Author/s:
Shao, Y;Bishop, CH

Title:
Improving CMIP6 projections of daily precipitation using a mean-adjusted time variability correction technique

Date:
2025-06-01

Citation:
Shao, Y. & Bishop, C. H. (2025). Improving CMIP6 projections of daily precipitation using a mean-adjusted time variability correction technique. Climate Dynamics, 63 (6), pp.260-. https://doi.org/10.1007/s00382-025-07731-7.

Persistent Link:
https://hdl.handle.net/11343/362106

Climate Dynamics (2025) 63:260
https://doi.org/10.1007/s00382-025-07731-7

Improving CMIP6 projections of daily precipitation using a mean-adjusted time variability correction technique

Yawen Shao[1,2] · Craig H. Bishop[1,2]

Received: 26 December 2024 / Accepted: 12 May 2025 / Published online: 14 June 2025

Abstract
Daily precipitation time series exhibit intermittent periods of high variability separated by periods of no rain, posing challenges to correct projected precipitation. To improve projected changes in probabilities of flooding and drought, it is critically important to improve temporal correlations of the precipitation time series. Previous work introduced a Time Variability Correction (TVC) method, which quantified and corrected time variability errors at differing time scales. This study extends TVC to post-process daily precipitation projections from 28 CMIP6 models over Australia, introducing a new mean adjustment procedure to eliminate negative precipitation values while ensuring that both the mean and variability of the final series aligns with the observations in the historical training period. The new TVC mean-adjusted (TVC-ma) method preserves each model's projected change in timescale covariances, and our analysis reveals interesting differences among CMIP6 projections of changes in time-scale-dependent variances. TVC-ma is evaluated using a leave-one-out model-as-truth setup. Results reveal that, in most cases, TVC-ma significantly improves the mean, variance, lag correlations, and projections of climate indices related to persistent, heavy, and low rainfall extremes compared to raw models. When applied to future precipitation projections for Australia, TVC-ma projects pronounced increases in prolonged dry periods and maximum 1-day and 5-day precipitation amounts under the high-emission scenario relative to the low-emission scenario. Compared to the historical period, corrected projections under the high-emission scenario show drier conditions in parts of Western Australia, greater variability, extended durations of consecutive dry days and increased multi-day precipitation extremes across most regions of the continent.

Keywords Statistical bias correction · Daily precipitation extremes · Climate change and variability · Global climate modelling

1 Introduction

Precipitation amounts vary significantly across the globe and throughout the seasons. Accurate and reliable projections of daily precipitation from climate models are crucial for downstream applications, such as driving hydrological models (Peter et al. 2023), and conducting subsequent assessments of climate change impacts (Michalek et al. 2024; Nguyen et al. 2020). However, raw precipitation projections, especially those from global climate models (GCMs), often exhibit substantial systematic biases, necessitating post-processing before they are released to users.

Compared to quasi-normally distributed variables like temperature, precipitation presents several distinct post-processing challenges: (1) Precipitation time series has an intermittent nature with a natural lower bound of zero that requires a special treatment of zero values during post-processing; (2) Precipitation is highly variable and has a skewed distribution, with its variation often associated with large uncertainties (Hartmann et al. 2013); and (3) Precipitation has lower temporal autocorrelation than temperature, yet accurately simulating persistence attributes is crucial for predicting extreme events such as floods or droughts.

Various statistical techniques have been developed and employed by research centers worldwide to correct daily precipitation projections. These techniques range from the basic delta change approach (Hay et al. 2000; Lenderink et al.

✉ Yawen Shao
yawen.shao@unimelb.edu.au

[1] ARC Centre of Excellence for Climate Extremes, The University of Melbourne, Parkville, VIC, Australia

[2] School of Geography, Earth and Atmospheric Sciences, The University of Melbourne, Parkville, VIC, Australia

2007) to complex multivariate bias correction approaches (Mehrotra and Sharma 2016, 2019; Vrac et al. 2022). The simplest delta change method involves multiplying projections of future precipitation time series by the ratio between the climatological mean of the raw historical model series divided by the mean of the observed series. Nevertheless, this approach does not, in general, correct model variance. A frequently explored bias adjustment category is distribution-based methods (Cannon et al. 2015; Chen et al. 2013; Vrac et al. 2016), represented by Quantile Mapping (QM), where the model series is mapped to the observed series by matching their cumulative distribution functions (CDF). Numerous extended univariate methods have been developed from or are relevant to QM for post-processing daily precipitation projections, including detrended QM (Cannon et al. 2015), quantile delta mapping (Kim et al. 2021; Tong et al. 2021; Zhao et al. 2024), nonparametric and parametric quantile mapping (Ghimire et al. 2019), Cumulative Distribution Function-transform (CDF-t) (Michelangeli et al. 2009; Vrac et al. 2016), semi-parametric quantile mapping (Rajulapati and Papalexiou 2023) and so on. Recent advances in multivariate quantile mapping bias correction methods have been introduced to account for inter-variable dependencies and spatiotemporal sequencing in climate model outputs (Cannon 2016, 2018). A limitation of these methods is their inability to correct time-scale-dependent variability errors and subsequent autocorrelation of the time series, which is essential for predicting climate extremes.

More complicated approaches aim to correct both the distribution and persistence attributes of precipitation series across different time scales, such as Multivariate Recursive Nesting Bias Correction (MRNBC) and Multivariate Recursive Quantile-matching Nested Bias Correction (MRQNBC) (Mehrotra et al. 2018; Mehrotra & Sharma 2016), or to correct the time series in the frequency domain (Nguyen et al. 2019). When post-processing precipitation simulations, these methods utilise the full time series without specially accounting for zero precipitation records, which may lead to suboptimal results if the time series contains many zeros.

In a previous study, we developed a novel time variability correction (TVC) method designed to diagnose time-scale-dependent variance errors and correct the covariance across different time scales (Shao et al. 2024). The study showed that TVC improved the temporal variability and persistence attributes of CMIP6 (The Coupled Model Intercomparison Project Phase 6) maximum temperature projections on a global scale. Here we provide an adjustment to this technique that makes it better suited for post-processing daily precipitation projections.

In brief, we first apply the original TVC to time series of daily precipitation. While this improves some aspects of the time series, it also produces a few spurious negative values, as TVC treats the data as continuous. Simply converting these negative values to zeros would artificially inflate the mean, leading to an overestimation of the time-averaged precipitation. In this work, we introduce a novel approach to removing these negatives that involves applying a negative offset to the TVC post-processed time series that compensates for the increase in the mean that occurs when spurious negative values are set to zero. Our approach aligns conceptually with constrained optimization frameworks. Specifically, our heuristic implementation is analogous to convex optimization, aiming to minimize deviations from the target mean while ensuring all adjusted values remain non-negative. The methodologies and convergence assurances detailed in Boyd and Vandenberghe (2004) provide theoretical support for such iterative procedures. Furthermore, Hess et al. (2023) employed physically constrained generative adversarial networks (GANs) to correct biases in CMIP6 models, explicitly enforcing non-negative precipitation values by penalizing any negative values produced by the generator. These studies collectively demonstrate the effectiveness of iterative correction schemes that function as constrained optimization algorithms within stochastic precipitation modeling frameworks.

Our new mean-adjusted TVC model, hereafter named TVC-ma, is evaluated on daily precipitation projections extracted from 28 CMIP6 models (Eyring et al. 2016) across the Australian continent using a model-as-truth setup (Abramowitz et al. 2019; Vrac et al. 2022).

2 Study data

This study utilises gridded daily precipitation amounts from 28 CMIP6 models. These models are selected because each provides data for historical runs as well as for low-emission SSP1-2.6 (ssp126) and high-emission SSP5-8.5 (ssp585) emission scenarios. Due to computational constraints, only one variant from each model is evaluated, and intermediate scenarios (e.g., SSP2-4.5 and SSP3-7.0) were not examined in this study. Details on the selected models and variants are provided in Table 1. The method is applied and evaluated across the Australian continent. The evaluation covers the historical period from 1950 to 2014, and the projection

Table 1 Summary of the CMIP6 model variants used in this study

No.	Institution	Model	Variant
1	CSIRO-ARCCSS	ACCESS-CM2	r1i1p1f1
2	CSIRO	ACCESS-ESM1-5	r1i1p1f1
3	AS-RCEC	TaiESM1	r1i1p1f1
4	BCC	BCC-CSM2-MR	r1i1p1f1
5	CAMS	CAMS-CSM1-0	r2i1p1f1
6	CAS	FGOALS-g3	r1i1p1f1
7	CCCma	CanESM5	r1i1p1f1
8	CMCC	CMCC-CM2-SR5	r1i1p1f1
9	CMCC	CMCC-ESM2	r1i1p1f1
10	NCAR	CESM2-WACCM	r1i1p1f1
11	CNRM-CERFACS	CNRM-CM6-1	r1i1p1f2
12	CNRM-CERFACS	CNRM-CM6-1-HR	r1i1p1f2
13	CNRM-CERFACS	CNRM-ESM2-1	r1i1p1f2
14	EC-Earth-Consortium	EC-Earth3	r1i1p1f1
15	EC-Earth-Consortium	EC-Earth3-Veg	r1i1p1f1
16	EC-Earth-Consortium	EC-Earth3-Veg-LR	r1i1p1f1
17	INM	INM-CM4-8	r1i1p1f1
18	INM	INM-CM5-0	r1i1p1f1
19	IPSL	IPSL-CM6A-LR	r1i1p1f1
20	KIOST	KIOST-ESM	r1i1p1f1
21	MIROC	MIROC6	r1i1p1f1
22	MIROC	MIROC-ES2L	r1i1p1f2
23	MPI-M	MPI-ESM1-2-LR	r1i1p1f1
24	MRI	MRI-ESM2-0	r1i1p1f1
25	NCC	NorESM2-LM	r1i1p1f1
26	NCC	NorESM2-MM	r1i1p1f1
27	NOAA-GFDL	GFDL-ESM4	r1i1p1f1
28	NUIST	NESM3	r1i1p1f1

period from 2015 to 2099. The observational dataset used is the Australian Gridded Climate Data (AGCD) product, which has a spatial resolution of approximately 5 km × 5 km (Australian Bureau of Meteorology 2020). Although the accuracy of the AGCD gridded product is not uniform in space or time, it represents the current 'best-available' estimate of the recent history of precipitation over Australia. Since the number of in-situ observations increased significantly since 1950s, we select the period of 1950–2014 for training and evaluation.

For the sake of unified comparison and computational efficiency, all CMIP6 and AGCD datasets are re-gridded to a 1.5° × 1.5° spatial resolution using the conservative remapping method, resulting in a total of 312 grid cells for evaluation.

3 Methods

3.1 The mean adjusted TVC model

This section introduces a new mean-adjusted time variability correction (TVC-ma) model for refining the temporal variability in daily precipitation projections. We first provide a brief overview of the original TVC method, followed by a detailed explanation of the strategy used to adjust the mean of the corrected series.

3.1.1 Time variability correction

Suppose we have a daily observation series $\mathbf{y}^T = (y_1, y_2, ..., y_t, \dots y_n)$ where y_t is the rainfall accumulation on the *tth* day or time step of the time series. In TVC, we use a sequential process that involves taking backward looking time averages of differing lengths of remnant time series obtained by removing their average. The *tth* backward looking smoother of the time series at the *tth* time step is formulated as,

$$y_t^i = \sum_{j=0}^{P_i-1} w_j^i y_{t-j} = \left(\mathbf{w}_j^i\right)^T \mathbf{y}_{t-P_i+1:t} \tag{1}$$

where $P_i (i = 1, 2, \dots, k)$ is the number of days in the *tth* smoother, and w_j^i are the averaging weights, i.e., $w_j^i = \frac{1}{P_i}$, so that $\sum_{j=0}^{P_i-1} w_j^i = 1$. In this work, P_i are set as 365, 183, 92, 46, 23, 12, 6, 3, 2, representing $k = 9$ levels of smoothers ranging from annual, seasonal, monthly to synoptic time scales of evolution.

