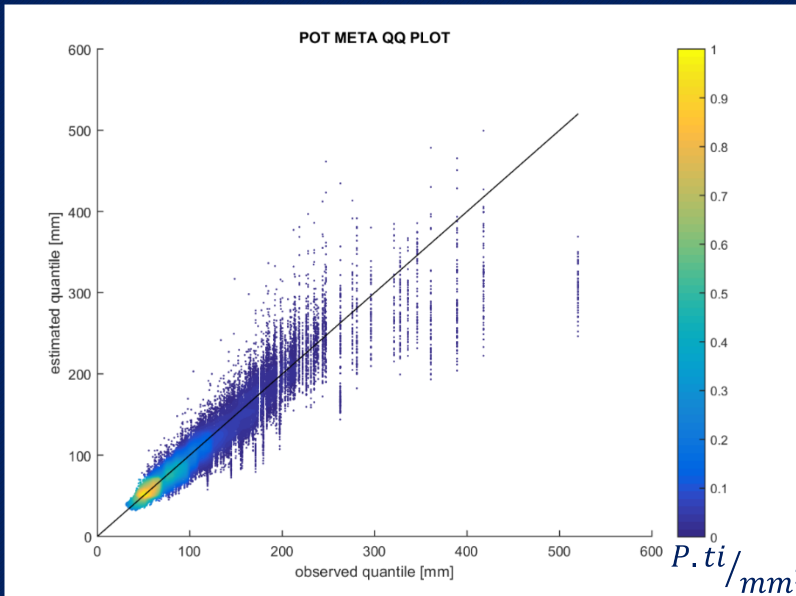
- links extremes to “ordinary” events: large samples, interannual variability/trends, physical processes

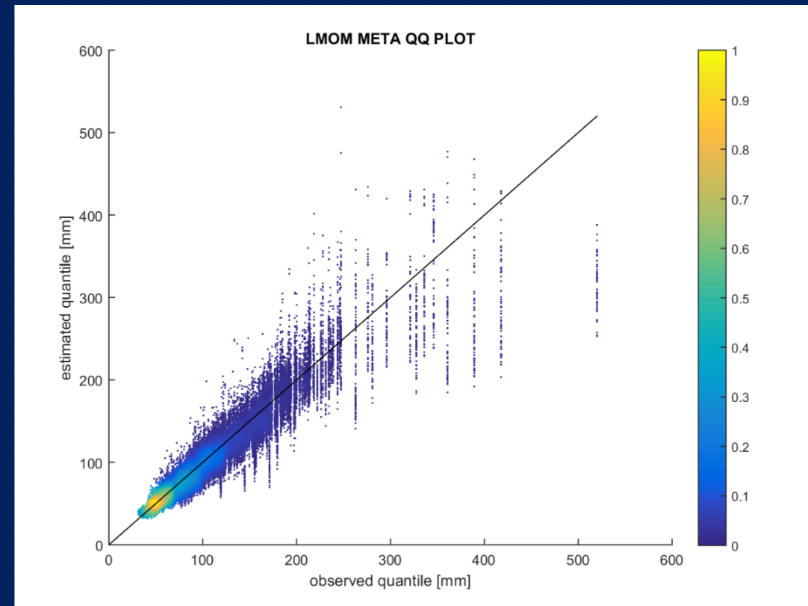
Time to move beyond limiting distributions and Poisson arrivals…

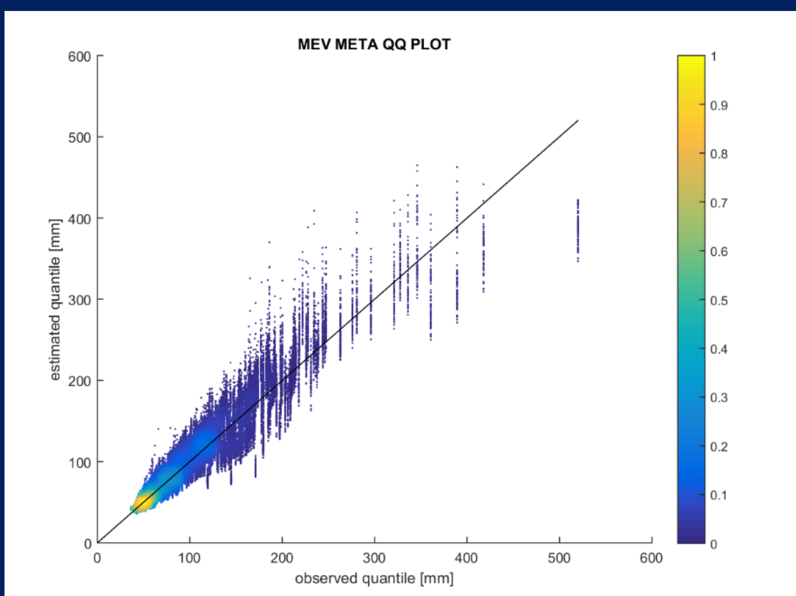
Tr=100 daily rainfall from TRMM observations (17 yrs)



Global QQ plots



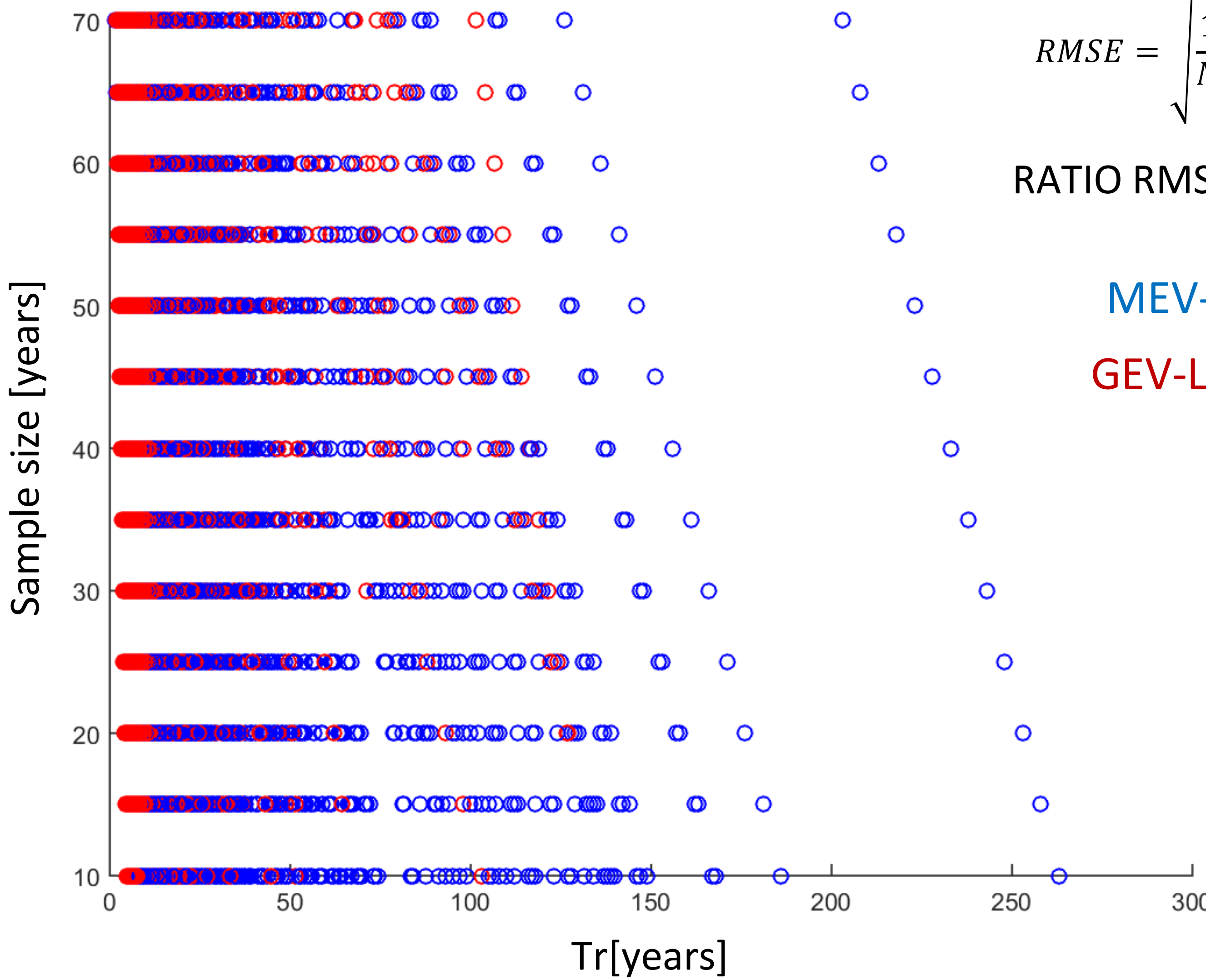




GEV/ POT are a good fit for the calibration sample but they fail in describing the stochastic process from which the sample has been generated

MEV allows a better description of the underlying process; less variance in high quantile estimation

