

Direct and Indirect Application of Univariate and Multivariate Bias Corrections on Heat-stress Indices based on Multi-RCM Simulations

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

Source		Spatial Resolution	Calendar	Coverage	
ERA5 Reanalysis		0.25° × 0.25°	Gregorian	Spatial: 125.25°-130.75°E, 33.25-38.5°N (South Korea land only) Temporal: 1979-2014, 3 hourly	
CORDEX- East Asia	Driving Model	RCM	Calendar		
	UKESM1-0-LL (historical, r1i1p1) in CMIP6 dataset	CCLM	0.22° × 0.22°		360-day
		GRIMs	25km × 25km		360-day
		HadGEM3-RA (HG3RA)	0.22° × 0.22°		Gregorian
		RegCM4	25km × 25km		360-day
		WRF	25km × 25km		Gregorian

5 Text S1 Statistical metrics used in this study

$$RMSE = \sqrt{\frac{\sum_{n=1}^N (X_n - OBS_n)^2}{N}} \quad (B1)$$

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.

$$MAE_n = |X_n - OBS_n| \quad (B2)$$

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

10 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.

15 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:

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

20 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

25 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:

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

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

30
$$T_{REF/FUT}^{VA} = T_{REF/FUT}^* \times \left[\frac{\sigma_m(V_{OBS})}{\sigma_m(V_{REF}^*)} \right] + \mu_m(V_{REF/FUT}^{LS}) \quad (6)$$

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

3) Empirical Quantile Mapping

35 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.

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

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.

4) Quantile Delta Mapping (QDM)

40 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 ($\Delta_{FUT}(t)$) between the reference period and the future time t :

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

45 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):

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

5) 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
50 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 \mathbf{X}_{OBS} , reference \mathbf{X}_{REF} and future \mathbf{X}_{FUT} datasets:

$$\tilde{\mathbf{X}}^{[j]} = \mathbf{X}^{[j]} \mathbf{R}^{[j]} \quad (10)$$

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

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

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

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