<u>Responses to comments from the reviewers</u>

We are thankful to the Editor and Associate Editor for handling this manuscript, and the anonymous reviewer for the comments and suggestions. We have carefully considered each of the comments when preparing the revised manuscript. We believe that all these comments are being addressed. Below we provide a response to each of the comments and describe in detail how the original manuscript was modified. Note that all the changes to the manuscript were implemented using Track Changes in Word. Furthermore, lines numbering is according to the manuscript under Simple Markup visualization of the changes.

In this response letter, we have added **comments from both reviewers** with **our detailed answers**, a version of the modified **manuscript without Tack Changes**, as well as a version of the **manuscript with Track Changes**.

Reviewer 1:

This article presents a software package that consolidates multiple quantile mapping methods into one common interface, and makes it available in the three most common coding languages used by climate scientists and others who use climate projections. While limited to quantile mapping type approaches, this will be a useful software package for many people exploring climate impacts on a local or regional scale. My suggestions below are largely aimed at providing a more complete context for this package relative to available resources and methods.

1) Line 34, following the sentence on RCMs, it might be worth mentioning that there are also hybrid methods (e.g., Walton et al., 2015, https://doi.org/10.1175/JCLI-D-14-00196.1) that combine dynamical and statistical downscaling.

Response: We thank the reviewer for the comment; to address it we have added more detail and a description of hybrid methods. This has been done by dividing the first paragraph of the Introduction (Section 1) into three paragraphs to clearly differentiate between downscaling and bias correction methods (lines 28 to 45 in the updated manuscript). Also, we now mention that there are statistical, dynamic, and hybrid downscaling methods, including the proposed reference.

2) Lines 34-36, While the software being presented is only for QM variations, the sentence listing "several bias correction (or bias adjustment) methods" should be a little more explicit in what

these consist of (like analogues). You might also mention that analogues would not be amenable to the approach of this software, where all methods essentially work on point-to-point bias correction. Also, current development of machine-learning approaches to downscaling is emerging (e.g., Soares et al., 2024, https://doi.org/10.5194/gmd-17-229-2024) and should be noted.

Response: We appreciate the comment; thanks to it we have strengthened the Introduction (Section 1). In the first three new paragraphs of the Introduction, we included a more detailed description of downscaling and bias correction methods to clarify that the climQMBC package only includes a set of bias correction methods. We also included the machine-learning approach and the proposed reference. Additionally, based on a new Appendix focused on execution times (Appendix B in the updated manuscript; line 709), we now included a brief discussion at the end of the Conclusions to acknowledge potential machine-learning contributions in execution time (line 605 in the updated manuscript).

3) Lines 58-61, a little more discussion would help characterize the uncertainties. The Lafferty et al. (2023) reference would provide a good framing of this (https://doi.org/10.1038/s41612-023-00486-0).

Response: We appreciate the comment and have extended the paragraph of the Introduction (Section 1) to include more details about the uncertainties in the climate change impacts assessment and the contribution of downscaling and bias correction based on the proposed reference and complementary articles (lines 78 to 98 in the updated manuscript). This comment, with others, has helped strengthen the introduction and better frame the need for a package to easily compare different bias correction methods.

4) Line 85, Add a short paragraph discussing the availability of downscaled data, which is increasingly being used by stakeholders to avoid having to do downscaling at all. A few examples of CMIP6 global downscaled data sets are those from the Climate Impacts Lab (Gergel et al., 2024, https://doi.org/10.5194/gmd-17-191-2024), School of Geography and Environmental Science (2022, https://doi.org/10.5285/c107618f1db34801bb88a1e927b82317) and the NASA-NEX

archive (Thrasher et al., 2022, https://doi.org/10.1038/s41597-022-01393-4). At regional/continental scales there are many more data sets. Of course, each uses its own observational baseline data, training period, etc., to attributing differences to specific sources is not possible, which means there is still value in facilitating statistical downscaling.

Response: We thank the reviewer for the comment; it has helped us to improve the Introduction (Section 1). We added a brief paragraph in the Introduction addressing the relevance and scope of available downscaled, both global and regional (lines 99 to 110 in the updated manuscript).

5) Line 285, the example script (Figure 4) and the linked example notebook on github are helpful illustrations of how to use the package for a point. It may exist on the github site but I could not find an example of applying this to a gridded (e.g., netCDF) data set. Since that would be a relatively common application, developing an example of that would make this more complete.

Response: We now included a working example and sample dataset for gridded products in the tester scripts available in the GitHub repository for each programming language (climQMBC_tester_Python.py, climQMBC_tester_R.R and climQMBC_tester_Matlab.m) and at the end of the Example_notebook which is available in the main page of the Github repository (https://github.com/saedoquililongo/climQMBC/blob/SA/Example_notebook.ipynb). This is now mentioned in lines 368 to 376 of updated manuscript. We thank the reviewer for the comment as now we have a more complete set of real examples that users and readers can explore.

Reviewer 2:

The manuscript introduces climQMBC, a coding package implemented in Python, R, and MATLAB that bundles five univariate quantile mapping bias correction methods (QM, DQM, QDM, UQM, SDM) for point-based daily, monthly, and annual climate series. The authors emphasize three contributions: a unified mathematical formulation of the methods, identical functionality across three languages, and a built-in summary report function that compares methods by their ability to preserve raw GCM trends in the mean and standard deviation. Performance is assessed using a synthetic exercise in which gamma-distributed precipitation series are perturbed by controlled biases in the mean and standard deviation across a range of coefficients of variation and projected-change magnitudes. The authors conclude that UQM and DQM are the most reliable methods overall, while acknowledging that QM remains useful for its simplicity.

Major Comments:

1) There is no real-world demonstration of bias correction. For a paper that motivates its tool by appealing to hydrology, wildfire, and heatwave applications, the complete absence of any application to actual GCM/RCM output paired with station or gridded observations is striking. Figure 4 shows a code snippet plotting a single monthly near-surface air temperature ('tas') series drawn from the package's bundled sample data; Figures 5 and 6 illustrate the report function on a precipitation example whose source, region, and driving model are not described. Neither constitutes a real-world application in the sense of documented climate-model output paired with observations for a stated location.

Without a real case the reader cannot assess whether the package handles realistic complications such as elevation-dependent biases, monsoonal bimodality, drizzle-dominant tropical regimes, or arid regions where the random-near-zero replacement may dominate the corrected series. This omission also makes it impossible to evaluate computational scalability, which matters because the moving-window strategy multiplies the cost of each method substantially relative to the original single-window formulations.

Response: We appreciate the comment; thanks to it we have better presented our package. First, to bring the reader to a real-world case of bias correction based on freely available datasets, we added the location of the point analyzed (Cocora Valle in Colombia) and both the observational (ERA5-5-Land) and modeled (CMIP6 MPI-ESM1-2-HR run by SSP 5-8.5) products used in Section 3.2 (lines 356 to 362 in the updated manuscript). Additionally, we added a brief annotation that comes out of the example shown where the UQM and QM methods result in different temperature projections (lines 378 to 382 in the updated manuscript).

The new version of Section 3.2 is now based on that dataset and example. As such, we updated Figures 4, 5 and 6 and the respective description to be consistent with the dataset used. We also updated the sample dataset of the repository to the one used in Section 3.2, so that any reader can verify the results. We have also added a grid-based application case to the repository, which is also mentioned now in Section 3.2 (lines 368 to 376 in the updated manuscript).

Regarding the concerns of realistic complications, the article presents a unified notation of all methods and describes the different assumptions of each method, along with strategies to solve some of the methods problems, like the random near-zero replacement or the moving windows. The specific analysis of each of these procedures and method for all potential complications is out of the scope of this work, and we expect the climQMBC package to support these future efforts. To complement the Introduction (Section 1), we included references to works that have made explicit the use of the climQMBC (lines 142 to 143 in the updated manuscript). Regarding the computational scalability, we created an Appendix B (starting at line 709 of the updated manuscript) to include a summary of the performance of the methods in terms of execution time relative to the QM method. We also mention the comparison of execution times in lines 247 to 250 of the updated manuscript. Additionally, we included a discussion on computational scalability and potential challanges in the Conclusions (Section 5; lines 601 to 607 of the updated manuscript).

Please note that although the moving-window strategy has some computational cost, there are significant benefits in terms of properly capturing the GCM changes. We have added Appendix A to explain this in the new manuscript (from line 608 of the updated manuscript). We would like to thank the reviewer for highlighting this; it has helped us strengthen the manuscript in important issues.

2) NSE and KGE are inappropriate as primary metrics for distribution-based bias correction. NSE and KGE measure point-by-point temporal agreement. Quantile mapping methods, by construction, do not aim at point-by-point reconstruction of the "true" observed projected series; they aim at distributional consistency and trend preservation. The bias-corrected series is sequenced according to the modeled time series, so a high NSE only arises when the model already has the correct temporal sequencing of values, which has nothing to do with the bias correction itself. The authors should justify why these metrics are appropriate at all, or replace them with metrics that directly probe distributional similarity (Wasserstein distance, integrated

quantile difference, tail-weighted scoring rules). As presented, the cumulative NSE/KGE curves in Figure 9 conflate distributional skill with the trivial fact that, in the synthetic design, the modeled and observed series share the same underlying random draws.

Response:

We appreciate the comment; with it we have realized that some important parts of the manuscript were not clear enough in the previous version. As the reviewer properly mentions, neither the NSE, nor the KGE are appropriate metrics to assess Quantile Mapping methods in real life scenarios. One of the contributions of our manuscript is to propose a simple way to verify the performance of bias correction methods. For this, we have added bias to a synthetically generated time series (Section 4.1) and then performed the bias correction, hence we know beforehand the objective time series without bias. In other words, the synthetic exercise considers a "observed" historical and future series, while for the "modeled" series we added bias (i.e., in the mean and standard deviation) to these original series. As such, the modeled series has a consistent sequence of events with the objective series (i.e., the "observed" one). In those conditions, in lines 469-470 of the updated manuscript we made explicit that "As such, in this synthetic exercise the bias corrected series should be the same as the synthetic observed historical or future objective series". In this context, the use of the NSE is appropriate as it is expected a point-by-point reconstruction of the objective series and the KGE is also appropriate because it considers the bias in the first two statistical moments and if the sequence of events is consistent with the objective series. Note, that this approach allows objectively comparing which bias correction method is having a better performance in removing the "added bias" in the synthetic exercise. To avoid misunderstandings of the exercise and on the relevance of the used metrics, we restructured the explanation of the synthetic exercise in Section 4.1 aiming 1) for a better explanation of the construction of the series and clarifications of what is expected from the bias correction process and 2) to explicit the appropriateness of the NSE and KGE as metrics and the scope and limitations that these have in real life applications. Additionally, we included complementary notes on the applicability of these NSE and KGE in the Conclusions (Section 5; lines 587 to 592 in the updated manuscript).

