Supplementary information

September 8, 2025

Effect of bias correction on precipitation regimes

Effects on future regimes

In this section, we investigate the effects which bias correction implies on wet-days and all-days regimes. Figure 1 compares the regimes computed between 1970–2000 and 2070–2100, on a subset of 6 out of 15 models from our ensemble, in three different set-ups: models corrected by ERA5 reanalysis with 1971-2005 as the reference period, exactly as in the current paper; models corrected with EOBS gridded observation dataset instead (Cornes et al., 2018); and the original raw models, without any bias correction. Note that the correction using EOBS, which is based solely on rain gauges data, provides values only on certain continental regions.

A first result emerging from the figure is the very close resemblance of the spatial patterns for all-days regimes, between the raw models and those bias-corrected. Though the models corrected with EOBS do show a light lack of consensus regarding a "all decrease" regime in the southern Mediterranean, this is probably explained by the smaller spatial availability of the EOBS data, as station data is relatively reduced in this region on the 1971-2005 period. The similarity in the all-days patterns between the three bias-correction set-ups is reinforced by the almost identical position of the transition zone between "all increase" and "U-shape" regimes.

A second result is the significant difference for wet-days, between the correction using EOBS and using ERA5 or the raw models: the northern Mediterranean is classified as a "U-shape" regime for the EOBS correction, rather than an "all increase" regime. This difference occurs on a large region spanning from Spain to Turkey, going through UK, France, Italy among others. This difference is explained in the following section.

As a sum up, we showed in a reduced setting experiment that the tripartite pattern of all-days regimes is unaffected by the use of bias correction, contrary to the wet-days patterns which can be significantly affected.

Analytical explanation: unaffected all-days regimes

Here, we explain the large similarity observed in all-days regimes, by coming back to the definition of bias correction and how this will affect quantiles changes. In this paper, the CDF-t method was applied, combined with the SSR tool, as presented in section 4. The CDF-t method stems from the quantile mapping bias correction, on which we will focus in a first time.

The principle of quantile mapping is very simple: it consists in equating, within the historical period, the two cumulative distribution functions (CDFs) of the reference observation and the model, noted $F_{o,h}$ and $F_{m,h}$ respectively. We use here the subscript o for the observations taken as the reference for bias correction, m for the model to correct, h for the historical period and f for the future period (as in Cannon et al. (2015)). Therefore, if x denotes the physical variable (in our case, precipitation daily intensity in mm/day), then $x_{o,h}$ is the observed historical data and $x_{m,h}$ the modeled one.

The quantile bias correction $\hat{x}_{m,h}$ of the value $x_{m,h}$ is defined as the unique positive value such as: $F_{o,h}(\hat{x}_{m,h}) = F_{m,h}(x_{m,h})$, which exists since both CDFs functions are invertible and bijective from the physical space to $[0, 1]$. This gives the expression for the correction:

$$\hat{x}_{m,h} = F_{o,h}^{-1}\left(F_{m,h}(x_{m,h})\right) \tag{1}$$

For the future period, where no observation is available, quantile mapping applies the same transformation as for the past period, to define the corrected future value $\hat{x}_{m,f}$:

