Can sub-daily multivariate bias correction of regional climate model boundary conditions improve simulation of the diurnal precipitation cycle?

Youngil Kim[1], Jason Peter Evans[1], and Ashish Sharma[1]

[1]University of New South Wales

May 8, 2023

Abstract

The diurnal cycle is often poorly reproduced in global climate model (GCM) simulations, particularly in terms of rainfall frequency and amplitude. While improvements in the regional climate model (RCM) with bias-corrected boundaries have been reported in previous studies, they assumed that diurnal patterns are simulated correctly by the GCM, potentially leading to inaccuracies in the maximum rainfall timing and magnitude within the RCM domain. Here we provide the first examination of improvements to the diurnal cycle, within a RCM domain, achieved through the use of sophisticated bias-corrected lateral and lower boundary conditions. Results show that the RCMs with bias-corrected boundaries generally present improvement in capturing both rainfall timing and magnitude, particularly in northern Australia, where a strong diurnal pattern in rainfall is prevalent. We show that correcting systematic sub-daily multivariate bias in RCM boundaries improves the diurnal rainfall cycle, which is particularly important in regions where short-term intense precipitation occurs.

Hosted file

962758_0_art_file_10960456_rx68hw.docx available at https://authorea.com/users/615110/articles/641679-can-sub-daily-multivariate-bias-correction-of-regional-climate-model-boundary-conditions-improve-simulation-of-the-diurnal-precipitation-cycle

Hosted file

962758_0_supp_10960517_rs674q.docx available at https://authorea.com/users/615110/articles/641679-can-sub-daily-multivariate-bias-correction-of-regional-climate-model-boundary-conditions-improve-simulation-of-the-diurnal-precipitation-cycle

Can sub-daily multivariate bias correction of regional climate model boundary conditions improve simulation of the diurnal precipitation cycle?

Youngil Kim[1], Jason P. Evans[2], Ashish Sharma[1*]

[1]School of Civil and Environmental Engineering, University of New South Wales, Sydney, NSW 2052, Australia

[2]Climate Change Research Centre and ARC Centre of Excellence for Climate Extremes, University of New South Wales, Sydney, New South Wales, Australia

Corresponding author: Ashish Sharma (a.sharma@unsw.edu.au), School of Civil and Environmental Engineering, University of New South Wales, Sydney, NSW 2052, Australia

Key points

- RCM with uncorrected boundaries poorly represented 3-hourly precipitation with only 20 ~ 30% agreements compared to observations.
- Sub-daily bias correction of RCM lateral boundaries improves the sub-daily representation of fields entering the RCM domain.
- Sub-daily multivariate bias correction on the boundaries resulted in additional improvements, especially in northern Australia.

Abstract

The diurnal cycle is often poorly reproduced in global climate model (GCM) simulations, particularly in terms of rainfall frequency and amplitude. While improvements in the regional climate model (RCM) with bias-corrected boundaries have been reported in previous studies, they assumed that diurnal patterns are simulated correctly by the GCM, potentially leading to inaccuracies in the maximum rainfall timing and magnitude within the RCM domain. Here we provide the first examination of improvements to the diurnal cycle, within a RCM domain, achieved through the use of sophisticated bias-corrected lateral and lower boundary conditions. Results show that the RCMs with bias-corrected boundaries generally present improvement in capturing both rainfall timing and magnitude, particularly in northern Australia, where a strong diurnal pattern in rainfall is prevalent. We show that correcting systematic sub-daily multivariate bias in RCM boundaries improves the diurnal rainfall cycle, which is particularly important in regions where short-term intense precipitation occurs.

Plain language summary

Global climate models have limitations in simulating precipitation characteristics at sub-daily time scales, leading to significant bias in capturing rainfall timing and magnitude in regional climate model simulations that use global climate model data as input. Previous studies have applied various bias correction approaches to the boundary conditions of regional climate models, but they often assume that global climate model sub-daily rainfall patterns are simulated correctly, leading to a significant bias in maximum rainfall timing and magnitude within the domain. This study investigates the consequences of multivariate sub-daily bias correction on the boundary conditions in order to evaluate the model performance for short-term precipitation events. The results show that regional climate models with bias-corrected boundary conditions exhibit improvement across Australia, especially in northern Australia, which regularly experiences intense sub-daily rainfall.

1. Introduction

Predicting short-term precipitation events and their possible change into the future is of significant interest to water resource managers and stakeholders for evaluating the frequency of extreme precipitation and convective storms. Although Global Climate Models (GCMs) are capable of capturing precipitation patterns at daily or longer time scales (Westra et al., 2014), their ability to represent changes in sub-daily precipitation is questionable due to the errors introduced by the deficiencies with regard to their discretization in time (Dai (2006); Rosa and Collins (2013)) and space (Lee et al. (2007); Wang et al. (2007)). In terms of the diurnal cycle, GCM simulations generally have poorly reproduced rainfall frequency and the amplitude (Stephens et al., 2010). This is due to multiple factors including the limitations of the convective scheme that can make the models produce moist convection several hours early, and can be dependent on the horizontal model grid spacing (Wang et al. (2007); Wang et al. (2011)).

Recent studies have used Regional Climate Models (RCMs), forced with GCM datasets through the input boundary conditions, to overcome some of the deficiencies noted above. These simulations show evidence that higher-resolution climate models can offer improvement at finer time scales especially when correct boundary inputs are used (Maraun et al., 2010). Evans and Westra (2012) showed that a RCM with a 10 km resolution has capable of reproducing the diurnal variability of rainfall occurrences, which is consistent with results from other studies (Gutowski et al. (2003); Yamada et al. (2012); Prein et al. (2013)).

Although the GCM-driven RCM simulations generally show similar diurnal variation to that of the observation in the mentioned studies, it, unfortunately, suffers from inherent limitations of the input boundary conditions driven by the GCM dataset, which contain systematic biases (Kim et al., 2020). These improper boundary conditions introduce bias in time and space that can be propagated into RCM outputs (Kim et al., 2021).

To overcome the scale gap and to reduce systematic bias, bias correction approaches have routinely been applied to the boundary conditions. Different techniques are in use for impact assessment, ranging from simple climatological correction (Xu and Yang (2012); Bruyere et al. (2014)) to a sophisticated method, potentially extended to temporal persistence and inter-variable relationships (Rocheta et al. (2017); Kim et al. (2020); Kim et al. (2023b)).