We thank Reviewer 2 for noting this aspect, which led us to include the synthetic example as a functionality in the climQMBC package to evaluate quantitatively the most suitable method for their specific conditions. This is mentioned in the Introduction (lines 138 to 140 of the updated manuscript), the Synthetic Example Section (lines 443 to 450 of the updated manuscript), the Conclusions (lines 581 to 592 of the updated manuscript) and the Abstract (lines 22 to 25 of the updated manuscript).

3) The novelty claim is thin for a methods paper at GMD. The five implemented methods are all from prior literature. The unified notation and the report function are useful but incremental. The strongest selling point, multi-language availability with identical functionality, is a software-engineering contribution rather than a scientific one. Major competing packages are dismissed in a single paragraph without quantitative comparison: xclim and ibicus are not benchmarked, and MBC is dismissed for lacking trend-preserving methods without acknowledging that it contributes the entire multivariate dimension that climQMBC lacks. A side-by-side benchmark on a common test case, demonstrating either superior accuracy, speed, or usability, would substantiate the contribution claim. Without it, the paper reads as documentation rather than as a research article.

Response: We appreciate the comment and agree with the reviewer that the previous version of the manuscript needed to better highlight its contribution. Regarding the novelty of the work, despite all efforts on the development of bias correction methods, 1) recent works (for example, Lafferty et al. (2023); https://doi.org/10.1038/s41612-023-00486-0) show that the contribution of bias correction methods to uncertainty of the climate change assessment is highly heterogeneous and can contribute to around 30% of the total uncertainty, 2) recent works (for example, Fowler et al, 2025; https://doi.org/10.1007/978-3-031-85542-9_18) still suggest the need for comparison of bias correction methods to discern the most appropriate for each analysis, and 3) there are no tools to allow users to do such a comparison of methods in a comparative framework. This work is a step to close this gap in terms of a tool to compare methods and two functions to do the comparison (report function and the synthetic example, recently added thanks to reviewer two comments). We updated the Introduction to make these gaps clear and to better present climQMBC package contribution.

Please note that this type of contribution is well within the scope of GMD journal. In particular, this work presents a novel tool and approach to compare bias correction methods (several methods in a unified framework and available in widely used programming languages, and both a report function and a synthetic comparison of controled conditions), alongside with a comparable description of the methods available in the climQMBC package.

Regarding the comparison with other packages, while we ensured that every process is as numerical as possible instead of based on loops and conditionals, the objective of the climQMBC package does not rely on being the fastest, but the more useful for consultancy, academia and research so that users can evaluate the trade-offs that each method has.

In terms of accuracy, there is a gap in understanding and measuring the impacts of different decisions on the implementation of bias correction methods. On the other hand, as the contribution of the bias correction process to uncertainty is highly heterogeneous, an extensive analysis would be a work on itself, considering that there is not yet a consensus on how this analysis should be performed. We expect that this work and tool will support the development of such types of studies in the future. Additionally, in the Introduction (Section 1) we already mention the scarcity of packages with several methods and available in more than one programming language, which needed to close the gap to make the comparison of bias correction methods and uncertainty analysis accessible to other studies. As such, a side-by-side benchmark on a common test case is beyond the scope and relevancy of the presented work.

4) The package is strictly univariate and point-based, which limits its relevance for modern impact assessment. Compound events (for example hot-dry conditions, wind-fire coupling, and precipitation-temperature co-variation in snowpack modelling) are increasingly recognized as the dominant climate-risk pathway, and they require methods that preserve inter-variable and inter-site dependence. The manuscript acknowledges this gap only obliquely. Users who adopt climQMBC and apply it independently to coupled variables risk introducing spurious decoupling,

particularly in snow-rain transitions or fire-weather indices where multivariate consistency matters. This limitation deserves explicit discussion and a warning to users, not a passing remark.

Response:

We agree on the need for multi-variable bias correction methods as univariate point-based bias correction methods do not consider the local inter-site and inter-variable dependance. However, existing algorithms that aim bridging this gap, like the MBCn (Cannon, 2017) and ISIMIP3BASD (Lange, 2019), do not always correctly represent local conditions other than the correlations and fail to preserve the raw GCM trends (Aedo-Quililongo and Chong, 2024; Aedo-Quililongo and Chong, 2025). Unfortunately, this conclusion is hardly reported as articles tend to focus on new methods rather than comparisons of methods and raising limitations.

We incorporated a paragraph in the Introduction (Section 1 lines 70 to 77 in the updated manuscript) and in the Conclusions (lines 573 to 580 in the updated manuscript) acknowledging this gap. Thank you for the comment, it has helped us acknowledge more explicitly that the climQMBC package in its current version works with univariate methods.

5) Methodological choices in distribution selection and zero-precipitation handling are under-justified. Fitting candidate distributions by the method of moments and then selecting among them by the Kolmogorov-Smirnov statistic is internally inconsistent. KS is insensitive to tail behaviour, which is precisely where the extremes most relevant to impact studies live. There is also no discussion of fit-quality reporting to the user, so a user could obtain a heavily biased correction without diagnostic information. The random-near-zero replacement procedure, while pragmatic, introduces stochasticity that breaks reproducibility and whose sensitivity is admitted but never quantified. For arid regions, where the fraction of zero-precipitation days can exceed 90 percent, this procedure may govern the correction more than the actual distribution fitting. The moving-window design is not stress-tested. The authors recommend windows of 20 to 40 years following Chadwick et al. (2023), but no sensitivity analysis is presented within this paper to validate that recommendation under the synthetic conditions explored. For long projections (for example to 2100), the window for the early decades of the projected period still contains

historical data, which dampens any abrupt nonstationarity. The synthetic experiment, which uses stationary gamma distributions in the projected period, cannot detect this issue at all.

Response: We appreciate the comment; it highlights several issues where improving our explanation of the method is required. We have significantly added more detail in several of these issues. Regarding the random near zero replacement, it is a procedure that has already been proposed in the literature (see Vrac et al., 2016; https://doi.org/10.1002/2015JD024511) by considering one threshold to define the values to be replaced and the range of the random values; however, in Chadwick et al. (2023) a variation with two parameters is presented, which is the variation implemented in the climQMBC package. The first parameter is a threshold that represents the physically zero precipitation value (defined as default as 1 mm, independently of the temporal scale). Every value below this threshold will be replaced. The second parameter is a factor to scale the physically zero precipitation value (1/100 as default) and defines the maximum possible replacement value. For example, considering the default values, every precipitation value below 1 mm will be replaced by random values between 0 and 1/100. If by any means, the projected change increases the value 10 times (change that would be unrealistic according to GCM projections), the maximum new value would be 1/10 (smaller than 1). As such, it would be removed at the end of the correction process. In Section 3.1.1 we indicate that these parameters should be analyzed as exceptional conditions resulting from observational data, or model results may require adjustments. However, we noted that the two parameters are not completely described in the article as in this response. We modified the last paragraphs in Section 3.1.1 to make the process clear (lines 218 to 242 in the updated manuscript).

Regarding the applicability or success of the random near zero replacement, several studies have successfully worked with the climQMBC in semi-arid to arid regions in Chile. For example, Blin (2023; https://repositorio.uc.cl/server/api/core/bitstreams/acf55587-dc5f-41fa-8d7e-e23cba55082d/content), Martínez et al. (2023; https://doi.org/10.1016/j.foreco.2023.121169), Vicuña et al. (2025; https://doi.org/10.1007/978-3-031-85040-0_16), Gonzáles-Leiva et al. (2026; https://doi.org/10.1080/02626667.2026.2625979) and Guzmán et al (2026; https://doi.org/10.2166/wcc.2026.267). We included these references in the Introduction

Section 1 of the updated manuscript (lines 142-143 in the updated manuscript). We thank the reviewer as this inclusion helped complement the manuscript.

Regarding the concerns of the use of the Kolmogorov-Smirnov test, following other comments of Reviewer 2, we included a disclosure of the limitations of the test in lines 205 to 207 of the updated document. We also included in text that the package prints a message of the times that the test fails for all available distributions in the library; however, the bias correction procedures continue based on the best distribution (lines 210 to 213 in the updated manuscript). Additionally, we mention now that the user could also force or define the desired distribution function to be used and that the parameters are estimated with the method of moments.

To clarify the logic behind using moving windows, we have included Appendix A (line 608 of the updated manuscript) with three sections analyzing different issues. These sections are based on synthetic examples to address the impacts of sample size on the accuracy of mean and standard deviation estimates, impacts on the trend estimates and pseudo-stationarity bias when considering moving windows of different lengths or considering a unique period for the analysis.

Regarding the synthetic example presented in Section 4.1, it focuses solely on the method rather on the random near zero replacement or the moving windows capabilities. We modified the example description explaining how it was implemented and applied with the climQMBC package: the future period is duplicated (concatenation of historical, future and future) so that the moving windows of the second future period consider only information of the future period and omit any information of the historical period. Through this process the climQMBC package can evaluate a future without capturing the historical period. This clarification has been added in the current version of the manuscript in Section 4.1 (lines 476-480 in the updated manuscript).

6) The conclusions overgeneralize. Statements such as "UQM and DQM are the more stable ones" and "there is always a trend-preserving method that is more stable than the QM method" are presented as general findings, but they are conditional on the specific synthetic design. Users in a region where the GCM bias is concentrated in extreme quantiles rather than in the mean may find QDM or SDM more appropriate, and nothing in the paper would warn them otherwise. The recommendation should be reframed as "in our synthetic exercise targeting moment biases,

methods that target moments perform best," which is closer to a tautology than to a usable guideline.

