A parametric approach for simultaneous bias correction and high-resolution downscaling of climate model rainfall for practical applications

Antonios Mamalakis[1,2] , **Andreas Langousis**[1] , Roberto Deidda[3], and Marino Marrocu[4]

[1]Department of Civil Engineering University of Patras, Greece (andlag@alum.mit.edu).
[2]Department of Civil and Environmental Engineering, University of California , Irvine, California, USA.
[3]Dipartimento di Ingegneria Civile, Ambientale e Architettura, University of Cagliari, Cagliari, Italy.
[4]CRS4, Centro di Ricerca, Sviluppo e Studi Superiori in Sardegna, Loc. Piscina Manna, Pula (CA), Italy.

1. Introduction

Among different approaches to bias correct climate model (CM) results, distribution mapping has been identified as the most efficient one in reproducing the statistics of rainfall at a regional level, and hydrologically relevant temporal scales (e.g. daily). However:

- Non-parametric (empirical) distribution mapping is sensitive to sample length variations, the presence of outliers, and may lead to significant biases, especially in extreme rainfall estimation.
- Bias-corrected results are constrained by the nominal resolution of the RCM used (i.e. ~ 25 km), which does not suffice to accurately resolve the statistical structure of rainfall at a basin level.

Here (see also Mamalakis *et al.*, 2017) we propose a two step parametric approach that addresses the aforementioned two issues:

✓ **Fit a two-component theoretical distribution model** (i.e. Exponential (Exp.) and Generalized Pareto (GP)) to the historical and CM rainfall series ⇨ ***robustness in extreme rainfall modeling***

✓ **Interpolate the distribution parameters** over a user-defined high-resolution grid, using kriging for uncertain data (KUD) ⇨ ***Downscaling***

2. Region of study – Available Data

We compare the proposed parametric method with its non-parametric variant, in bias correcting daily rainfall products from 4 CMs (i.e. KNMI, MPI, C4I, SMHE), over Sardinia (Italy) for the period 1951-2008.

Data used:
243 stations with more than 30 years of recordings

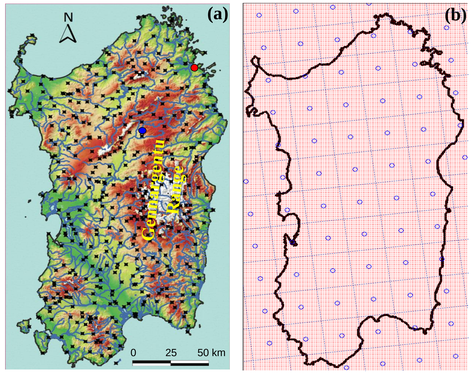


Figure 1: a) The island of Sardinia in Italy. Points indicate rainfall stations, while circles indicate the stations used in Figure 2: #110 (blue) and #288 (red). b) High-resolution grid (i.e. 1 km resolution; red dots) and CM grid (25 km resolution; blue circles).

3. Fitting the two-component distribution model

Use a **GP** model for rainfall intensities above a specified threshold u^*, and an **exponential** model for lower rainrates.

GP model
- **Distribution fitting** using the *Multiple Threshold Method* (MTM) of Deidda (2010) with probability weighted moments (PWM) ⇨ ***robust parameter estimation***
- **Standardization of distribution parameters** to zero threshold ⇨ ***prerequisite for regional frequency analysis***

Exp. model ⇨ **Distribution fitting using** *Least Squares* (LS) to account for all available information on low rainrates.

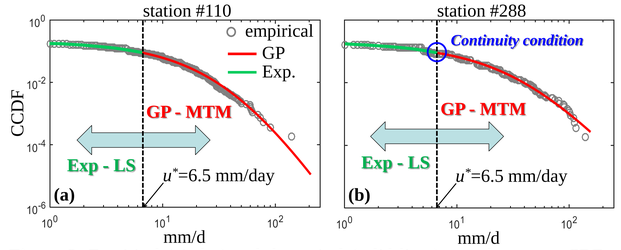


Figure 2: Empirical (circles) and theoretical (solid lines) complementary CDFs for the selected stations shown in Figure 1.a, in the control period 1951-1965.