While a detailed explanation of the approach is given in Shao et al. (2024), the pseudo-code in the Appendix provides a step-by-step flow of how we harness different levels of backward-looking averaging smoothers to sequentially transform time series into different scales. Consequently, the original time series is transformed as,

$$\mathbf{y}_{n_9:n} = \tilde{\mathbf{y}}_{n_9:n}^9 + \sum_{k=1}^{9} \bar{\mathbf{y}}_{n_9:n}^k. \tag{2}$$

In other words, each original daily precipitation time series is given as the sum of 10 components. Note that since the smoothers are operated backward looking in time, some data from the first two years of the time series becomes unavailable after post-processing. However, no information is lost during this time-scale filtering process. Specifically, if

the original daily time series for a fixed geographical location contained in the vector $\mathbf{y}_{1:n}$ goes from day 1 to day n, the backward smoothing in our treatment leaves a reduced time series $\mathbf{y}_{n_9:n}$, which goes from day $n_9 = -8 + \sum_{i=1}^{9} P_i = 724$ to day n in the original daily time series of precipitation.

Equation (2) enables us to concisely express the original time series as,

$$\mathbf{y}_{n_9:n} = \left[\tilde{\mathbf{y}}^9_{n_9:n}, \bar{\mathbf{y}}^9_{n_9:n}, \bar{\mathbf{y}}^8_{n_9:n}, \cdots, \bar{\mathbf{y}}^2_{n_9:n}, \bar{\mathbf{y}}^1_{n_9:n}\right]\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} = \mathbf{Y}\mathbf{1}_{10} \quad (3)$$

where $\mathbf{1}_{10}$ is a column vector of 10 ones. The matrix of filtered observation series $\mathbf{Y}$ has $N_h = n + 1 - n_9 = n + 9 - \sum_{i=1}^{9} P_i$ rows, and 10 columns. At any time ($n_9 \le t \le n$), the true state y_t is precisely equal to the sum of ten corresponding values from filtered timescale series, i.e., $y_t = \mathbf{Y}_t\mathbf{1}_{10}$.

Now we have split the original time series into 10 time series, each associated with a different time scale, we can take measures of the observed variabilities in terms of time scales. For example, the variance of the elements $\bar{\mathbf{y}}^1$ given by

$$\sigma^2\left(\bar{\mathbf{y}}^1\right) = \frac{1}{N_h - 1}\left(\bar{\mathbf{y}}^1\right)^T\left(\mathbf{I} - \frac{\mathbf{1}_{N_h}\mathbf{1}^T_{N_h}}{N_h}\right)\left(\mathbf{I} - \frac{\mathbf{1}_{N_h}\mathbf{1}^T_{N_h}}{N_h}\right)^T\bar{\mathbf{y}}^1 \quad (4)$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{1}_{N_h}$ is a column vector with N_h elements, all of which are equal to 1. Thus, the quantity $\sigma^2\left(\bar{\mathbf{y}}^1\right)$ gives the variance of the time series at annual time scales whereas,

$$\sigma^2\left(\bar{\mathbf{y}}^9\right) = \frac{1}{N_h - 1}\left(\bar{\mathbf{y}}^9\right)^T\left(\mathbf{I} - \frac{\mathbf{1}_{N_h}\mathbf{1}^T_{N_h}}{N_h}\right)\left(\mathbf{I} - \frac{\mathbf{1}_{N_h}\mathbf{1}^T_{N_h}}{N_h}\right)\bar{\mathbf{y}}^9 \quad (5)$$

quantifies variance at the 2-day time scale after longer time scale variances have been removed. One would expect a slowly varying time series with large multi-day lag correlations to have large variance at long time scales and small variance at short time scales compared with a rapidly varying time series that had equal variance at all time scales.

More generally, we can define the covariance of the observed time series across all time scales using the time scale covariance matrix,

$$\mathbf{C}_\mathbf{Y} = \frac{1}{N_h - 1}\mathbf{Y}^\mathrm{T}\left(\boldsymbol{I} - \frac{\mathbf{1}_{N_h}\mathbf{1}^T_{N_h}}{N_h}\right)\left(\boldsymbol{I} - \frac{\mathbf{1}_{N_h}\mathbf{1}^T_{N_h}}{N_h}\right)\mathbf{Y} \quad (6)$$

Disregarding sampling errors, one would expect a time series produced by a perfect model would have an identical time series covariance matrix as the observed time series. As shown below, our TVC method provides a transformation that ensures that an imperfect climate model's corrected time series has the same time scale covariance matrix over the historical period as the observed time series.

A model time series $\mathbf{x}^T = \left[\mathbf{x}_h^T, \mathbf{x}_f^T\right]$ produced by an (imperfect) model climate simulation has both a historical $\mathbf{x}_h^T$ and future $\mathbf{x}_f^T$ period associated with it. Assume that $\mathbf{x}_h^T$ and $\mathbf{x}_f^T$ have N_h and N_f elements in them, respectively. We can apply the exact transform smoothers to the historical and future model time series as we applied to the observed time series and thus obtain,

$$\begin{aligned}\mathbf{x} = \begin{bmatrix}\mathbf{x}_h\\\mathbf{x}_f\end{bmatrix} &= \begin{bmatrix}\left(\tilde{\mathbf{x}}^9_h, \bar{\mathbf{x}}^9_h, \bar{\mathbf{x}}^8_h, \ldots, \bar{\mathbf{x}}^2_h, \bar{\mathbf{x}}^1_h\right)\\\left(\tilde{\mathbf{x}}^9_f, \bar{\mathbf{x}}^9_f, \bar{\mathbf{x}}^8_f, \ldots, \bar{\mathbf{x}}^2_f, \bar{\mathbf{x}}^1_f\right)\end{bmatrix}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}\\ &= \begin{bmatrix}\mathbf{X}_h\\\mathbf{X}_f\end{bmatrix}\mathbf{1}_{10} = \mathbf{X}\mathbf{1}_{10}\end{aligned} \quad (7)$$

In this work, we assume that the mean and the covariance of different time scales remain time-invariant, and no corrections are applied to the model's precipitation trend. The mean of the model time series is adjusted to match the observed mean during the historical period using,

$$\mathbf{X}^c(:, l) = \left[\mathbf{X}(:, l) + \mathbf{1}_{N_h+N_f}\left(\mu_l^\mathbf{Y} - \mu_l^{\mathbf{X}_h}\right)\right] \quad (8)$$

where $\mu_l^\mathbf{Y}$ and $\mu_l^{\mathbf{X}_h}$ give the climatological mean of the l^{th} column of $\mathbf{Y}$ and $\mathbf{X}_h$, respectively, $N_f = n_f + 1 - n_9 = n_f + 9 - \sum_{i=1}^{9} P_i$, and $\mathbf{1}_{N_h+N_f}$ is a column vector with $N_h + N_f$ elements, all of which are equal to 1.

To correct the covariance of $\mathbf{X}_h^c$, so that its elements covary in time in the same way as the elements of $\mathbf{Y}$, we compute the time scale covariance matrix of $\mathbf{X}_h$ using

$$\mathbf{C}_\mathbf{X} = \left(\frac{1}{N_h - 1}\right)\mathbf{X}_h^{c^T}\left(\mathbf{I} - \frac{\mathbf{1}_{N_h}\mathbf{1}^T_{N_h}}{N_h}\right)\left(\mathbf{I} - \frac{\mathbf{1}_{N_h}\mathbf{1}^T_{N_h}}{N_h}\right)\mathbf{X}_h^c. \quad (9)$$

The corrected timeseries matrix of differing time scale components is given by

$$\mathbf{Z}_h = \frac{\mathbf{1}_{N_h}\mathbf{1}^T_{N_h}}{N_h}\mathbf{X}_h^c + \left(\mathbf{I} - \frac{\mathbf{1}_{N_h}\mathbf{1}^T_{N_h}}{N_h}\right)\mathbf{X}_h^c\mathbf{C}_\mathbf{X}^{-1/2}\mathbf{C}_\mathbf{Y}^{1/2} \quad (10)$$

where $\mathbf{C}_\mathbf{X}^{1/2}$ and $\mathbf{C}_\mathbf{Y}^{1/2}$ are the symmetric square roots of $\mathbf{C}_\mathbf{X}$ and $\mathbf{C}_\mathbf{Y}$, respectively.

We can simply extend the correction of the historical time series to a correction of the future model time series using

$$\mathbf{Z}_f = \frac{\mathbf{1}_{N_f}\mathbf{1}_{N_f}^T}{N_f}\mathbf{X}_f^c + \left(\mathbf{I} - \frac{\mathbf{1}_{N_f}\mathbf{1}_{N_f}^T}{N_f}\right)\mathbf{X}_f^c\mathbf{C}_\mathbf{X}^{-1/2}\mathbf{C}_\mathbf{Y}^{1/2} \tag{11}$$

And the final TVC corrected historical and future time-series are simply given by,

$$\mathbf{z}_h = \mathbf{Z}_h\mathbf{1}_{10} \quad \text{and} \quad \mathbf{z}_f = \mathbf{Z}_f\mathbf{1}_{10} \tag{12}$$

This TVC corrected model time series has an inter-timescale covariance matrix that is identical to that of the observed time series $\mathbf{C}_\mathbf{Y}$ in the historical period. Its covariance matrix in the future period may differ from the historical period if the climate model's future time scale covariance matrix differs from that in the historical period. To be precise, let $\mathbf{C}_\mathbf{X}$ and $\mathbf{C}_\mathbf{X}^f$ denote the time scale covariance matrices of the raw model time series in the historical and some future period, respectively. The climate change ratio matrix of time scale covariances $\Lambda_\mathbf{X} = \mathbf{C}_\mathbf{X}^{-1/2}\mathbf{C}_\mathbf{X}^f\mathbf{C}_\mathbf{X}^{-1/2}$ is a measure of the model predicted change in time scale covariances due to changes in greenhouse gas (GHG) concentrations that is present in the model time series $\boldsymbol{x}$. For the corrected time series,

$$\begin{aligned}\Lambda_\mathbf{Z} &= \mathbf{C}_\mathbf{Y}^{-1/2}\mathbf{C}_\mathbf{Z}^f\mathbf{C}_\mathbf{Y}^{-1/2} = \mathbf{C}_\mathbf{Y}^{-1/2}\frac{1}{N_f-1}\mathbf{Z}_f^T\left(\mathbf{I} - \frac{\mathbf{1}_{N_f}\mathbf{1}_{N_f}^T}{N_f}\right)\left(\mathbf{I} - \frac{\mathbf{1}_{N_f}\mathbf{1}_{N_f}^T}{N_f}\right)\mathbf{Z}_f\mathbf{C}_\mathbf{Y}^{-1/2} \\ &= \mathbf{C}_\mathbf{Y}^{-1/2}\mathbf{C}_\mathbf{Y}^{1/2}\mathbf{C}_\mathbf{X}^{-1/2}\left[\frac{1}{N_f-1}\left(\mathbf{X}_f^c\right)^T\left(\mathbf{I} - \frac{\mathbf{1}_{N_f}\mathbf{1}_{N_f}^T}{N_f}\right)\left(\mathbf{I} - \frac{\mathbf{1}_{N_f}\mathbf{1}_{N_f}^T}{N_f}\right)\mathbf{X}_f^c\right]\mathbf{C}_\mathbf{X}^{-1/2}\mathbf{C}_\mathbf{Y}^{1/2}\mathbf{C}_\mathbf{Y}^{-1/2} \\ &= \mathbf{C}_\mathbf{X}^{-1/2}\mathbf{C}_\mathbf{X}^f\mathbf{C}_\mathbf{X}^{-1/2} = \Lambda_\mathbf{X}\end{aligned} \tag{13}$$

Equation (13) proves that the TVC method preserves the raw model's predicted change in time scale covariances due to changes in GHG concentrations. As such, it retains the model's climate change signal. Hence, an underlying assumption of the TVC-ma approach is that the required multiplicative time scale covariance correction matrix for the historical period is identical to that of the future.