Response: We appreciate the comment; we have modified the text accordingly. In particular, to avoid a generalization outside the presented exercise, in Section 4.3, in the paragraph after Figure 10 (lines 541 to 544 in the updated manuscript) we added a phrase to remind the reader that the exercise considers a bias in the first two moments. Additionally, in the paragraph after Figure 11 (lines 557 to 559 in the updated manuscript) we reinforced that the exercise is based on adding bias in the first two statistical moments. In the Conclusions (Section 5), lines 593 to 597, we narrow the conclusons to the conditiions of the synthetic example.

Minor Comments:

7) Line 38: check the citation, like the "(Andres et al., 2014)" is inconsistent with the reference-list year (Line 520, dated 2013), and some name misspelled, such as Hageman and Hagemann, in more than one place.

Response: Thank you for your comment. We checked and updated the references and citations for consistency.

8) Line 74: the statement "there used to be a research lab with a package available in two programming languages, but one of these is no longer available" is vague and unverifiable. Please specify the lab and the package, or remove this argument from the motivation.

Response: As it is no longer available, we decided to remove the sentence. Thank you for the comment.

9) Line 92: the report function is restricted to the monthly temporal scale, yet the package supports daily, monthly, and annual data. Please clarify why daily and annual scales are not covered by the report function, and whether this is a planned extension.

Response: Effectively, as mentioned after Fig. 4 (lines 388 of the updated manuscript), the report function is only available on the monthly temporal scale. The report function has not been implemented for the annual temporal scale because its analysis is not expected to be a challenge for users. For the daily temporal scale, there are two main considerations: 1) the resulting figures of the report would be more cluttered as for monthly values, and 2) the answer to which are the minimum time series characteristics that need to be analyzed to consider a result robust and reliable both for general applications and for each line of analysis is not completely defined in the literature, and the possibilities are both extensive and depends on each area of analysis. However, the package has been and is still on continuous development both motivated by enabling complex methods to users in a comparative framework and by users request or need on their daily work on climate change impacts assessment. We consider that this explanation is not necessary to be included in the article and that disclosing the available capabilities is enough for the reader and user to understand the scope of the package. Please, if you consider this explanation should be in the manuscript let us know and we will add some clarification in this line.

10) Line 157: "by default chooses the distribution considering the one with the lowest error in the Kolmogorov-Smirnov test" reads awkwardly; rephrase, for example, as "by default selects the distribution with the lowest Kolmogorov-Smirnov test statistic". The authors should also acknowledge that the KS test is insensitive to tail behaviour, which is the regime most relevant for climate-extreme assessment.

Response: We thank the reviewer for the comment. We updated the sentence as proposed and added "...which is more sensitive to differences near the center of the distribution and less sensitive to differences in the tails" (lines 206 to 207 in the updated manuscript).

11) Line 167: "1 mm as default" lacks a time unit, whereas the SDM description (around Line 242) specifies "< 1 mm/day". Please make the units explicit and consistent at first mention, and clarify how this threshold is applied at coarser (monthly, annual) temporal resolutions.

Response: We appreciate the comment that was not clear enough in the previous version; we have improved this in the manuscript. The objective of this replacement is to avoid numerical problems in periods with large physically zero precipitation values. Typically, precipitation values of 1 mm/day are considered as a no-rain, but if is aggregated to monthly or annual scale, the values would be approximately 30 mm/month and 365 mm/year. If the threshold is fixed to 1 mm/day and the process is applied at a monthly scale, it could remove data associated with rain day values. As such, considering the objective of the replacement, the package considers a threshold of 1 mm, independently of the temporal scale (i.e. 1 mm/day, 1 mm/month and 1 mm/year for daily, monthly and annual precipitation, respectively). This also applies to the SDM method. We updated now the document with this explanation in the general description (lines 220 to 222 of the updated version) and for the SDM method (lines 310 to 311 of the updated version).

12) Equation (13): the subscript "RD_mh" is missing the comma separator and is inconsistent with "RD_{m,p}" used in the same equation. Please make subscript notation consistent throughout.
Response: We appreciate the comment and have made improvements accordingly. We checked this inconsistency across the document and finally updated two cases in equation 13 and the line before equation 13 (line 316 in the updated manuscript).

13) Equations (20)–(21): these are algebraically equivalent to the more transparent expression $x^b = \mu^o(1 + \Delta\mu_{bias}) + (x^o - \mu^o)(1 + \Delta\sigma_{bias})$, which makes explicit that the perturbation scales the mean by $(1 + \Delta\mu_{bias})$ and the standard deviation by $(1 + \Delta\sigma_{bias})$. Rewriting in this form would improve readability and would make the connection to Major Comment 1 transparent.
Response: We appreciate the comment; in effect this proposed change makes it more transparent. We replaced equations 20 and 21 with the proposed expression (line 468 in the updated manuscript).

14) Figure 4 (move to the Supplement). The worked example is presented as a screenshot of source code (panel a) and a single illustrative time series (panel b). Embedding code as an image is poor practice: it is not selectable, searchable, accessible, or verifiable. I recommend moving the worked example to the Supplement, and, if any code is retained in the manuscript, typesetting it as a formatted code listing rather than a screenshot, with the plot rendered as a separate figure.

Panel (b) is visually cluttered, with the QM, DQM, and observed lines overlapping heavily, so a shorter representative window or subplots would aid comparison.

Response: We appreciate the comment and agree that the way in which Figure 4 was presented in the previous version could be improved and have implemented changes in this version. Following another work from GMD (Davies et al. (2022; https://gmd.copernicus.org/articles/15/5127/2022/gmd-15-5127-2022.pdf) as an example, we included the code lines into a Listing element (and uploaded the corresponding pdf file) and replaced Figure 4 with the base plot and two close-ups to compare a window in the reference period and in the future period.

15) Figure 5: the "Future" header in the second column is not defined in the caption (it is explained only in the body text, around p. 15). Please define in the caption whether this column represents the ratio for the entire future period or for a specific centered period.

Response: We included in the caption the explanation for both the “Future” header and the periods shown, as it is included in the body text. Thank you for the comment.

16) Figure 6(a), right panel: the time-series figure is essentially unreadable because all five corrected series and the modeled series are overplotted as vertical bars in saturated colours. Please redesign using thin lines with transparency, or display only a representative subset.In panels (b) and (c), the legends obscure the curves to varying degrees.

Response: Thank you for the comment; it has helped us improve the visualization and presentation of our Figures. We separated Figure 6a and duplicated it in Figure 6b to show only one method as an example for better visualization and to show another capacity of the report function. What used to be Figure 6b and 6c were incorporated now as Figure 7a and 7b with modifications in the legend for better visualization.

17) Line 437: "the closer to the right the curve is, the better" is technically correct for the chosen cumulative-probability convention but may confuse readers; suggest rephrasing as "curves whose

mass is concentrated near NSE = 1 (that is, shifted toward the right edge of the plot) indicate better performance".

Response: We updated the sentence as proposed. Thank you, this helps with clarity and readability.

18) Figure 7: the in-cell percentages are described as "the percentage of synthetic time series that have some negative values that were replaced with random values below 0.01," which is unclear given that the synthetic series are gamma-distributed (strictly positive). Please clarify that the negative values arise after applying the bias formula (Eqs. 20–21), and specify the units of the 0.01 replacement threshold. Figure 7: the in-cell percentages are described as "the percentage of synthetic time series that have some negative values that were replaced with random values below 0.01," which is unclear given that the synthetic series are gamma-distributed (strictly positive). Please clarify that the negative values arise after applying the bias formula (Eqs. 20–21), and specify the units of the 0.01 replacement threshold.

Response: Thank you for the comment. We modified text before Figure 7 (Now Figure 8; lines 494 to 500 in the updated manuscript) to add an explanation to the negative values as proposed and specified that the negative values are associated to the biased synthetic series. This explanation was also included in the caption of Figure 7. For the synthetic example, we did not specify units to focus on the performance rather than on the variable.

19) Sections 4.3 and 5: the conclusion that UQM and DQM are the most reliable methods is conditional on the synthetic design, in which the bias is a simple shift in the first two moments; UQM and DQM are constructed precisely to preserve those moments, so the result is partly tautological. This caveat should be stated explicitly in the conclusion so that readers do not over-generalize the recommendation.

Response: We appreciate the comment; we have modified the text accordingly. To avoid a generalization outside the presented exercise, in Section 4.3, in the paragraph after Figure 10 (lines 541 to 544 in the updated manuscript) we added a phrase to remind the reader that the exercise considers a bias in the first two moments. Additionally, in the paragraph after Figure 11 (lines 557 to 559 in the updated manuscript) we reinforced that the exercise is based on adding bias in the first two statistical moments. In the Conclusions (Section 5), lines 593 to 597, we narrow the conclusons to the condiitions of the synthetic example.

20) Code-availability section: the expected input-data format is not stated in the manuscript. Please state briefly whether the input must be a flat array, a CSV with specific column headers, or a time-indexed series, so that potential users can assess compatibility without inspecting the repository.

Response: Thank you for the comment. We have added specifications regarding the data type and format, independently of the file type it comes from, at the end of the Code Availability section.