Rocheta et al. (2017) showed that more complex bias correction techniques generally produced a better representation of the rainfall pattern, and correcting the sea surface temperature plays a crucial role in improving the simulation of rainfall. Investigation into extremes using several univariate bias corrections was also addressed by Kim et al. (2020). They found that RCM simulation with a complex method better represents the seasonal extremes than the simple methods.

Previous work highlights the need for physical consistency, including inter-variable relationships, in the boundary conditions for regional climate models (Rocheta et al. (2014); Kim et al. (2021)). Recently, (Kim et al., 2023b) investigated the RCM with multivariate bias-corrected boundaries because the univariate bias correction techniques can produce physical inconsistencies among the variables. They showed that the complex bias correction approaches better present rainfall variability than the univariate bias correction approaches. While some degradations were shown after the generation of the boundaries, the physical relationships between the atmospheric variables along the boundaries were preserved inside the domain. They highlighted that RCM with multivariate bias correction generally better represents extreme events for three surface variables compared to the RCMs with univariate bias correction.

Although improvements in RCM with bias-corrected boundaries have been reported in previous studies, they corrected bias at daily and longer time scales, assuming that diurnal patterns are preserved properly inside the domain. This may impact the RCM simulated maximum rainfall timing and magnitude.

Observational studies that have identified key features of the global diurnal cycle characteristics note that precipitation typically peaks at mid-afternoon, showing a stronger diurnal cycle over land than over the ocean (Dai (2006); Watters et al. (2021)). In addition, the diurnal amplitude is stronger in summer, and the diurnal cycle of precipitation accumulation is related to its occurrence (Watters et al., 2021).

Therefore, the present study evaluates the RCMs focusing on the simulation of the diurnal cycle of precipitation for summer (DJF) over Australia using RCMs with corrected boundary inputs, using the natural resource management (NRM) Super-clusters (https://www.climatechangeinaustralia.gov.au/en/overview/methodology/nrm-regions/) as the regions where results are assessed. In all, four RCM simulations which differ based on the source and bias correction method of their lateral boundary conditions are assessed.

The rest of the paper is as follows. The datasets and methods are described in section 2. Section 3 presents the results, and section 4 provides a discussion and limitations of this work. Finally, conclusions are described in section 5.

2. Data and Method

2.1. Models and data

In this study the Australian Community Climate and Earth System Simulator Earth System Model Version 1.5 (ACCESS-ESM1.5) GCM simulation made available for contributing to the internationally coordinated Coupled Model Intercomparison Project Phase 6 (CMIP6) (Ziehn et al., 2020) was used. For the RCM simulations, we used the Weather Research and Forecasting model (WRF) with dynamical core (ARW), version 4.2.1 (Skamarock et al., 2019). The ACCESS-ESM1.5 has a resolution of approximately 1.875°EW x 1.25°NS with 38 vertical hybrid sigma levels to 40 km from the surface.

The ERA5, the fifth-generation model reanalysis of the global climate from the European Centre for Medium-range Weather Forecasts (ECMWF) (Hersbach et al., 2020) with a resolution of 31 km with 37 pressure levels, was treated as an observation for correcting GCM biases. Here, the ERA5-driven RCM simulation was considered a "perfect" simulation.

The five variables in the RCM lateral boundary conditions (zonal wind u (m/s), meridional wind v (m/s), specific humidity q (g/kg), temperature T (K)) and lower (sea surface temperature SST (K)) were corrected towards ERA5. To correct the variables, the ERA5 variables were first regridded using conservative remapping for the specific humidity and bilinear for the other four variables, toward those of the ACCESS-ESM1.5. The RCM boundary conditions were built after bias correction, and all other variables remained identical.

The following parameterizations, evaluated in previous studies Evans and McCabe (2010); Evans et al. (2012)), were used for WRF simulations: Mellor–Yamada–Janjic planetary boundary (Janjic, 1994); Betts–Miller–Janjic cumulus parameterization scheme (Janjic, 1994); WRF double-moment 5-class microphysics scheme (Lim and Hong, 2010); Dudhia shortwave radiation scheme (Dudhia, 1989); Rapid Radiative Transfer Model (RRTM) longwave radiation (Mlawer et al., 1997) and unified Noah land surface scheme (Tewari et al., 2004).

The Australasian Coordinated Regional Climate Downscaling Experiment (CORDEX) domain (http://www.cordex.org/community/domain-australasiacordex.html) at a resolution of 50 km with 50 vertical levels was used for downscaling.

Here, the model simulations cover 1982 – 2012 (31 years), and the first year was considered a spin-up period to remove issues related to the equilibrium state for the soil moisture (Cosgrove et al. (2003); Chen et al. (2007)). The spin-up year was not included in the subsequent analysis.

2.2. Bias correction approaches

Two bias correction approaches were applied to the regional climate model boundary conditions. The first corrected the multivariate relationship among atmospheric variables, such as specific humidity, temperature, zonal and meridional winds, and surface variable sea surface temperature toward those of the reanalysis data. The second further corrected the distribution of these variables using quantile mapping which adjusts the RCM input boundaries at a sub-daily time scale to match the reanalysis data.

2.2.1. Multivariate bias correction (DMBC)

The daily multivariate recursive bias correction (DMBC) was implemented on a daily GCM dataset to correct inter-variable relationships among the three-dimensional atmospheric and surface variables at the RCM input boundaries. This approach aims to correct climatological statistics as well as the lag0 and lag1 auto- and cross-correlation attributes and has been used in previous studies (Mehrotra and Sharma (2015), Kim et al. (2023b)). It has shown improvement in representing mean, variance, persistence, and physical consistency between the atmospheric variables at multiple time scales. DMBC presented improvement in the simulation of extreme events compared to the univariate bias corrections and generally better represented the rainfall characteristics (Kim et al., 2023b). The DMBC was applied with respect to ERA5 before downscaling over 31 years.