4. Kriging of distribution parameters

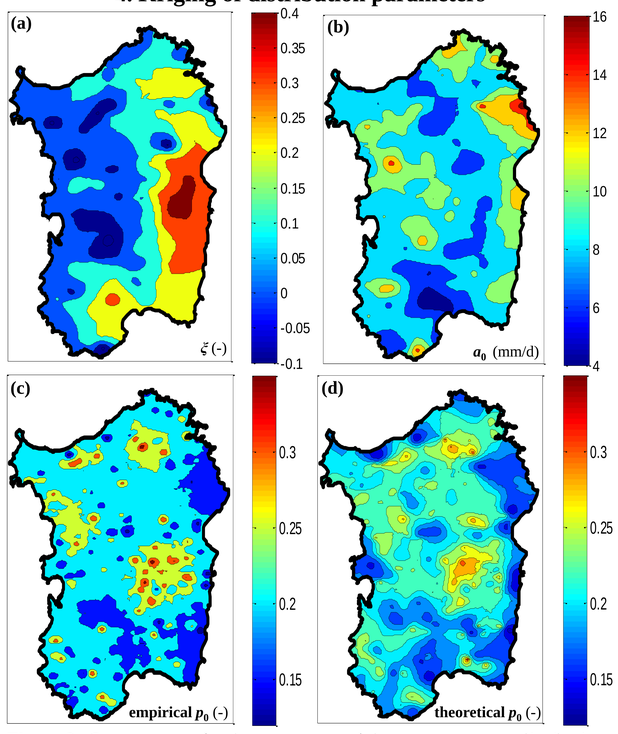


Figure 3: Contour maps for the parameters of the two-component distribution model in the control period 1951-1965, interpolated using KUD: a) GP shape parameter ξ (-), b) standardized to zero threshold scale parameter a_0 (mm/d), c) empirically estimated fraction of wet days p_0 (-), d) theoretically calculated p_0 (i.e. honoring continuity condition; see Figure 2).

5. Simultaneous bias correction and downscaling

$$x_{cor}^{(m,l)}(t)=\begin{cases} Q^{-1}_{KUD,l}\{Q_{m,k}[x_{CM}^{(m,k)}(t)]\}\ , & x_{CM}^{(m,k)}(t) > 0 \\ Q^{-1}_{KUD,l}(p_k) & , \ x_{CM}^{(m,k)}(t) \le 0 \end{cases}$$

$x_{CM}(t)$ denotes the rainfall estimate at grid cell k (blue circles in Figure 1.b) on day t from CM m; $x_{cor}(t)$ is the corresponding bias corrected and downscaled rainfall product at location l (red dots in Figure 1.b); Q is the two-component distribution model, and p_k is uniformly distributed in the interval $[0, Q_{m,k}(0)]$.

6. KNMI results (control period 1951-1965)

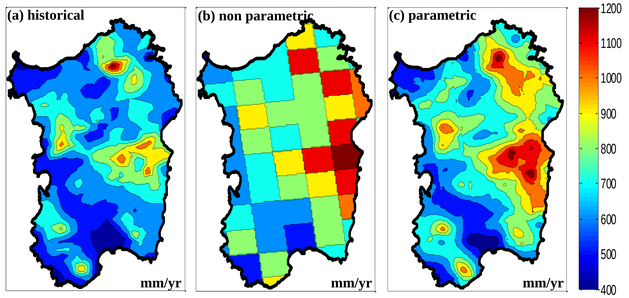


Figure 4: Estimates of the mean annual rainfall depth (mm/yr) for the 43-year validation period 1966-2008, using: a) historical rainfall observations, b) bias corrected KNMI rainfall products with non-parametric distribution mapping, and c) parametric distribution mapping, with parameters drawn from Figure 3.

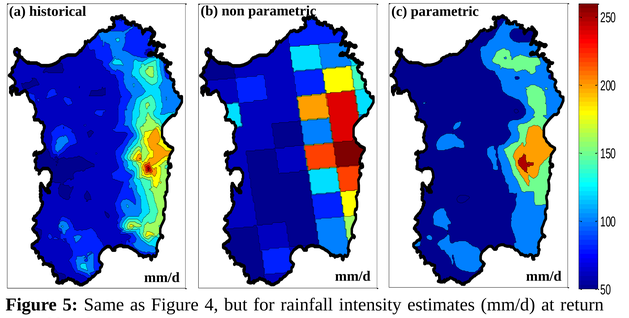


Figure 5: Same as Figure 4, but for rainfall intensity estimates (mm/d) at return period level T =10 years.

