Supplement of Earth Syst. Dynam., 14, 507–517, 2023
https://doi.org/10.5194/esd-14-507-2023-supplement

Earth System Dynamics
Open Access

Supplement of

Direct and indirect application of univariate and multivariate bias corrections on heat-stress indices based on multiple regional-climate-model simulations

Liying Qiu et al.

Correspondence to: Eun-Soon Im (ceim@ust.hk) and Liying Qiu (liying.qiu@connect.ust.hk)

Table S1 Information of the climate simulation and reanalysis data used in this study

<table>
<tr><th colspan="3">Source</th><th>Spatial Resolution</th><th>Calendar</th><th>Coverage</th></tr>
<tr><td colspan="3">ERA5 Reanalysis</td><td>0.25° × 0.25°</td><td>Gregorian</td><td rowspan="8">Spatial: 125.25°-130.75°E, 33.25-38.5°N (South Korea land only)

Temporal: 1979-2014, 3 hourly</td></tr>
<tr><td rowspan="7">CORDEX-East Asia</td><td>Driving Model</td><td>RCM</td><td>Spatial Resolution</td><td>Calendar</td></tr>
<tr><td rowspan="6">UKESM1-0-LL (historical, r1i1p1) in CMIP6 dataset</td><td>CCLM</td><td>0.22°× 0.22°</td><td>360-day</td></tr>
<tr><td>GRIMs</td><td>25km ×25km</td><td>360-day</td></tr>
<tr><td>HadGEM3-RA (HG3RA)</td><td>0.22°× 0.22°</td><td>Gregorian</td></tr>
<tr><td>RegCM4</td><td>25km ×25km</td><td>360-day</td></tr>
<tr><td>WRF</td><td>25km ×25km</td><td>Gregorian</td></tr>
</table>

Text S1 Statistical metrics used in this study

$$RMSE = \sqrt{\frac{\sum_{n=1}^{N}(X_n - OBS_n)^2}{N}} \quad \text{(S1)}$$

where X_n and OBS_n is the data on the land grid *n* and *N* is the total number of land grids in the domain.

$$MAE_n = |X_n - OBS_n| \quad \text{(S2)}$$

MAE_n is the mean absolute error on the land grid *n*.

The Kolmogorov–Smirnov method (K–S test) is for testing the goodness of fit by comparing the maximum distance between the cumulative distribution functions (CDFs). The maximum discrepancy is defined as

$$D_{n,n^\tau} = sup\left|F_{1,n}(x) - F_{2,n^\tau}(x)\right| \quad \text{(S3)}$$

where $F_{1,n}(x)$ and $F_{2,n^\tau}(x)$ are the two distributions to be compared and sup is the supermum function.

Text S2 Introduction of BC methods

Five BC methods are applied in this study, including Linear Scaling (LS), Variance Scaling (VA), Empirical Quantile Mapping (EQM), Quantile Delta Mapping (QDM), and Multivariate Bias Correction with the N-dimensional probability density function (MBCn). The five methods cover a varying range of complexity in the transformation algorithm. We include a brief description of the five BC methods below.

1) Linear Scaling

The LS method is designed by simply removing the bias of monthly mean between the simulated variables in the reference period and the observation:

$$V_{REF/FUT}^{\mathrm{LS}} = V_{REF/FUT} + [\mu_m(V_{OBS}) - \mu_m(V_{REF})] \quad (4)$$

V is the variable to be corrected. REF and FUT stand for the reference and future periods in the correction, which in this case are calibration and validation periods, respectively. OBS is the observation or reanalysis data used as the standard for BC. μ_m is the monthly mean. LS is easy to use but it can only correct the monthly mean values, so it may cause large problem in the distribution of corrected output (Teutschbein & Seibert, 2012).

2) Variance Scaling

As the advance of LS method, VA is designed to correct the variance in addition to the mean by Chen and Dudhia (2001). The LS-corrected variables are shifted to a zero mean first:

$$V_{REF/FUT}^{*} = V_{REF/FUT}^{LS} - \mu_m(V_{REF/FUT}^{\mathrm{LS}}) \quad (5)$$

Then, the standard deviations are scaled based on the ratio of OBS and REF and then shifted back by the corrected mean:

$$T_{REF/FUT}^{\mathrm{VA}} = T_{REF/FUT}^{*} \times \left[\frac{\sigma_m(V_{OBS})}{\sigma_m(V_{REF}^{*})}\right] + \mu_m(V_{REF/FUT}^{\mathrm{LS}}) \quad (6)$$

, σ_m is the monthly standard deviation. VA guarantees that the adjusted variables have the same mean and standard deviation as OBS.