RMSE ratios for several Sample sizes and return times

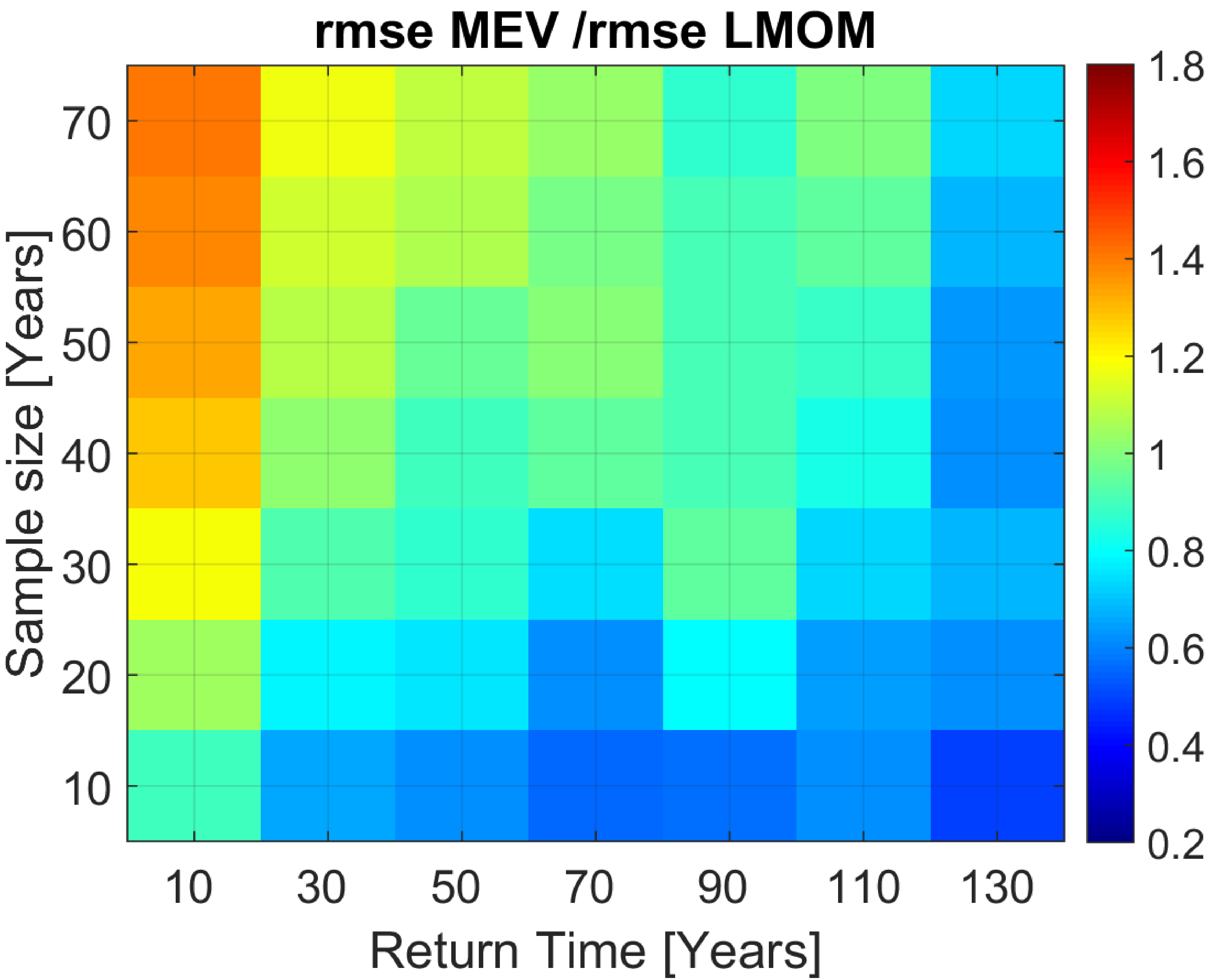


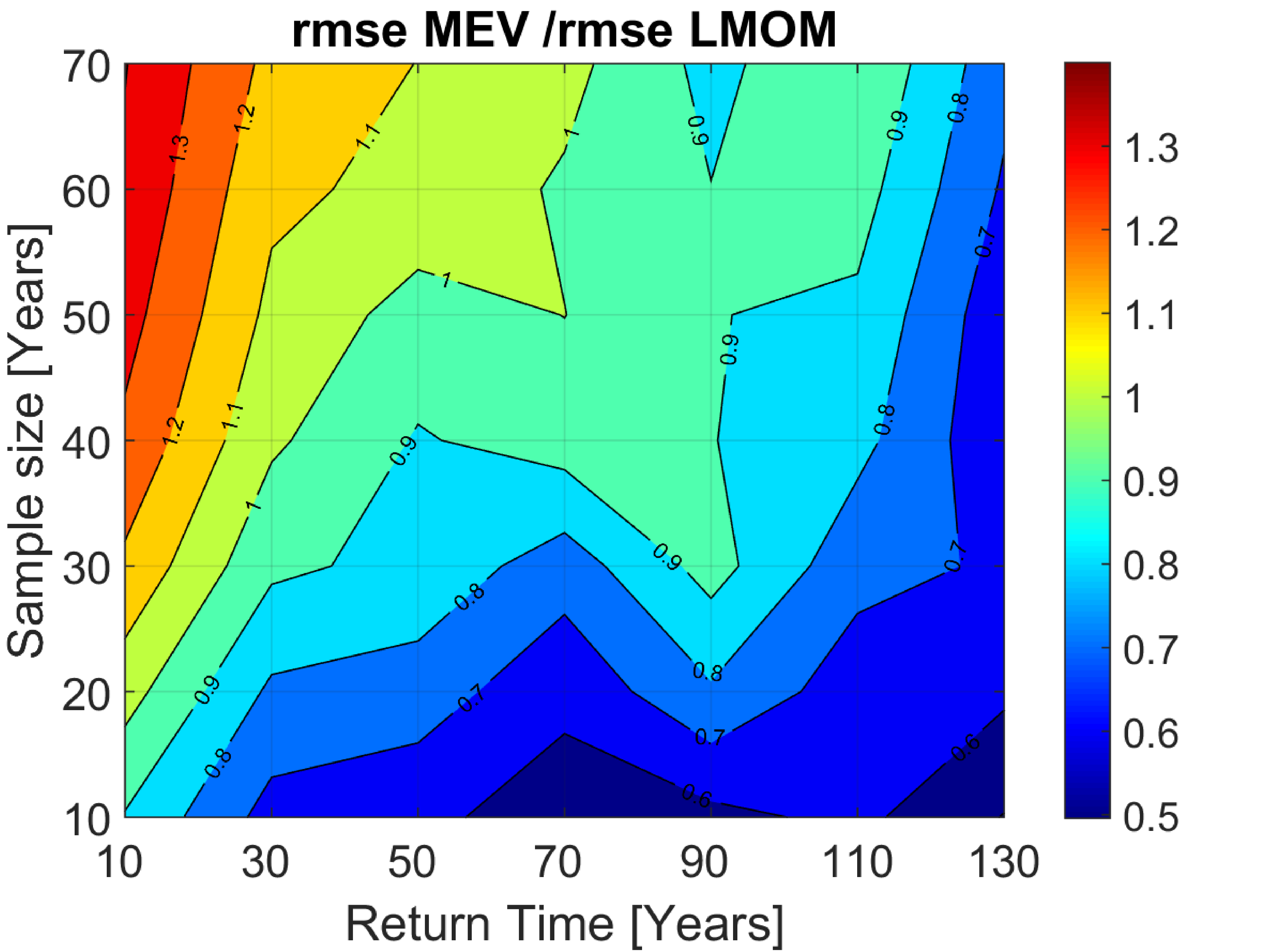
$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\hat{x} - x_{obs}}{x_{obs}}\right)^2}$$

RATIO RMSE MEV/ LMOM

MEV-PWM

GEV-LMOM


rmse MEV /rmse LMOM
Sample size [Years]
70
60
50
40
30
20
10
Return Time [Years]
10
30
50
70
90
110
130
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2


rmse MEV /rmse LMOM
Sample size [Years]
Return Time [Years]
10
30
50
70
90
110
130
20
40
60
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5

[Milly et al, Science, 2008]

POLICY FORUM

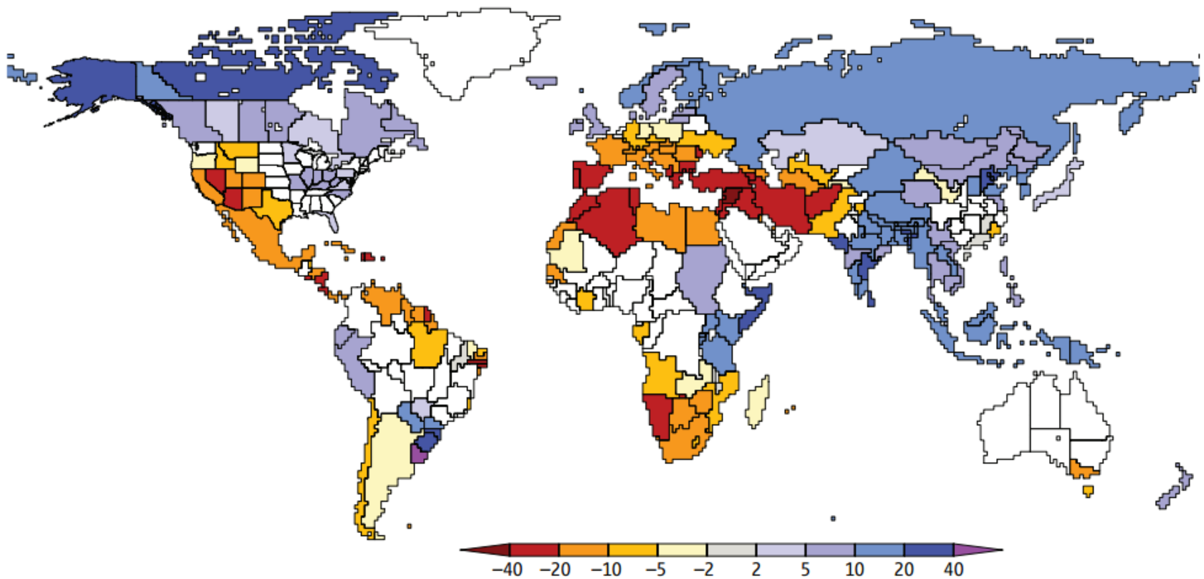
CLIMATE CHANGE

Stationarity Is Dead: Whither Water Management?

