

Bias Adjustment of ESPO-R5 v1.0

Pascal Bourgault, Travis Logan, Trevor J. Smith, Juliette Lavoie

November 21, 2022

Temperature and precipitation data from the simulations were first extracted over an area covering North America and, if necessary, converted into daily values. Then using the ESMF software, were accessed through its python xESMF interface, all the extracted simulation data is interpolated bilinearly to the ERA5-Land grid.

The ESPO-R5 v1.0 bias adjustment procedure uses algorithms provided by xclim to adjust simulation bias following a quantile mapping procedure. In particular, the algorithm used is inspired by the "Detrended Quantile Mapping" method described by [2]. The procedure is univariate, bipartite, acts differently on the trends and the anomalies, and is applied iteratively on each day of the year (grouping) as well as on each grid point.

1 Variables

Adjustments were applied separately for each of the three (3) variables. Adjusting `tasmax` and `tasmin` independently can lead to physical inconsistencies in the final data (i.e. cases with `tasmin > tasmax`; [4], [1]). To ensure a physically consistent dataset (for this aspect at least), we computed the daily temperature range (or amplitude; `dtr = tasmax - tasmin`) and adjusted this variable, in addition to the `tasmax` and `pr` variables. `tasmin` is reconstructed after the bias-adjustment and this is the variable we store.

While `tasmax` has no physical bounds in practice, this is not the case for `pr` and `dtr` where a lower bound, at zero (0), exists. Because of this, the adjustment process explained below exists in two flavours: additive and multiplicative. In the latter, it is mathematically impossible for adjusted data to descend below zero (0).

2 Bias-adjustment

The bias adjustment acts independently on each day of the year as well as each grid point. To render the procedure more robust, a window of 31 days around the current day of year was included in the inputs of the calibration (training step). For example, the adjustment for February 1 (day 32) was calibrated using data from January 15 to February 15, over the 30 years of the reference period; For leap years, this would mean that there are four (4) times fewer data points for the 366th day of the year. To circumvent this issue, we converted all inputs to a "no leap" calendar by dropping data on the 29th of February, except for simulations using the "360_day" calendar. In the latter case, the simulations were untouched but the reference data was converted to that calendar by dropping extra days taken at regular intervals¹

2.1 Detrending

For each day of the year and each grid points, we first computed the averages and "anomalies" of the reference data and the simulations over the 1981-2010 reference period. Depending on the variable, anomalies are either taken additively or multiplicatively:

$$Y_r = \begin{cases} \overline{Y_r} + Y'_r & \text{tasmax} \\ \overline{Y_r} \cdot Y'_r & \text{dtr, pr} \end{cases} \quad (1)$$

and similarly for X_c , $\overline{X_c}$ and X'_c .

Instead of a simple moving mean, X_s was detrended with a locally weighted regression (LOESS; [3]). We chose this method for its slightly heavier weights given at the center of the moving window, reducing impacts of abrupt

¹In a normal year, February 6th, April 20th, July 2nd, September 13th and November 25th are dropped. For a leap year, it is January 31st, April 1st, June 1st, August 1st, September 31st and December 1st.

inter-annual changes on the trend and anomalies. It also has a more robust handling of the extremities of the timeseries. The LOESS window had a 30-year width and a tricube shape, the local regression was of degree 0 and only one iteration was performed. The detrending was applied on each day of the year but after averaging over the 31-day window, it yielded the trend $\overline{X_s}$ and the residuals X'_s . Here again, the process can be additive or multiplicative.

2.2 Adjustment of the residuals

With $F_{Y'_r}$ and $F_{X'_c}$ the empirical cumulative distribution functions (CDF) of Y'_r and X'_c respectively, an adjustment factor function was first computed:

$$A_+(q) := F_{Y'_r}^{-1}(q) - F_{X'_c}^{-1}(q) \quad A_\times(q) := \frac{F_{Y'_r}^{-1}(q)}{F_{X'_c}^{-1}(q)} \quad (2)$$

Where q is a quantile (in range $[0, 1]$), $A_+(q)$ is the additive function used with **tasmax** and $A_\times(q)$ the multiplicative one, used with **pr** and **dtr**. The CDFs were then estimated from the 30 31-day windows. In the implementation, maps of A were saved to disk by sampling q with 50 values, going from 0.01 to 0.99 by steps of 0.02. The adjustment was then as follows:

$$X'_{ba} = X'_s + A_+(F_{X'_c}(X'_s)) \quad X'_{ba} = X'_s \cdot A_\times(F_{X'_c}(X'_s)) \quad (3)$$

Nearest neighbor interpolation was used to map $F_{X'_c}(X'_s)$ to the 50 values of q . Constant extrapolation was used for values of X'_s outside the range of X'_c .

2.3 Adjustment of the trend

In the training step, a simple scaling or offset factor was computed from the averages:

$$C_+ = \overline{Y_r} - \overline{X_c} \quad C_\times = \frac{\overline{Y_r}}{\overline{X_c}} \quad (4)$$

This factor was then applied to the trend in the adjustment step:

$$\overline{X_{ba}} = \overline{X_s} + C_+ \quad \overline{X_{ba}} = \overline{X_s} \cdot C_\times \quad (5)$$

2.4 Final scenario

Finally, the bias-adjusted timeseries for this day of year, grid point, and variable was as follows:

$$X_{ba} = \overline{X_{ba}} + X'_{ba} \quad (6)$$

3 Pre-processing of precipitation

It should be noted that the multiplicative mode is prone to division by zero, especially with precipitation where values of zero are quite common. This problem was avoided by modifying the inputs of the calibration step where the zeros of precipitation were replaced by random values between zero (excluded) and 0.01 mm/d. The **dtr** timeseries were not modified as it is almost impossible to have zeros for that variable and the few that appeared were dissolved by the aggregations of the calibration step.

As observed by [5], when the model has a higher dry-day frequency than the reference, the calibration step of the quantile mapping adjustment will incorrectly map all dry days to precipitation days, resulting in a wet bias. The frequency adaptation method finds the fraction of "extra" dry days:

$$\Delta P_{dry} = \frac{F_{X_c}(D) - F_{Y_r}(D)}{F_{X_c}} \quad (7)$$

Where D is the dry-day threshold, taken here to be 1 mm/d. This fraction of dry days was transformed into wet days by injecting random values taken in the interval $[D, F_{Y_r}^{-1}(F_{X_c}(D))]$.

Both pre-processing functions were applied only on the calibration step inputs (Y_r and X_c) before the division between average and anomalies. As such, only the adjustment factors were impacted while there were no explicitly injected precipitation values in the final scenarios.

References

- [1] Agbazo, M. N., & Grenier, P. (2020). Characterizing and avoiding physical inconsistency generated by the application of univariate quantile mapping on daily minimum and maximum temperatures over Hudson Bay. *International Journal of Climatology*, 40(8), 3868–3884. <https://doi.org/10.1002/joc.6432>
- [2] Cannon, A. J., Sobie, S. R., & 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. <https://doi.org/10.1175/JCLI-D-14-00754.1>
- [3] Cleveland, W. S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*, 74(368), 829–836. <https://doi.org/10.1080/01621459.1979.10481038>
- [4] Thrasher, B., Maurer, E. P., McKellar, C., & Duffy, P. B. (2012). Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrology and Earth System Sciences*, 16(9), 3309–3314. <https://doi.org/10.5194/hess-16-3309-2012>
- [5] Themeßl, M. J., Gobiet, A., & Heinrich, G. (2012). Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal. *Climatic Change*, 112(2), 449–468. <https://doi.org/10.1007/s10584-011-0224-4>