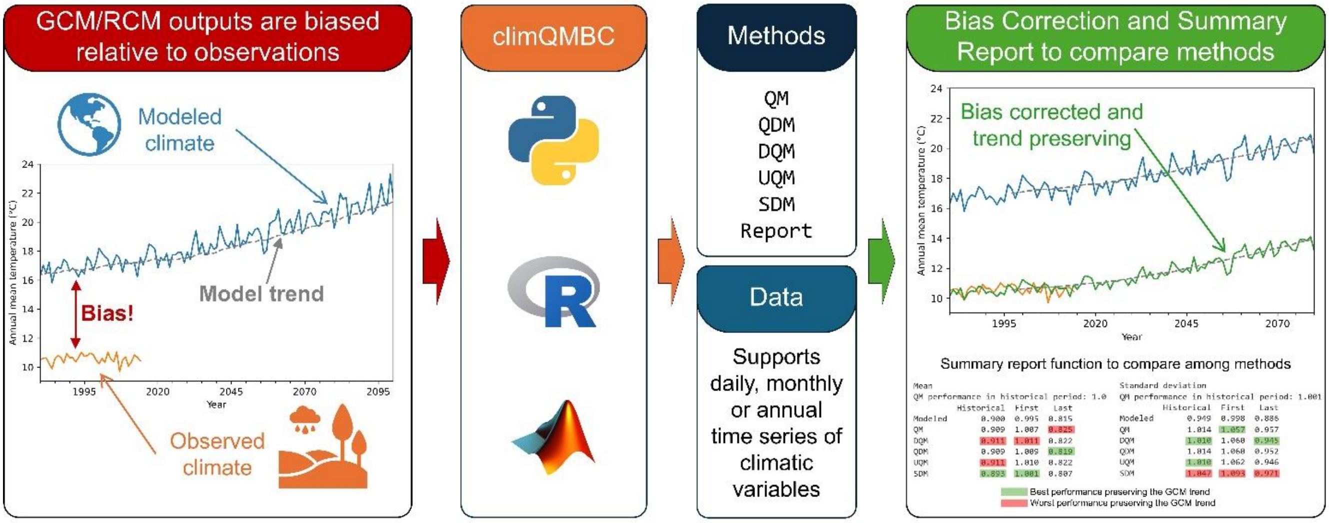
climQMBC: A package with multiple bias correction methods of GCM climatic variables at daily, monthly and annual scale, developed in Python, R and MATLAB

Sebastián Aedo-Quililongo[1], Cristián Chadwick[2], Fernando González-Leiva[3], Jorge Gironás[3,4,5]

[1]Stockholm Environment Institute, Latin America Centre, Bogotá, 110231, Colombia
[2]Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibánez, Santiago, 7941169, Chile
[3]Departamento de Ingeniería Hidráulica y Ambiental, Pontificia Universidad Católica de Chile, Santiago, 7820436, Chile
[4]Centro de Cambio Global UC, Pontificia Universidad Católica de Chile, Santiago, 7820436, Chile
[5]Centro de Desarrollo Urbano Sustentable ANID CEDEUS CIN250009, Santiago, 7820436, Chile

Correspondence to: Cristián Chadwick (cristian.chadwick@uai.cl)

Abstract. Climate change projections are studied using General Circulation Models (GCMs). GCMs simulate climate on a broad scale, hence they cannot be directly used in local impact studies, such as, hydrological studies. GCM results must be downscaled and bias corrected, to adjust their results in terms of spatial scale and reduce their bias before being used locally. However, downscaling and bias correction contributes to climate change impacts assessment uncertainty. Quantile Mapping (QM) is one of the most widely used approaches in bias correcting. However, standard QM assumes a time-invariant correction function, which might add bias. This has motivated the development of trend-preserving variations, which perform QM while aiming to preserve raw GCM changes. Unfortunately, choosing which variation to use is not straight-forward. We present the climQMBC package (https://github.com/saedoquililongo/climQMBC or https://doi.org/10.5281/zenodo.18392899) as an easy-to-use tool to compare point-to-point quantile mapping approaches. climQMBC is available in Python, R and MATLAB, and contains the standard QM method and four trend-preserving variations: Detrended Quantile Mapping (DQM), Quantile Delta Mapping (QDM), Unbiased Quantile Mapping (UQM) and Scaled Distribution Mapping (SDM). This package has a built-in summary report that allows comparing methods in terms of their capability of preserving raw GCM trends and a synthetic tester of bias correction methods. The synthetic tester develops times series to which bias is added allowing testing bias correction methods, given that one knows the objective time series before the bias correction. The synthetic tester showed that the UQM and DQM methods are the most reliable capturing statistical moments changes.



1 Introduction

Results from General Circulation Models (GCMs) are often used to assess climate change impacts on water resources, crop productivity, heat waves, among others climate risks (Fowler et al., 2007; Cannon, 2017; Portalanza et al., 2024). However, GCMs spatial resolution is not suitable for local-scale analysis as these models have biases associated with the parametrization of sub-processes, and their results correspond to mean values over large area (~ hundreds of km) (Mauritsen et al., 2012; Yuval and O'Gorman, 2020). To fulfill the limitations of climate models and allow the assessment of climate change impacts at the local scale a downscaling and bias correction (or bias adjustment) process is required, where downscaling is most commonly associated with spatial disaggregation (Aboitiz et al., 2026) and other processes that allow using GCM information locally and bias correction is the process where climate model results are adjusted to be consistent with observations in terms of magnitude, variability, among other characteristics (Hakala et al., 2019).

Downscaling methods are generally classified into dynamical and statistical methods (Fowler et al., 2025). Dynamical downscaling commonly use Regional Climate Models (RCMs), which are climate models with higher spatial resolution, better representing topography and physical processes than GCMs in a limited area (Christensen et al., 2008; Hakala et al., 2019; Bayissa et al., 2021; Petrovic et al., 2024), driven by GCM results as boundary conditions. Statistical downscaling consists of building empirical relationships between GCM or RCM results and local climate variables (Hakala et al., 2019; Fowler et al., 2025), including transfer functions, weather typing, weather generators, among others. Bias correction can also be considered a statistical downscaling method as it, in some cases closes the spatial scale gap. It is worth noting the emergence of hybrid approaches, combining dynamical and statistical downscaling methods, and machine learning approaches (Walton et al., 2015; Slater et al., 2023; Soares et al., 2024) for downscaling.

Several bias correction methods of climatic variables have been developed to allow the assessment of climate change impacts at the local scale (e.g. Maraun et al., 2010; Teutschbein and Seibert, 2012; Gutiérrez et al., 2013; Maraun, 2016, Bedia et al., 2020). Bias correction methods range from simply adjusting the mean to sophisticated adjustments of the distribution (aiming for several statistical moments), from univariate to multivariate, and from time-invariant to trend-preserving. One approach that stands out is that of quantile mapping, given its easy implementation, low need of calibration parameters, low computational cost, and applicability to several climate variables (e.g., temperature and precipitation) (Andres et al., 2014; Kim et al., 2016).

The standard Quantile Mapping method (QM; Wood et al., 2002) removes the climate model bias by matching the historical model distribution to that of the observed climate. Then the same correction is applied to the projected period. This implies a time-invariant assumption where the transfer function obtained with information from the historical or reference period is valid for the projected periods (Christensen et al., 2008; Maraun, 2012; Themeßl et al., 2012; Brekke et al., 2013; Chen et al., 2013). Several studies have raised concerns towards this stationarity assumption of the QM method and its capability to preserve the climate models signal (Hagemann, 2011; Ehret et al., 2012; Maurer et al. 2013; Maurer and Pierce, 2014; Cannon et al., 2015; Pierce et al., 2015; Hnilica et al., 2017; Chadwick et al., 2018; Chadwick et al, 2023). As a response to this issue, several authors have developed trend-preserving versions of the quantile mapping approach focusing on preserving the raw changes or signals of the climate models (e.g. Bürger et al., 2013; Cannon et al., 2015; Switanek et al., 2017; Chadwick et al., 2023). Some of the trend-preserving methods are: the Detrended Quantile Mapping (DQM; Bürger et al., 2013), developed to preserve the projected changes in the mean; the Quantile Delta Mapping (QDM; Cannon et al., 2015), developed to preserve the projected changes in the quantiles; the Scaled Distribution Mapping (SDM; Switanek et al., 2017), developed to preserve the likelihood of events and changes in quantiles; and the Unbiased Quantile Mapping (UQM; Chadwick et al., 2023), developed to preserve the changes in statistical moments: mean and standard deviation. Several studies have shown that these modified versions outperform the standard QM method (e.g. Cannon et al., 2015; Switanek et al., 2017; Chadwick et al., 2023). Currently, the QM method and its trend-preserving variants are widely used in several areas of analysis (Fowler et al., 2007; Cannon, 2017; Portalanza et al., 2024).

Another relevant dimension for bias correction methods is the capability to capture local interrelations or correlations between variables or in space like the MBCn (Cannon, 2017) and the ISIMIP3BASD (Lange, 2019), which are based on trend-preserving variations of the QM, also incorporating the relationships between variables (or points in space). These approaches are relevant to account for compound events like hot-dry conditions, precipitation-temperature co-variations in snowpack modeling, interrelations that univariate methods do not (Meyer et al., 2019; Menapace et al., 2025). However, while successfully capturing local correlations between variables, multi-variable bias correction methods may fail to capture local conditions and raw GCM trends (Aedo-Quililongo and Chong, 2024), currently existing a trade-off in the representativity of these methods.

Despite the efforts over the last decades, it has been challenging to determine the most robust bias correction strategy, with no consensus regarding the most adequate for each specific condition, adding a layer of uncertainty to the climate change impacts assessment (Hakala et al., 2019; Lafferty and Sriver, 2023; Fowler et al., 2025). Other uncertainties in the climate change impacts assessment, before the assessment model (e.g., a hydrological model), include: 1) GCM structure, spatio-temporal resolution and parametrization, leading to different responses to the same radiative forcing; 2) RCM boundary conditions, which may be inherited from GCM results; 3) scenarios, which define a broad range of potential drivers of GCMs; 4) internal variability, related to the chaotic nature of climate and represented by small perturbations of the GCM or RCM initial conditions; and 5) downscaling, specifically in the spatial disaggregation. In many cases, downscaling and bias correction can be the main source of uncertainty, with heterogeneity across space, time and climate metrics (Lafferty and Sriver, 2023).

Lafferty and Sriver (2023) compares uncertainty contribution of GCM, scenario, internal variability and downscaling and bias correction at the global scale and suggest that downscaling and bias correction contributes around 30% of total uncertainty for annual and extreme precipitations between 2020 and 2099, with higher contributions in areas with complex topographies. The higher uncertainty in complex topography is caused by atmospheric dynamics being not well represented by GCM due to coarse-resolution, while internal variability contributes around 50 to 60%. On the other hand, downscaling and bias correction contribution to uncertainty for temperatures contributes around 40% in the short term and decreases in time as GCM and scenario contribution increases. This means that choosing a bias correction method is not a simple decision and should be considered carefully to avoid leading to decisions that are not robust as different downscaled and bias correction methods can result in different products and, in consequence, different hazard characterizations (Hanel et al., 2017; Aedo-Quililongo, 2024; Aedo-Quililongo and Chong, 2024). Due to the complexity of these methods and application strategies, there is a need for development of benchmarks and comparison of different methods to discern the most appropriate for each specific condition (Fowler et al., 2025).

