Ground-truthing flood risk estimates from global models

Ignazio Giuntoli[1,2,*], Ilaria Prosdocimi[3], David M. Hannah[1]

[1] University of Birmingham, UK [2] Politecnico di Torino, Italy [3] University of Bath, UK * i.giuntoli@bham.ac.uk

RATIONALE

Projections of runoff from global models can be used for assessing future changes in hydrological extremes (e.g., Dankers et al., 2014; Prudhomme et al., 2014). The uncertainties that cascade along the different components of the modeling chain can be large and may undermine the utility of estimates (e.g. peak flows return periods) that are relevant for the engineering/water management practice.

We explore the reliability of these estimates over selected catchments of the conterminous United States. To this aim, we seek to ground-truth a global multi-model ensemble (MME) in its ability to simulate annual maxima (AMax). We thus compare the MME AMax historical runs to the observed Amax at corresponding streamflow gauges. Having fitted AMax distributions to an extreme value distribution we then assess how return periods change from past to future.

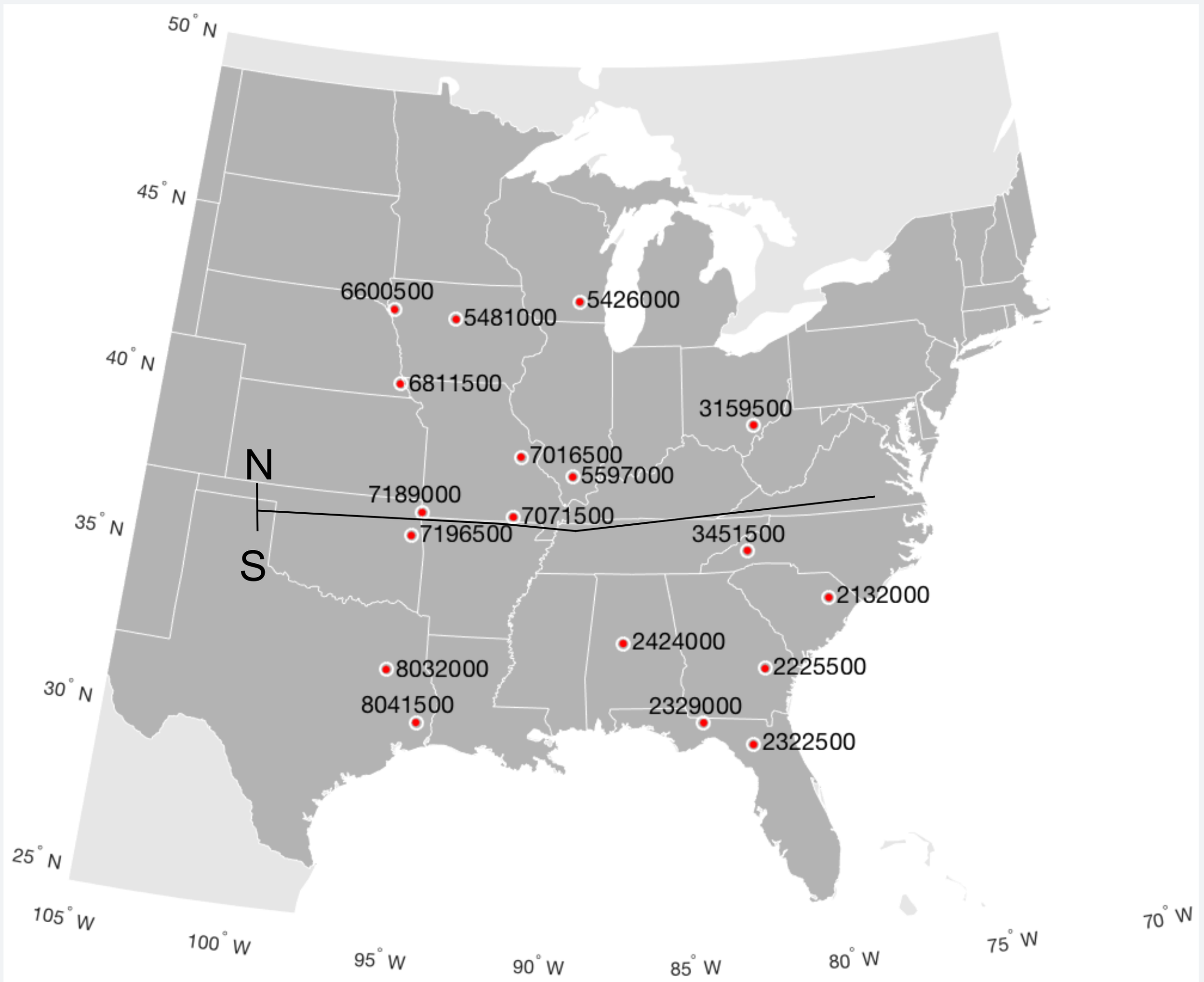
RESEARCH QUESTIONS

- Are simulated peak flows' distributions consistent with observed ones?
- How do simulated peak flows change in the future?