3) Quantile Delta Mapping (QDM)

QM methods are the most popular BC technique. They generally follow the basic principle of making the Cumulative Distribution Function (CDF) of the modeled data equal to that of the observation after correction.

$$V_{REF}^{\mathrm{EQM}} = F_{OBS}^{-1}[F_{REF}(V_{REF/FUT})] \quad (7)$$

F is the CDF function and F^{-1} is the inverse of CDF. EQM (Gudmundsson et al., 2012) solves Equation 7 using the empirical CDF of observed and modelled values instead of assuming parametric distributions. QDM is a more advanced EQM method that seeks to preserve the projected relative change in the quantiles for the future period by removing and adding the future trend before and after EQM (Cannon et al., 2015). First, it detrends the quantiles in the future model outputs and obtains a relative change in quantiles ($\Delta_{FUT}(\mathrm{t})$) between the reference period and the future time t:

$$\Delta_{FUT}(\mathrm{t}) = \frac{F_{FUT}^{(t)-1}[\tau_{FUT}(t)]}{F_{REF}^{-1}[\tau_{FUT}(t)]} = \frac{V_{FUT}(t)}{F_{REF}^{-1}[\tau_{FUT}(t)]} \;, \; \tau_{FUT}(t) \in \{0,1\} \quad (8)$$

τ is the non-exceedance probability of the value at time t. Then, the projected trends are reapplied to the quantiles obtained in EQM (Equation 4):

$$x_{FUT}^{\mathrm{QDM}} = x_{REF}^{\mathrm{EQM}} \cdot \Delta_{FUT}(\mathrm{t}) \quad (9)$$

4) Multivariate Bias Correction with the N-dimensional probability density function (MBCn)

MBCn is developed by Cannon (2018) with the goal of correcting the multivariate dependence in addition to individual variable adjustments. It applies random rotation to the multivariate data distribution iteratively, then corrects the rotated data with QDM until the corrected data distribution has converged to the observation. First, it rotates the observation $\boldsymbol{X}_{OBS}$, reference $\boldsymbol{X}_{REF}$ and future $\boldsymbol{X}_{FUT}$ datasets:

$$\widetilde{\boldsymbol{X}}^{[j]} = \boldsymbol{X}^{[j]} R^{[j]} \quad (10)$$

$\boldsymbol{X}$ is the $I \times N$ matrix where there are N variables to be corrected together. $[j]$ stands for the *j*th iteration to construct the $N \times N$ random orthogonal rotation matrix $\boldsymbol{R}^{[j]}$. Then, QDM is applied to each variable in $\widetilde{\boldsymbol{X}}_{REF}^{[j]}$ and $\widetilde{\boldsymbol{X}}_{FUT}^{[j]}$, using $\widetilde{\boldsymbol{X}}_{OBS}^{[j]}$ as reference. The output $\widehat{\boldsymbol{X}}_{REF}^{[j]}$ and $\widehat{\boldsymbol{X}}_{FUT}^{[j]}$ will be rotated back by

$$\widehat{\boldsymbol{X}}^{[j]} = \widehat{\boldsymbol{X}}^{[j]} \boldsymbol{R}^{[j]^{-1}} \quad (11)$$

The iteration is repeated until $\widehat{\boldsymbol{X}}$ converges to the $\widehat{\boldsymbol{X}}_{\boldsymbol{OBS}}$.

MBCn is flexible in considering the joint dependence of the variables and can also preserver the changes between the periods.