Freely available global and regional downscaled datasets prevent users or stakeholders from performing a downscaling and bias correction process in their analysis. Examples of global downscaled products are the Global Downscaled Projections for Climate Impacts Research (GDPCIR; Gergel et al., 2024) or the NASA Globaly Daily Downscaled Projections CMIP6 (NEX-GDDP-CMIP6; Thrasher et al., 2022), and examples of regional downscaled products are the Climate Change Dataset for Brazil (CLIMBra; Ballarin et al., 2023) or the High-resolution climate projection dataset based on CMIP6 for Perú and Ecuador (BASD-CMIP6-PE; Fernandez-Palomino et al., 2024). These products consider global or regional products as surrogate of observations and consider both a specific reference period and strategy for the downscaling and bias correction process. Additionally, even if validated, at local scale, the extent of the validation may not be appropriate for all uses and need to be analyzed with caution (Lanzante et al., 2018; Aedo-Quililongo and Chong, 2024). As such, these downscaled products are a good alternative for a preliminary or regional analysis, while for local studies both the analysis of different methods and the application of downscaling and bias correction must be part of the workflow to reduce the aforementioned biases and uncertainties.

Several efforts have been made to provide users with different versions of the quantile mapping approach for their climate change impacts assessment workflow. The publicly available coding packages for bias correction of climatic variables based on the approach have the following characteristics: 1) the existing coding packages are available mainly in R (MBC, https://github.com/cran/MBC; downscaleR, https://doi.org/10.5281/zenodo.5070432; snowQM, https://doi.org/10.5281/zenodo.10257951), Python (pyCAT, https://github.com/wegener-center/pyCAT; xclim, https://doi.org/10.5281/zenodo.18352200; ibicus, https://doi.org/10.5281/zenodo.8101842) and MATLAB (GCD, https://github.com/thomasmosier/GCD), and a few in Julia (ClimateTools.jl, https://doi.org/10.5281/zenodo.8047681), NCAR programming language (QQmap, https://doi.org/10.5281/zenodo.4009101), Excel (ClimateScenarioAnalysisToolbox, https://github.com/mxgiuliani00/ClimateScenarioAnalysisToolbox), among others; 2) at this moment, none of the explored tools is available in more than one programming language with the exact structure and functionality to allow a unique framework for different user needs or capabilities; 3) the more sophisticated packages are usually frameworks, like the xclim (Bourgault et al., 2023) and downscaleR (Iturbide et al., 2019), that aim for a more complete tool that escapes from the scope of users that only require the downscaling process in their workflow; thus, a deep understanding of the workflow and more structured inputs are needed, which could lead to avoid using that package; and 4) except for the MBC package, which includes several multivariate versions of quantile mapping, but not several trend-preserving bias correction methods, to the best of our knowledge, no package includes more than three contrasting trend preserving bias correction methods to systematically compare their performance in a simple and user-friendly framework, comparison needed to reduce uncertainties regarding bias correction (Fowler et al., 2025). Thus, there is the need for comprehensive coding packages that implement and allow comparing a wide range of quantile mapping methods to facilitate the obtention of local climatic projections and discern the most appropriate one for specific applications.

This paper presents and assesses the climQMBC coding package (Aedo-Quililongo et al., 2026), an easy-to-use tool written in Python, R and MATLAB to perform point-to-point bias correction of climatic variables with the QM, DQM, QDM, UQM and SDM methods for daily, monthly and annual time series. This package aims to make available the QM and several trend-preserving QM methods to facilitate the performance assessment of these methods, making it easier to incorporate the bias correction process as a standard part of climate change impact assessment of final users of local climatic projections that might not be directly related to climate sciences. The climQMBC package also includes a summary report utility to easily compare the performance of the implemented methods at the monthly temporal scale, allowing the user to identify the most appropriate method for their specific situation, along with a synthetic tester of bias correction methods. The synthetic tester develops times series to which bias is added (i.e., in the mean and standard deviation) allowing testing bias correction methods, given that one knows the objective time series before the bias correction.

While application examples for all challenging climatic conditions for bias correction methods are extensive and considering that the climQMBC package has already been used in several studies (Blin, 2023; Martínez et al., 2023; Nguyen et al., 2024; Kassahun et al., 2025; Vicuña et al., 2025; Wang et al., 2025; González-Leiva et al., 2026; Guzmán et al., 2026), for demonstration purposes and as a proposed benchmark, we present an analysis based on the developed synthetic tester to show the capabilities of the climQMBC package to systematically bias correct climatic variables and to show the general performance of the implemented methods. Additionally, we show with a few lines of code how to bias correct real data with the climQMBC package and compare the performance of the implemented methods with the summary report utility.

2 Climate change impact assessment

Climate change assessment at local scales based on projections from climate models (GCMs or RCMs) requires a downscaling and/or bias adjustment/correction process to overcome the differences in resolution and biases of climate models and local climate. The bias corrected climatic variables are then used to assess climate change impacts on water availability, crop productivity, wildfire, or heat waves occurrence, among many others. Figure 1 presents a general workflow of a typical assessment study of climate change impacts, highlighting in grey where the climQMBC package is involved.

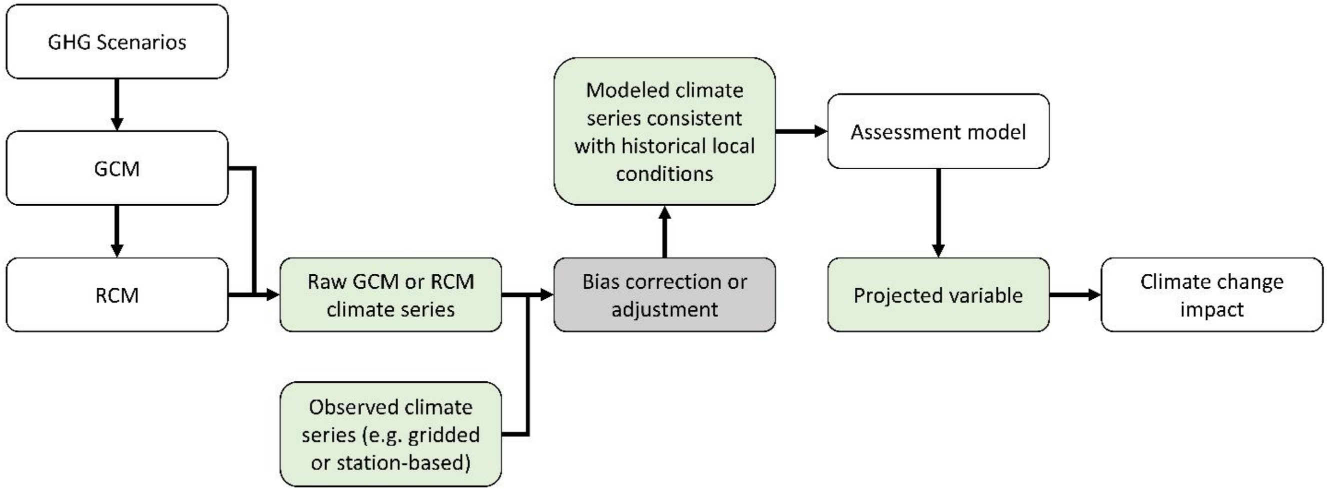


Figure 1: General workflow of a typical assessment study of climate change impacts. Green boxes indicate data or time series. Grey box shows where the climQMBC package is involved in the workflow.

3 climQMBC package

3.1 Package characteristics and functionalities

The climQMBC package is a coding package available for Python, R and MATLAB that contains five point-to-point bias correction methods: QM, DQM, QDM, UQM, and SDM. These methods were implemented for variables associated to both relative (or multiplicative) and absolute (or additive) changes (e.g., precipitations and temperature, respectively). The package supports point or station-based data at daily, monthly or annual time step. The bias correction process and a unified description of methods implemented in the climQMBC package are summarized in the following subsections.

3.1.1 Bias correction process in the climQMBC package

The climQMBC package takes as input (1) an observed and modeled series at one specific point, (2) a flag indicating if the variable allows negative values, (3) a flag indicating if the variable is related to relative or absolute changes (does not apply for the QM method), and (4) a flag indicating if the series frequency is daily, monthly or annual. In the SDM method, a single flag is used to indicate if the variable allows negative values and is related to absolute changes or if the variable does not allow negative values and is related to relative changes, as this method was built explicitly for temperature and precipitation. On the other hand, the modeled series is a single series that merges the historical and future projected periods starting on the same date as the observed series. In the climQMBC package, the trend-preserving bias correction methods capture the transient model changes by a continuum approach using moving windows, as recommended by Chadwick et al. (2018), to capture the non-stationary signals of the GCM. Hence, for each value x of the modeled future period, the climQMBC package considers moving windows with length equal to the number of observed years, ending in the value to be corrected (Fig. 2). The moving window strategy is an improvement implemented in Chadwick et al. (2023), as the original DQM, QDM and SDM consider future data as a single window.

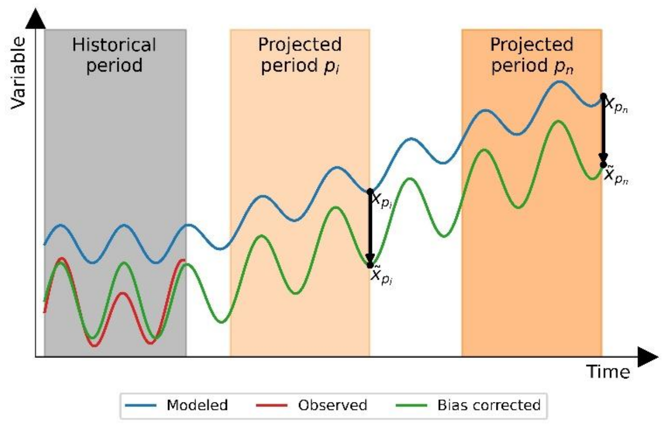


Figure 2: Schematic illustration of the trend-preserving bias correction (i.e., all but the QM) continuum approach considered in the climQMBC package. Each projected period p ends in the value x which is corrected to a value $\tilde{x}$, considering the change rates between the projected and historical period of the modeled series.

As an example of the moving window, if the observed period is 1985-2014, then the moving window will consider 30 years, and when correcting a future year, e.g., 2030, the moving window will cover 2001-2030. Each moving window is used to compute the change rates of the modeled series and the probability distribution function (if applies), and to modify the modeled value x to a corrected value $\tilde{x}$. Periods of 20 to 40 years are recommended (see Appendix A for a brief analysis of the trade-off of selecting larger or smaller windows). Shorter periods might not capture correctly the variable statistics, while longer ones might dampen the signal of the variable or induce a "pseudo-stationary bias" of the transient changes (Chadwick et al., 2023). Note that during the historical period, DQM, QDM and UQM methods become the standard QM method as the change rates are 0 for additive changes and 1 for relative changes (Chadwick et al., 2023).