$$\hat{x}_{m,f} = F_{o,h}^{-1}\left(F_{m,h}(x_{m,f})\right) \tag{2}$$

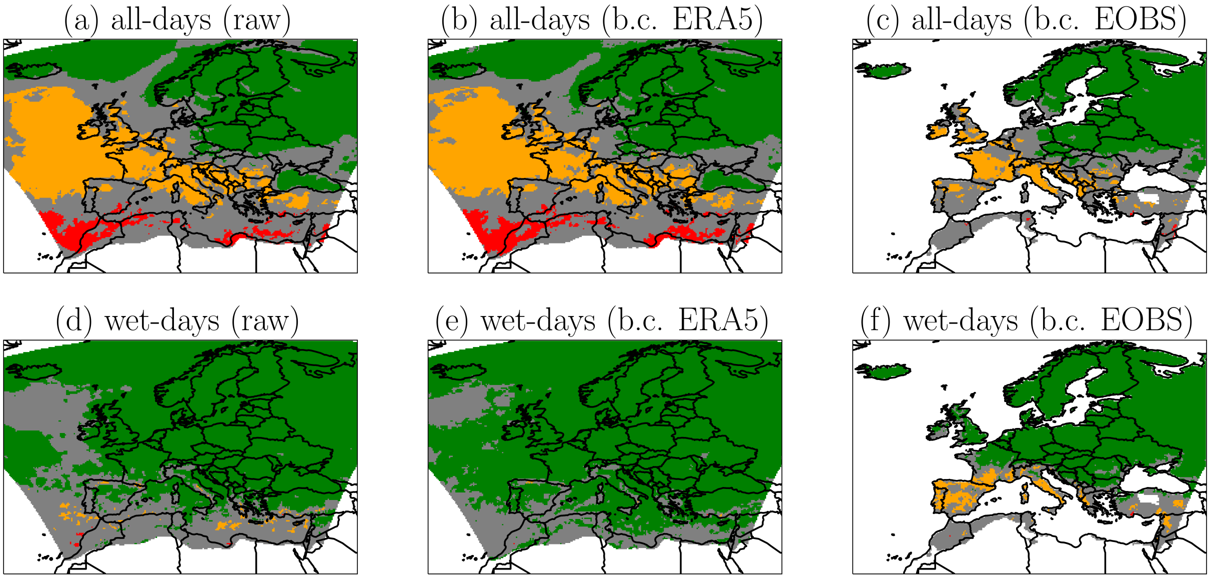

Figure 1: Subfigures (a), (b), (c): Multimodel all-days regimes between 1970–2100 and 2070–2100, for 6 Euro-CORDEX models, either without any correction (a), with bias correction using ERA5 as the reference dataset (b) or using EOBS instead (c). Subfigures (d), (e), (f): Same, but on wet-days regime maps. White color displays the border of the Euro-CORDEX simulation domain and desert areas. In addition, it shows the EOBS gridded dataset limits for figures (c) and (f), defined as the grid-points with less than 80% available precipitation data over the 1971-2005 reference period. Grey color shows locations with available data where fewer than 4 out of the 6 models agree on the regime.

In other words, the quantile mapping boils down to applying a single correction function, $F_{o,h}^{-1}F_{m,h} : \mathbb{R}^+ \to \mathbb{R}^+$, to any given precipitation value.

The effect on quantile trends can be derived analytically. As the correction transformation is bijective from $\mathbb{R}^+$ to $\mathbb{R}^+$, we can write the correction on a given quantile value as a multiplication by a strictly positive factor $c(p_a) \geq 0$ for any all-days percentile p_a in $[0, 1]$:

$$\hat{Q}_a(p_a) = c(p_a)Q_a(p_a). \tag{3}$$

Thus the quantile trends between a period 1 (in the past) and 2 (in the future) can be written as:

$$\begin{aligned}
\Delta\hat{Q}_a(p_a) &= \hat{Q}_{a,2}(p_a) - \hat{Q}_{a,1}(p_a) && (4)\\
&= c(p_a)\left(Q_{a,2}(p_a) - Q_{a,1}(p_a)\right) && (5)\\
&= c(p_a)\Delta Q_a(p_a). && (6)
\end{aligned}$$

with a single factor $c(p_a) \geq 0$, as the quantile mapping uses the same correction function for the historical and future periods. This shows that the quantile trends conserve their sign through quantile mapping. Therefore, all-days regimes, directly based on the signs of the lower-to-medium and extreme quantiles trends, should be the same whether raw of corrected with quantile mapping, whatever the reference dataset.

CDF-t methods differs from quantile mapping in such that the correction on the future is a more complex function, which allows for an evolution of the CDF with time. The corrected future value is defined as:

$$\hat{x}_{m,f} = \left(F_{m,f}^{-1}F_{m,h}F_{o,h}^{-1}F_{m,f}\right)(x_{m,f}) \tag{7}$$

With our notations, this means that a different correction factor is applied for the future values and historical values: $c_f(p_a) \neq c_h(p_a)$. Therefore,

$$\begin{aligned}
\Delta\hat{Q}_a(p_a) &= c_f(p_a)Q_{a,2}(p_a) - c_h(p_a)Q_{a,1}(p_a) && (8)\\
&\neq c(p_a)\ \Delta Q_a(p_a).
\end{aligned}$$

This potentially results in different signs for the quantile change and its corrected values, and the all-days regime trends could thus be different. However, we will suppose this difference to be small, as we have seen that the difference in distribution are of small order, especially compared to the change of occurrence frequency in Southern Europe and the Mediterranean. This is consistent with our empirical observations (previous section) that all-days regimes patterns across Europe and the Mediterranean are almost identical in three different set-ups of bias correction with CDF-t-SSR.