Climate change undermines a basic assumption that historically has facilitated management of water supplies, demands, and risks.

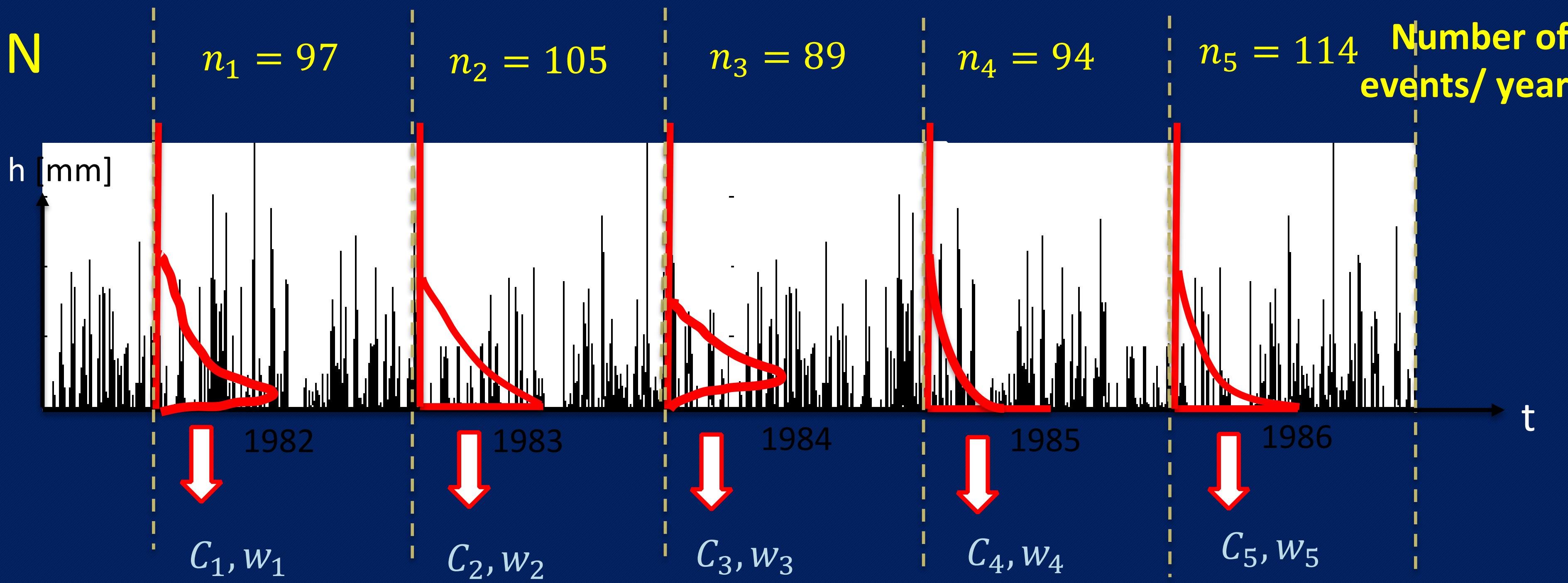
P. C. D. Milly,[1*] Julio Betancourt,[2] Malin Falkenmark,[3] Robert M. Hirsch,[4] Zbigniew W. Kundzewicz,[5] Dennis P. Lettenmaier,[6] Ronald J. Stouffer[7]

Systems for management of water throughout the developed world have been designed and operated under the assumption of stationarity. Stationarity—the idea that natural systems fluctuate within an unchanging envelope of variability—is a foundational concept that permeates training and practice in water-resource engineering. It implies that any variable (e.g., annual streamflow or annual flood peak) has a time-invariant (or 1-year–periodic) probability density function (pdf), whose properties can be estimated from the instrument record. Under stationarity, pdf estimation errors are acknowledged, but have been assumed to be reducible by additional observations, more efficient estimators, or regional or paleohydrologic data. The pdfs, in turn, are used to evaluate and manage risks to water supplies, water-



The MEV distribution

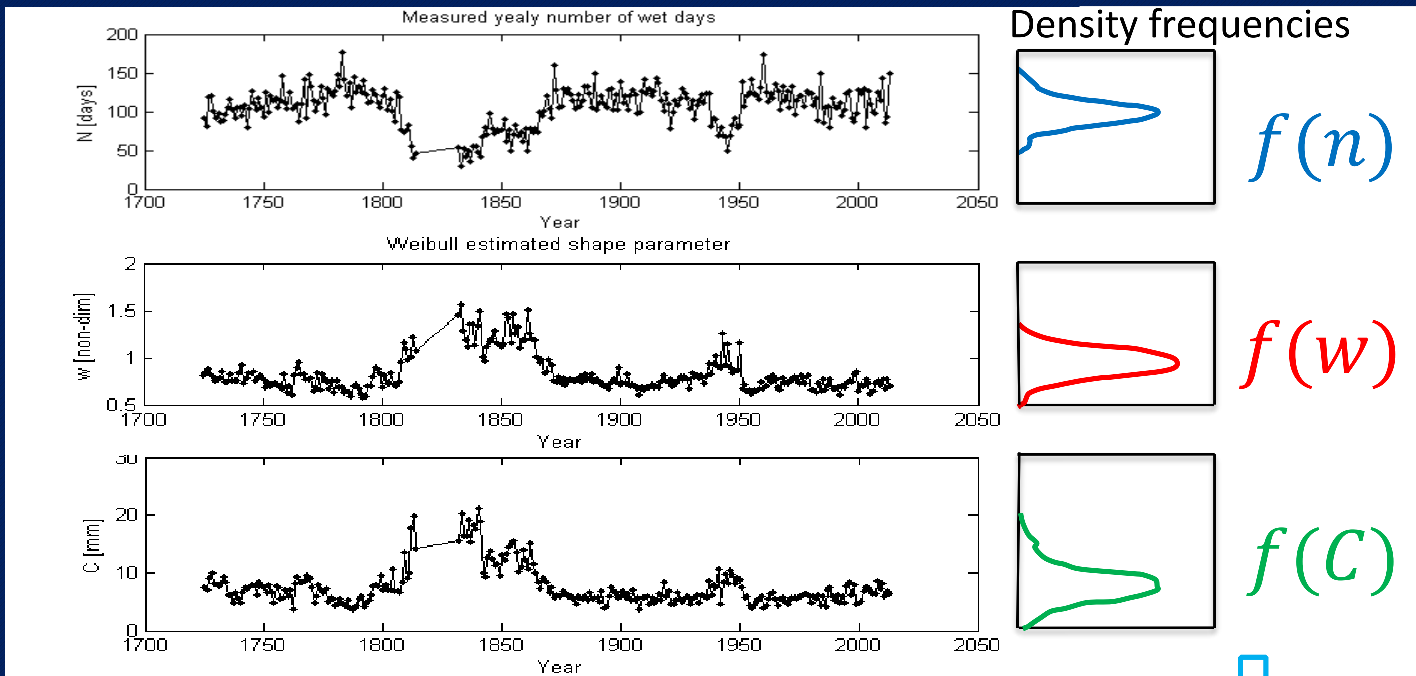
- **Assuming Weibull as a pdf for daily rainfall**
- **Fit performed using Probability Weighted Moments (Greenwood et al, 1979)**