When the frequency is set to monthly, the bias correction process is performed for each month of the year independently (e.g., each January is corrected based on the information of January in the historical and projected period). If the frequency is set to daily, the bias correction process is performed for each day of the year considering a centered moving window of length defined by the user (e.g., if the window is of 31 days, then the 16th of July uses the days between July 1st to July 31st to compute the required statistics for the bias correction).

The values of the historical period and each projected period are used to fit a probability distribution function. Furthermore, changes in the mean (δ_{μ}), standard deviation (δ_{σ}) or quantiles (Δ), depending on each method, between the projected and historical period of the modeled series are computed with the following equations:

$$\delta_{\mu} = \begin{cases} \dfrac{\overline{x_{m,p}}}{\overline{x_{m,h}}} & \text{, relative change variable} \\ \overline{x_{m,p}} - \overline{x_{m,h}}, & \text{absolute change variable} \end{cases}, \tag{1}$$

$$\delta_{\sigma} = \begin{cases} \dfrac{\widehat{x_{m,p}}}{\widehat{x_{m,h}}} & \text{, relative change variable} \\ \widehat{x_{m,p}} - \widehat{x_{m,h}}, & \text{absolute change variable} \end{cases}, \tag{2}$$

$$\Delta = \begin{cases} \dfrac{F_{m,p}^{-1}(\tau_p)}{F_{m,h}^{-1}(\tau_p)} & \text{, relative change variable} \\ F_{m,p}^{-1}(\tau_p) - F_{m,h}^{-1}(\tau_p) & \text{, absolute change variable (all but SDM),} \\ \dfrac{\widehat{x_{o,h}}}{\widehat{x_{m,h}}}\left(F_{m,p}^{-1}(\tau_p) - F_{m,h}^{-1}(\tau_p)\right), & \text{absolute change variable (SDM)} \end{cases} \tag{3}$$

with

$$\tau_p = F_{m,p}(x), \tag{4}$$

Subindexes o and m refer to the observed and modeled series, respectively, whereas subindexes h and p refer to the historical and projected period, respectively. Accents $\bar{}$ and $\tilde{}$ denote mean and standard deviation, respectively. F and F^{-1} are the cumulative distribution function (CDFs) and its inverse CDF (invCDFs), respectively; and τ is the non-exceedance probability. The climQMBC package assigns a probability distribution function for each day or month and period based on the method of moments and by default selects the distribution with the lowest Kolmogorov-Smirnov test statistic, which is more sensitive to differences in the center of the distribution and less sensitive to differences in the tails. The exception is the SDM method in which a gamma or normal distribution is fitted by the maximum likelihood method for variables associated with precipitation and temperature, respectively (Switanek et al., 2017). The available distribution functions in the climQMBC package are the Normal, Log-normal, Two-parameter Gamma, Three-parameter Gamma, Log-Gamma, Gumbel, and Exponential. If the Kolmogorov-Smirnov test fails for all the available distributions, the climQMBC package prints a warning and continues with the procedure. Additionally, the user can define the desired distribution function to be used, and the package will estimate the parameter with the method of moments, skipping the Kolmogorov-Smirnov test.

For variables that do not allow negative values, the climQMBC package internally filters those distributions that allow negative values (Normal, Three-parameter Gamma, Gumbel and Exponential distributions). On the other hand, if the series has negative values, the climQMBC package discards those distributions that are strictly positive (Log-normal, Two-parameter Gamma and Log-Gamma distributions).

To ensure stability when computing the invCDFs, τ values are constrained to $0.001<\tau<0.999$. Additionally, to improve the management of non-precipitation days, the climQMBC package follows the two parameters “random near zero replacement” presented in Chadwick et al. (2023). The first parameter is a threshold that represents physically zero precipitation values (1 mm as default, independently of the temporal scale. Hence, 1 mm/day, 1 mm/month and 1 mm/year for daily, monthly and annual precipitation, respectively) so that all values below this threshold are replaced with random numbers that are barely positive. The second parameter is a scaling factor (1/100 as default) of the physically zero precipitation values that defines the magnitude of the “barely positive” replacement values. After performing the bias correction methods, the corrected values below the physically zero precipitation values are replaced with zeros.

As an example of the “random near zero replacement” considering the default values in the climQMBC package, all values bellow 1 mm/day (for daily values) would be replaced with random values between 0 and 1/100 mm/day. If as a result of the trend-preserving procedures one of these values is scaled by a factor of 5, the corrected value would be 5/100 mm/day and then would be replaced by a zero value at the end of the process. The two parameters scheme ensure that even for unrealistic changes with factors of 10, the replaced values will still be removed after the correction, avoiding these values to create false rain-day values. On the other hand, if all values are below the physically zero precipitation value (or only zero values), the correction would result in no precipitation values.

As such, this procedure allows fitting distributions and performing bias correction to avoid numerical problems in periods with large number of physically zero precipitation values, or periods with only zero values. Nevertheless, the threshold and the magnitude of the random values should be analyzed to enhance the methods performance depending on the specific conditions.

For daily data, the number of wet days is corrected as most models have a drizzling bias that generates rainy days but with intensity lower than observed (Chen et al., 2021). To correct the number of rainy days, in all methods but SDM, an internal threshold representing physically zero precipitation value for the modeled series is fitted to match the number of rainy days of the observed series in the reference period. Then, only rainy days are corrected (values above the physically zero precipitation threshold), but, if the number of rainy days to fit the distributions is less than 30, some of the near-zero replaced values are considered to complete 30 values (these would later be removed as physically zero precipitation values). The SDM has its own method of correcting rainy days, as described in Switanek et al. (2017).

3.1.2 Description of the bias correction methods implemented in the climQMBC package

The bias correction methods available in the climQMBC package are now briefly described using a unified mathematical notation. Their main objectives and corresponding considerations are summarized in Table 1, while Fig. 3 shows a schematic illustration of the procedure performed in the projected periods and the CDF involved, highlighting the main differences among methods. In terms of execution times, considering the implementation in the climQMBC package, the QM method is the fastest, followed by the DQM (twice the QM execution time for annual and monthly data, and 7 times for daily data), being the QDM, SDM and UQM the slowest (around 7 to 10 times the QM execution time for annual data, 22 to 27 times for monthly data and 32 to 39 times for daily data). Appendix B presents further details of the execution time comparison.

Table 1: Summary of the bias correction methods available in the climQMBC package

Method	Abbreviation	Summary	Reference
Quantile Mapping	QM	Considers a transfer function built with information from the observed and modeled historical period and applies it to the modeled series. Corrected series values will match its CDF with the observed series, maintaining the model occurrence of the events and limiting future scenarios exceeding this range.	Wood et al. (2002)
Detrended Quantile Mapping	DQM	Considers the same transfer function as in the QM method but applies it to a detrended modeled series, to then reimpose the trends. Corrected series tends to preserve changes in the mean of the modeled series.	Bürger et al. (2013)
Quantile Delta Mapping	QDM	The QM method is applied to the modeled future to detrend the modeled values. Then, the change between the future value and its respective quantile in the modeled historical period is used to rescale the series. Corrected series tends to preserve changes in the quantiles of the modeled series.	Cannon et al. (2015)
Unbiased Quantile Mapping	UQM	The QM method is applied to the modeled future, and the changes of the moments (e.g. mean and standard deviation) are imposed in the observed CFD to scale the transfer function. Corrected series tends to preserve changes in the moments of the modeled series.	Chadwick et al. (2023)
Scaled Distribution Mapping	SDM	Scales the observed CDF by the ratio between the future and historical modeled CDFs to correct the recurrence interval. A scaling factor like in the QDM method is then applied to the scaled probability distribution. Finally, the corrected values are allocated accordingly to likelihood of events. For precipitation, the SDM method additionally adjusts rain-day frequency. Corrected series tends to preserve changes in quantiles (as in the QDM method) and likelihood of events.	Switanek et al. (2017)

Note: In the DQM, QDM, UQM and SDM methods the future period is divided in moving windows as an improvement implemented in Chadwick et al. (2023) to prevent pseudo-stationary biases.