3.1.2 Mean-adjusting TVC

The following outlines detailed procedures for extending TVC to adjust the time mean while post-processing daily rainfall projections for both in-sample and out-of-sample periods.

(a) In-sample training periods

Step 1: Apply the TVC method to post-process the raw historical model series $\mathbf{x}_h$ towards the observed series $\mathbf{y}$ following Eqs. (2–10) and generate corrected time series $\mathbf{z}_h$. The parameters $\boldsymbol{\mu}^{\mathbf{X}_h}$,$\boldsymbol{\mu}^\mathbf{Y}$, $\mathbf{C}_\mathbf{X}$ and $\mathbf{C}_\mathbf{Y}$ are derived from this training period.

Step 2: Find a negative value of the scalar $\alpha(\alpha < 0)$ such that the non-negative time series $\mathbf{z}_h^*$ generated by,

$$\mathbf{z}_h^* = \begin{cases} \mathbf{z}_h + \alpha\mathbf{1} & when\ \ \mathbf{z}_h + \alpha\mathbf{1} \geq 0 \\ 0 & when\ \ \mathbf{z}_h + \alpha\mathbf{1} < 0 \end{cases} \tag{14}$$

satisfies $\left|\overline{\mathbf{z}}_h^* - \overline{\mathbf{y}}_{n_9:n}\right| < 10^{-4}$ mm/day, where $\overline{\mathbf{y}}_{n_9:n}$ and $\overline{\mathbf{z}}_h^*$ are the means of the time series $\mathbf{y}_{n_9:n}$ and $\mathbf{z}_h^*$, respectively, $\mathbf{1}^T = [1, 1, ..., 1]$ is used to appropriately vectorise the scalar α. Note that the step in Eq. (14) that sets negatives to zero increases the mean $\overline{\mathbf{z}}_h^*$ of $\mathbf{z}_h^*$ relative to the mean $\overline{\mathbf{z}}_h$ of $\mathbf{z}_h$ while the other step decreases $\overline{\mathbf{z}}_h^*$ relative to $\overline{\mathbf{z}}_h$. Hence, we seek the α value such that the mean decreasing effect of the operation $\mathbf{z}_h^* = \mathbf{z}_h + \alpha\mathbf{1}$ *when* $\mathbf{z}_h + \alpha\mathbf{1} \geq 0$ is exactly balanced by the mean increasing effect of the operation $\mathbf{z}_h^* = \mathbf{z}_h + \alpha\mathbf{1}$ *when* $\mathbf{z}_h + \alpha\mathbf{1} < 0$.

Here, we obtain α through a binary, iterative, half interval search type procedure that is outlined in Fig. 1. The initial guess for α was set equal to the difference between the mean $\overline{\mathbf{y}}_{n_9:n}$ of the time series $\mathbf{y}_{n_9:n}$ and the mean $\overline{\mathbf{z}}_h^0$ of the non-negative time series $\mathbf{z}_h^0 = \frac{1}{2}\left(\left|\mathbf{z}_h\right| + \mathbf{z}_h\right)$ (i.e.,$\alpha = \overline{\mathbf{y}}_{n_9:n} - \overline{\mathbf{z}}_h^0$).

Step 3: Output the final post-processed TVC-ma time series $\mathbf{z}_h^*$ as shown in the right triangle in Fig. 1.

In this work, we set the tolerance threshold as 10^{-4} mm/day, with convergence typically achieved in fewer than 20 iterations. The selection of this threshold balances computational efficiency and accuracy. A stricter threshold (e.g., 10^{-5}) increases runtime, whereas a more relaxed threshold (e.g., 10^{-3}) reduces the statistics accuracy of the post-processed time series. For the model-as-truth results in our work, 90% of the α values fall within the range of [-1.3, 0] mm/day—which is a barely rain-gauge measurable range. The largest magnitude α value obtained is -8.5 mm/day.

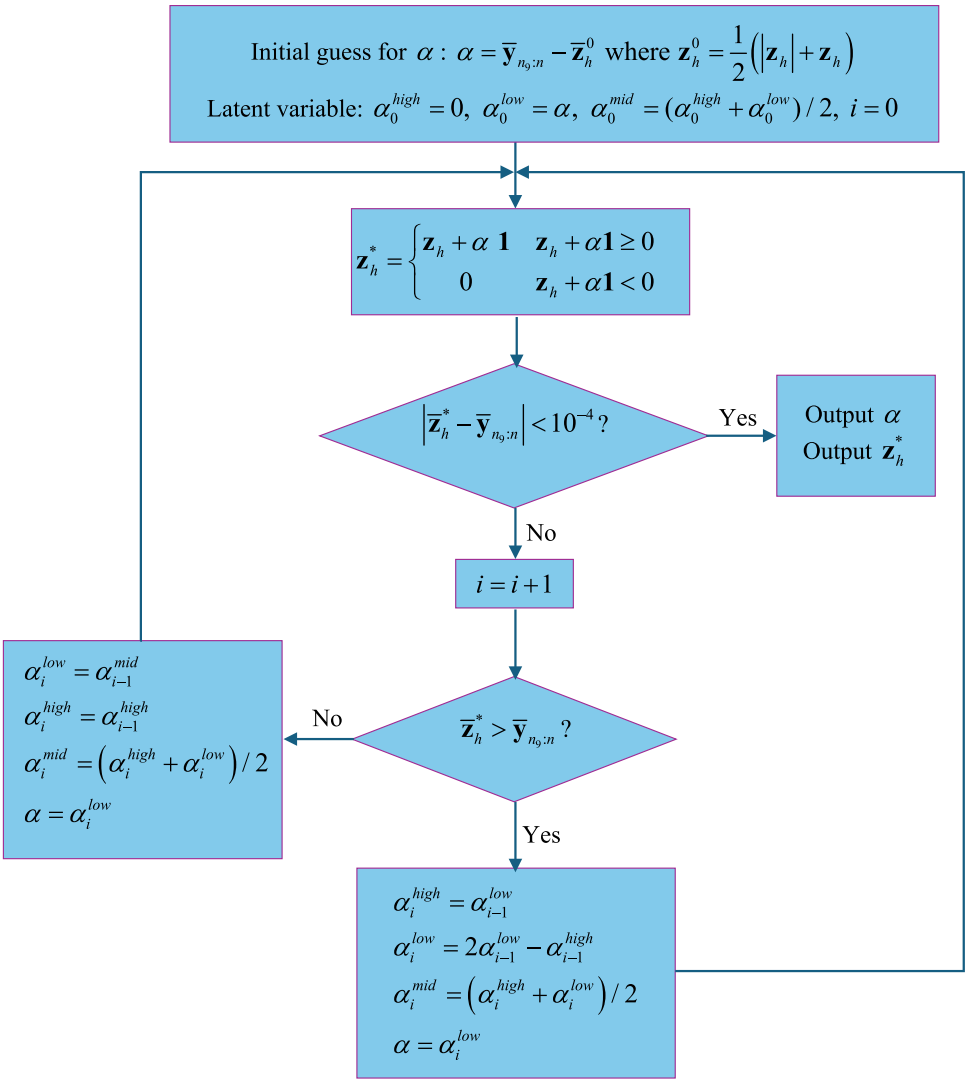


Fig. 1 Procedures for estimating the value the negative scalar α. Note that the overbars in $\bar{\mathbf{y}}_{n_9:n}$ and $\bar{\mathbf{z}}_h^0$ indicate the time series average. We also emphasize that the $\mathbf{z}_h$ vector is never updated in the above loop

(b) Out-of-sample validation periods

Step 1: Separate $\mathbf{x}_f$ into different time scales as in Eq. (7) and correct the means and covariance of $\mathbf{X}_f$ using Eqs. (8,11).

Step 2: Add α to TVC corrected time series $\mathbf{z}_f$ (see Eq. (12)), convert any negative values to zero, and obtain the final TVC-ma time series $\mathbf{z}_f^*$ during the out-of-sample period.

3.2 Model evaluation

3.2.1 Evaluation metrics

For individual time series, we calculate the mean, sample variance, single time-lag autocorrelation (lag-1 correlation), and 5-day lag autocorrelation of 5-day averages (lag-5 correlation) for the in-sample historical and out-of-sample projection periods respectively.

Sample variance measures the variability over the evaluation period, which is computed as,

$$\sigma^2 = \frac{\sum_{j=1}^{J} \left(x_j - \bar{x}_{1:J}\right)^2}{J-1} \tag{15}$$

where J is the length of the time series, and $\bar{x}_{1:J}$ is the average of the entire time series during the evaluation period.

Lag-1 correlation refers to the correlation between values in a time series that are one day apart, while lag-5 correlation is defined as the correlation between 5-day averages in a blocked time series that are 5 days apart. Note that the 5-day average blocked series is 1/5 of the length of the original series. The correlation is formulated as,

$$\rho = \frac{\sum_{j=1}^{J-1}\left(x_j - \bar{x}_{1:J-1}\right)\left(x_{j+1} - \bar{x}_{2:J}\right)}{\sqrt{\sum_{j=1}^{J-1}\left(x_j - \bar{x}_{1:J-1}\right)^2 \sum_{j=1}^{J-1}\left(x_{j+1} - \bar{x}_{2:J}\right)^2}} \tag{16}$$

where $\bar{x}_{1:J-1}$ and $\bar{x}_{2:J}$ are the time means of the series starting from the first- and second-time steps, respectively.

We also compute four precipitation-based climate indices defined by the Expert Team on Climate Change Detection and Indices (ETCCDI) (Zhang et al. 2011). These indices are evaluated because they capture different aspects of precipitation variability, from low to extreme rainfall. We expect that improvements in the lag correlations of the time series will lead to the enhanced predictions of these extreme indices.

The R10mm index is defined as the count of days when the daily precipitation amount is no less than 10 mm over a given period. The Consecutive Dry Days (CDD) index represents the maximum number of consecutive days within a certain period with daily precipitation below 1 mm. R10mm (CDD) is an indicator of heavy (low) rains in the time series. In this work, we calculate both indices on an annual basis, with the results averaged/pooled over the evaluation period.

Additionally, we compute the extreme rainfall indicators Rx1day and Rx5day, which represent, respectively, the maximum 1-day precipitation amount and maximum precipitation accumulated over five consecutive days within a certain period. Both indices are also determined annually, and then averaged over the evaluation period.

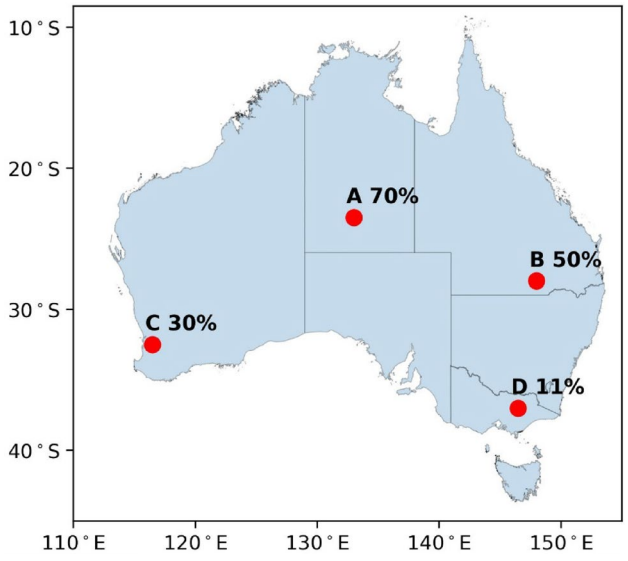


Fig. 2 Locations of four selected grid cells. The percentage of no-rain days over the historical period is annotated for each cell

3.2.2 Case study analysis

To understand the performance of the TVC-ma model, we select four grid cells (A-D in Fig. 2) for detailed evaluations. These cells are close to weather stations with historical records of daily precipitation from 1950 to 2014. Furthermore, these cells exhibit diverse percentages of no-rain days (i.e., zero rainfall amounts) over the 65-year period.