$$F(x) = 1 - e^{\left(\frac{x}{C}\right)^{w}}$$

1. Sampling n from the distribution p(n|C,w)
2. Fit Weibull to the single *years* → C_i, w_i

The Metastatistical Extreme Value Distribution

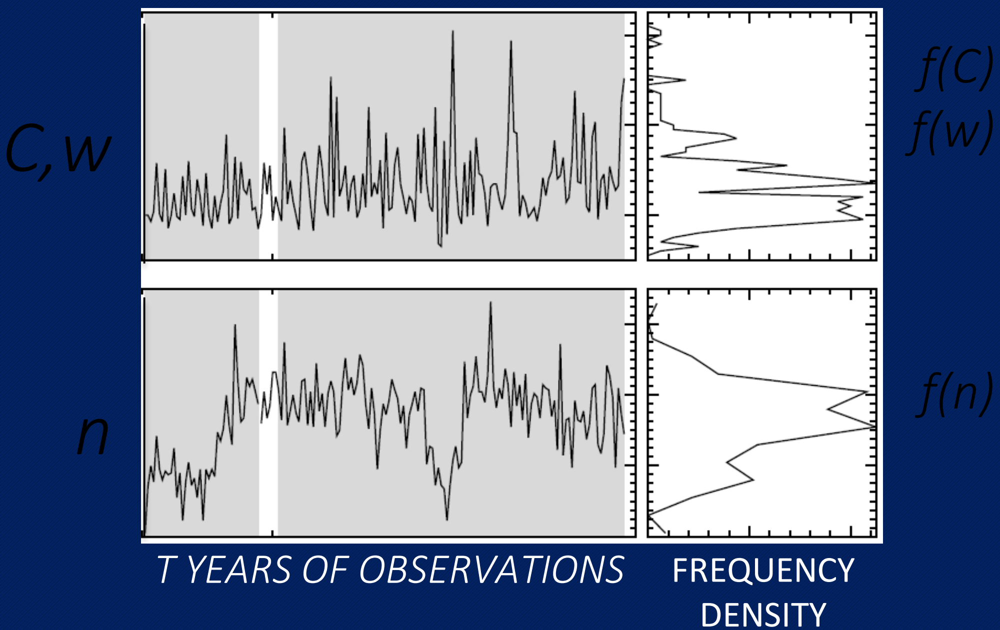


$$\zeta(x) = \sum_{n=1}^{\infty} \iint_{C\,w} g(n,C,w)\left[1 - e^{\left(\frac{x}{C}\right)^{w}}\right]^{n} dC\,dw$$

- Weibull parameters $\vec{\theta} = [C, w]$ and N are random variables themselves
- The CDF of annual maximum is the mean on all their possible realizations

A discrete expression for MEV

$$\int p(x) f(x) \cong \sum_i f(x_i) \quad if \quad x_i \sim p(x)$$



$$\zeta(x) \cong \frac{1}{T} \sum_{j=1}^{T} \left\{1 - e^{\left(\frac{x}{C_j}\right)^{w_j}}\right\}^{n_j}$$

Instead of an analitical model $for\ g(n, C, w)$ we use the observed frequencies

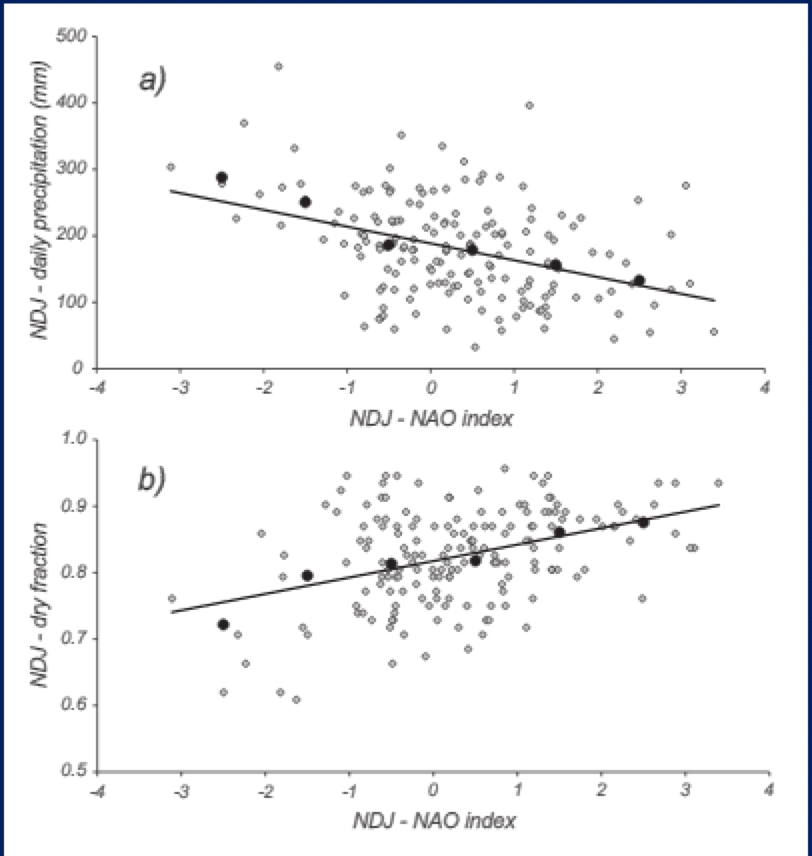
An alternative approach consists in semplifying the MEV espression (e.g. Hypothesizing independence among N,C,w:

$$\zeta(x) = 1 - \sum_n n \cdot p(n) \cdot e^{\left(\frac{x}{C}\right)^{w}}$$

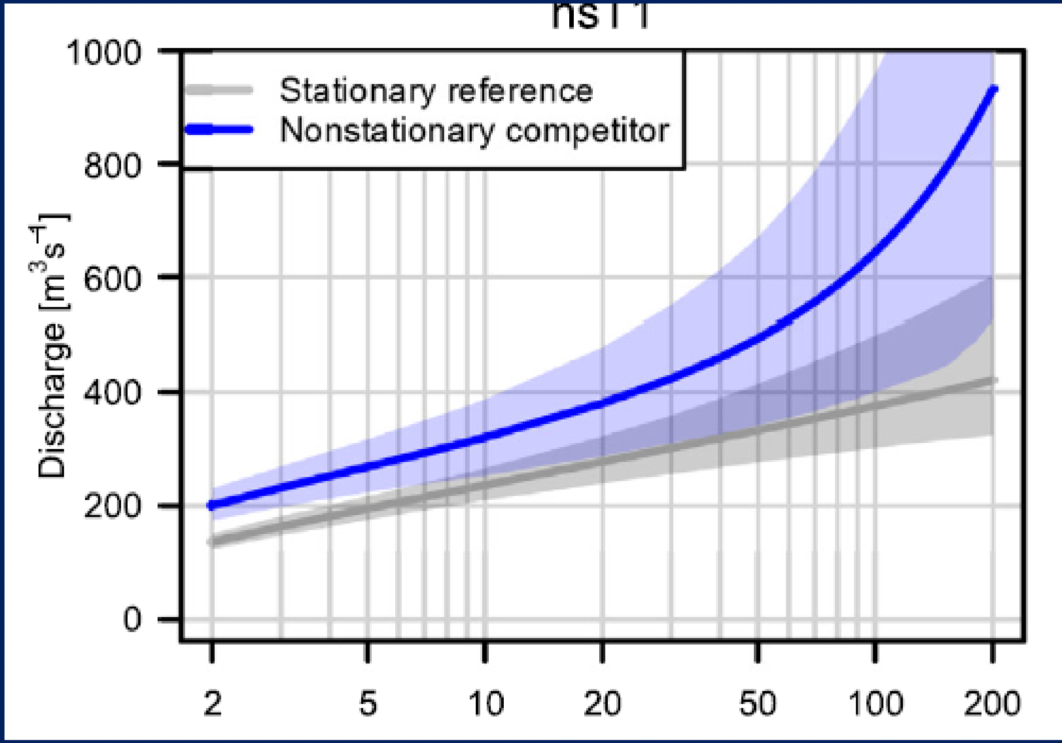
Penultimate Approximation

What to do next?

Nonstationary EV analysis: Dependence of the parameters {N,C,w} on climatological parameters (downscaling)



[Marani and Zanetti, 2014]