Denoting the standardized daily GCM and observed vectors with zero mean and unit variance are denoted $Z_{()}^{g}$and $Z_{()}^{o}$, respectively, and using bold capital letters indicate matrices, a simplified Multivariate first-order AutoRegressive model (MAR1) that forms the basis for DMBC can be expressed as (Salas, 1980):

$$Z_t^o = \mathbf{C}Z_{t-1}^o + \mathbf{D}\varepsilon_t \quad (1a)$$

and

$$Z_t^g = \mathbf{E}Z_{t-1}^g + \mathbf{K}\varepsilon_t, \quad (1b)$$

where $\mathbf{C}$ and $\mathbf{D}$ are the coefficient matrices that contain the lag0 cross-correlations of the reanalysis data. $\mathbf{E}$ and $\mathbf{K}$ are the coefficient matrices of the GCM data. A simplified model used here considers $\mathbf{C}$ and $\mathbf{E}$ as diagonal matrices to avoid many parameters that may cause estimation errors (Kim et al., 2023b). Eq. (1b) can be simplified for the random vector ε_t as follows:

$$\varepsilon_t = \mathbf{K}^{-1}\left[Z_t^g - \mathbf{E}\dot{Z}_{t-1}^g\right], \quad (2)$$

where ε_t is a standardized vector after removing the correlation attributes from the GCM data series. Eq. (2) is then used to obtain a modified $\dot{Z}_t^g$ that maintains the observed lag0 and lag1 attributes as follows:

$$\dot{Z}_t^g = \mathbf{C}\dot{Z}_{t-1}^g + \mathbf{D}\mathbf{K}^{-1}Z_t^g - \mathbf{D}\mathbf{K}^{-1}\mathbf{E}Z_{t-1}^g. \quad (3)$$

Adding back the means and standard deviations of observed data provides bias-corrected attributes with appropriate means, standard deviations, lag1 auto-, and lag0 cross-dependence.

Matrices $\mathbf{C}$ and $\mathbf{E}$ or $\mathbf{D}$ and $\mathbf{K}$ can be derived as follows (Matalas, 1967):

$$\mathbf{C} \vee \mathbf{E} = \mathbf{B}_1\mathbf{B}_0^{-1} \quad (4a)$$

and

$$\mathbf{D}\mathbf{D}^T \vee \mathbf{K}\mathbf{K}^T = \mathbf{B}_0 - \mathbf{B}_1\mathbf{B}_0^{-1}\mathbf{B}_1^T, \quad (4b)$$

where $\mathbf{B}_0$ and $\mathbf{B}_1$ are the lag0 and lag1 cross-correlation matrices, and the elements of $\mathbf{D}$ and $\mathbf{K}$ can be found by singular value decomposition of $\mathbf{D}\mathbf{D}^T$ and $\mathbf{K}\mathbf{K}^T$, respectively. The elements of $\mathbf{C}$ or $\mathbf{E}$, and $\mathbf{D}$ or $\mathbf{K}$ of simplified MAR1 corresponding to variables i and j are defined as:

$$\left.\begin{array}{ll} c^{i,j} \vee e^{i,j} = b_1^{i,j}, & \text{if } i = j \\ = 0, & \text{otherwise} \end{array}\right\} \quad (5a)$$

and

$$d^{i,j} \vee k^{i,j} = b_0^{i,j}\left(1 - b_1^{i,j}b_1^{i,j}\right). \quad (5b)$$

The elements of $\mathbf{B}_0$ and $\mathbf{B}_1$ corresponding to variables i and j can be derived by the sets of standardized time series as:

$$b_0^{i,j} = \frac{1}{N}\sum_{t=1}^{N} z_t^i z_t^j, \quad (6a)$$

and

$$b_1^{i,j} = \frac{1}{N-1}\sum_{t=2}^{N} z_t^i z_{t-1}^j, \quad (6b)$$

where N is the total number of datasets.

MAR1 can be defined with periodic parameters that are derived for each period separately.

$$\dot{Z}_{y,\tau}^g = \mathbf{C}_\tau\dot{Z}_{y,\tau-1}^g + \mathbf{D}_\tau\mathbf{K}_\tau^{-1}Z_{y,\tau}^g - \mathbf{D}_\tau\mathbf{K}_\tau^{-1}\mathbf{E}_\tau Z_{y,\tau-1}^g. \quad (7)$$

Matrices $\mathbf{C}$ and $\mathbf{E}$ or $\mathbf{D}$ and $\mathbf{K}$ are as:

$$\mathbf{C}_\tau \vee \mathbf{E}_\tau = \mathbf{B}_{1,\tau}\mathbf{B}_{0,\tau-1}^{-1} \quad (8a)$$

and

$$\mathbf{D}_\tau\mathbf{D}_\tau^T \vee \mathbf{K}_\tau\mathbf{K}_\tau^T = \mathbf{B}_{0,\tau} - \mathbf{B}_{1,\tau}\mathbf{B}_{0,\tau-1}^{-1}\mathbf{B}_{1,\tau}^T, \quad (8b)$$

where τ is time internal and y is year. $Z_{y,\tau}^{g}$ indicates the GCM periodic series having a zero mean and unit variance. The elements of **D** or **K** of simplified periodic MAR1 corresponding to variables i and j are defined as:

$$d_{\tau}^{i,j} \vee k_{\tau}^{i,j} = b_{0,\tau}^{i,j}\left(1 - b_{1,\tau}^{i,j} b_{0,\tau-1}^{i,j} b_{1,\tau}^{i,j}\right). \quad (9)$$

The outputs were then combined with the nesting approach (Kim et al., 2023b), indicating the daily GCM data incorporated the effect of bias correction at longer time scales: month, season, and annual. This means that the bias-corrected GCM values exhibit the same persistence-related attributes compared to observed values. The bias-corrected variables at multiple time scales can be employed as a form of the weighting factor for the raw daily GCM data (Srikanthan and Pegram (2009)):

$$\dot{X}_{d}^{g} = \left(\frac{\dot{X}_{d,m,s,y}^{g}}{X_{d,m,s,y}^{g}}\right) \times \left(\frac{\dot{X}_{m,s,y}^{g}}{X_{m,s,y}^{g}}\right) \times \left(\frac{\dot{X}_{s,y}^{g}}{X_{s,y}^{g}}\right) \times \left(\frac{\dot{X}_{y}^{g}}{X_{y}^{g}}\right) \times X_{d}^{g}, \quad (10)$$

where $\dot{X}_{d}^{g}$ is the bias-corrected daily value and X_{d}^{g} is the original daily value. Subscript *d*, *m*, *s*, *y* indicate day, month, season, and year, respectively. $X_{m,s,y}$, $X_{m,s,y}$, $X_{s,y}$, X_{y} indicate the aggregated values from daily to monthly, seasonal, and yearly, respectively.