For each cell, we apply the hybrid model to post-process raw ACCESS-ESM1-5 daily precipitation simulations against the corresponding AGCD data over the in-sample period, 1950–2014. The statistical metrics and climate indices presented in Sect. 3.2.1 are computed for the observations, raw ACCESS-ESM1-5, and post-processed simulations for comparison.

3.2.3 Model-as-truth framework

We employ the model-as-truth setup (Abramowitz and Bishop 2015; Abramowitz et al. 2019) to evaluate and compare the performance of raw models and TVC-ma post-processed results. Specifically, one CMIP6 model is designated as the 'truth', while the remaining 27 CMIP6 models are considered as raw simulations with equal weighting. Each of these raw models is corrected against the 'truth' over the in-sample historical period, 1950–2014. Subsequently, the trained TVC-ma model is applied to post-process projections for the out-of-sample period, 2015–2099. Each CMIP6 model serves as the 'truth', resulting in 756 (28×27) raw and TVC-ma time series for each grid cell and scenario analysed.

For each model-as-truth case, we compute the percentage improvements in the averaged mean absolute errors (MAE) for all statistics outlined in Sect. 3.2.1, comparing the raw models and TVC-ma data. The percentage improvement is calculated as,

$$\mathrm{MAE} = \frac{1}{27}\sum_{m=1}^{27}\left|\mathrm{Model}_m - \mathrm{Obs}\right| \tag{17}$$

$$\%\ \text{improvement} = \frac{\mathrm{MAE}_{\mathrm{RAW}} - \mathrm{MAE}_{\mathrm{TVC\text{-}ma}}}{\mathrm{MAE}_{\mathrm{RAW}}} \times 100\% \tag{18}$$

3.2.4 Application to the future emission scenarios

We apply the TVC-ma model to post-process Australian climate change projections by training on the AGCD data from the historical period of 1950–2014. Subsequently, the trained model is utilised to post-process the emission scenarios, ssp126 and ssp585, spanning the period from 2015

to 2099. We compute and present results for the verification metrics.

4 Results

4.1 Testing case results

To demonstrate the effectiveness of the new TVC-ma method, we first apply it to correct four selected grid cells (A–D) extracted from the raw ACCESS-ESM1-5 model against AGCD observed data over the in-sample historical period of 1950–2014.

Table 2 displays metric results for observations, raw ACCESS-ESM1-5, and TVC-ma post-processed time series for the selected grid cells. Compared to AGCD observed data, the raw ACCESS-ESM1-5 model tends to overestimate precipitation and exhibit higher variability in relatively dry regions, while underestimating precipitation and variability in regions with more rain days. After post-processing, both precipitation mean, and variance are closer to observed values across all test cases.

Regarding single-day lag correlation, improvements are observed in three cells (A, C, D) after TVC-ma post-processing, except for cell B where the raw model performs better. The degradation for cell B is unsurprising given that the raw model's 1-day lag correlation was almost perfect in this case with just a 2% error. Interestingly, cells with more no-rain days tend to exhibit larger lag-1 correlation, likely due to more frequent consecutive zero occurrences. Lag-5 correlations of the TVC-ma post-processed time series are improved in cells A–C, while a very slight degradation is observed in cell D.

For climate indices, improvements are seen in simulating R10mm values across all cells. Averaged CDD values are better simulated in cells A–C, while results worsen in cell D. Similarly, both Rx1day and Rx5day scores are degraded in cell D, possibly due to lower mean, and underestimation of variance in the raw model compared to observations. TVC-ma increases the mean and variability of the model series, leading to an overestimation of precipitation extremes while significantly reducing the number of zero and low precipitation values in the post-processed results.

To obtain more statistically robust results to those of the above and to test whether improvements in the historical period can be translated to improvements in the future period, in the next section we test TVC-ma in a multitude of "leave-one-out-model-as-truth" experiments.

4.2 Model-as-truth results

Here, we further evaluate TVC-ma for rainfall projections within a model-as-truth framework.

Figure 3 presents box plots summarising the improvement (%) in MAE of the mean and variance for all grid cells over the Australian land region for each 'truth'. During the in-sample historical period, 100% improvements in precipitation means (Fig. 3a) are observed in all grid cells because TVC-ma precisely corrects the mean in the in-sample period. For variance (Fig. 3b), the percentage improvement is closer to 100% over the historical period, as TVC directly corrects the covariance of all the model series from differing time scales to match the observed covariance. The mean adjustment and conversion of negative values to zero after TVC post-processing slightly lowers the variance.

In the out-of-sample projection periods, larger improvements are observed in most model-as-truth cases for the ssp126 scenario compared to the ssp585 scenario. For both mean and variance, improvements are evident across more than half of the grid cells in all the model-as-truth cases. Specifically, for the ssp126 scenario, 75% of the 312 grid cells show improvements over raw predictions in 100% and 96.4% of the model-as-truth cases for mean and variance respectively, while for the ssp585 scenario, 75% of the grid

Table 2 Statistics of observed, raw ACCESS-ESM1-5, and TVC-ma post-processed results for four grid cells over the in-sample historical period

	Cell A			Cell B			Cell C			Cell D		
	Obs	Raw	TVC-ma	Obs	Raw	TVC-ma	Obs	Raw	TVC-ma	Obs	Raw	TVC-ma
Mean (mm/day)	0.81	1.32	0.81	1.29	2.04	1.29	1.89	1.00	1.89	2.97	1.87	2.97
Variance (mm^2/day^2)	13	27.9	12.5	18.8	35.7	18	22	9.05	21.6	37.6	22.7	37.6
Lag-1 correlation	0.537	0.507	0.514	0.428	0.429	0.414	0.341	0.290	0.322	0.306	0.302	0.308
Lag-5 correlation	0.232	0.190	0.195	0.172	0.210	0.185	0.321	0.181	0.274	0.109	0.124	0.125
R10mm (day/year)	7.79	12.7	6.92	13.7	21.2	13.3	21.8	6.78	14.3	33.8	15.4	22.9
CDD (day/year)	90.5	76.7	93.1	46.9	41.3	50.3	45.3	50.2	43.7	20.5	24	14.7
Rx1day (mm/year)	37.6	57.6	38.6	40.8	53.4	38.4	38.3	31.2	47.7	47.2	45.6	59.2
Rx5day (mm/year)	80.1	119.3	80.3	80.5	110.4	77.4	74.4	50.3	79.6	89	77.9	104

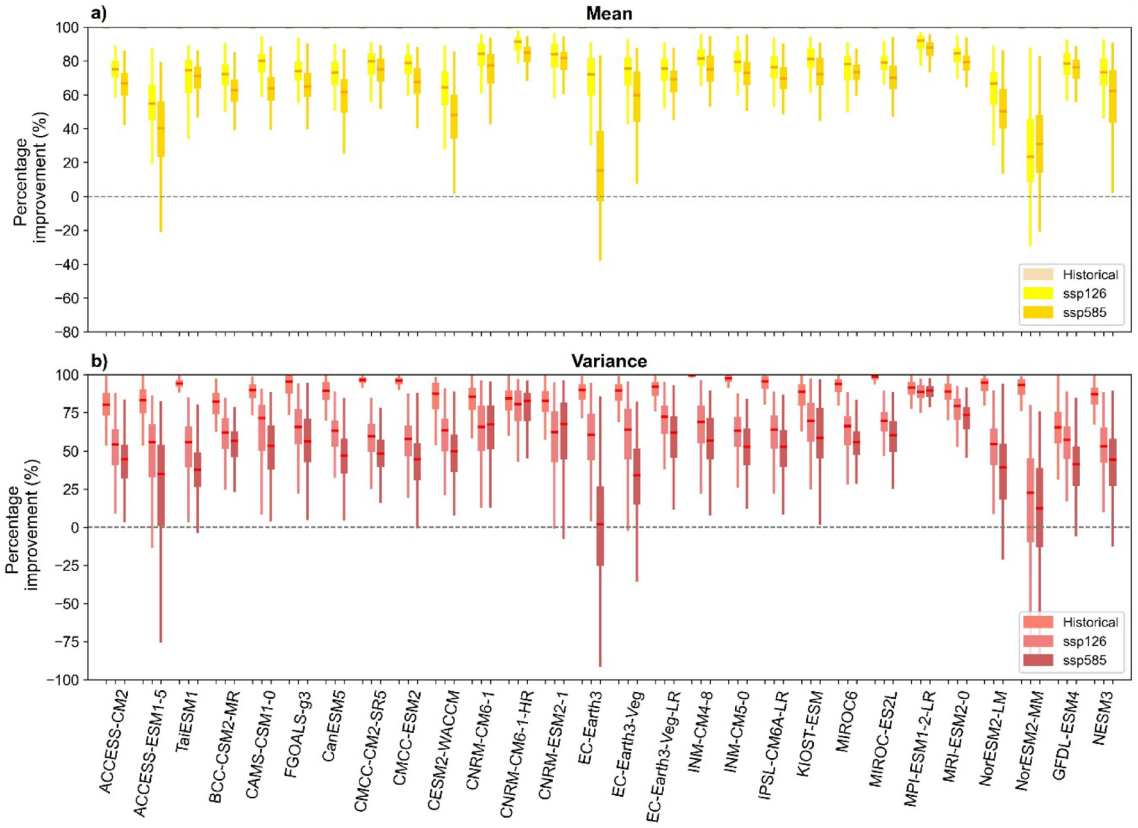


Fig. 3 Percentage improvement in MAE of precipitation a) mean and b) variance for each model-as-truth over the in-sample historical period of 1950–2014, and the out-of-sample projection period, ssp126 and ssp585, of 2015–2099. The middle horizontal line represents the median value. The inter-quartile range, shown by the middle box, spans from the 25th percentile (Q1) to the 75th percentile (Q3). The upper whisker extends to $Q3+1.5\times(Q3-Q1)$, and the lower whisker extends to $Q1-1.5\times(Q3-Q1)$. Outliers beyond the whiskers are not plotted

cells exhibit improvements in 96.4% and 92.9% of the truth cases for mean and variance respectively. However, degradation is also observed in many cells, particularly when EC-Earth3 and NorESM2-MM models are considered as 'truth'. Possible explanations for these less favourable results will be explored after all metric results are presented.

Figure 4 displays the percentage improvement in MAE for lag-1 and lag-5 correlation. Significant improvements are evident for both correlation metrics over the historical and projection periods, with very few cells performing worse results than the raw model series. Specifically, for both lag-1 and lag-5 correlations, more than 50% of the grid cells show at least 30% improvements compared to raw predictions across most model-as-truth cases, except for MPI-ESM1-2-LR. Overall, larger spreads of improvements are observed for lag-5 correlations.

The significant improvements in mean, variance, and lag correlations of the resulting series using the TVC-ma model also positively affect the predictions of R10mm, CDD, Rx1day and Rx5day indices, as shown in Figs. 5 and 6. For R10mm (Fig. 5a), improvements are generally observed in over half of the grid cells across most model-as-truth cases in all three scenarios, except for the EC-Earth3 model. Substantial improvements are evident in many model-as-truth cases, where 89.3% of the 'truth' cases show improvements in over 75% of the grid cells across all historical and projection periods compared to raw predictions. Regarding CDD (Fig. 5b), over half of the grid cells show improvements across all model-as-truth and all scenarios, with 27 out of 28 'truth' cases exhibiting improvements in over 75% of the grid cells, except for the BCC-CSM2-MR model. The improvement patterns are similar for both Rx1day (Fig. 6a) and Rx5day (Fig. 6b). For Rx1day, every model-as-truth and scenario show improvements in more than half of the grid cells. For Rx5day, 27 out of 28 'truth' models exhibit improvements in over half of the grid cells, with the exception of EC-Earth3 under the ssp585 scenario.