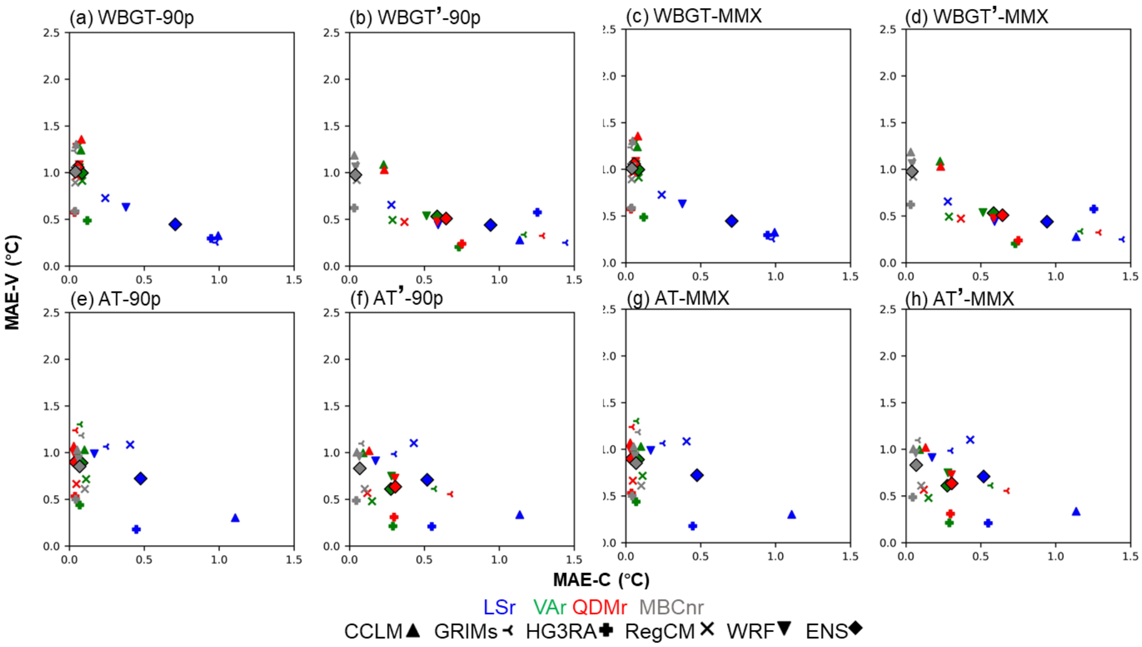

Figure S1: The MAE over South Korea (land only) for the calibration period (1997-2014, x-axis) and validation period (1979-1996, y-axis) in terms of the (a, b, e, f) 90p, and (c, d, g, h) MMX from (a, c) WBGT, (c, d) WBGT', (e, f) TW, and (g, h) TW'. The different colors stand for different BC methods, and the different markers stand for different RCMs. This result is from the reverse test.

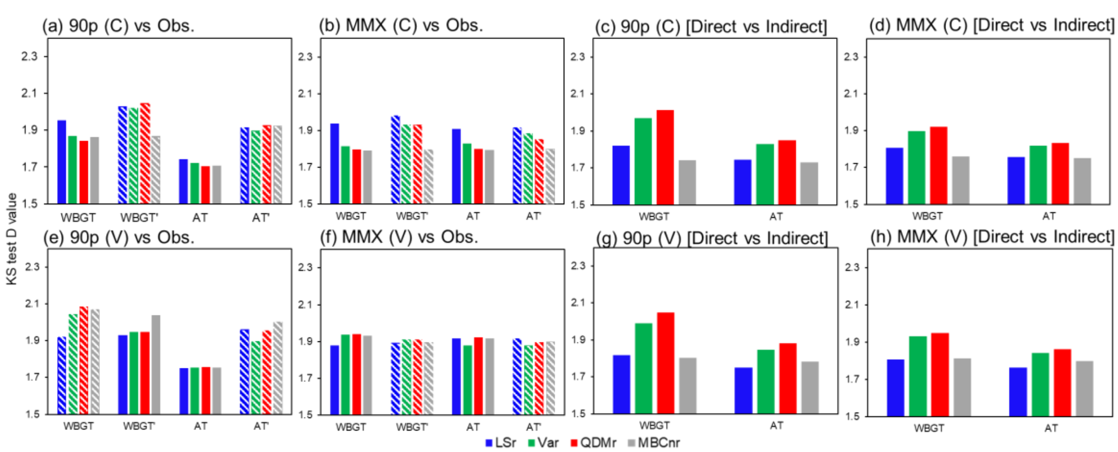

Figure S2: K-S test D value between bias-corrected output and observation for (a,e) 90p, and (b,f) MMX, and between direct and indirect corrected output for (c,g) 90p and (d,h) MMX. The D value is ensemble mean of 5 RCMs averaged over South Korea (land only). The different colors stand for different BC methods. The first row is for the Calibration period (C) and the second is for the Validation period. In (a, b, e, f), the solid and patterned fill is for the direct and indirect BC, respectively. This result is from the reverse test.