2.2.2. Sub-daily multivariate bias correction (SDMBC)

Bias at the sub-daily time scale, if present beyond what is corrected at the daily level using DMBC, is addressed by correcting the entire frequency distribution of the variables at a 6-hourly time scale after implementing the multivariate bias correction at a daily time scale (Figure D).

We first resampled 6-hourly variables to a daily time scale for multivariate bias correction and calculated sub-daily fractions (SF) to rescale the outputs to a 6-hourly time scale.

The daily corrected variables ($\dot{X}_{d}^{g}$) were then rescaled to 6 hours using SF, as noted above. For sub-daily bias correction, quantile-mapped bias-corrected SF was applied to the daily corrected variables so that the results can reproduce the observed distributional and dependence attributes – mean, standard deviation, lag1 auto-, and lag0 cross-correlation at multiple time scales.

Quantile mapping (QM) generally corrects the distribution of modeled data based on observed data. QM was designed to preserve long-term quantile changes based on the model's cumulative distribution functions (CDF). The transformation function used in this study can be defined as follows:

$$\dot{\mathrm{SF}}^{g} = F_{o}^{-1}(F_{g}(\mathrm{SF}^{g})), \quad (11)$$

where F_g and F_o are the CDF of the raw GCM and the reanalysis data, respectively. $F_{o}^{-1}$is the inverse CDF corresponding to the reanalysis data. SF^{g} and $\dot{\mathrm{SF}}^{g}$ are the uncorrected and quantile-mapped corrected sub-daily fractions. Thus, the 6-hourly simulations rescaled using SF^{g} and $\dot{\mathrm{SF}}^{g}$ indicate corrections based on DMBC and SDMBC, respectively.

2.3. Performance assessment

Several statistics were used to assess the RCM simulations with uncorrected and bias-corrected boundary conditions. The mean absolute error (MAE) was defined as:

$$\mathrm{MAE} = \frac{\sum_{n=1}^{I}|M_n - O_n|}{I} \tag{11}$$

where M_n and O_n are the models and observed data at each grid cell, and I is the total number of grid cells, respectively.

The bias in the means between the model and observed data was calculated at each grid cell for each vertical level and can be represented as:

$$\mathrm{Bias} = \frac{\sum_{n=1}^{I}(M_n - O_n)}{I}. \tag{12}$$

To determine whether two samples, model simulations and observation, come from populations with the same distribution, the two sample Kolmogorov-Smirnov (K-S) test was used. It employs the probability distribution of the quantity L, which is defined as the maximum absolute difference between the cumulative frequency distributions of the model (F_g) and observed (F_o) simulations as follows:

$$L = \max_x \left|F_g(x) - F_o(x)\right|. \tag{13}$$

The null hypothesis is both samples come from a population with the same distribution and is rejected if the test statistic (L) is greater than the critical value ($L_{critical} = c(\alpha)\sqrt{\frac{n_1+n_2}{n_1 n_2}}$), where n_1 and n_2 are the sample sizes of modeled and observation, respectively, or the p-value is lower than the threshold of α (here, 0.05).

3. Results

In this section, the bias correction approaches, DMBC and SDMBC, were evaluated along the boundary conditions in comparison to the reanalysis data. The RCM simulations were then compared to the ERA5-driven RCM simulation, considered the "perfect" boundary simulations, based on the K-S test across the Australasian CORDEX domain. Furthermore, we analyzed the magnitude of the range of the diurnal precipitation cycle in northern Australia. The outermost five grid cells were excluded as the relaxation zone.

The four RCM simulations used here are named RCM(ERA5), RCM(GCM), RCM(DMBC), and RCM(SDMBC), and indicate that RCM with reanalysis-driven boundary conditions, RCM with uncorrected GCM boundary conditions, RCM with daily multivariate bias-corrected boundary conditions, and RCM with sub-daily multivariate bias-corrected boundary conditions, respectively.

3.1. Is bias correction in RCM lateral boundary conditions at a sub-daily scale necessary?

We first assess the performance of the two bias correction approaches, DMBC and SDMBC, along the lateral boundary conditions at a sub-daily time scale. Both approaches correct inter-variable relationships as well as mean, variance, and persistence attributes at multiple time scales, which was previously shown to be effective in a study by (Kim et al., 2023b). Figures A1-4 present a scatter plot comparing the uncorrected and bias-corrected atmospheric variables to the ERA5 datasets at a daily time scale in terms of multivariate

relationships for all vertical levels over 30 years for each boundary. The results indicate that both DMBC and SDMBC effectively correct the inter-variable relationship between the three atmospheric variables, with points tending to cluster along the 45-degree line.

We then evaluate the approaches at a sub-daily time scale (6-hourly) using the K-S test along the boundaries. Figure 1 presents a pdf for temperature (K) along the western boundary of each model for four seasons, DJF, MAM, JJA, and SON. The black dotted vertical line represents the critical value at a p-value of 0.05. If the critical value calculated at a given grid cell is higher than the black dotted line, it means that the values for that model differ from those of ERA5. As seen in the figure, SDMBC is generally more similar to ERA5, with more than 76% of values in agreement. DMBC also shows improvement compared to the GCM, with over 19% of values in agreement, compared to 0.3 to 1.4% for GCM. Although sub-daily correction was applied to the boundary variables, the results do not match those of ERA5 perfectly. The possible reasons for this are discussed in Section 4. Other variables along the east, north, and south boundaries are provided in the supplementary material (Figures B1 to B11).

Figure 1. PDF of each model for temperature (K) along the western boundary over 30 years for all vertical levels for each season. SDMBC and DMBC presented here mean the sub-daily- and daily multivariate bias correction, respectively. The black dotted line indicates a critical

value calculated using the K-S test at each grid cell. The percentage represents a ratio between the number of grid cells that agree with ERA5 and the total number of grid cells.

3.2. Can the diurnal precipitation pattern be improved inside the RCM domain?

This section investigates the model performance using the K-S test inside the RCM domain at a sub-daily time scale to assess whether the impact of bias correction can be preserved. Figure 2 shows the percentage of grid cells showing agreement between the models and ERA5-driven RCM simulation for precipitation over the Australasian CORDEX domain at a 3-hourly time scale over 30 years for four seasons, DJF, MAM, JJA, and SON. The results show that RCMs with bias-corrected boundary conditions better represent the 3-hourly precipitation than RCM(GCM). RCM(GCM) produces relatively low agreement, ranging from 20% to 30% over the domain. This implies that bias corrections need to be used before downscaling to simulate diurnal patterns appropriately. In contrast, RCM(DMBC) and RCM(SDMBC) improve the percentage, ranging from 50% to 70%. From the result, we see that RCM(SDMBC) generally presents similar performance when compared to RCM(DMBC) in all seasons, suggesting that the impact of correcting the sub-daily distribution is relatively small compared to the impact of bias correcting daily and longer timescales. Of interest here is whether this similarity is consistent across regions and seasons, or becomes more enhanced in seasons and areas where the diurnal cycle is significant.

