Comparison of IDF Estimation Methods

Daniele Veneziano[1], Chiara Lepore[2],
Andreas Langousis[1], and Pierluigi Furcolo[2]

[1]Dept. of CEE, MIT, U.S.A.
[2]Dept. of CE, U. of Salerno,Italy

IDF Estimation Procedure

i(d,T) = average intensity in *d* with return period *T* (yr)

- **From Annual Maxima** *i(d,T) = upper 1/T quantile of* $I_{max}(d,T)$

- **From POT values** $T(i,d) = \dfrac{1}{\lambda(i^*) \cdot P[\,POT(d,i^*) > i - i^*]}$, *i* =threshold*

- **From Marginal Distribution** $\begin{cases} F_{I_{\max}(d)}(i) = [F_{I(d)}(i)]^{1/d} \\ i(d,T) = \text{see annual max} \end{cases}$

- **From *I(t)* process**

- **Hybrid** – *Combine marginal and annual -max information*

Other IDF Estimation choices

- **Parameterization of *i(d,T)***
  - Dependence on *T* ? $\Rightarrow$ Quantiles of I_{max}, POT, I
  - Dependence on *d?* Smooth distribution parameters?

    Smooth quantiles?

- **Parameter estimation procedure**

  MoM , PWM , ML , tail fitting …. ?

Evaluation Criteria

- *Bias, variance, RMS error*
- *Sensitivity to outliers*
- *Does the optimal method depend on d, T or record length D ?*

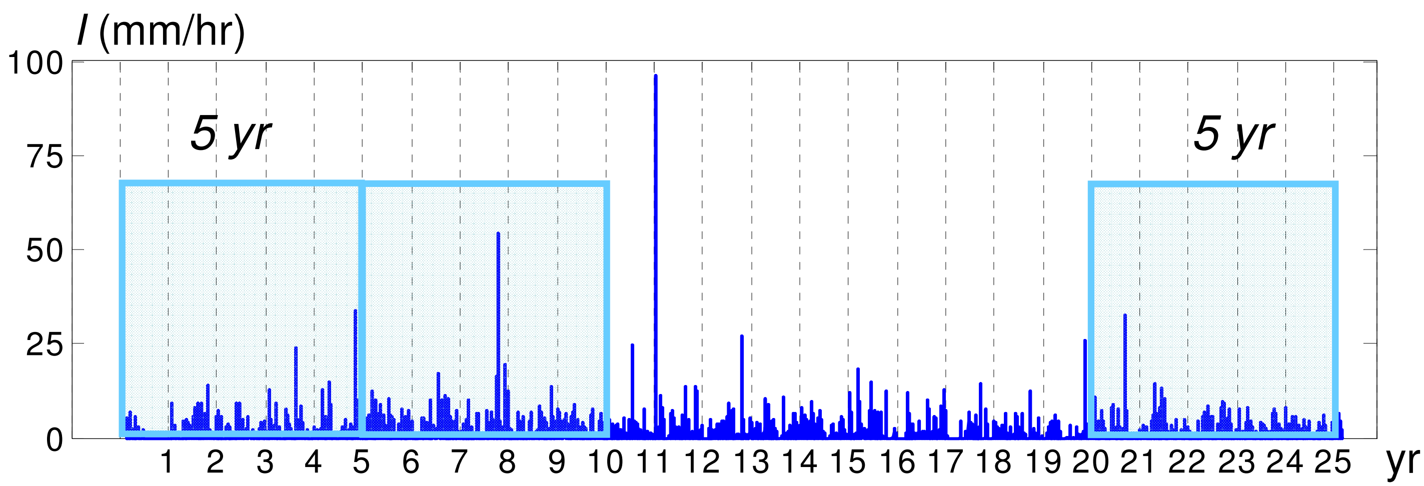
Methods Compared

Method		Parameters	Estimation Procedure
I_{max} *(AM)*	SPM	▪For each *d*: GEV parameters ▪For each *T*: smooth IDF values in *d* (pars. θ,η)	PWM for each *d* LS for each *T*
	CPM	▪Dependence on *d*: pars. θ,η ▪Dependence on *T*: pars. GEV	Koutzoyiannis et al. 1998
POT		▪For each *d*: threshold *i**, 3 GP pars.	PWM and LS
Marginal (MD)		▪For each *d*: 3 pars. for LN tail	MoM (1°, 2°, 3°)
I(t) process MF		▪K(q), 2 β-LN parameters, D_{max}	MoM (1°, 2°, 3°)
Hybrid (marginal+I_{max})		▪Same as marginal + calibration to mean annual max	

