# Predictive performance of flood frequency analysis approaches: a national comparison based on an extensive French dataset

Benjamin Renard, K. Kochanek, M. Lang, P. Arnaud, Yann Y. Aubert, T. Cipriani, Eric Sauquet

# Predictive performance of flood frequency analysis approaches: a national comparison based on an extensive French dataset


**Benjamin Renard[(1)], Krzysztof Kochanek[(2)], Michel Lang[(1)], Patrick Arnaud[(3)], Yoann Aubert[(3)] , Thomas Cipriani[(1)] , Eric Sauquet[(1)]**

(1) Irstea, UR HHLY Hydrology-Hydraulics, Lyon, France

(2) Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland

(3) Irstea, UR OHAX, Aix-en-Provence, France


## Introduction: Flood Frequency Analysis (FFA)

**Importance of FFA in engineering**

Central in risk assessment and management:
- Design of civil engineering structures
- Inundation maps

**An abundance of approaches**
- Local estimation of a distribution
- Regional implementations
- Continuous simulation approaches

**Objectives**

Compare the predictive performance of FFA implementations
- Presentation of the comparison framework
- Application to an extensive dataset of French stations

## A data-based comparison framework [1]

### Spirit of the game

**General principles**

Focus is on **predictive** (as opposed to descriptive) performance => split-sample evaluation

*Would you rather build a dam that will withstand upcoming floods, or one that would have withstood past floods?*

Should be applicable to any FFA family (local, regional, mixed local-regional, continuous simulation)

**Complements (but not replaces!):**

Monte-Carlo evaluations

*The truth is known, but how realistic is the Monte-Carlo setup?*

Statistical testing

*Tests not available for many implementations*

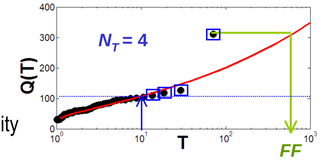
⚠ *General-purpose tests exist, but they assume known parameters!*

### Rules of the game

Compute on validation data:

$N_T$ : number of exceedances of the estimated T-year flood

$FF$ : non-exceedance probability of the largest observation




**For a reliable implementation:**

$N_T \sim Bin(n, 1/T)$, $pr(FF \le z) = z^n$

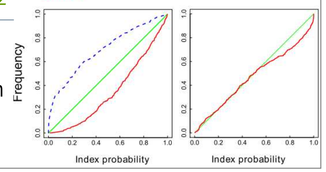
Repeat on many sites…

… and evaluate adequacy with the theoretical distribution

*Note: due to its discrete nature, a "randomization" trick is required for $N_T$*

*Tendency to over / under estimation*

*Predictive failures*

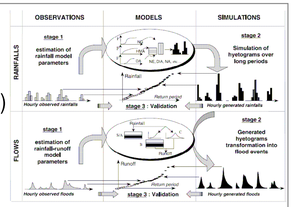



## Application: FFA implementations in France

### Competing teams

***The local league: using at-site data only***

1. Gumbel distribution (LOC_GUM)
2. GEV distribution (LOC_GEV)
3. A continuous simulation approach: SHYPRE [2]

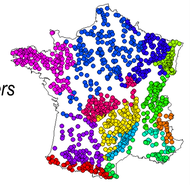



***The regional league: estimation at ungauged sites***

1. Gumbel distribution (REG_GUM)
2. GEV distribution (REG_GEV)
3. SHYREG, regionalized version of SHYPRE

*1-2: Region-specific regressions between parameters and covariates (Catchment size, 10-year rainfall, mean elevation, drainage density) [3]. Constant shape parameter for GEV*

**Hydro-eco-regions**



***The local-regional league: at-site estimation using regional information***

1. Gumbel distribution (L+R_GUM)
2. GEV distribution (L+R _GEV)

*Bayesian approach, prior = regional, likelihood = local* [4]

### Data

1076 stations, 20 years +

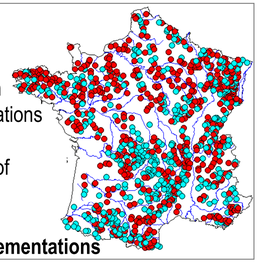
Catchment size: 10-2000 km²

**Calibration-validation decomposition**

Red: calibration of regional implementations

Blue (>40 years):
- 20 years (random) = calibration of local implementations
- All remaining years = validation