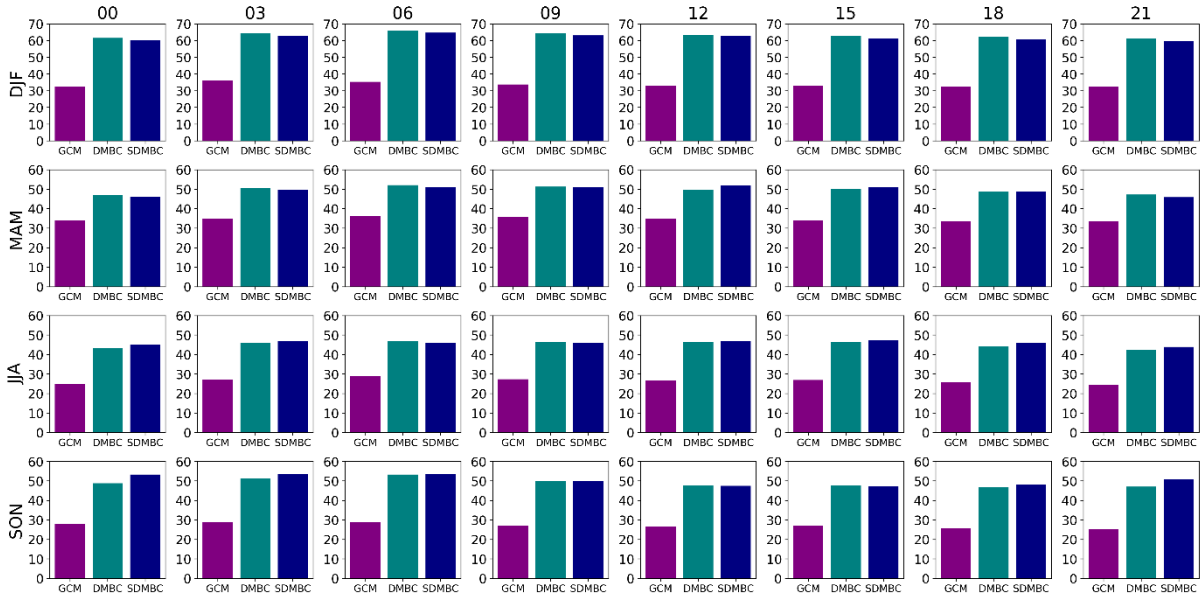

Figure 2. The percentage of agreement based on the K-S test for precipitation of each model compared to RCM(ERA5) for four seasons, DJF, MAM, JJA, and SON, at a 3-hourly time scale over 30 years across the Australasian CORDEX domain. The models, GCM, DMBC, and SDMBC, presented here indicate that RCM(GCM), RCM(DMBC), and RCM(SDMBC).

3.3. Can the timing of the maximum and the range of diurnal precipitation be better represented?

Precipitation typically peaks at mid-afternoon, showing a stronger diurnal cycle over land. However, the GCM simulations often poorly reproduce the diurnal rainfall frequency and

amplitude. This suggests that RCM simulations using boundary conditions generated from GCM datasets may have a bias in the simulation of precipitation at a sub-daily scale. Here we examine whether explicitly correcting this diurnal bias at the boundary improves the RCM simulation within its domain. Figure 3 shows the mean absolute error of maximum 3-hourly precipitation timing and magnitude of the diurnal range (hereafter magnitude range) of 3-hourly precipitation (maximum minus minimum in a day) averaged over 30 years across the regions. The results show that RCMs with bias-corrected boundaries perform similarly in the simulation of maximum precipitation timing. Although bias still can be seen even after correcting bias in the boundaries, the results show the effectiveness of bias correction for the diurnal precipitation cycle. RCM(DMBC) and RCM(SDMBC) generally show lower bias than RCM(GCM) except for maximum timing averaged over all seasons in northern Australia. RCM(SDMBC) better represents the magnitude range of the diurnal precipitation cycle compared to other models, showing the lowest bias, particularly in northern Australia. From the map presented in Figure C, RCM(SDMBC) produces bias close to zero, particularly in the northeast regions.