DATA

Simulated: Gridded (0.5 degree) daily runoff from 9 GIMs forced by 5 CMIP5 GCMs – from the ISIMIP Fast Track (Warszawski et al., 2014) – in the control period (1971-2005) and under RCP2.6 and RCP8.5 scenarios (2065-2099).

Observed: daily runoff [1970-2005] from 18 streamflow gauges (HCDN, low anthropogenic influence) from catchments of comparable size to that of the grid-cells.



Dankers, R., Arnell, N. W., Clark, D. B., Falloon, P. D., Fekete, B. M., Gosling, S. N., Heinke, J., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y., and Wisser, D.: First look at changes in flood hazard in the Inter-Sectoral Impact Model Intercomparison Project ensemble, P. Natl. Acad. Sci. USA, 111, 1–5, doi:10.1073/pnas.1302078110, 2013.

Prudhomme, C., Giuntoli, I., Robinson, E. L., Clark, D. B., Arnell, N. W., Dankers, R., Fekete, B. M., Franssen, W., Gerten, D., Gosling, S. N., Hagemann, S., Hannah, D. M., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y., and Wisser, D.: Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment, P. Natl. Acad. Sci. USA, 111, 3262–3267, doi:10.1073/pnas.1222473110, 2014.

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., and Schewe, J.: The Inter-Sectoral Impact Model Intercompar- ison Project (ISI-MIP): project framework., P. Natl. Acad. Sci. USA, 111, 3228–3232, doi:10.1073/pnas.1312330110, 2014.

METHODS

INDEX block annual maximum runoff (peak discharge), AMax.

DIFFERENCES IN THE DISTRIBUTION

AMax series are compared for 18 streamflow gauges and corresponding grid-cells.

- Non-parametric two sample approach:
  - i. Same distribution (Kolmogorov-Smirnoff)
  - ii. Equal median (Wilcoxon Rank Sum)
  - iii. Equal variance (Ansari-Bradley)

EXTREME VALUE FIT An EV distribution is fitted to the AMax times series.

$$\text{Gumbel: } F(Q) = \Pr(Q<q) = \exp\{-\exp\{(q-\mu)/\sigma\}\}$$

- Estimation of the parameters via *joint maximum likelihood*
  - μ location
  - σ scale
- Estimation of confidence intervals via *profile likelihood*

DESIGN EVENT The largest event in T-years on average.

Return levels for the one in 30 years and one in 100 years events (extrapolation) on the basis of the fitted model.

$$T = 1 / (1-p)$$

30 yr return period → Prob_exceedance 1-p = 0.033 (non exceedance probability p = 0.967)