To explore how TVC-ma affects the inter-annual distribution of indices, we analyze the probability density functions (PDFs) of R10mm and CDD for each truth model and scenario. Specifically, we pool and bin the annual values from the truth, raw, and TVC-ma results, and plot their respective

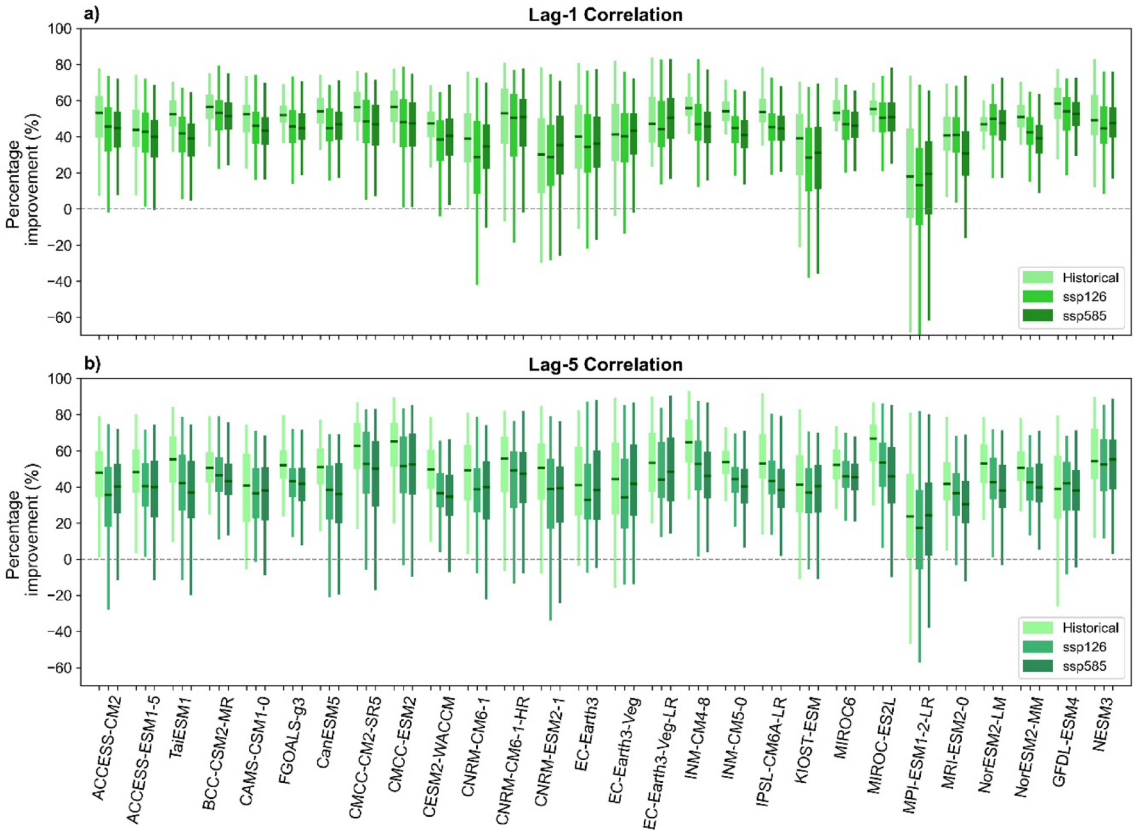


Fig. 4 The same as Fig. 3, but for **a** lag-1 correlation, **b** lag-5 correlation

PDFs based on histogram distributions. Figure 7 illustrates the PDFs for three selected truth models: CNRM-CM6-1, GFDL-ESM4, MIROC6, while Fig. S6 and S7 present the PDFs for all truth models.

As shown in Fig. 7, for R10mm in CNRM-CM6-1 and MIROC6, and for CDD in CNRM-CM6-1 and GFDL-ESM4, the TVC-ma PDFs (blue) align more closely with the observed PDFs (red) than the raw PDFs (black) during both in-sample and out-of-sample periods. However, when considering all truth models (Figs. S6 and S7), no clear pattern emerges regarding whether TVC-ma consistently improves the PDFs of R10mm or CDD.

To quantify these differences, we compute the total variance distance (TVD), which measures the distance between the probability mass functions of the observed indices and those from the raw or TVC-ma indices, formulated as,

$$\mathrm{TVD} = \frac{1}{2}\sum_{x} |O(x) - M(x)| \tag{19}$$

where $O(x)$ is the probability measure of the observed index, $M(x)$ is the probability measure of the raw or TVC-ma indices, and x is the possible value of R10mm or CDD. Table S1 summarizes the TVD values between the observed PDFs and the PDFs derived from raw and TVC-ma results. For R10mm, TVC-ma exhibits lower TVD values than the raw models in at least 75% of the truth models, while for CDD, TVC-ma outperforms the raw models in over 60% of the cases. These findings suggest that TVC-ma has the potential to enhance the long-term distribution of the R10mm and CDD indices.

Some specific models show degradation. For example, in many grid cells, the TVC-ma model does not enhance: (1) the mean, variance, R10mm and Rx5day values for the EC-Earth3 model, (2) the variance for the NorESM1-MM and ACCESS-ESM1-5 model, (3) the lag-1 and lag-5 correlations for the MPI-ESM1-2-LR model, (4) the CDD values for the BCC-CSM2-MR model when these models are treated as 'truth'.

The degradation observed in some grid cells can be better understood along with the TVC-only results. In the TVC-only approach, we apply TVC to generate in-sample $\mathbf{z}_h$ time series and out-of-sample $\mathbf{z}_f$ projections (see Eq. (12) and convert negative values to zero. Figs. S2-S5 illustrate the counterparts of Figs. 3, 4, 5, and 6, showing the percentage improvement across all statistics for the TVC-only results. By comparing these Figures with the TVC-ma results, one can better discern whether the issues are associated with

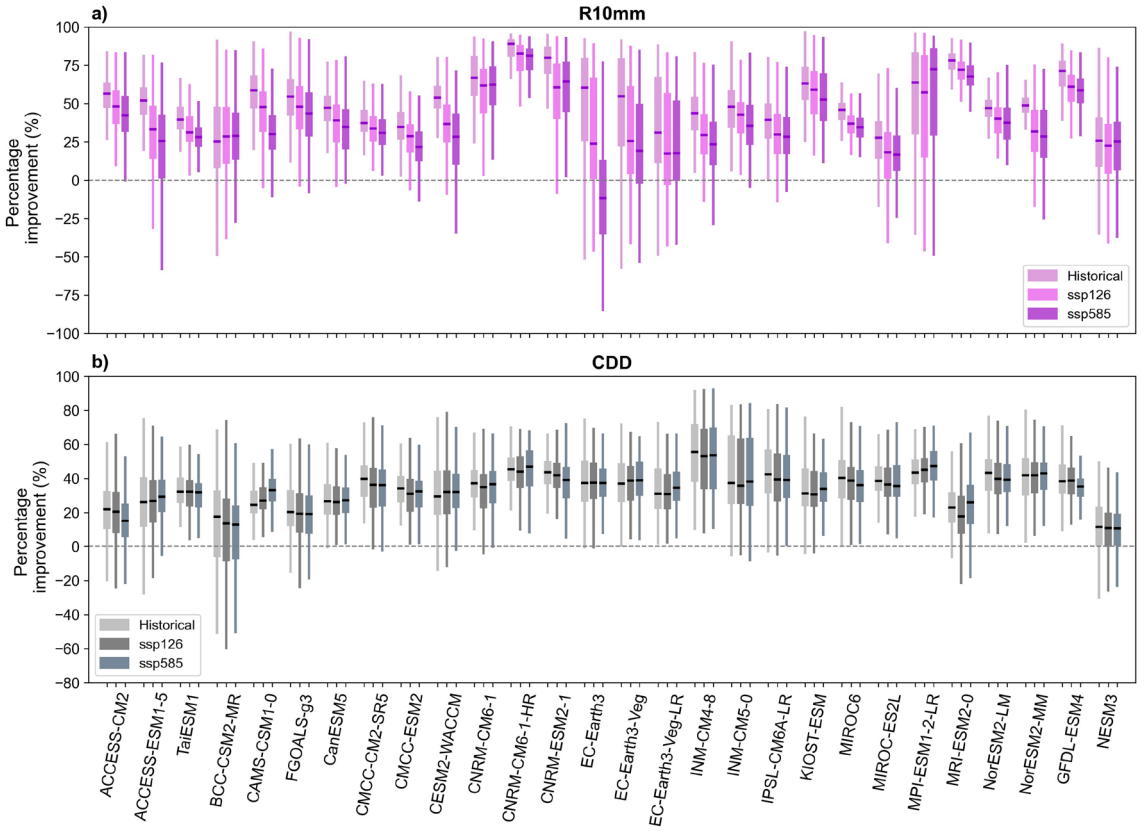


Fig. 5 The same as Fig. 3, but for **a** R10mm, **b** consecutive dry days (CDD)

TVC post-processing itself or with the subsequent mean adjustment.

Interestingly, there are no major differences between Fig. 3b and Fig. S2b for variance improvement, suggesting that the mean adjustment following TVC post-processing has minimal impact on variance. However, the mean adjustment slightly worsens the variance for EC-Earth compared to the TVC-only results during the projection period. This finding is also evident for mean improvement, where the mean adjustment leads to a smaller improvement in mean when EC-Earth is treated as 'truth' (Fig. S2a compared to Fig. 3a).

To explain variance degradation, because we assume that the time scale covariance error remains consistent between the past and future and we apply the same scaling factors to the projection period, poorer results are observed if the model's behaviour differs between the training and validation periods. Furthermore, if the climate change ratio matrix of time scale covariances $\Lambda_{\mathbf{X}} = \mathbf{C}_{\mathbf{X}}^{-1/2}\mathbf{C}_{\mathbf{X}}^{f}\mathbf{C}_{\mathbf{X}}^{-1/2}$ of the truth model is very different from all the models used in the ensemble then TVC is unlikely to improve the variance of the projection because, as shown in Eq. (13), TVC does not change this aspect of the model time series.

Figure S1 in the Supplementary Information plots the diagonal elements of $\Lambda_{\mathbf{X}}$, representing the model predicted change in variance at each time scale. The Figure shows that, on average, the EC-Earth model has an exceptionally large ratio across all time scales, while NorESM2-MM and ACCESS-ESM1-5 each have exceptionally small ratios. Since TVC preserves these ratios, it is unsurprising that when TVC is applied to model time series with substantially different $\Lambda_{\mathbf{X}}$ values to the chosen truth model's $\Lambda_{\mathbf{X}}$, the TVC imparted improvement is less reliable. This lack of improvement when EC-Earth, NorESM2-MM and ACCESS-ESM1-5 are used as the truth models is clearly evident in percentage improvement in variance for both TVC-only results (Fig. S2b) and TVC-ma results (Fig. 3b).

Regarding the lack of improvement in CDD values when BCC-CSM2-MR is treated as the 'truth' (Fig. 5b), compared to the TVC-only results (Fig. S4b), the mean adjustment results in an increase in the number of consecutive low rainfall days which degrades the CDD score for TVC-ma relative to TVC-only.