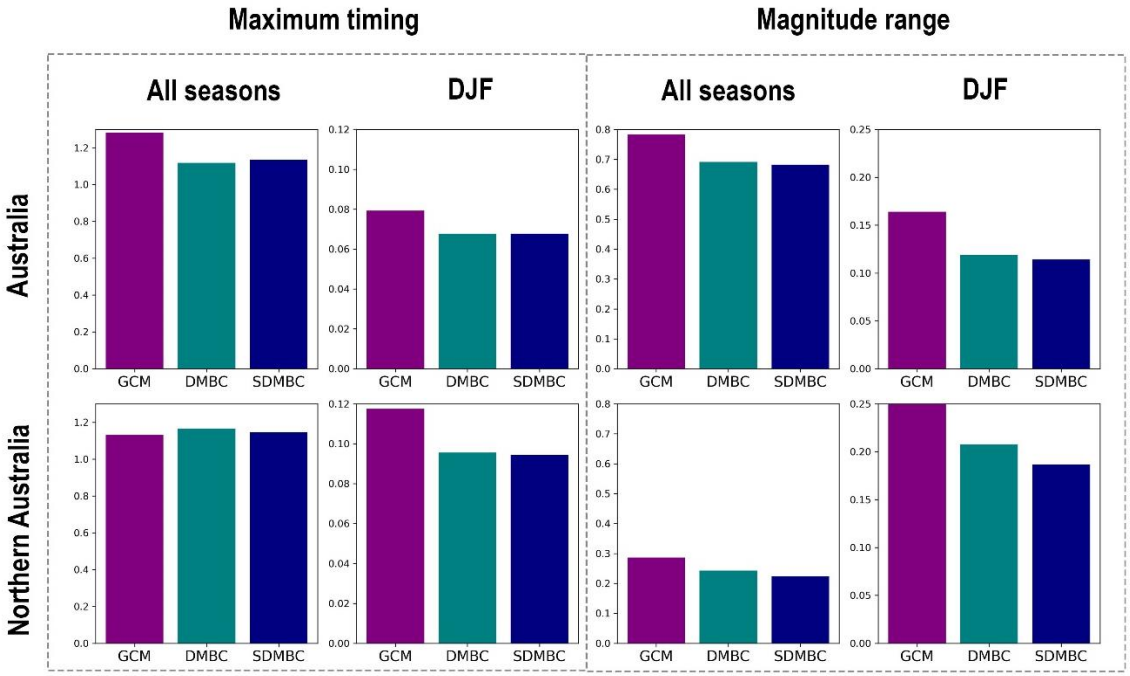

Figure 3. The MAE of maximum timing and magnitude range of 3-hourly precipitation in a day over 30 years for all seasons and DJF across Australia (first row) and northern Australia (second row), following the natural resource management (NRM) Super-clusters. The models, GCM, DMBC, and SDMBC indicate that RCM(GCM), RCM(DMBC), and RCM(SDMBC), respectively.

4. Discussion and Conclusions

Global climate models (GCMs) often exhibit limitations in simulating sub-daily precipitation, in part due to errors introduced by the deficiencies of the convection scheme, leading to producing moist convection several hours earlier than the observations. While some

363 improvements in reproducing the diurnal variability of rainfall occurrences using the regional
364 climate models (RCMs) have been noted in previous studies, their application is hindered by
365 systematic biases inherent in the GCM data (Kim et al., 2023a). Although various mathematical
366 approaches have been applied to the RCM boundary conditions, no study has yet been
367 conducted to address the sub-daily correction, which is important for accurately predicting the
368 maximum sub-daily rainfall timing and magnitude.

369 With this in mind, we conducted the first ever study of how well more sophisticated
370 alternatives for correcting systematic bias at the sub-daily time scale on RCM boundary
371 conditions improves the diurnal precipitation patterns within the domain, particularly in
372 northern Australia, where intense sub-daily rainfall is often present (Guerreiro et al. (2018);
373 Minobe et al. (2020)).

374 We find that sub-daily bias correction of the lateral boundary conditions is effective at
375 improving the sub-daily representation of the RCM input variables. We also find that RCMs
376 with bias-corrected boundary conditions improved the simulation of sub-daily precipitation,
377 and the findings can be summarized as follows.

378 RCM with uncorrected boundary conditions, RCM(GCM), represented a significant
379 bias in simulating 3-hourly precipitation across the Australasian CORDEX domain, with only
380 0.3 ~ 1.4% agreement for temperature according to the K-S test compared to the ERA5-driven
381 RCM outputs. This indicates that systematic bias introduced by the GCM datasets through the
382 input boundary conditions was not sufficiently reduced in the relaxation zone (five grid cells
383 from the outermost zones of each boundary), causing a significant bias in simulating sub-daily
384 precipitation within the domain. In contrast, the RCMs with bias-corrected boundary
385 conditions, RCM(DMBC) and RCM(SDMBC), showed improvement with 19.9 ~ 46.8% and
386 76.6 ~ 84% agreement with the ERA5-driven RCM outputs, respectively. Despite the fact that
387 biases related to short-term precipitation were reduced, there are still discrepancies between
388 the models and the ERA5-driven RCM outputs. Further investigation is necessary to enhance
389 the model's ability to simulate the diurnal precipitation cycle.

390 RCM(DMBC) and RCM(SDMBC) showed generally similar performance over the
391 domain, despite the latter being further corrected using sub-daily quantile mapping along the
392 boundaries. This similarity indicates that most of the bias exists at daily and longer time scales
393 and is corrected without consideration of the sub-daily biases. Also contributing to this
394 similarity is the tendency for the relaxation zone to moderate the corrections made to the
395 boundary conditions (Rocheta et al., 2020). Since the additional sub-daily correction is often
396 quite small it may not survive through the relaxation zone. The end result is that SDMBC often
397 only provides a relatively small improvement over DMBC.

398 With regard to the model simulation for the precipitation maximum timing and
399 magnitude, the RCM(DMBC) showed improvement compared to RCM(GCM), but the
400 RCM(SDMBC) generally showed further improvement exhibiting the lowest bias for
401 magnitude range. An interesting finding from the results was that RCM(GCM) showed a
402 relatively low bias in simulating the maximum precipitation timing, even during high rainfall
403 season (DJF). This aligns with a previous study that found that most RCMs tend to present
404 more accurate diurnal range simulations during DJF-SON (Di Virgilio et al., 2019). This
405 implies that the state-of-the-art GCM simulation used in this study may have the ability to
406 capture the diurnal nature of monsoon rainfall such as that present in Northern Australia.

407 The results demonstrate that sub-daily correction on the boundary conditions can
408 improve sub-daily precipitation patterns and generally showed better performance, especially
409 in northern Australia, which experiences a strong diurnal cycle in precipitation.

Although sub-daily correction using quantile mapping was applied to the boundary conditions, the distributions of the bias-corrected variables were not perfectly aligned with the reanalysis data (as shown in Figure 1). This suggests that the multivariate bias correction may undermine the effects of sub-daily correction, as the daily bias-corrected variables are converted into 6-hourly variables using a fraction factor determined by quantile mapping prior to multivariate bias correction. Moreover, while sub-daily multivariate bias correction showed notable improvement in the lateral boundary conditions, its impact on the diurnal precipitation cycle was reduced far from the boundaries, indicating that the RCM plays a significant role in driving the output variables. Despite these factors, sub-daily multivariate bias correction of RCM lateral boundary conditions consistently produced the best simulated climate within the RCM domain and should be considered for use within regional climate projection projects like CORDEX (Evans et al., 2021) or NARCliM (Ji et al., 2022).

Acknowledgments

Youngil Kim is supported by the UNSW Scientia PhD Scholarship Scheme. J.P.E. was supported via the ARC Centre of Excellence for Climate Extremes (CE170100023) and the Australian Government under the National Environmental Science Program. This research was undertaken using the computational cluster Katana supported by Research Technology Services at UNSW Sydney. This research includes computations with the assistance of resources from the National Computational Infrastructure (NCI Australia), an NCRIS-enabled capability supported by the Australian Government.