RESULTS – Distributions

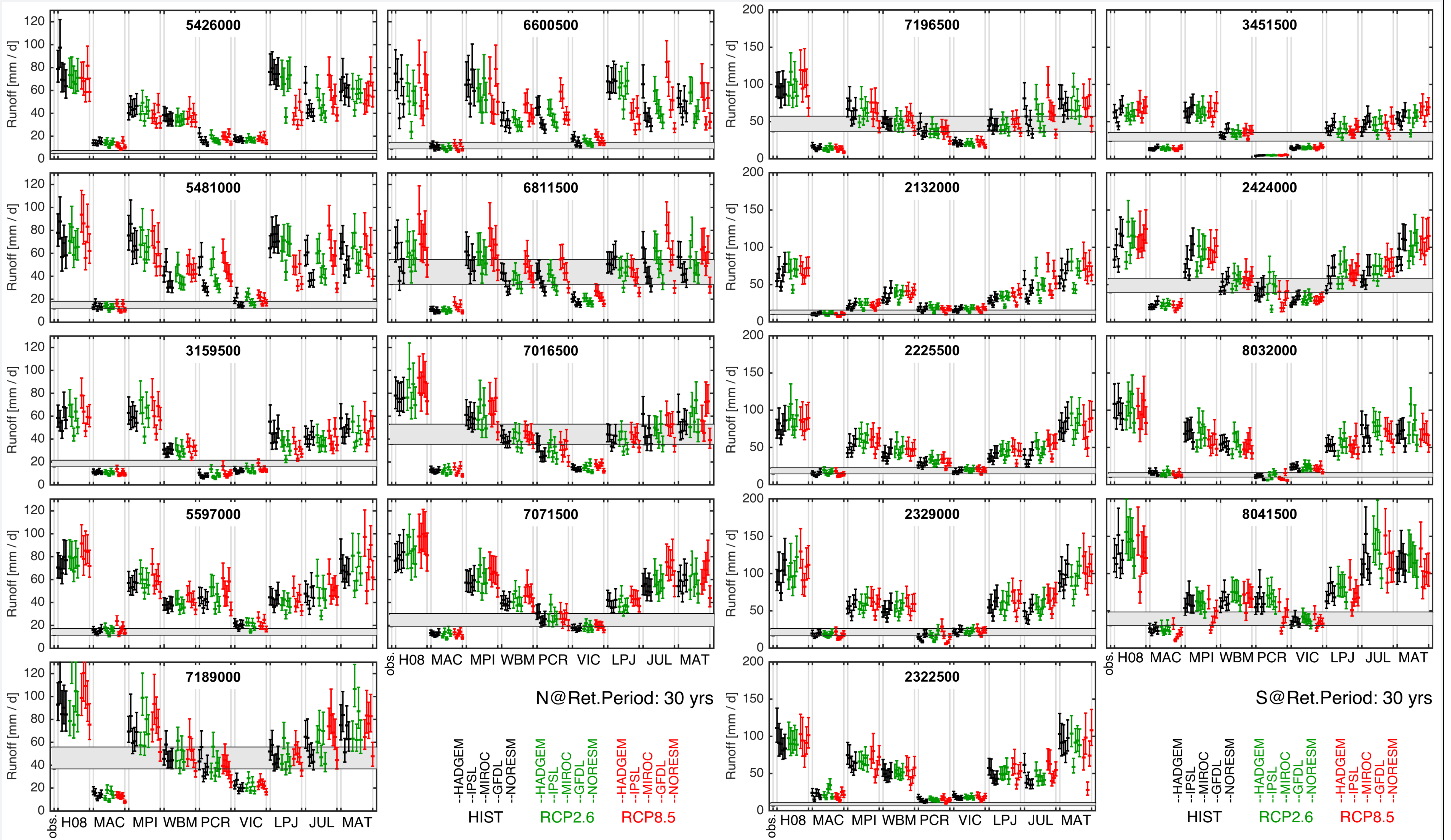
For the 18 gauge-gridcell pairs:

	Runs [%]	OVERLAP
Same distribution – KS test		
observed vs 45 modeled-hist	**9.3%**	→ Little
modeled-hist vs RCP2.6	86.3%	→ High
modeled-hist vs RCP8.5	65.9%	→ Considerable
RCP2.6 vs RCP8.5	69.5%	→ " "
Equal medians – W-ranksum test		
observed vs 45 modeled-hist	**11.9%**	→ Little
modeled-hist vs RCP2.6	88.6%	→ High
modeled-hist vs RCP8.5	68.8%	→ Considerable
RCP2.6 vs RCP8.5	69.3%	→ " "
Equal variances – AB test		
observed vs 45 modeled-hist	84.4%	→ High
modeled-hist vs RCP2.6	96.3%	→ "
modeled-hist vs RCP8.5	89.8%	→ "
RCP2.6 vs RCP8.5	91.7%	→ "

There is a large gap observed-modeled in distribution and medians, while the same is not true for the variances.

RESULTS – Design event (T = 30 yr)

Stations / Gridcells



The horizontal gray band corresponds to the 95% confidence interval of the observed data fit.

CONCLUSIONS

Runoff annual maxima from global models differ systematically from observed data in terms of distribution and medians. With only few exceptions, the majority of the models struggle to reproduce return period ranges (confidence intervals) of observed Amax even at time spans for which extrapolations are small (30 yrs).

There is an evident scale discrepancy between observed and modeled data. Bearing in mind that grid-cell runoff is difficult to benchmark with observed data for the models' incapability in reproducing many of the catchment scale physical processes, efforts should go into finding common metrics for comparing results across ensembles and evaluating model runs against observations. The choice of the metric should be based on the assessment of its variability/uncertainty, which can be very high and therefore may bring about misleading results.

Extreme value theory represents a valuable way for benchmarking simulated runoff via observed streamflow data. It will also allow (ongoing work), via the modeling of the distribution parameters, to assess whether the changes in future peak flows are statistically significant.