Data sets used for evaluation

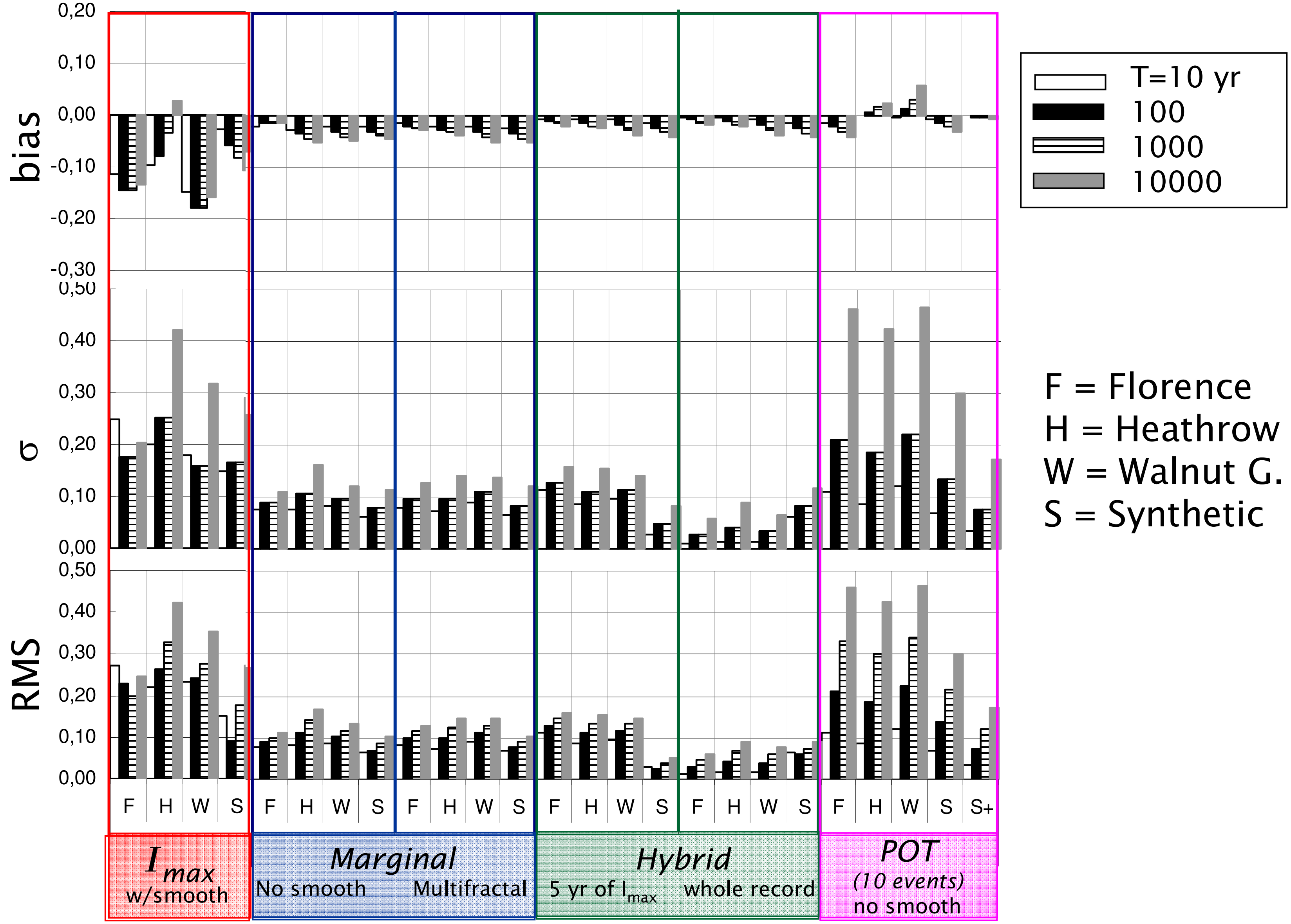
	n° years	mean Intensity mm/hr	rainy fraction %	mean I|I>0 mm/hr	*Comments*
Florence	24	0.087	7.4	1.18	outliers: 1966, (long *d*)
Heathrow	51	0.068	8.7	0.78	outliers: 1959, 1970, (short *d*)
Walnut Gulch	49	0.035	2.3	1.53	-
Synthetic	1000	0.102	9.6	1.07	β-LN cascade C_β = 0.46, C_{LN}= 0.06, D_{max} =9.56 *days*