Data availability

ERA5 data is available through the European Centre for Medium-Range Weather Forecasts (ECMWF) online data archive (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels?tab=overview). The authors acknowledge the modeling groups, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the WCRP's Working Group on Coupled Modeling (WGCM) for their roles in making available the WCRP CMIP6 multimodel data set (https://esgf-node.llnl.gov/search/cmip6/). Support of this data set is provided by the Office of Science, U.S. Department of Energy.

References

BRUYERE, C. L., DONE, J. M., HOLLAND, G. J. & FREDRICK, S. 2014. Bias corrections of global models for regional climate simulations of high-impact weather. *Climate Dynamics*, 43**,** 1847-1856.

CHEN, F., MANNING, K. W., LEMONE, M. A., TRIER, S. B., ALFIERI, J. G., ROBERTS, R., TEWARI, M., NIYOGI, D., HORST, T. W., ONCLEY, S. P., BASARA, J. B. & BLANKEN, P. D. 2007. Description and evaluation of the characteristics of the NCAR high-resolution land data assimilation system. *Journal of Applied Meteorology and Climatology*, 46**,** 694-713.

COSGROVE, B. A., LOHMANN, D., MITCHELL, K. E., HOUSER, P. R., WOOD, E. F., SCHAAKE, J. C., ROBOCK, A., SHEFFIELD, J., DUAN, Q. Y., LUO, L. F., HIGGINS, R. W., PINKER, R. T. & TARPLEY, J. D. 2003. Land surface model spin-up behavior in the North American

Land Data Assimilation System (NLDAS). *Journal of Geophysical Research-Atmospheres,* 108.

DAI, A. 2006. Precipitation characteristics in eighteen coupled climate models. *Journal of Climate,* 19**,** 4605-4630.

DI VIRGILIO, G., EVANS, J. P., DI LUCA, A., OLSON, R., ARGUESO, D., KALA, J., ANDRYS, J., HOFFMANN, P., KATZFEY, J. J. & ROCKELL, B. 2019. Evaluating reanalysis-driven CORDEX regional climate models over Australia: model performance and errors. *Climate Dynamics,* 53**,** 2985-3005.

DUDHIA, J. 1989. Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model. *Journal of the Atmospheric Sciences,* 46**,** 3077-3107.

EVANS, J. P., DI VIRGILIO, G., HIRSCH, A. L., HOFFMANN, P., REMEDIO, A. R., JI, F., ROCKEL, B. & COPPOLA, E. 2021. The CORDEX-Australasia ensemble: evaluation and future projections. *Climate Dynamics,* 57**,** 1385-1401.

EVANS, J. P., EKSTROM, M. & JI, F. 2012. Evaluating the performance of a WRF physics ensemble over South-East Australia. *Climate Dynamics,* 39**,** 1241-1258.

EVANS, J. P. & MCCABE, M. F. 2010. Regional climate simulation over Australia's Murray-Darling basin: A multitemporal assessment. *Journal of Geophysical Research-Atmospheres,* 115.

EVANS, J. P. & WESTRA, S. 2012. Investigating the Mechanisms of Diurnal Rainfall Variability Using a Regional Climate Model. *Journal of Climate,* 25**,** 7232-7247.

GUERREIRO, S. B., FOWLER, H. J., BARBERO, R., WESTRA, S., LENDERINK, G., BLENKINSOP, S., LEWIS, E. & LI, X. F. 2018. Detection of continental-scale intensification of hourly rainfall extremes. *Nature Climate Change,* 8**,** 803-+.

GUTOWSKI, W. J., DECKER, S. G., DONAVON, R. A., PAN, Z. T., ARRITT, R. W. & TAKLE, E. S. 2003. Temporal-spatial scales of observed and simulated precipitation in central U.S. climate. *Journal of Climate,* 16**,** 3841-3847.

HERSBACH, H., BELL, B., BERRISFORD, P., HIRAHARA, S., HORANYI, A., MUNOZ-SABATER, J., NICOLAS, J., PEUBEY, C., RADU, R., SCHEPERS, D., SIMMONS, A., SOCI, C., ABDALLA, S., ABELLAN, X., BALSAMO, G., BECHTOLD, P., BIAVATI, G., BIDLOT, J., BONAVITA, M., DE CHIARA, G., DAHLGREN, P., DEE, D., DIAMANTAKIS, M., DRAGANI, R., FLEMMING, J., FORBES, R., FUENTES, M., GEER, A., HAIMBERGER, L., HEALY, S., HOGAN, R. J., HOLM, E., JANISKOVA, M., KEELEY, S., LALOYAUX, P., LOPEZ, P., LUPU, C., RADNOTI, G., DE ROSNAY, P., ROZUM, I., VAMBORG, F., VILLAUME, S. & THEPAUT, J. N. 2020. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society,* 146**,** 1999-2049.

JANJIC, Z. I. 1994. The Step-Mountain Eta Coordinate Model - Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure Schemes. *Monthly Weather Review,* 122**,** 927-945.

JI, F., NISHANT, N., EVANS, J. P., DI VIRGILIO, G., CHEUNG, K. K. W., TAM, E., BEYER, K. & RILEY, M. L. 2022. Introducing NARCliM1.5: Evaluation and projection of climate extremes for southeast Australia. *Weather and Climate Extremes,* 38.

KIM, Y., EVANS, J. P. & SHARMA, A. 2023a. Correcting Systematic Biases in Regional Climate Model Boundary Variables for Improved Simulation of High-Impact Compound Events. *Available at SSRN 4366152*.

KIM, Y., EVANS, J. P. & SHARMA, A. 2023b. Multivariate bias correction of regional climate model boundary conditions. *Climate Dynamics*.

KIM, Y., EVANS, J. P., SHARMA, A. & ROCHETA, E. 2021. Spatial, Temporal, and Multivariate Bias in Regional Climate Model Simulations. *Geophysical Research Letters,* 48.

KIM, Y., ROCHETA, E., EVANS, J. P. & SHARMA, A. 2020. Impact of bias correction of regional climate model boundary conditions on the simulation of precipitation extremes. *Climate Dynamics,* 55**,** 3507-3526.

LEE, M. I., SCHUBERT, S. D., SUAREZ, M. J., HELD, I. M., KUMAR, A., BELL, T. L., SCHEMM, J. K. E., LAU, N. C., PLOSHAY, J. J., KIM, H. K. & YOO, S. H. 2007. Sensitivity to horizontal resolution in the AGCM simulations of warm season diurnal cycle of precipitation over the United States and northern Mexico. *Journal of Climate,* 20**,** 1862-1881.