Comparison of Fig. 4 (TVC-ma) with Fig. S3 (TVC-only) reveals that the mean adjustment procedure significantly degraded the lag correlation results when MPI-ESM1-2-LR was taken as the 'truth'. Since the mean adjustment

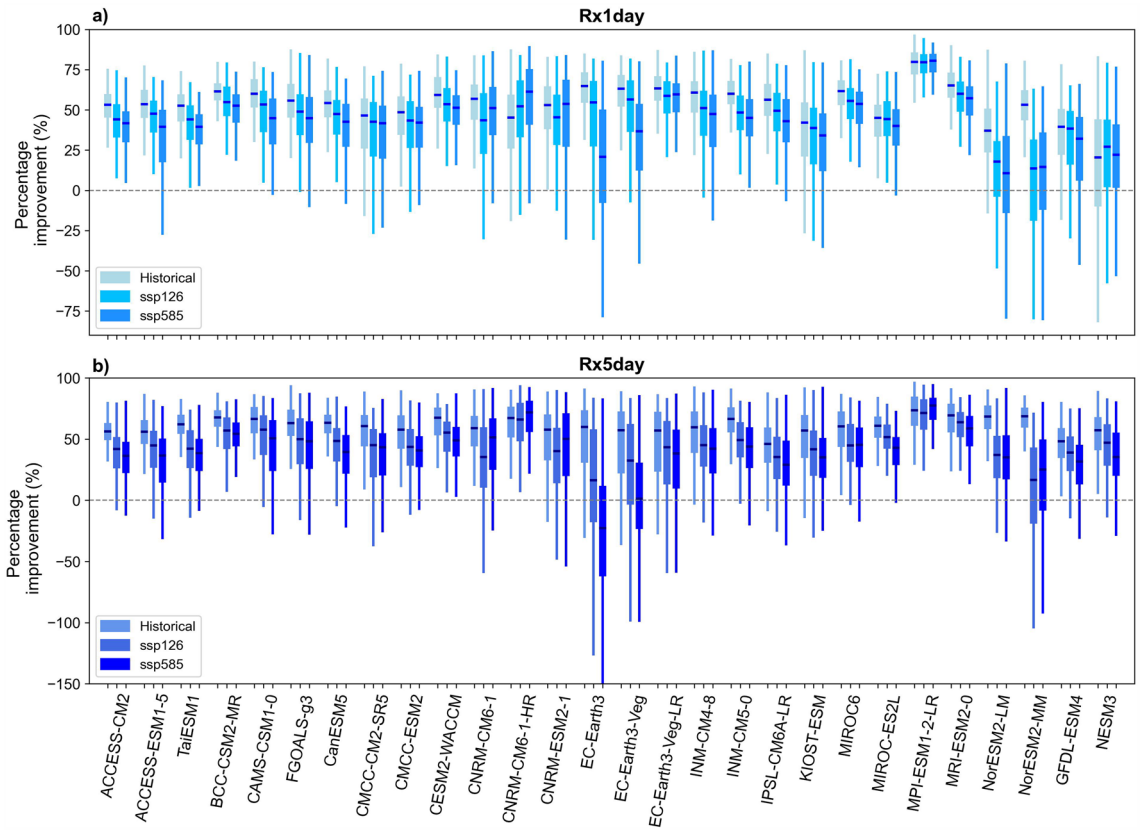


Fig. 6 The same as Fig. 3, but for **a** Rx1day, **b** Rx5day

procedure introduces a higher frequency of zero rainfall days, we can speculate that, for some reason, when MPI-ESM1-2-LR is taken as the truth, the mean adjustment is leading to a much larger number of zero rain days in the TVC-ma than in the TVC-only time series for this case. Increasing the number of zero-rain days would eventually degrade the time scale covariance correction of TVC-only and hence degrade the lag correlations of the time series.

Despite the discrepancies, Figs. 3, 4, 5 and 6 show that the overall improvement produced by TVC-ma is significant. Specifically, averaging across the results for all "model-as-truth" experiments, there is over 64, 48, 40, 39, 35, 32, 35, 37% improvement for mean, variance, lag-1 correlation, lag-5 correlation, R10mm, CDD, Rx1day and Rx5day scores, respectively across all scenarios.

4.3 Use of TVC-ma with real observations over Australia

In this section, the TVC-ma method is applied to post-process the projections of 28 CMIP6 models from the ssp126 and ssp585 emission scenarios across the Australian continent, trained on the AGCD data from the historical period. We evaluate the metrics for mean, variance, R10mm, CDD, Rx1day and Rx5day, averaging them across all models for AGCD, raw, and TVC-ma post-processed predictions during the in-sample historical, as well as raw and TVC-ma results for the out-of-sample projection periods. These multi-model mean (MMM) results are plotted in Figs. 8 and 9, respectively.

During the historical period, the AGCD and raw MMM results show distinct spatial patterns across all metrics. The AGCD data indicate mean daily precipitation exceeding 2 mm/day (Fig. 8a) and variances (Fig. 8d) over 30 mm^2/day^2 in northern Australia, along the eastern coastal regions, and in Tasmania. In contrast, central Australia, being relatively arid, exhibits near-zero mean and variance values. However, raw model means (Fig. 8b), and variances (Fig. 8e) tend to overestimate precipitation compared to the observed data across much of the continent.

For extreme indices, the spatial distribution of AGCD (Fig. 8g), raw (Fig. 8h), and post-processed R10mm value (Fig. 8i) closely aligns with their corresponding mean values. On average, northern Australia, eastern coastal regions, and Tasmania experience no fewer than 20 days annually with precipitation exceeding 10 mm. Regarding CDD, the raw models (Fig. 8k) often project a higher number of consecutive dry days in northern Australia, with

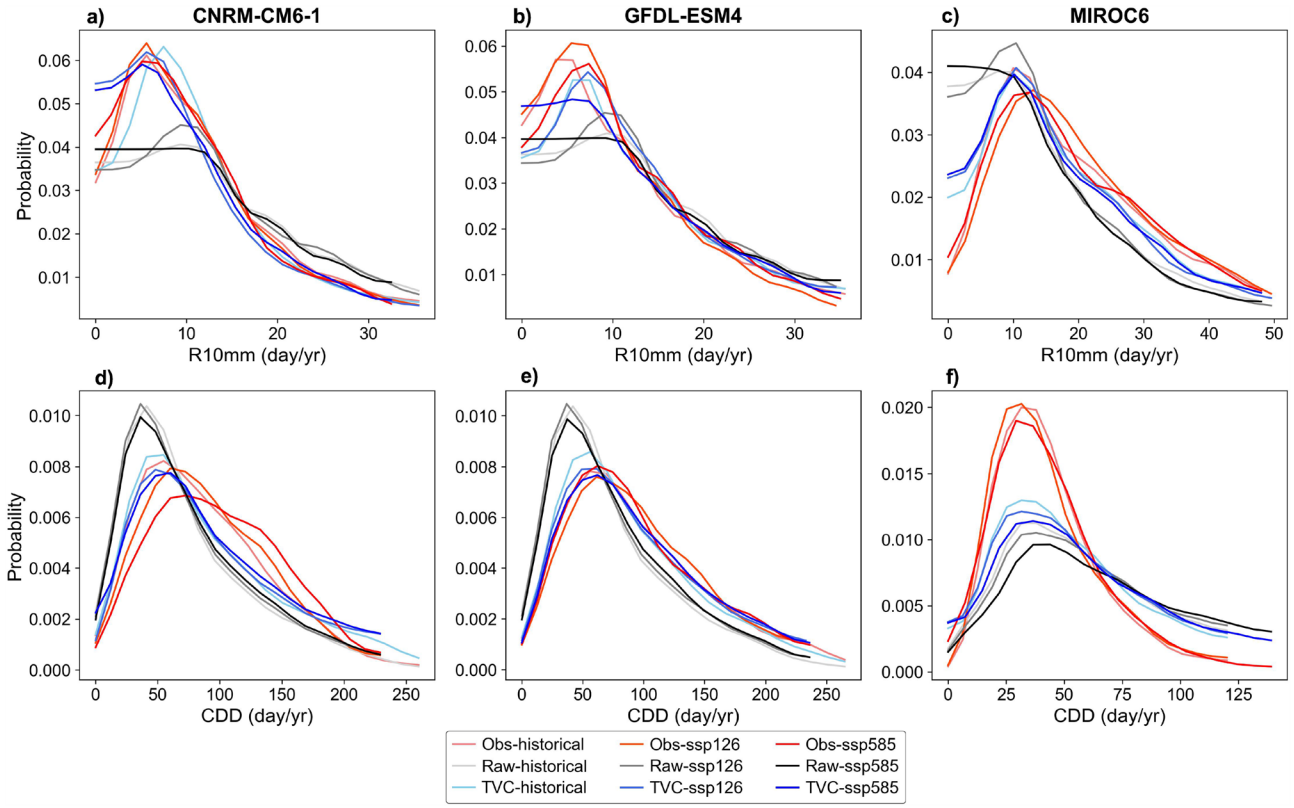


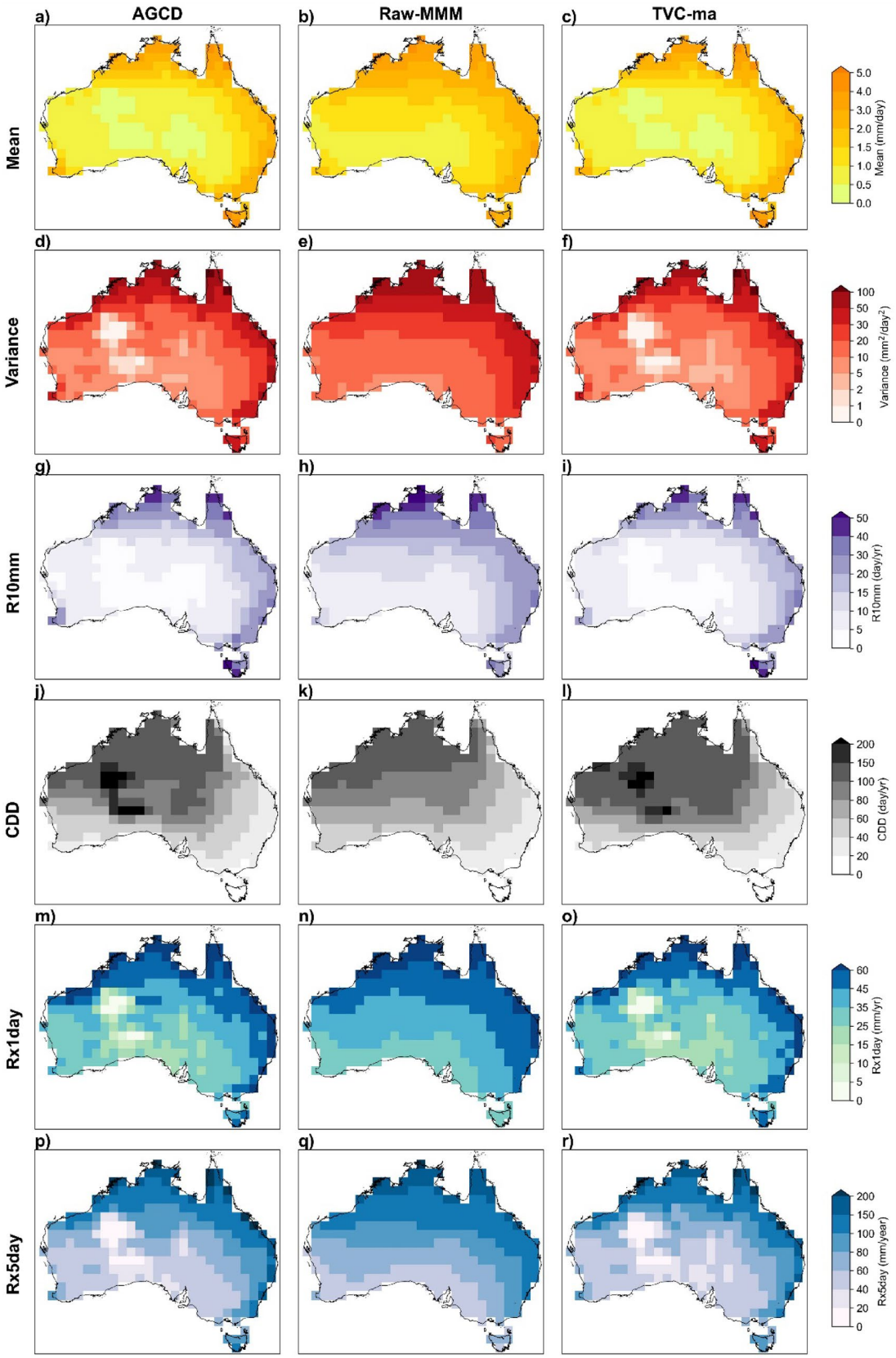
Fig. 7 PDF of R10mm and CDD for CNRM-CM6-1 (**a**, **d**), GFDL-ESM4 (**b**, **e**), and MIROC6 (**c**, **f**) obs, raw and TVC-ma results over the in-sample historical period of 1950–2014, and the out-of-sample projection period, ssp126 and ssp585, of 2015–2099

CDD values generally decreasing as latitude increases. The AGCD (Fig. 8j) and TVC-ma (Fig. 8l) results, on the other hand, show a much higher number of CDDs in central Australia, while CDDs are slightly higher elsewhere except in the coastal regions. For both Rx1day (Fig. 8n) and Rx5day (Fig. 8q), the raw models produce maximum 1-day (and 5-day) precipitation amounts exceeding 45 mm (100 mm) per year in northern Australia and the eastern coastal regions, with these amounts generally decreasing with the increasing latitude. In comparison, AGCD (Fig. 8m, p) and TVC-ma (Fig. 8o, r) results reveal larger precipitation amounts along the northern and eastern coastlines as well as Tasmania, while central Australia exhibits lower values.