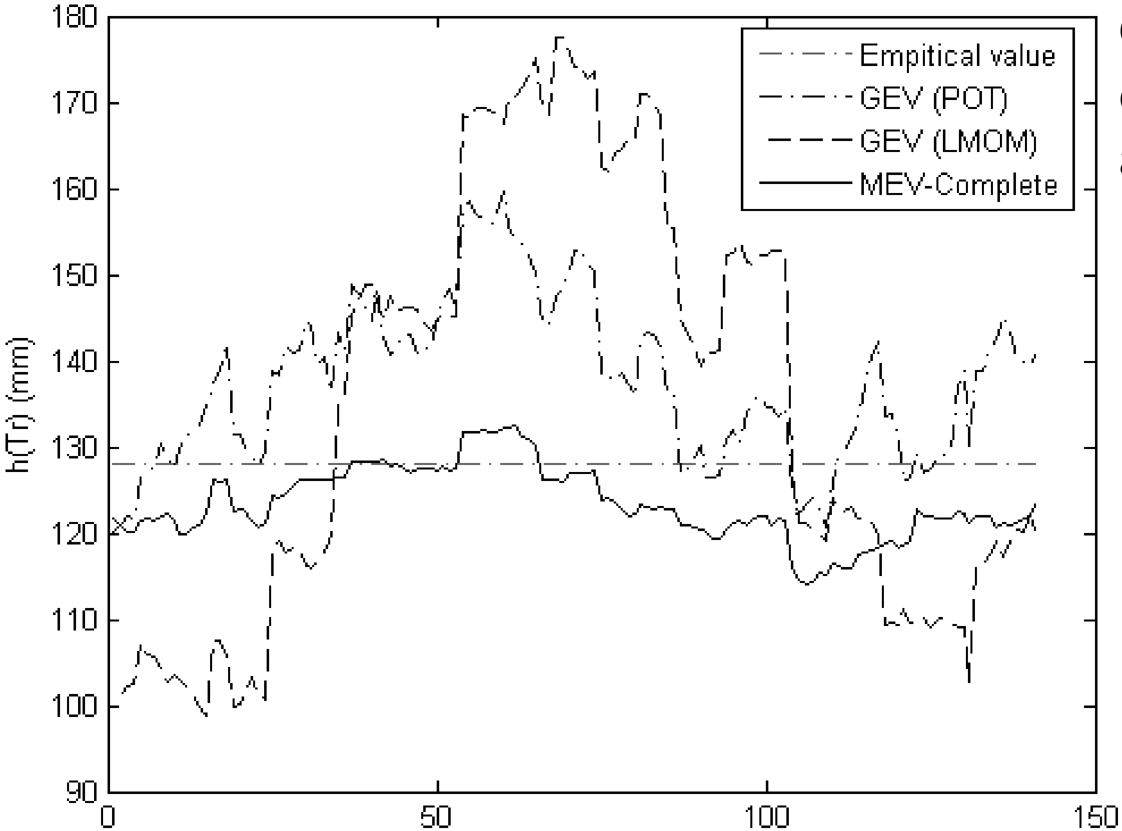


[Serinaldi and Kilsby, 2015]

Non stationary analysis

Bologna (Italy) randomly reshuffled time series
Sliding and overlapping windows analysis

Tr=100 years



i-th window

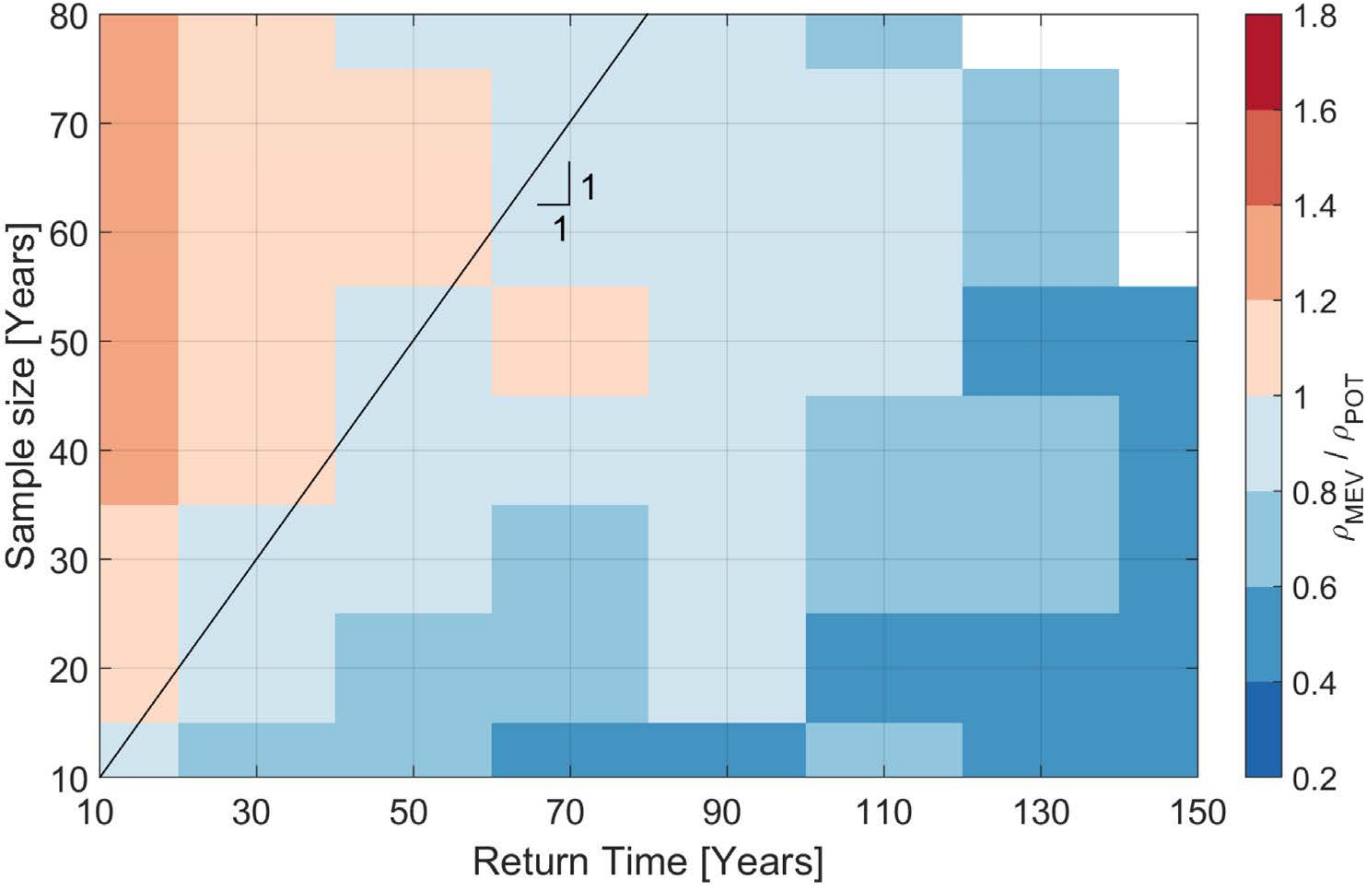
GEV and POT estimated quantil es show oscillations with same amplitude

⇩

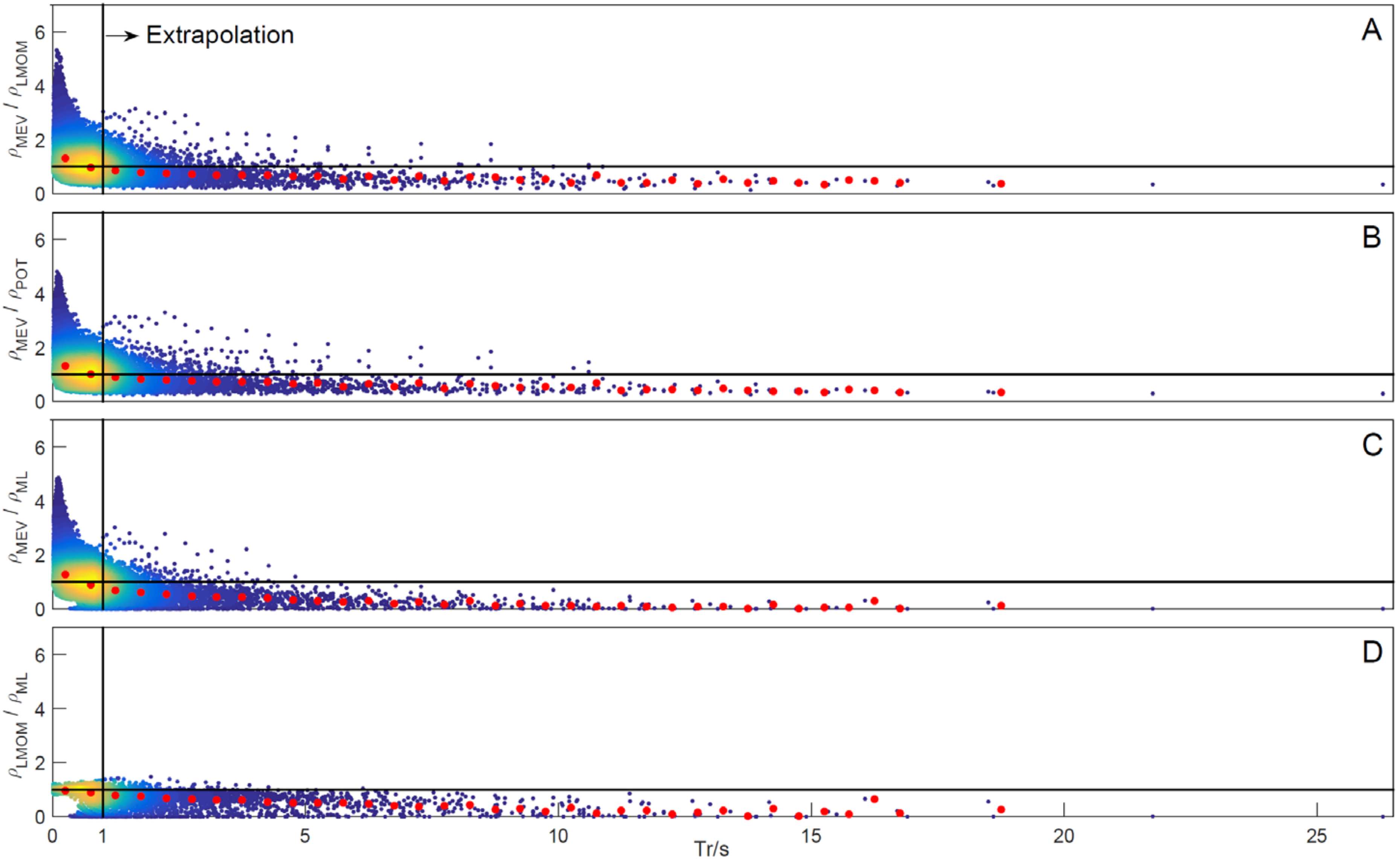
Due to the variance in the parameter estimates