LIM, K. S. S. & HONG, S. Y. 2010. Development of an Effective Double-Moment Cloud Microphysics Scheme with Prognostic Cloud Condensation Nuclei (CCN) for Weather and Climate Models. *Monthly Weather Review,* 138**,** 1587-1612.

MARAUN, D., WETTERHALL, F., IRESON, A. M., CHANDLER, R. E., KENDON, E. J., WIDMANN, M., BRIENEN, S., RUST, H. W., SAUTER, T., THEMESSL, M., VENEMA, V. K. C., CHUN, K. P., GOODESS, C. M., JONES, R. G., ONOF, C., VRAC, M. & THIELE-EICH, I. 2010. Precipitation Downscaling under Climate Change: Recent Developments to Bridge the Gap between Dynamical Models and the End User. *Reviews of Geophysics,* 48.

MATALAS, N. C. 1967. Mathematical assessment of synthetic hydrology. *Water Resources Research,* 3**,** 937-945.

MEHROTRA, R. & SHARMA, A. 2015. Correcting for systematic biases in multiple raw GCM variables across a range of timescales. *Journal of Hydrology,* 520**,** 214-223.

MINOBE, S., PARK, J. H. & VIRTS, K. S. 2020. Diurnal Cycles of Precipitation and Lightning in the Tropics Observed by TRMM3G68, GSMaP, LIS, and WWLLN. *Journal of Climate,* 33**,** 4293-4313.

MLAWER, E. J., TAUBMAN, S. J., BROWN, P. D., IACONO, M. J. & CLOUGH, S. A. 1997. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research-Atmospheres,* 102**,** 16663-16682.

PREIN, A. F., HOLLAND, G. J., RASMUSSEN, R. M., DONE, J., IKEDA, K., CLARK, M. P. & LIU, C. H. H. 2013. Importance of Regional Climate Model Grid Spacing for the Simulation of Heavy Precipitation in the Colorado Headwaters. *Journal of Climate,* 26**,** 4848-4857.

ROCHETA, E., EVANS, J. P. & SHARMA, A. 2014. Assessing atmospheric bias correction for dynamical consistency using potential vorticity. *Environmental Research Letters,* 9.

ROCHETA, E., EVANS, J. P. & SHARMA, A. 2017. Can Bias Correction of Regional Climate Model Lateral Boundary Conditions Improve Low-Frequency Rainfall Variability? *Journal of Climate,* 30**,** 9785-9806.

ROCHETA, E., EVANS, J. P. & SHARMA, A. 2020. Correcting lateral boundary biases in regional climate modelling: the effect of the relaxation zone. *Climate Dynamics,* 55**,** 2511-2521.

ROSA, D. & COLLINS, W. D. 2013. A case study of subdaily simulated and observed continental convective precipitation: CMIP5 and multiscale global climate models comparison. *Geophysical Research Letters,* 40**,** 5999-6003.

SALAS, J. D. 1980. *Applied modeling of hydrologic time series,* Water Resources Publication.

SKAMAROCK, W. C., KLEMP, J. B., DUDHIA, J., GILL, D. O., LIU, Z., BERNER, J., WANG, W., POWERS, J. G., DUDA, M. G. & BARKER, D. M. 2019. A description of the advanced research WRF model version 4. *National Center for Atmospheric Research: Boulder, CO, USA,* 145**,** 550.

SRIKANTHAN, R. & PEGRAM, G. G. S. 2009. A nested multisite daily rainfall stochastic generation model. *Journal of Hydrology,* 371**,** 142-153.

STEPHENS, G. L., L'ECUYER, T., FORBES, R., GETTELMAN, A., GOLAZ, J. C., BODAS-SALCEDO, A., SUZUKI, K., GABRIEL, P. & HAYNES, J. 2010. Dreary state of precipitation in global models. *Journal of Geophysical Research-Atmospheres,* 115.

TEWARI, M., CHEN, F., WANG, W., DUDHIA, J., LEMONE, M., MITCHELL, K., EK, M., GAYNO, G., WEGIEL, J. & CUENCA, R. Implementation and verification of the unified NOAH land surface model in the WRF model. 20th conference on weather analysis and forecasting/16th conference on numerical weather prediction, 2004. 2165-2170.

WANG, B., KIM, H. J., KIKUCHI, K. & KITOH, A. 2011. Diagnostic metrics for evaluation of annual and diurnal cycles. *Climate Dynamics,* 37**,** 941-955.

WANG, Y. Q., ZHOU, L. & HAMILTON, K. 2007. Effect of convective entrainment/detrainment on the simulation of the tropical precipitation diurnal cycle. *Monthly Weather Review,* 135**,** 567-585.

WATTERS, D., BATTAGLIA, A. & ALLAN, R. P. 2021. The Diurnal Cycle of Precipitation according to Multiple Decades of Global Satellite Observations, Three CMIP6 Models, and the ECMWF Reanalysis. *Journal of Climate,* 34**,** 5063-5080.

563 WESTRA, S., FOWLER, H. J., EVANS, J. P., ALEXANDER, L. V., BERG, P., JOHNSON, F.,
564 KENDON, E. J., LENDERINK, G. & ROBERTS, N. M. 2014. Future changes to the intensity
565 and frequency of short-duration extreme rainfall. *Reviews of Geophysics*, 52**,** 522-555.
566 XU, Z. & YANG, Z.-L. 2012. An Improved Dynamical Downscaling Method with GCM Bias
567 Corrections and Its Validation with 30 Years of Climate Simulations. *Journal of Climate*, 25**,**
568 6271-6286.
569 YAMADA, T. J., LEE, M. I., KANAMITSU, M. & KANAMARU, H. 2012. Diurnal Characteristics
570 of Rainfall over the Contiguous United States and Northern Mexico in the Dynamically
571 Downscaled Reanalysis Dataset (US10). *Journal of Hydrometeorology*, 13**,** 1142-1148.
572 ZIEHN, T., CHAMBERLAIN, M. A., LAW, R. M., LENTON, A., BODMAN, R. W., DIX, M.,
573 STEVENS, L., WANG, Y. P. & SRBINOVSKY, J. 2020. The Australian Earth System Model:
574 ACCESS-ESM1.5. *Journal of Southern Hemisphere Earth Systems Science*, 70**,** 193-214.

575