Under the future ssp126 emission scenario (Fig. 9a–f), the spatial patterns of all the metrics from the raw averaged projections are similar to those presented during the historical period (Fig. 8b, e, h, k, n, q), but with a lower precipitation mean in most regions, except for the north, a much larger variance in northern Australia and the southeast, a shorter period of heavy rainfall across the continent, except for Tasmania, a longer period of continent-wide dry days, and more intense maximum 1-day and 5-day precipitation events in northern Australia.

Since we apply the same mean and covariance scaling factor obtained in the historical period to the future projections, projected changes are largely preserved in the TVC-ma metric results. Figure 9g–l show that the main differences between the TVC post-processed projections and the raw projections in the future period have a similar geographic shape and magnitude to the differences over the historical period. For both emission scenarios, after post-processing, precipitation means and variances greater than 2 mm/day and 30 mm^2/day^2 are evident in northern Australia, Tasmania and along the eastern coastal region, while precipitation means and variances lower than 0.5 mm/day and 2 mm^2/day^2 are observed in central Australia. The R10mm value greater than 30 days per year is seen along the northern and eastern coastal regions, and in Tasmania. The CDD values longer than 100 days per year are distributed in northern, western, and central Australia. The Rx1day and Rx5day values exceeding 45 mm and 100 mm per year, respectively, are prominent in northern Australia, Tasmania, and the eastern coastal regions.

Under the higher emission ssp585 scenario (Fig. 9s–x), compared to the ssp126 scenario, post-processed projections, on average, exhibit noticeably higher precipitation means and R10mm values in the north and along the eastern


a) AGCD
b) Raw-MMM
c) TVC-ma
Mean
d)
e)
f)
Variance
g)
h)
i)
R10mm
j)
k)
l)
CDD
m)
n)
o)
Rx1day
p)
q)
r)
Rx5day
Mean (mm/day)
Variance (mm²/day²)
R10mm (day/yr)
CDD (day/yr)
Rx1day (mm/yr)
Rx5day (mm/year)

◂**Fig. 8** Mean (**a–c**), variance (**d–f**), R10mm (**g–i**), CDD (**j–l**), Rx1day (**m–o**), Rx5day (**p–r**) of observed, raw multi-model mean (raw-MMM) and corresponding TVC-ma corrected precipitation predictions across the Australian continent over the in-sample historical period in Dec 1951–2014. The metric results are averaged value across all CMIP6 models

coastline, while showing lower values for these metrics in parts of western Australia. Furthermore, increased variability is observed from the northern to the eastern regions, and longer periods of consecutive dry days are present across most parts of the continent, except in a few areas in northern and central Australia. Higher maximum 1-day and 5-day precipitation amounts are observed across most of the continent, except for a few regions in western Australia.

Compared to the historical period, ssp585 projections post-processed with the TVC-ma method show significantly lower precipitation mean and R10mm vales in parts of western Australia, with a slight increase in precipitation mean in parts of northern Australia. Across the rest of the continent, particularly in northern areas, precipitation displays greater variability, except for a few regions in western Australia. Prolonged periods of dry days are evident across much of the continent, with only a few scattered cells in northern and central Australia experiencing shorter dry spells. The entire continent, except for parts of western Australia, experiences increased maximum 1-day and 5-day precipitation amounts, with the most pronounced increases observed in northern Australia.

5 Discussion

The TVC-ma approach inherits several key features from the TVC method (Shao et al. 2024), including: (1) preserving the time sequencing of the precipitation amounts in the raw model series (not shown); (2) improving the persistence attributes of daily precipitation projections, leading to enhanced estimates of prolonged extreme events, as clearly reflected in the CDD and Rx5day scores computed in this study; and (3) maintaining projections of changes in time-invariant mean and variability by applying the same covariance scaling factor obtained in the in-sample historical period to the out-of-sample projection period. Feature 3 means that TVC-ma cannot reliably improve projections of time variability when the true model predicts a change in time scale covariances that is significantly different to that of the models being adjusted by TVC-ma.

Our work highlights that some models' projections of the change in precipitation time variability under increased GHGs is significantly different compared to others. Since precipitation variance typically increases with mean precipitation amounts, these differences may be linked to corresponding differences in the projected changes in precipitation rates. Figure S8 illustrates the ratio of the mean for the 365-day time scale, representing the model predicted change in the mean of the 365-day averaged time series. Other time scales are not evaluated, as the mean of the remnant time series at these scales is nearly zero. Consistent with the results for the predicted change in time variability, the EC-Earth model shows the largest ratio under the ssp585 scenario, while NorESM2-MM and ACCESS-ESM1-5 exhibit the smallest ratios, compared to the other models.

After applying TVC to post-process lower-bounded precipitation time series, simply converting negative amounts to zero results in an inflated mean. To address this, we introduced a negative offset to the TVC results to ensure that the mean of the corrected time series matches the observed mean during the historical period. This offset is generally small and does not significantly alter the characteristics of the original time series that includes negative values. The correlation between TVC results (with negative values) and the TVC-ma results is typically greater than 0.99. This additive correction approach effectively adjusts the mean while largely preserving the improvements in variance and lag correlations achieved through the TVC-only method. We also explored alternative approaches, including: (1) adjusting the mean by multiplying the TVC series by the ratio of the observed mean to the TVC-only mean, and (2) using a linear relationship between the observation and TVC-only series to align both the mean and variance with the observed values. Results reveal that method 1, a multiplicative correction approach, improves the mean without negatively impacting lag correlations, but significantly reduces variance. Method 2 enhances both the mean and variance, yet it notably degrades lag correlations. In comparison, the mean adjustment approach presented in this work strikes a balance among these key metrics, providing a more effective solution.

To assess the impact of mean adjustment, we calculate the averaged percentage improvement of all metrics across all grid cells and model-as-truth cases, as summarized in Table S2. During both the in-sample historical and out-of-sample projection periods, mean adjustment significantly enhances mean predictions, yielding at least an 11% improvement difference. Additionally, it slightly improves the projections of R10mm and CDD values. While TVC-ma marginally reduces the improvements in variance, lag correlations, Rx1day and Rx5day scores (with differences below 7%), its overall enhancement of all metrics remain substantial compared to raw projections. Without mean adjustment, the corrected precipitation series would systematically overestimate mean values, compromising the accuracy of post-processed projections.

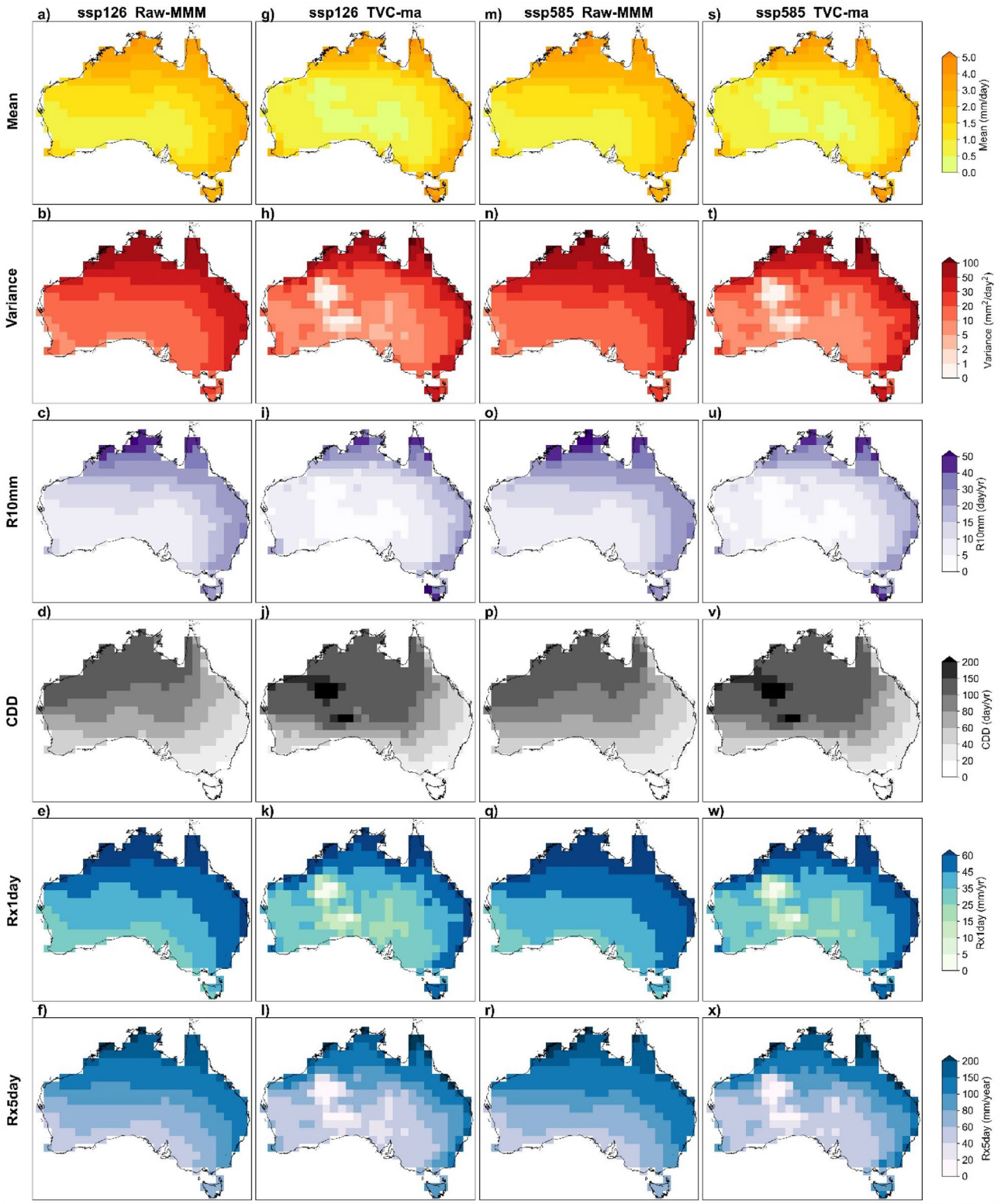


Fig. 9 The same as Fig. 8, but for raw-MMM and TVC-ma results across ssp126 and ssp585 scenarios

This work evaluates the application of the TVC-ma method across the Australian continent. Recent regional studies have similarly focused on employing various bias correction techniques to improve both mean and extreme

precipitation projections. For example, Kim et al. (2021) introduced a new regional quantile delta mapping method for correcting precipitation projections in South Korea, finding that the largest increase in precipitation changes occurred in the southwestern area during 2071–2100 under the high emission scenario. Wu et al. (2022) analyzed bias-corrected and spatially downscaled precipitation projections from five CMIP6 models for the upper Yangtze River basin (UYRB), revealing that lower emissions of aerosols and greenhouse gases may increase flood risks, with the lower basin facing a greater risk of extreme precipitation. Similarly, Babaousmail et al. (2022) evaluated projected changes in mean and extreme precipitation over the Mediterranean and Sahara regions using quantile-mapped daily precipitation data, concluding that the wet regions are expected to become wetter while dry regions become drier, with reductions in mean precipitation and increases in dry days likely exacerbating droughts and aridity.