In the case of MEV the variance of estimated quantiles is much smaller; Stationary behaviour

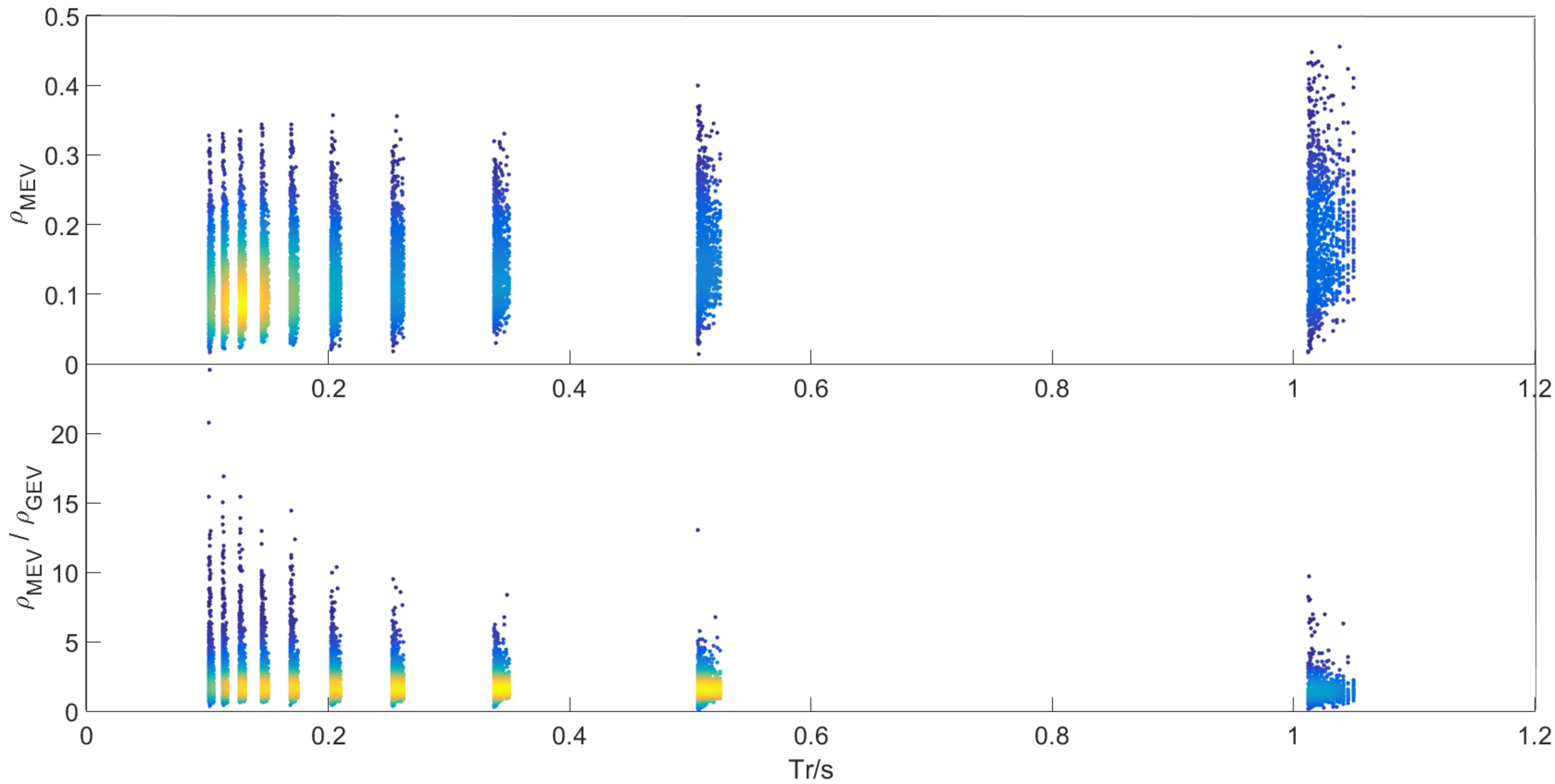
Ratio of MEV estimation error to GEV-POT error



Estimation error as a function of Tr/(sample size)



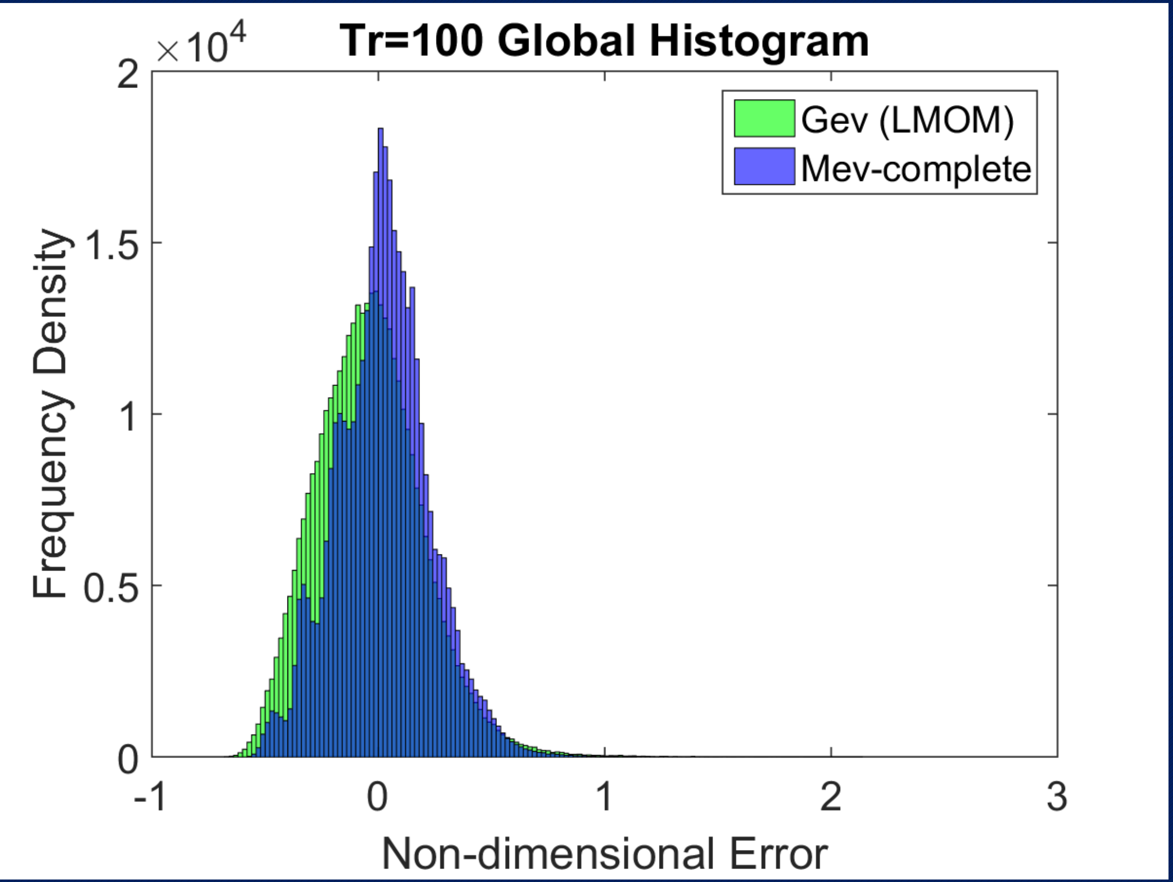
Performance when testing sample = calibration sample



Error distribution

Distribution of the error computed over 1000 random reshuffling, for all the analyzed datasets.

Quantiles (Tr=100 yrs) estimated by GEV, POT, MEV calibrated over 30-years samples