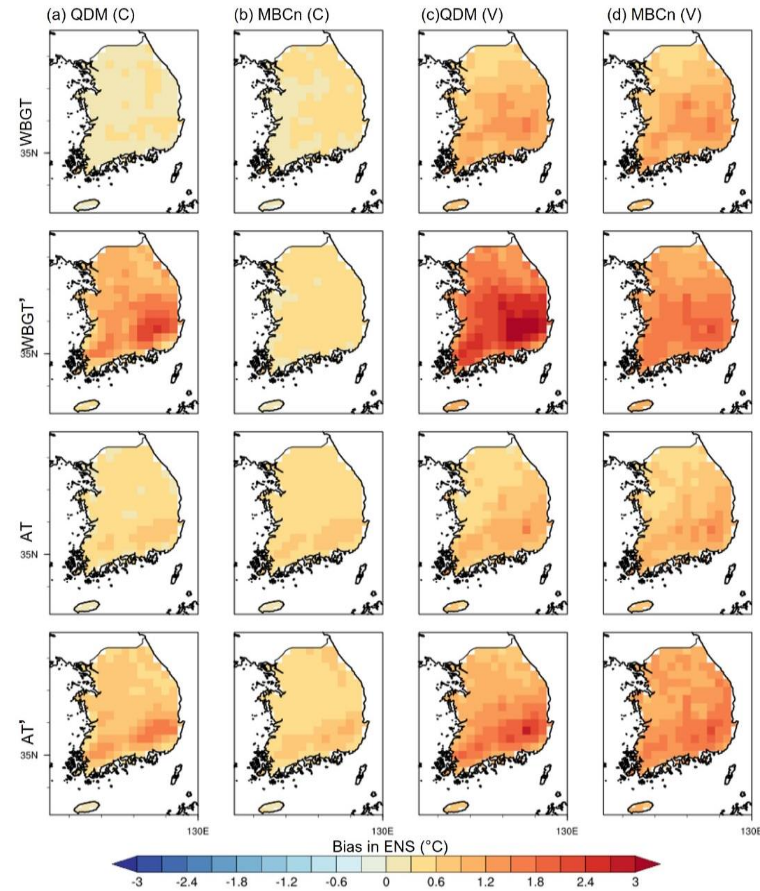

Figure S3: Spatial maps of the bias in the MMX during the calibration period (C) and validation (V) period corrected by QDM and MBCn in ENS. The first and third rows are the directedly corrected WBGT and AT. The second and fourth rows are the WBGT' and AT' calculated by the corrected T and RH.

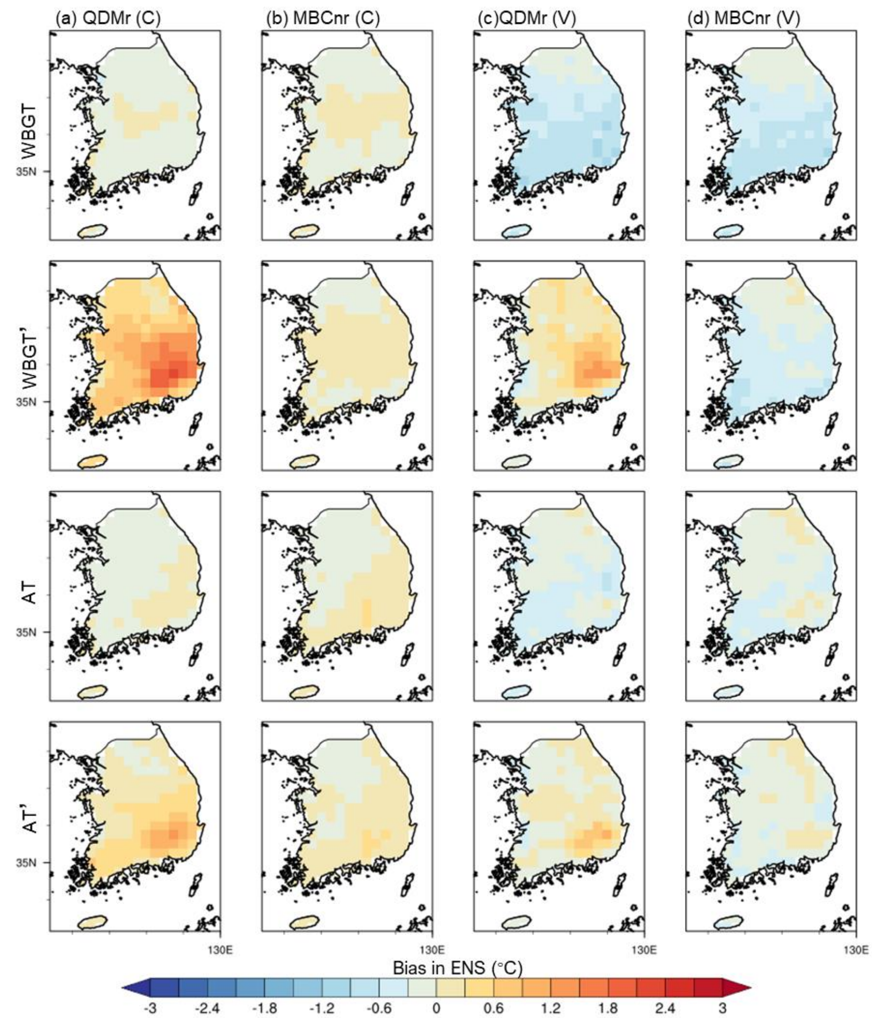

Figure S4: Same as Fig. S3 but for the reverse test.