To demonstrate TVC-ma's transferability to other global regions, we apply it under a model-as-truth setup to five cases across different continents: Asia, North America, Europe, Africa, and Antarctica (see Fig. S9). The percentage improvement for each case is calculated as outlined in Eqs. (17–18). As shown in Table S3, all verification metrics and extreme indices show improvements across all cases during both in-sample historical and out-of-sample projection periods. These results are consistent with the hypothesis that the application of TVC-ma to locations outside of Australia would yield improvements in precipitation projections similar to those found over Australia. TVC-ma's potential for application to larger regions is enhanced by its computational efficiency. In this work, training a single TVC-ma model typically takes less than 0.1 s. With parallel computing, it becomes feasible to train models for global grid cells within an hour.

The TVC approach for continuous temperature variables and the TVC-ma approach for non-continuous precipitation projections were both developed in a univariate setting for post-processing at the same spatial resolution. Theoretically, there are opportunities to extend the TVC-based methods for multivariable post-processing and downscaling. It might be possible to correct compound events in the future. Future research will extend the evaluation of TVC-ma across all emission scenarios, including SSP2-4.5 and SSP3-7.0, to investigate whether its performance varies under different forcing scenarios.

Similar to the temperature work, we assume a time-invariant bias in precipitation post-processing. This assumption may be more reasonable for precipitation than for temperature, as precipitation trends under global warming remain uncertain and may not be significant in many regions. However, in the regions where a significant precipitation trend is present, it may be more appropriate to assume that the trend difference between historical observations and raw model simulations extends into future projections. According to the IPCC AR6 report (Pörtner et al. 2022), there is high confidence that global warming will intensify global extreme precipitation conditions, including monsoon precipitation, and very wet and very dry weather and climate events and seasons under continued greenhouse gas emissions. Therefore, gauging the uncertainty of future precipitation projections is crucial. Ongoing research will focus on combining TVC-based methods and model constraining approaches to improve both individual model and multi-model ensemble projections.

6 Conclusion

In this work, we extended the time variability correction (TVC) method to post-process daily precipitation projections under climate change scenarios. To address negative values in the precipitation time series and ensure alignment with observed means, we introduced a mean adjustment step, resulting in the new TVC-ma method. This new approach was tested on four selected grid cells in Australia during the in-sample historical period, demonstrating substantial improvements in the mean and variance, better alignment with observed persistence attributes, and enhanced performance for the R10mm, consecutive dry day (CDD), Rx1day and Rx5day indices in some cases.

Subsequently, TVC-ma was evaluated in the model-as-truth framework, where each of 28 CMIP6 models was treated as the reference observation in turn. Results indicate that, in most model-as-truth cases, TVC-ma substantially improves the mean, variance, single-day time-lag-correlation, 5-day lag correlation of 5-day averages, R10mm, CDD, Rx1day and Rx5day compared to raw precipitation projections in most grid cells and in both in-sample historical and out-of-sample projection periods. Significant enhancements are particularly seen for lag correlations and CDD indices across 27 in 28 model-as-truth cases. Furthermore, TVC-ma effectively enhances the inter-annual distribution of the R10mm and CDD indices in many cases.

Another key finding from this work is that the projected changes in time-scale-dependent variability due to increased greenhouse gases (GHGs) differ significantly across models. These discrepancies are likely driven by corresponding differences in the projected changes in the mean precipitation amounts.

When applied to climate change projections under the ssp126 and ssp585 emission scenarios trained on the AGCD data, TVC-ma effectively reproduces the differences seen between observations and model outputs during the historical period. The corrected precipitation projections show higher means and variances in northern Australia, Tasmania and

along the eastern coastal region, while means and variances are lower in central Australia over the historical period. More heavy rain days are evident along the northern and eastern coastal regions, and in Tasmania, while consecutive dry days are more prevalent in northern, western, and central Australia. Additionally, more intense maximum 1-day and 5-day precipitation events are concentrated in northern Australia.

Under the high emission scenario, northern Australia appears wetter and has more heavy rain days compared to the low emission scenario. Regions from the north to the east are more variable, while most parts of the continent tend to have longer periods of consecutive dry days and more intense maximum 1-day and 5-day precipitation events. Compared to the historical period, corrected projections under the high emission scenario indicate drier conditions in parts of Western Australia, increased variability, longer periods of consecutive dry days, and higher maximum 1-day and 5-day precipitation amounts across most parts of the continent.

TVC-ma post-processing was also applied to selected land grid points globally. The results showed that the improvements in projections were comparable to those observed over the Australian region. Future work will extend TVC-based methods for multivariable post-processing and combine the methods with model constraining approaches to produce more reliable ensemble climate change projections at the global scale.

Appendix

Pseudo-code for the Time Scale Transform

#Initialise remnant sequence

$\tilde{y}_t^0 = y_t$, **for all** t **in the time series**.**This is the original time series**.

#Define the length of backward looking time averages

$$\left[P_1, P_2, ..., P_9\right] = [365, 183, 92, 46, 23, 12, 6, 3, 2]$$

#Each backward time − scale − dependent average is

for $i = 1 : 9$

$$t_0 = \sum_{k=1}^{i} P_k - (i-1)$$

for $t = t_0 : n$

$$\overline{y}_t^i = \sum_{j=0}^{P_i-1} w_j^i \tilde{y}_{t-j}^{i-1} = \left(\mathbf{w}_j^i\right)^T \tilde{\mathbf{y}}_{t-P_i+1:t}^{i-1} \quad w_j^i = \frac{1}{P_i} \text{ and } \sum_{j=0}^{P_i-1} w_j^i = 1$$

end

$$\tilde{\mathbf{y}}_{t_0:n}^i = \tilde{\mathbf{y}}_{t_0:n}^{i-1} - \overline{\mathbf{y}}_{t_0:n}^i$$

end

$$\tilde{\mathbf{y}}_{t_0:n}^9 = \mathbf{y}_{t_0:n} - \sum_{k=1}^{9} \overline{\mathbf{y}}_{t_0:n}^k \quad \textbf{where } t_0 = \sum_{k=1}^{9} P_k - 8$$

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00382-025-07731-7.

Acknowledgements We thank National Computational Infrastructure (NCI) for supplying supercomputing systems for computational analyses.

Author contributions YS and CHB conceptualized the methodology. YS led the experimental design, conducted the analyses, plotting figures, and drafted the manuscript. CHB revised manuscript and provided supervision.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. This work was funded by the Australian Research Council Centre of Excellence for Climate Extremes (CLEX; CE170100023).

Data availability statement The CMIP6 data used in this work can be accessed at: https://esgf-node.llnl.gov/search/cmip6/. The AGCD datasets (Australian Bureau of Meteorology 2020) are archieved on NCI, and are available upon registering for local access. The scripts for the TVC-ma method, data analyses, and figure plotting are available at https://doi.org/10.5281/zenodo.15572295 (Shao and Bishop, 2025).

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

References

Abramowitz G, Bishop C (2015) Climate model dependence and the ensemble dependence transformation of CMIP projections. J Clim 28(6):2332–2348

Abramowitz G, Herger N, Gutmann E et al (2019) ESD reviews: model dependence in multi-model climate ensembles: weighting, subselection and out-of-sample testing. Earth Syst Dyn 10(1):91–105

Australian Bureau of Meteorology (2020) Australian gridded climate data (AGCD)/AWAP; v1.0.0 Snapshot (1900-01-01 to 2020-06-30). https://doi.org/10.25914/6009600786063

Babaousmail H, Hou R, Ayugi B et al (2022) Future changes in mean and extreme precipitation over the Mediterranean and Sahara regions using bias-corrected CMIP6 models. Int J Climatol 42(14):7280–7297

Boyd SP, Vandenberghe L (2004) Convex optimization. Cambridge University Press

Cannon AJ (2016) Multivariate bias correction of climate model output: matching marginal distributions and intervariable dependence structure. J Clim 29(19):7045–7064

Cannon AJ (2018) Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables. Clim Dyn 50(1):31–49

Cannon AJ, Sobie SR, Murdock TQ (2015) Bias correction of GCM precipitation by quantile mapping: how well do methods preserve changes in quantiles and extremes? J Clim 28(17):6938–6959

Chen J, Brissette FP, Chaumont D et al (2013) Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America. Water Resour Res 49(7):4187–4205

Eyring V, Bony S, Meehl GA et al (2016) Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. Geosci Model Dev 9(5):1937–1958

Ghimire U, Srinivasan G, Agarwal A (2019) Assessment of rainfall bias correction techniques for improved hydrological simulation. Int J Climatol 39(4):2386–2399

Hay LE, Wilby RL, Leavesley GH (2000) A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the United States 1. JAWRA J Am Water Resour Assoc 36(2):387–397

Hess P, Lange S, Schötz C et al (2023) Deep learning for bias-correcting CMIP6-class earth system models. Earth's Future 11(10):e2023EF004002

Kim S, Joo K, Kim H et al (2021) Regional quantile delta mapping method using regional frequency analysis for regional climate model precipitation. J Hydrol 596:125685

Lenderink G, Buishand A, Van Deursen W (2007) Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach. Hydrol Earth Syst Sci 11(3):1145–1159

Mehrotra R, Johnson F, Sharma A (2018) A software toolkit for correcting systematic biases in climate model simulations. Environ Model Softw 104:130–152

Mehrotra R, Sharma A (2016) A multivariate quantile-matching bias correction approach with auto-and cross-dependence across multiple time scales: implications for downscaling. J Clim 29(10):3519–3539

Mehrotra R, Sharma A (2019) A resampling approach for correcting systematic spatiotemporal biases for multiple variables in a changing climate. Water Resour Res 55(1):754–770

Michalek AT, Villarini G, Kim T (2024) Understanding the impact of precipitation bias-correction and statistical downscaling methods on projected changes in flood extremes. Earth's Future 12(3):e2023EF004179

Michelangeli PA, Vrac M, Loukos H (2009). Probabilistic downscaling approaches: application to wind cumulative distribution functions. Geophys Res Lett 36(11)

Nguyen H, Mehrotra R, Sharma A (2019) Correcting systematic biases across multiple atmospheric variables in the frequency domain. Clim Dyn 52(1):1283–1298

Nguyen H, Mehrotra R, Sharma A. (2020) Assessment of climate change impacts on reservoir storage reliability, resilience, and vulnerability using a multivariate frequency bias correction approach. Water Resour Res 56(2):e2019WR026022

Peter J, Vogel E, Sharples W et al (2023) Continental-scale bias-corrected climate and hydrological projections for Australia. Geosci Model Dev Discuss 2023:1–45

Pörtner HO, Roberts DC, Adams H et al (2022) Climate change 2022: impacts, adaptation and vulnerability

Rajulapati CR, Papalexiou SM (2023) Precipitation bias correction: a novel semi-parametric quantile mapping method. Earth Space Sci 10(4):e2023EA002823

Shao Y, Bishop C, Hobeichi S et al (2024) Time variability correction of CMIP6 climate change projections. J Adv Model Earth Syst 16(2):e2023MS003640

Shao Y, Bishop C (2025) Mean-Adjusted Time Variability Correction. Zenodo 1.0.0. https://doi.org/10.5281/zenodo.15572295

Tong Y, Gao X, Han Z et al (2021) Bias correction of temperature and precipitation over China for RCM simulations using the QM and QDM methods. Clim Dyn 57(5):1425–1443

Vrac M, Noël T, Vautard R (2016) Bias correction of precipitation through singularity stochastic removal: because occurrences matter. J Geophys Res Atmosph 121(10):5237–5258

Vrac M, Thao S, Yiou P (2022) Should multivariate bias corrections of climate simulations account for changes of rank correlation over time? J Geophys Res Atmosph 127(14):e2022JD036562

Wu H, Lei H, Lu W et al (2022) Future changes in precipitation over the upper Yangtze River basin based on bias correction spatial downscaling of models from CMIP6. Environ Res Commun 4(4):045002

Zhang X, Alexander L, Hegerl GC et al (2011) Indices for monitoring changes in extremes based on daily temperature and precipitation data. Wiley Interdisc Rev Climate Change 2(6):851–870

Zhao R, Zhou X, Li Y et al (2024) Quantile delta-mapped spatial disaggregation analysis for summertime compound extremes over China. Clim Dyn 62(9):8453–8473