Bias and Variance Comparison Using Subsets of the Record

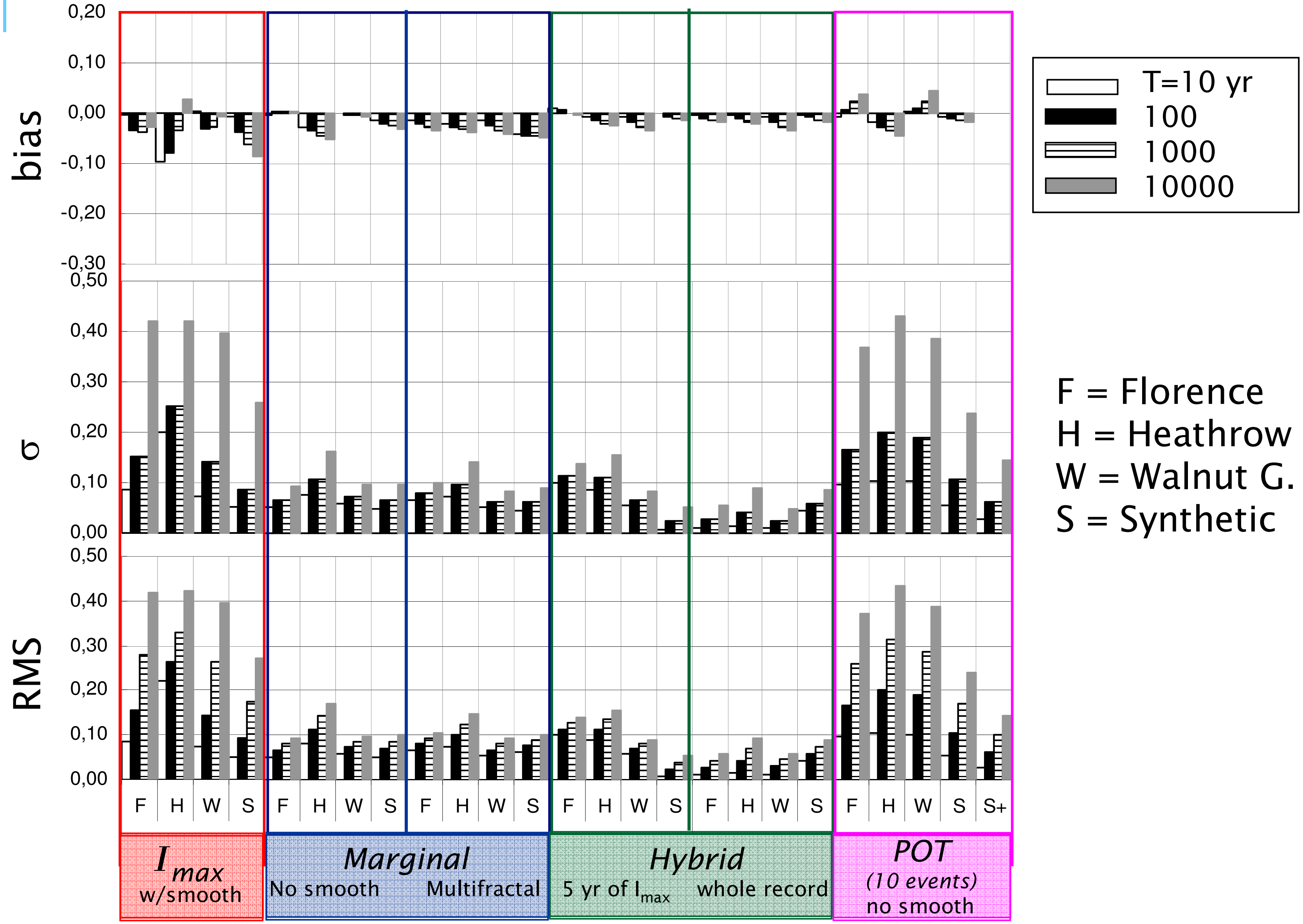


Subsets extraction → IDF estimation → BIAS, Variance, RMS error relative to the IDF from whole dataset

Bias, Variance, RMS error -1hr



Bias, Variance, RMS error -24 hr



Conclusions

➔ ***Bias, variance, RMS error, outliers sensitivity:*.**

- Marginal or Hybrid < POT < Annual-max

➔ ***Does the optimal method depend on d, T or record length D?***

- If I_{max} is available for many years $\rightarrow$ Hybrid is best;
- Expected dependence on *D*:
  - Annual-max best for very long *D*;
  - POT competitive for Intermediate *D* (a few decades);
  - Marginal best for short *D* (a few years).

These trends are qualitatively true, but Marginal is generally best for all *D*

- Models that use shape pars. (GEV, GP) have large error variance (for small D, large T) benefit from smoothing, but smoothing affect bias;
- Bias/σ are similar for all methods when T≈D; increasing T enhances error variance, whereas bias is less variable.

Thank you for your attention!

Aknowledgments

Università degli Studi di Salerno

M.I.T. Portugal Project