Supplementary Material for
On the evaluation of climate model simulated precipitation extremes

Andrea Toreti\textsuperscript{1} and Philippe Naveau\textsuperscript{2}
\textsuperscript{1}European Commission, Joint Research Centre, Ispra (VA), Italy
\textsuperscript{2}Laboratoire des Sciences du Climat et l’Environnement (LSCE) CNRS, Orme des Merisiers, Gif-sur-Yvette, France
E-mail: andrea.toreti@jrc.ec.europa.eu
1. Methods

As stated in the Appendix, the statistic $A$ can be derived as follows,

$$ A = \frac{nm}{N} \sum_{i=1}^{N-1} \left( \frac{(G_m - F_n)}{1 - \frac{i}{N}} \right)^2 \frac{1}{N} = \frac{1}{nmN} \sum_{i=1}^{N-1} \frac{(M_iN - ni)^2}{(N - i)} $$

(1)

where $M_i = nF_n \circ H_N^{-1}(\frac{i}{N})$, i.e. the number of $X$’s less equal the $i$-th smallest value in the pooled sample.

Under the null hypothesis of equally distributed samples and by noticing that $M_i$ is hypergeometric distributed (with population size equal to $N$, sample size given by $i$ and number of successes equal to $n$), the expected value of $A$ is given by

$$ \mathbb{E}(A) = \frac{1}{nmN} \sum_{i=1}^{N-1} N^2 \frac{\nabla(M_i)}{(N - i)} = \frac{1}{N} \sum_{i=1}^{N-1} \frac{i}{N - 1} = \frac{1}{2} $$

(2)

The variance of $A$ can derived as follows

$$ \nabla(A) = -\frac{1}{4} + \frac{N^2}{n^2m^2} \mathbb{E} \left( \left( \sum_{i}^{N-1} \frac{\bar{M}_i^2}{N - i} \right)^2 \right) $$

(3)

where $\bar{M}_i = M_i - \frac{ni}{N}$. Now, the expected value in (3) can be subdivided into two parts, namely $A_1$ and $A_2$. $A_1$ is given by

$$ A_1 = \mathbb{E} \left\{ \sum_{i}^{N-1} \frac{\bar{M}_i^4}{(N - i)^2} \right\} $$

$$ = \frac{nm}{N^2(N - 1)(N - 2)(N - 3)} \left( \frac{mnN^3}{4} - \frac{3mnN^2}{2} + \frac{11mnN}{4} - \frac{3mn}{2} \right) $$

(4)

while $A_2$ can be derived ($M_i$ and $M_j$ being independent) as follows

$$ A_2 = \mathbb{E} \left\{ \sum_{i}^{N-1} \sum_{j=1}^{N-1} \sum_{j 
eq i} \frac{2}{N - i} \frac{\bar{M}_i^2 \bar{M}_j^2}{N - j} \right\} $$

$$ = \frac{2n^2m^2}{N^4(N - 1)^2} \left( \frac{N^4}{4} - \frac{5N^3}{6} + \frac{3N^2}{4} - \frac{N}{6} \right) $$

(5)

Thus,

$$ \nabla(A) = -\frac{1}{4} + \frac{N^2}{n^2m^2}(A_1 + A_2) $$

$$ = -\frac{1}{4} + \frac{1}{(N - 1)(N - 2)(N - 3)} \left( \frac{N^3}{4} - \frac{3N^2}{2} + \frac{11N}{4} - \frac{3}{2} \right) $$

$$ + \frac{2}{N(N - 1)^2} \left( \frac{N^3}{4} - \frac{5N^2}{6} + \frac{3N}{4} - \frac{1}{6} \right) $$

(6)
The approximation of the upper tail area of the survival distribution function of the limiting distribution of $A$ suggest by Sinclair et al. (1990) is:

$$
\bar{F}(x) = 0.889 \sqrt{1.835} x e^{-1.835x}
$$

(8)

The Kullback Leibler directed divergence can be estimated as follows (Naveau et al. 2013):

$$
\hat{I}(f_e; g_e) = 1 + \frac{1}{N_e} \sum_{i=1}^{N_e} \log \left( \frac{\bar{G}(X^i_e/\hat{\mu}_0^X)}{G(0)} \right),
$$

(9)

where $N_e$ is the number of $X$-excesses and $G(t) = 1 - (m + 1)^{-1} \sum_{i=1}^{m} I_{\{Y^i_e/\hat{\mu}_0^Y \leq t\}}$ with $m$ equal to the number of $Y$-excesses.

The divergence method of Naveau et al. (2013) is based on:

$$
D(f, g) = I(f_e; g_e) + I(g_e; f_e),
$$

(10)

where $I$ can be estimated as in eq. (8).

The multiple testing approach of Genovese and Wasserman (2004) is based on the False Discovery Rate - FDR (Benjamini and Hochberg 1995). Given $m$ tests, let $M_0$ ($M_1$) be the number of true (false) null hypothesis. The FDR is defined as:

$$
FDR = \begin{cases} 
\mathbb{E} \left\{ \frac{V}{R} \right\} & \text{if } R > 0 \\
0 & \text{if } R = 0 
\end{cases}
$$

(11)

where $R$ represents the number of rejected null hypotheses and $V$ the true null hypotheses wrongly rejected. Associated to the $m$ null hypotheses, there are $m$ p-values $P_i$ ($i = 1, \ldots, m$) having a marginal distribution $G = (1 - a)U + aF$, where $U$ represents the uniform distribution on $(0, 1)$, $a$ is a coefficient varying in $[0, 1]$ and $F$ denotes an unknown distribution. Thus, the FDR can be controlled at a chosen level $\alpha$ (here, 0.1) by rejecting all null hypotheses associated with a p-value less equal a specific threshold $T(a, G)$ that can be estimated by:

$$
T(\hat{a}, \hat{G}) = \sup \left\{ t : \hat{G}(t) = \frac{(1 - \hat{a})t}{\alpha} \right\}
$$