Analytical explanation: modified wet-days regimes

We then want to explain why the wet-days regimes can be impacted by bias-correction. We can analytically express the effect of quantile mapping method on wet-days quantiles, by coming back to the corresponding all-days percentile. As shown previously, the relationship between all-days and wet-days percentiles is $p_a = (1 - f_d)p + f_d$, with f_d the dry-days frequency, p_a the all-days percentile and p the wet-days percentiles. This means that the corresponding all-days percentile to a fix wet-days rank is non-constant in time, as long as f_d is evolving. Equation (3) then gives:

$$\begin{aligned}\hat{Q}_w(p) &= \hat{Q}_a(p_a) = c(p_a)\ Q_a(p_a) \\ &= c(p_a)\ Q_w(p) \\ &= c(p_{a,f_d})\ Q_w(p).\end{aligned} \tag{9}$$

The correction factors thus depend on the occurrence frequency f_d, which can greatly vary with climate change, as shown in this paper. The bias-corrected wet-days quantiles change between given periods 1 and 2 is then:

$$\begin{aligned}\Delta\hat{Q}_w(p) &= \hat{Q}_{w,2}(p) - \hat{Q}_{w,1}(p) \\ &= c(p_{a,f_{d,2}})\ Q_{w,2}(p) - c(p_{a,f_{d,1}})\ Q_{w,1}(p) \\ &= c(p_{a,f_{d,2}})(Q_{w,2}(p) - c(p_{a,f_{d,1}})/c(p_{a,f_{d,2}})Q_{w,1}(p)) \\ &\neq \text{factor}\ \ (Q_{w,2}(p) - Q_{w,1}(p)) \\ &\neq \text{factor}\ \Delta Q_w(p)\end{aligned} \tag{10}$$

To go further, a Taylor development at first order on the correction factor function c gives a linear relationship with the dry-days frequency change Δf_d:

$$\begin{aligned}c(p_{a,f_{d,2}}) - c(p_{a,f_{d,1}}) &\approx \left(p_{a,f_{d,2}} - p_{a,f_{d,1}}\right)\ \frac{dc}{dp_a}(p_{a,f_{d,1}}) \\ &\approx \Delta f_d\ (1-p)\ \frac{dc}{dp_a}(p_{a,f_{d,1}})\end{aligned} \tag{11}$$

Therefore, the larger Δf_d the larger the difference on the corrected quantile change, $\Delta\hat{Q}_w(p)$, and the more likely a difference in wet-days regime becomes. This is why we observed a large difference between raw wet-days regimes and those corrected using EOBS on the Mediterranean, a region where the future change of precipitation occurrence is very large.

Summary of the effects of bias correction

In a nutshell, the effects of quantile mapping and CDF-t on precipitation regimes can be summarized by table 1:

	All-days regimes	Wet-days regimes
Quantile mapping	no effect	effect proportional to the change in dry-days frequency
CDF-t SSR	small effect, independent of the change in dry-days frequency	effect proportional to the change in dry-days frequency

Table 1: Summary of the effects of bias correction on all-days and wet-days regime, from the mathematical framework, at first order.

With these similarities and differences in mind, we decided to use the ERA5 dataset as the reference for bias correction, as mentioned in the main text. ERA5 is preferred to EOBS gridded dataset both for its global coverage other about seven decades, but also to avoid the large temporal inhomogeneities in EOBS gridded dataset and the resulting uncertainties in trends computed from them (Hofstra et al., 2009; Cornes et al., 2018; Bandhauer et al., 2022; André et al., 2024). The fact that the bias-correction with ERA5 gives the same results as the non-corrected models gives even more confidence in our choice of ERA5 as the reference dataset for bias correction.

References

André, J., d'Andrea, F., Drobinski, P., and Muller, C. (2024). Regimes of precipitation change over Europe and the Mediterranean. *Journal of Geophysical Research: Atmospheres*, 129(15):e2023JD040413.

Bandhauer, M., Isotta, F., Lakatos, M., Lussana, C., Båserud, L., Izsák, B., Szentes, O., Tveito, O. E., and Frei, C. (2022). Evaluation of daily precipitation analyses in e-obs (v19. 0e) and era5 by comparison to regional high-resolution datasets in european regions. *International Journal of Climatology*, 42(2):727–747.

Cannon, A. J., Sobie, S. R., and Murdock, T. Q. (2015). Bias correction of gcm precipitation by quantile mapping: how well do methods preserve changes in quantiles and extremes? *Journal of Climate*, 28(17):6938–6959.

Cornes, R., Schrier, G., Van den Besselaar, E., and Jones, P. (2018). An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. *Journal of Geophysical Research Atmospheres*, 123.

Hofstra, N., Haylock, M., New, M., and Jones, P. D. (2009). Testing e-obs european high-resolution gridded data set of daily precipitation and surface temperature. *Journal of Geophysical Research: Atmospheres*, 114(D21).