**Validation data identical for all implementations**



## Results: comparison of FFA implementations in France

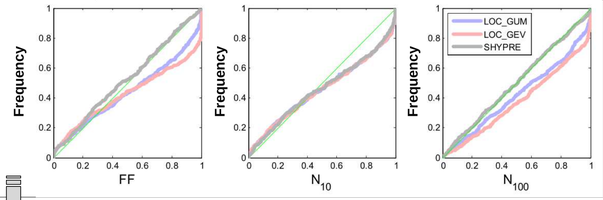
### The local league

SHYPRE is the most reliable

*Especially for high quantiles, cf. indices FF and $N_{100}$*

Poor performance of locally-estimated GUM/GEV

*Especially GEV: for 20% of stations, what is considered as "impossible" actually occurs in validation!*



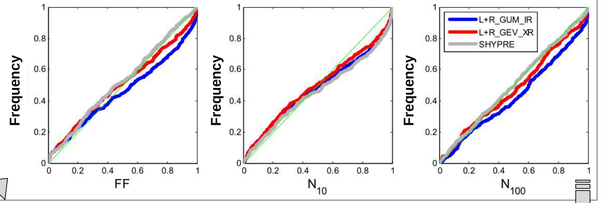

### The local-regional league

All implementations fairly reliable

*SHYPRE ~ L+R_GEV*

*L+R_GUM less reliable*

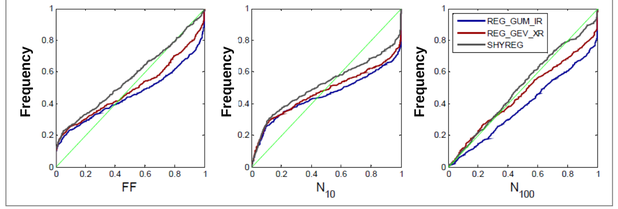
In particular, note the clear improvement for the GEV distribution, compared with both purely local and purely regional estimations




### The regional league

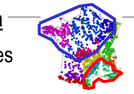
Reliability is poor for all implementations

SHYREG most reliable, REG_GUM least reliable



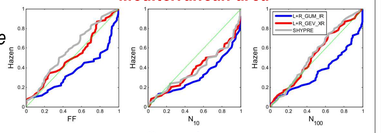

### Results by climatic area

Marked regional differences



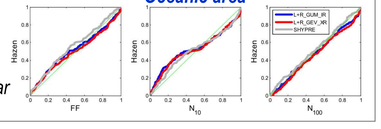
***Mediterranean area***

Mediterranean:
- *GUM clearly inappropriate*
- *SHYPRE ~ L+R_GEV*
- *Slight over-estimation for SHYPRE?*




***Oceanic area***

Oceanic:
- *No evidence that GUM is inappropriate*
- *All implementations similar*




## Conclusions

**Main conclusions**
- Two winners: SHYPRE and L+R_GEV
- The reliability of regional implementations is in general quite poor
- Purely local estimation of a GEV distribution is dangerous

**Using more information…**
- In general, purely local implementations are not sufficiently reliable
- Benefit of additional information: rainfall (SHYPRE) or regional (L+R_GEV)
- Perspective: combine more diverse sources of information


[1] Renard, B., et al. **2013**. Data-based comparison of frequency analysis methods: A general framework, Water Resources Research, 49

[2] Arnaud, P., and J. Lavabre. **2002.** Coupled rainfall model and discharge model for flood frequency estimation, Water Resources Research, 38(6).

[3] Cipriani, T., T. Toilliez, and E. Sauquet. **2012.** Estimating 10 year return period peak flows and flood durations at ungauged locations in France, La houille blanche

[4] Ribatet, M., et al. **2006.** A regional Bayesian POT model for flood frequency analysis, Stoch. Environ. Res. Risk Assess., 21(4)



**Acknowledgments.** This work was funded by the French Research Agency (ANR) through the project EXTRAFLO (https://extraflo.cemagref.fr/)