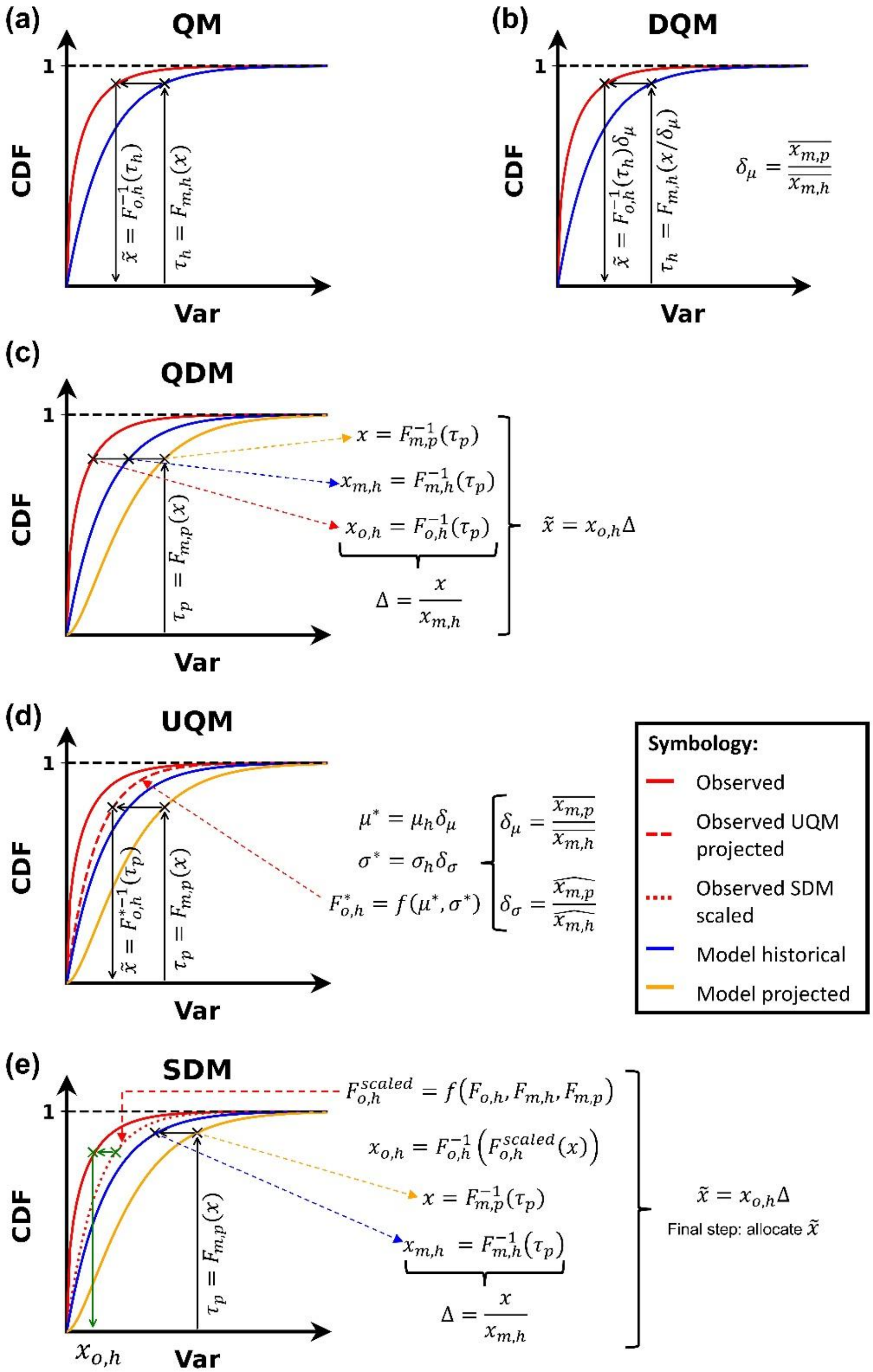


Figure 3: Schematic illustration of the main processes computed in the (a) QM, (b) DQM, (c) QDM, (d) UQM and (e) SDM methods available in the climQMBC package, at each projected period for variables related to relative changes (e.g. precipitation). For variables related to absolute changes (e.g. temperatures) the deltas (*δ* and *Δ*) are computed and applied additively, rather than multiplicatively.

Quantile Mapping (QM):

The QM method (Fig. 3a) bias corrects the modeled series by evaluating the modeled values (x) in the CDF of the modeled historical period ($F_{m,h}$) to get the associated probabilities of each value and then evaluates them in the invCDF of the observed historical period ($F_{o,h}^{-1}$) to obtain a vector of corrected values ($\tilde{x}$). This method aims to correct each value of the modeled series by matching its quantile computed on its historical period with the quantile of the observed values, and is formulated as:

$$\tilde{x} = F_{o,h}^{-1}\big(F_{m,h}(x)\big), \tag{5}$$

The QM method assumes that the correction process is time-invariant and in some cases may result in distortion of the raw model projected change rates. In most cases distortions are caused by projected climate being outside the range of historical modeled climate.

Detrended Quantile Mapping (DQM):

The DQM method (Fig. 3b) detrends the modeled series by computing the projected change in the mean of the modeled series for each rolling window (δ_μ, Eq. 1) and scaling the last value of each rolling window (x) to a range within the historical values. Then, the same transfer function described in the QM method is applied to the detrended values. Finally, the bias corrected values are rescaled multiplicatively or additively to obtain the corrected value ($\tilde{x}$). The procedure for each projected period is formulated as follows:

$$\tilde{x} = \begin{cases} F_{o,h}^{-1}\left(F_{m,h}(x/\delta_\mu)\right)\delta_\mu\,, & \text{relative change variable} \\ F_{o,h}^{-1}\left(F_{m,h}(x-\delta_\mu)\right)+\delta_\mu\,, & \text{absolute change variable} \end{cases}, \tag{6}$$

Quantile Delta Mapping (QDM):

In the QDM method (Fig. 3c) the non-exceedance probability τ_p (Eq. 4) of each value x to be corrected is calculated based on its associated moving window. Then, τ_p is evaluated in the observed invCDF to obtain the modeled projected values after correction, $x_{m,p\to o,h}$, as follows:

$$x_{m,p\to o,h} = F_{o,h}^{-1}(\tau_p), \tag{7}$$

given that τ_p depends on a moving window, $x_{m,p\to o,h}$ does not have the model trend. Finally, the percentile change Δ (Eq. 3) of the projections is estimated as the ratio between the projected climate and its respective quantile in the historical modeled period. Then, Δ is applied to $x_{m,p\to o,h}$ to obtain the corrected value $\tilde{x}$ as follows:

$$\tilde{x} = \begin{cases} x_{m,p\to o,h}\Delta\,, & \text{relative change variable} \\ x_{m,p\to o,h}+\Delta\,, & \text{absolute change variable'} \end{cases} \tag{8}$$

When Δ is estimated for relative changes, it can get indetermined as a result of a division by a near-zero value. To avoid this, the climQMBC package considers an upper limit for Δ when dividing by a physically no-precipitation value. This strategy mimics the implementation of the QDM method in the MBC package (Multivariate Bias Correction of Climate Model Outputs; Cannon, 2017), available in CRAN.

Unbiased Quantile Mapping (UQM):

In the UQM method (Fig. 3d) the projected change in the moments (i.e., mean and standard deviation) of the modeled series is first computed for each moving window (δ_μ and δ_σ, respectively; Eq. 2 and Eq. 3, respectively). Then, these change rates are used to project the observed historical mean (μ_h) and standard deviation (σ_h) into the future to obtain future values μ^* and σ^* as follows:

$$\mu^* = \begin{cases} \mu_h \delta_\mu, & \text{relative change variable} \\ \mu_h + \delta_\mu, & \text{absolute change variable'} \end{cases} \tag{9}$$

$$\sigma^* = \begin{cases} \sigma_h \delta_\sigma, & \text{relative change variable} \\ \sigma_h + \delta_\sigma, & \text{absolute change variable'} \end{cases} \tag{10}$$

Subsequently, μ^* and σ^* are used to produce future projections of the selected distribution for the historical observed values, $F^*_{o,h}$, as follows:

$$F^*_{o,h} = f(\mu^*, \sigma^*), \tag{11}$$

Finally, the non-exceedance probability of the modeled values x in each moving window (τ_p, Eq. 4) are evaluated in the inverse projected distribution invCDF, $F^{*-1}_{o,h}$, to obtain the bias corrected value $\tilde{x}$ as follows:

$$\tilde{x} = F^{*-1}_{o,h}(\tau_p), \tag{12}$$

Scaled Distribution Mapping (SDM):

The SDM method (Fig. 3e) considers several steps (Switanek et al., 2017; Chadwick et al., 2023). Note that they are presented in a sequence different than the originally depicted by Switanek et al. (2017), with the intention to better illustrate the implications and structure of the method.

For temperatures, the first step is to linearly detrend both the observed and modeled series for each period. For precipitation, the first step is to remove all non-precipitation values (1 mm as default, independently of the temporal scale. Hence, 1 mm/day, 1 mm/month and 1 mm/year for daily, monthly and annual precipitation, respectively) to account only for the rain values in each period. Then, the expected number of rain events for the corrected series (RD^{scaled}) is calculated by scaling multiplicatively (linear scaling) the number of rain events of each moving window ($RD_{m,p}$) with the ratio between the frequency of rainy days in the observed historical period ($RD_{o,h}/TD_{o,h}$) and in the modeled historical period ($RD_{m,h}/TD_{m,h}$) as follows:

$$RD^{scaled} = RD_{m,p} \times \frac{RD_{o,h}/TD_{o,h}}{RD_{m,h}/TD_{m,h}}, \tag{13}$$

where TD represents the total number of days, including rain and non-rain days. Hereafter, the SDM method is applied to a vector (bold notation) of sorted detrended temperature or non-zero precipitation values, to which the normal and gamma probability distribution functions are fitted by the maximum likelihood method, respectively.

In the second step the sorted vectors of the observed historical ($\boldsymbol{x}_{o,h}$), modeled historical ($\boldsymbol{x}_{m,h}$) and modeled projected period or moving window ($\boldsymbol{x}_{m,p}$) series are evaluated in their respective fitted CDF to obtain their associated probability vectors ($F_{o,h}(\boldsymbol{x}_{o,h})$, $F_{m,h}(\boldsymbol{x}_{m,h})$, $F_{m,p}(\boldsymbol{x}_{m,p})$). In the climQMBC package, each CDF is fitted considering a gamma distribution for positive precipitation and normal distribution for detrended temperature, using the maximum likelihood, following the SDM implementation in the pyCAT package. These probability vectors are used to estimate each recurrence intervals $\boldsymbol{RI}_{i,j}$ as follows:

$$\boldsymbol{RI}_{i,j} = \begin{cases} \frac{1}{1-F_{i,j}(x_{i,j})} & \text{, for precipitation} \\ \frac{1}{0.5-|0.5-F_{i,j}(x_{i,j})|}, & \text{for temperature} \end{cases} \quad ; \quad i = \{m, o\}, j = \{h, p\}\ , \tag{14}$$

In the third step, the recurrence interval of the observed historical series is linearly scaled by the ratio of the recurrence intervals of the model projected period and the model historical period (i.e., capturing the changes in recurrence intervals in the model) to build a scaled recurrence interval $\boldsymbol{RI}^{scaled}$ as follows:

$$\boldsymbol{RI}^{scaled} = \max\left(1, \boldsymbol{RI}_{o,h} \circ \boldsymbol{RI}_{m,p} \circ \frac{1}{\boldsymbol{RI}_{m,h}}\right)\ , \tag{15}$$

$\boldsymbol{RI}^{scaled}$ is built using the Hadamard product for vectors (i.e., multiplying element by element in a vector and preserving their dimensions), which requires equally sized vectors. Hence, prior to computing the precipitation recurrence intervals, the probability vectors are linearly interpolated to match the number of rain day events of the modeled projected period. The linear interpolation stretches or contracts the series to match the number of values in the modeled vector, keeping the original range and distribution of values.