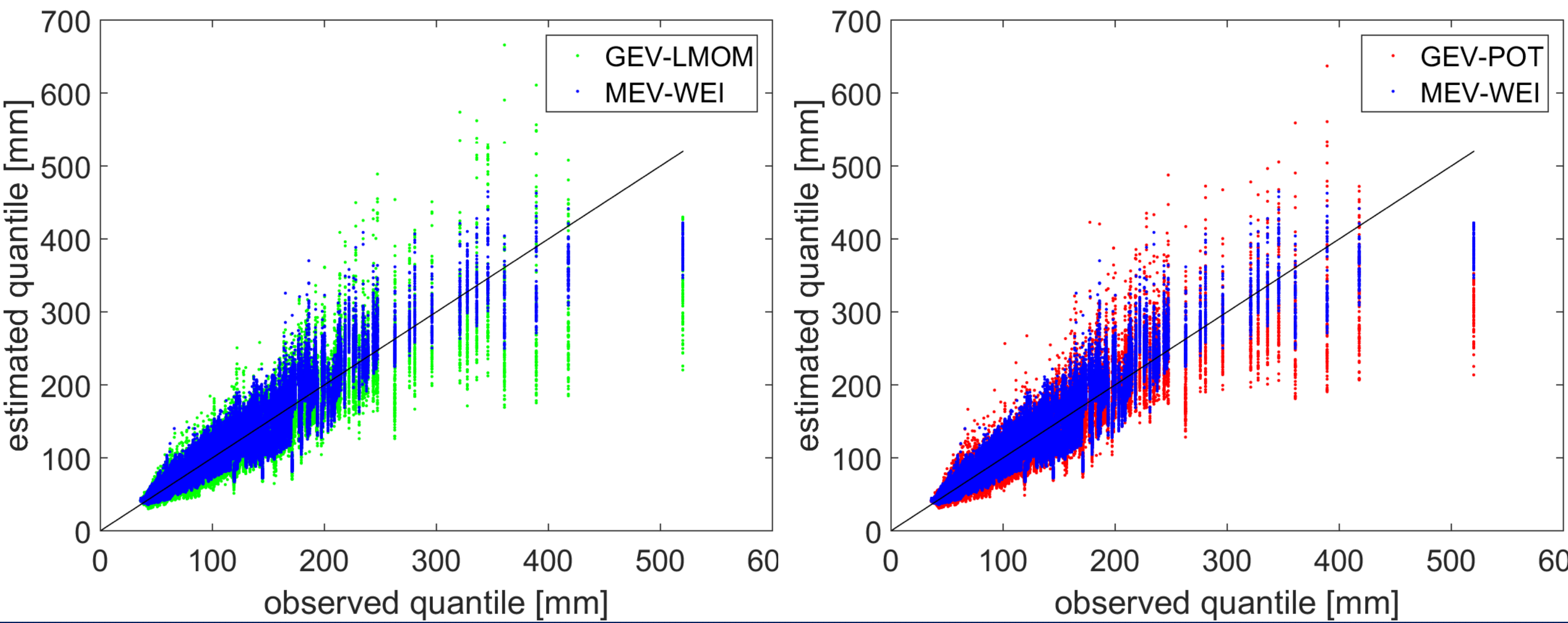


$$\epsilon = \frac{\hat{x} - x_{obs}}{x_{obs}}$$

$$\hat{x} = F^{-1}\left\{1 - \frac{1}{Tr_i}\right\}$$

x_{obs} from the observational (independent) sample

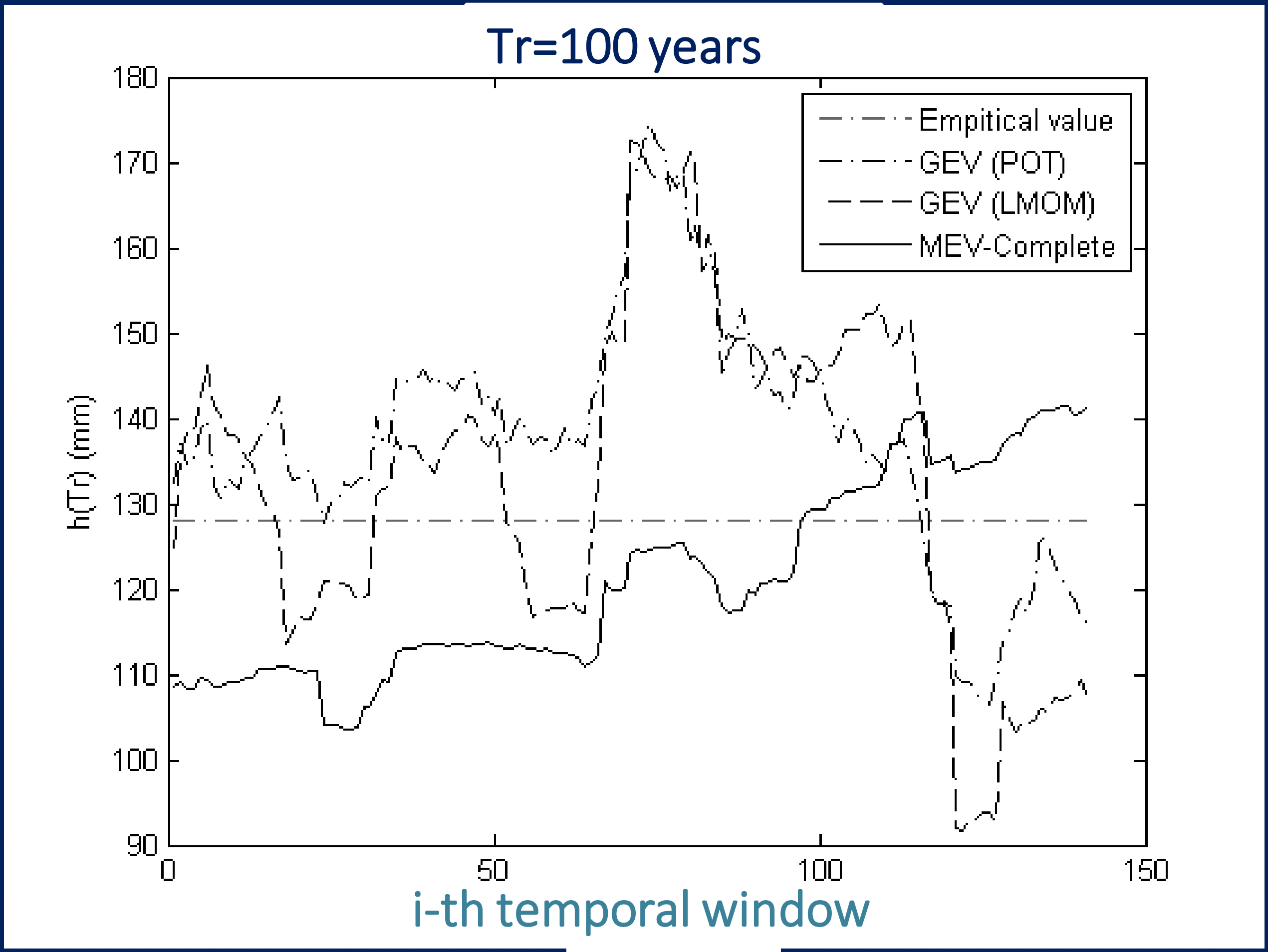
Global QQ-Plots



Sample size=45 years 100 random reshuffling

Some thoughts on non stationarity

Bologna (Italy) original 180 years time-series
Sliding and overlapping windows analysis



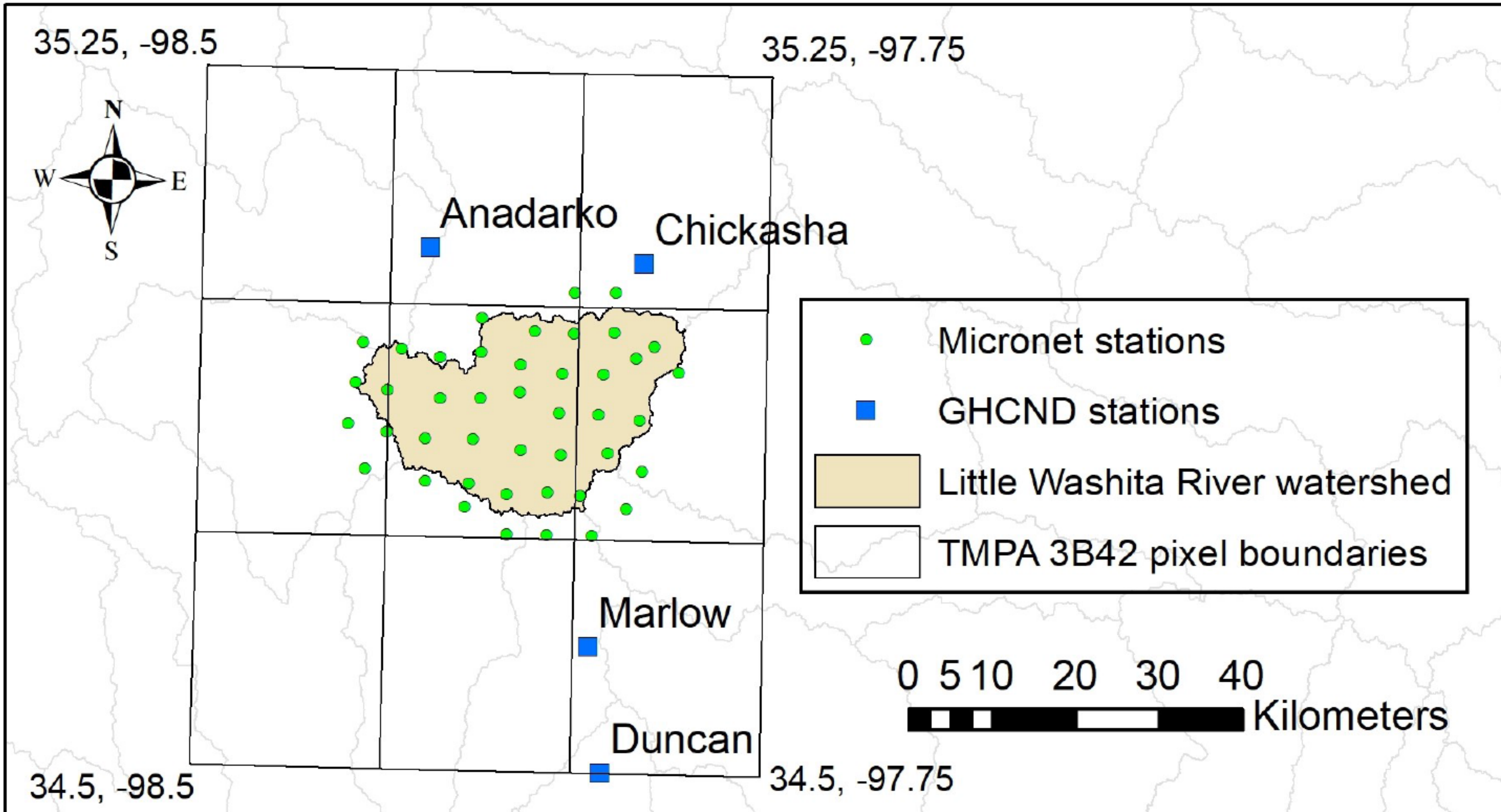
GEV and POT estimated q uantile shows higher vari ance