(12)

where $\hat{G}$ is the empirical cumulative distribution function of the $P_i$ and the estimation of $a$ is given by:

$$
\hat{a} = \max_t \frac{\hat{G}(t) - t - \epsilon_m}{1 - t}, \quad \epsilon_m = \sqrt{\frac{1}{2m} \log \left( \frac{2}{\alpha} \right)}
$$

(13)
Table 1: Number of wrong decisions (based on 95% critical values) derived by testing (10^3 times) samples coming from three different tails (see the bold red values) against samples from tails having a different ξ (see bold numbers on rows) and using three sample sizes (10^2,10^3,10^4). Values in percentage.

| ξ = 0.1 | 0.05 | 0.08 | 0.1 | 0.15 | 0.2 |
|---------|------|------|-----|------|-----|
| 100     | 94.6 | 95.6 | 5.2 | 95.4 | 94.8|
| 1000    | 90.9 | 91.1 | 4.2 | 90.8 | 67.5|
| 10000   | 18.8 | 85.5 | 3.3 | 24.6 | 0   |

| ξ = −0.1 | 0.2 | -0.15 | -0.1 | -0.08 | -0.05 |
|----------|-----|-------|------|-------|-------|
| 100      | 92.5 | 94.5 | 5.7  | 95.5  | 95.2  |
| 1000     | 50.5 | 87.4 | 3.3  | 94.1  | 90.1  |
| 10000    | 0    | 6.7  | 4    | 80    | 9.7   |

| ξ = 0    | -0.4 | -0.3 | -0.2 | -0.1 | 0     |
|----------|------|------|------|------|-------|
| 100      | 32.7 | 62.3 | 81.5 | 92.6  | 5     |
| 1000     | 0    | 0    | 0.7  | 55.1  | 3.3   |
| 10000    | 0    | 0    | 0    | 0     | 4.2   |

| ξ = 0    | 0 | 0.1 | 0.2 | 0.3 | 0.4 |
|----------|---|-----|-----|-----|-----|
| 100      | 5 | 92.8 | 85.3 | 73 | 56.6 |
| 1000     | 3.3 | 64.6 | 6.3 | 0 | 0 |
| 10000    | 4.2 | 0 | 0 | 0 | 0 |

Table 2: CMIP5 Global Climate Models used in this study. We acknowledge the World Climate Research Programmes Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in the table) for producing and making available their model output. For CMIP the U.S. Department of Energys Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

| Model                  | Institution                                                                 |
|------------------------|-----------------------------------------------------------------------------|
| CMCC-CM (CMCC)         | Centro Euro-Mediterraneo sui Cambiamenti Climatici                           |
| CNRM-CM5 (CNRM)        | Centre National de Recherches Meteorologiques - Centre Europeen de Recherche et Formation Avancées en Calcul Scientifique |
| HadGEM2-CC (HadCC)     | Met Office Hadley Centre                                                     |
| HadGEM2-ES (HadES)     | Met Office Hadley Centre - Instituto Nacional de Pesquisas Espaciais         |
| INM-CM4 (INM)          | Institute for Numerical Mathematics                                         |
| IPSL-CM5A-MR (IPSL)    | Institut Pierre-Simon Laplace                                               |
| MIROC5 (MIROC)         | Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology |
| MRI-CGCM3 (MRI)        | Meteorological Research Institute                                           |

References

Benjamini Y and Y Hochberg 1995 Controlling the False Discovery Rate: a practical and powerful approach to multiple testing J. Roy. Stat. Soc. B 57 289–300
Genovese C and L Wasserman 2004 A stochastic process approach to false discovery control *Ann. Stat.* **32** 1035–1061

Naveau P, Guillou A and Rietsch T 2013 A non-parametric entropy-based approach to detect changes in climate extremes *J. Roy. Stat. Soc. B* in press

Sinclair C D, Spurr B D and M I Ahmad 1990 Modified Anderson-Darling test *Comm. Stat. A* **19** 3677–3686
Figure 1: Spearman-based spatial correlation matrix of the tail scaling factors, estimated for the eight GCMs, $\hat{\mu}_0^{\text{model}}$, and the gridded observations E-OBS, $\hat{\mu}_0^{\text{obs}}$, in the autumn period 1966-2005. The colors and the shape of the ellipses are associated with the correlation values. The last column refers to the same analysis without the southern part of the domain (south of 38.25 degrees North).
Figure 2: Rescaled-tail comparison of model simulations during the historical autumn period and E-OBS. Colors are associated with the values of the 2-sample modified Anderson-Darling statistic with the sign given by the estimated KLD-divergence. Blank areas are associated with non-significant values.
precipitation extremes

Figure 3: Boxplots of the ratio between the estimated conditional means of the excesses for the future winter time periods (2020-2059: blue; 2060-2099: green) and the historical simulations, $\hat{\mu}_{0}^{\text{scenario}} / \hat{\mu}_{0}^{\text{hist}}$, derived for each grid point in the domain.
Figure 4: Results of the rescaled-tail comparison of the autumn period 2020-2059 w.r.t. the historical simulation (1966-2005). Colors are associated with the values of the 2-sample modified Anderson-Darling statistic with the sign given by the estimated KLD-divergence. Blank areas are associated with non-significant values.
Figure 5: As for Figure 4, but for 2060-2099.