7. Sensitivity to control-validation periods and CM used

$$DRMSE = \left(\frac{1}{N} \sum_{i=1}^{N} \left(\frac{z_i' - z_i}{z_i} \right)^2 \right)^{1/2}$$

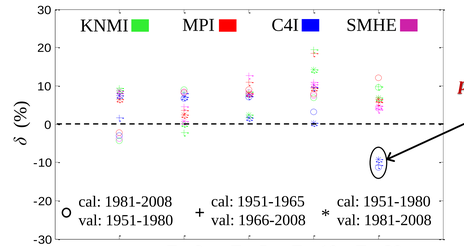
z_i and z_i' are the historical and model-based estimates of statistic Z, respectively.

Table 1: Dimensionless RMSEs (%) for different combinations of control-validation periods, based on Q-Q corrected KNMI rainfall products using non-parametric (i.e. empirical; emp.) and parametric (par.) distribution mapping; **bold** values indicate cases where the non-parametric approach performs better.

control period	validation period	mean annual depth		T = 2 years		T = 5 years		T = 10 years		T = 30 years	
		emp	par	emp	par	emp	par	emp	par	emp	par
1951-1980	1981-2008	33.19	24.46	19.14	19.02	21.36	19.00	35.45	21.30	31.90	25.02
1951-1965		39.09	30.15	**18.91**	**25.09**	27.29	24.61	44.08	26.54	36.84	28.96
1951-1957		37.63	25.98	**27.13**	**27.83**	38.44	26.21	49.22	29.83	40.84	31.17
1951-1965	1966-2008	29.37	19.96	**16.91**	**19.15**	26.02	17.64	40.97	21.49	34.79	25.03
1951-1957		29.98	20.56	23.95	22.47	37.24	21.54	46.54	24.20	37.40	28.72
1981-2008	1951-1980	**23.89**	**28.22**	25.71	16.77	28.84	21.11	30.49	23.61	37.02	27.39

Similar findings when
- ✓ upscaling the results of the parametric approach to the CM grid
- ✓ randomly eliminating half of the stations

✓ ***parametric approach performs better****, especially for extreme rainfall*



pronounced biases in C4I rainfall products; see Mamalakis et al. (2017)

Figure 6: Difference δ = DRMSE$_n$ – DRMSE$_p$ for the statistics in Table 1, obtained by applying the non-parametric bias correction approach (DRMSE$_n$) and parametric distribution mapping (DRMSE$_p$) to 4 different CMs, for 3 combinations of non-overlapping control and validation periods. Positive values of δ indicate superiority of the parametric approach.

8. Conclusions

While **parametric bias correction** of climate model results cannot eliminate the signature of pronounced biases in raw CM rainfall products:

- It demonstrates **significant skill** in modeling the effects of topography and local climatology on the magnitude, general tendencies, and high-resolution spatial structure of rainfall statistics, including rainfall extremes
- It is rather **insensitive** to the characteristics of the calibration period, including its length, and the climate model used.

References

Deidda, R. (2010) A multiple threshold method for fitting the generalized Pareto distribution to rainfall time series, *Hydrol. Earth Syst. Sci.*, **14**(12), 2559-2575.

Mamalakis, A., A. Langousis, R. Deidda, and M. Marrocu (2017) A parametric approach for simultaneous bias correction and high-resolution downscaling of climate model rainfall, *Water Resour. Res.*, **53**, doi:10.1002/2016WR019578.

Acknowledgements

The work of Antonios Mamalakis was supported by LIFE (Linked Institutions for Future Earth) funded under NSF's Science Across Virtual Institutes (SAVI) program (NSF grant EAR-1242458). The work of Andreas Langousis was supported by the Onassis Foundation under the "Special Grant and Support Program for Scholars' Association Members". The work conducted by Roberto Deidda was funded under the Sardinian Regional Law 7/2007 (funding call 2013). CRS4 highly acknowledges the contribution of the Sardinian regional authorities.