MEV shows a positive trend in est. quantiles

⇩

Due to trend in the par ameters of Weibull **C, w and n**

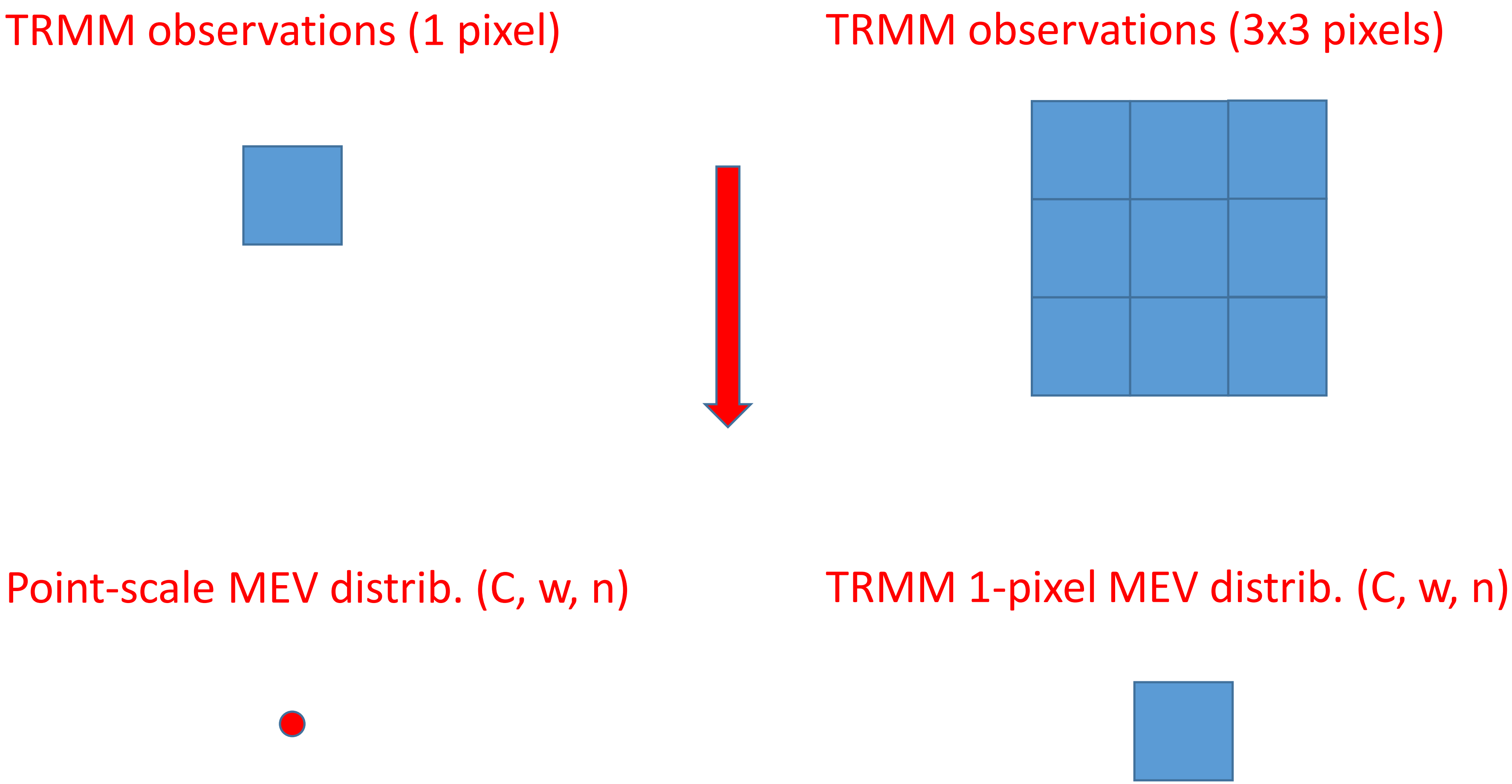
A case study over Oklahoma



[Elliot et al., 1993]

Zorzetto and Marani, WRR, 2018

Parameters of rainfall pdf downscaled using closed-form relations from space-time stochastic rainfall models (Marani, 2003)



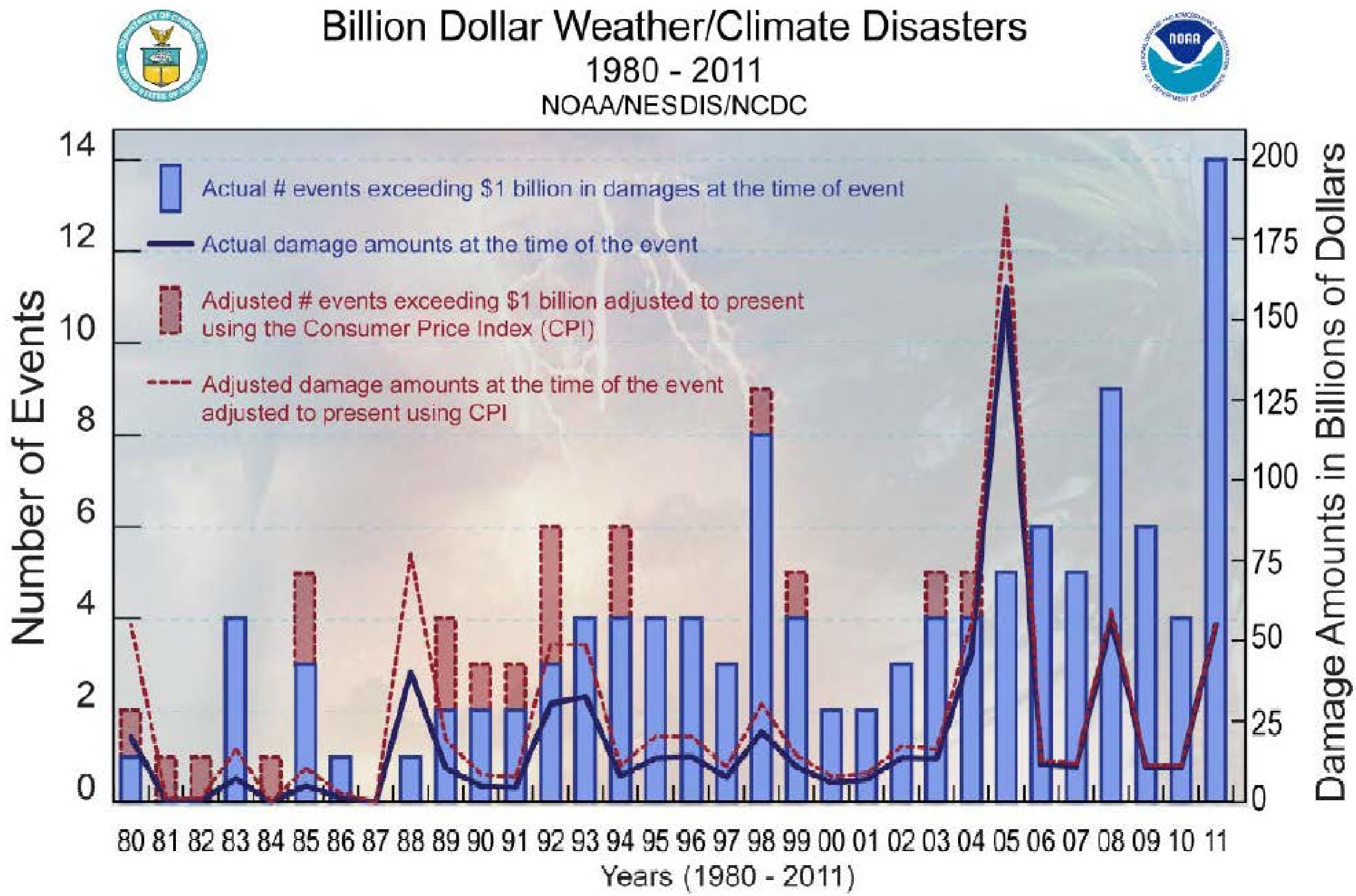


Fig. 1 US Billion-dollar Weather and Climate Disaster time series from 1980-2011 indicates the number of annual events exceeding $1 billion in direct damages, at the time of the event and also adjusted to 2011 dollars using the Consumer Price Index (CPI)

Smith and Katz, 2013