In the fourth step, $\boldsymbol{RI}^{scaled}$ is used to go back into a scaled probability vector $\boldsymbol{F}_{o,h}^{scaled}$, which represents the scaled distribution function vector of the observed historical period as follows:

$$\boldsymbol{F}_{o,h}^{scaled} = \begin{cases} 1 - \frac{1}{\boldsymbol{RI}^{scaled}} & \text{, for precipitation} \\ \left(1 - \frac{1}{\boldsymbol{RI}^{scaled}}\right) \times sgn\left(F_{o,h}(\boldsymbol{x}_{o,h}) - 0.5\right) & \text{, for temperature} \end{cases}, \tag{16}$$

In the fifth step, $\boldsymbol{F}_{o,h}^{scaled}$ is sorted in descending order for precipitation, then, for both precipitation and temperature, it is evaluated in the invCDF of the observed historical period to get the associated value in the historical observed series domain $\boldsymbol{x}_{scaled\rightarrow o,h}$ as follows:

$$\boldsymbol{x}_{scaled\rightarrow o,h} = F_{o,h}^{-1}\left(\boldsymbol{F}_{o,h}^{scaled}\right), \tag{17}$$

In the sixth step, the vector with the projected change in the quantiles $\boldsymbol{\Delta}$ (Eq. 3. for SDM; similar to that of the QDM) is computed for all the values of the projected series and it is used to scale the vector computed in the previous step as follows:

$$\widetilde{\boldsymbol{x}}_{temp} = \begin{cases} \boldsymbol{x}_{scaled\rightarrow o,h}\boldsymbol{\Delta}, & \text{for precipitation} \\ \boldsymbol{x}_{scaled\rightarrow o,h} + \boldsymbol{\Delta}, & \text{for temperature} \end{cases}, \tag{18}$$

For precipitation, if the number of rainy days of the projected period ($RD_{m,p}$) is greater than the expected number of rain events (RD^{scaled}), the frequency of rain events is adjusted by linearly interpolating $\widetilde{\boldsymbol{x}}_{temp}$ to a length equal to RD^{scaled}.

As a final step for precipitation, each value of the $\widetilde{\boldsymbol{x}}_{temp}$ is placed in the location of the rainy values of the modeled projected temporal series, from the highest to the lowest. In case where RD^{scaled} is greater than $RD_{m,p}$, the remaining values are replaced with zeros. For temperatures the final step is to place back the $\widetilde{\boldsymbol{x}}_{temp}$ values in the correct temporal location based on the modeled time series, process that is done in descending order. Then, the linear trend of the modeled projected temporal series is added back. Because the climQMBC package considers a continuum approach to capture transient changes from the models, the corrected value $\tilde{x}$ of the projected period corresponds to the value of the last year of the corrected series; then, the complete procedure is repeated for the following year.

3.2 climQMBC package example and report function

This section is based on an example that uses sample data available in the package repository (https://github.com/saedoquililongo/climQMBC/tree/main/Sample_data). The sample data considers daily precipitation and mean temperature for the coordinate 4.64N and 75.48W, located near the Cocora Valley, Quindío, Colombia. The observational ERA5-Land reanalysis series was developed by the European Centre for Medium-Range Weather Forecasts (Muñoz-Sabater et al., 2021) and for the modeled series considers the GCM MPI-ESM1-2-HR (Müller et al., 2018), under the SSP 5-8.5 scenario of the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (Eyring et al., 2016). The observational and modeled time series were extracted with the nearest point method.

As an easy-to-use tool, the climQMBC package requires familiarity with importing and processing arrays, matrices, or vectors. Four steps are needed for its use: (1) Import the climQMBC package and its functions; (2) load the observed and modeled series (the package assumes that both series begin in the same date and that the reference period corresponds to the period of the observed series); (3) define the type of variable (i.e. negative or strictly non-negative, and variable associated with additive or multiplicative changes) and data frequency (daily, monthly or annually); and (4) apply the selected bias correction method or methods. These steps can be applied to multiple locations or grids cells to systematically downscale a region of interest (this example is included in the climQMBC tester scripts available in GitHub for Python, R and MATLAB at https://github.com/saedoquililongo/climQMBC/blob/main/Python/climQMBC_tester_Python.py, https://github.com/saedoquililongo/climQMBC/blob/main/R/climQMBC_tester_R.R and https://github.com/saedoquililongo/climQMBC/blob/main/Matlab/climQMBC_tester_Matlab.m, respectively). Also, different bias correction methods can be used and compared to identify the better ones in preserving the changes from raw GCM/RCMs for the specific conditions under study. Detailed demonstration of the capabilities of the package is provided in an example notebook available in its repository (https://github.com/saedoquililongo/climQMBC/blob/main/Example_notebook.ipynb).

Listing 1 illustrates a simple Python code to apply the five bias correction methods available in the climQMBC package to temperature sample data available in the repository and then plot the QM and UQM results. Figure 4 presents the plots of Listing 1 and a close-up to the reference period and to the period when QM and UQM diverge, showing how the corrected series match the magnitude and variability of the observed series (Fig. 4b) and the potential differences in their results (Fig. 4c) which could lead to bias in a further climate change impacts assessment. More specifically, Fig. 4a shows the inability of the QM method to capture trends when the projected values are outside the range of the values of the reference period.

```
# 1. Import the climQMBC package and its functions
from climQMBC.methods import QM, DQM, QDM, UQM, SDM

# 2. Load the observed and modeled series
import pandas as pd
obs = pd.read_csv('sample_data/obs_tas_M.csv', index_col=0, parse_dates=True)
mod = pd.read_csv('sample_data/mod_tas_M.csv', index_col=0, parse_dates=True)

## Aggregate to annual scale and format to column vector
obs = obs[['tas']].resample('YE').mean().values
mod = mod[['tas']].resample('YE').mean().values

# 3. Define the type of variable (temperature)
frq = 'A'                # Annual data
mult_change = 0          # Additive changes
allow_negatives = 1      # Allow negative values
SDM_var = 0              # Temperature variable for SDM

# 4. Apply the selected bias correction method or methods
qm_series  = QM(obs, mod, allow_negatives=allow_negatives, frq=frq)
dqm_series = DQM(obs, mod, mult_change=mult_change, allow_negatives=allow_negatives, frq=frq)
qdm_series = QDM(obs, mod, mult_change=mult_change, allow_negatives=allow_negatives, frq=frq)
uqm_series = UQM(obs, mod, mult_change=mult_change, allow_negatives=allow_negatives, frq=frq)
sdm_series = SDM(obs, mod, SDM_var=SDM_var, frq=frq)

# 5. Plot time series
import matplotlib.pylab as plt
plt.figure()
plt.title('Example plot of the time series')
plt.plot(mod, label='Modeled')
plt.plot(qm_series, label='QM')
plt.plot(uqm_series, label='UQM')
plt.plot(obs, label='Observed')
plt.xlabel('Years since starting date')
plt.ylabel('°C')
plt.legend(loc='upper left')
plt.grid()
plt.show()
```

Listing 1: Python script showing the usage of the climQMBC package.

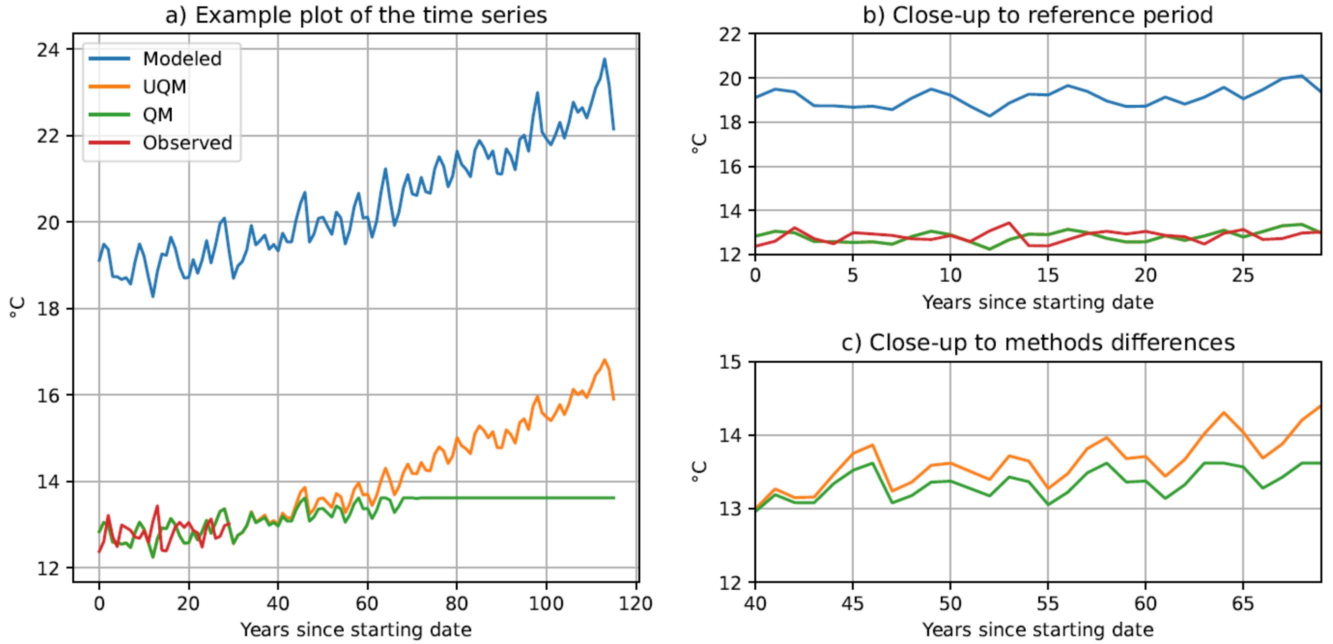


Figure 4: Example plot of the example Python code presented in Listing 1. a) Complete time series; b) close up to the reference period (note that, in the reference period, UQM is the same as QM); and c) close-up to the period when UQM and QM diverge.

For monthly values, the climQMBC package has a summary report function to compare the performance or assess trade-offs between bias correction methods based on the preservation of the changes in the mean and standard deviation of the modeled series. This function outputs two reports: (1) a summary table with the overall performance of the bias correction methods in the historical period and in future periods for the mean and standard deviation (Fig. 5), and (2) a set of figures to visualize the main differences among methods in terms of monthly values and quantiles (Fig. 6 and Fig. 7). The report function requires the same inputs as the bias correction methods functions. Additional inputs of the report functions are the bias correction methods and future periods to be displayed in the summary tables and figures.

The structure of the summary tables for the mean and standard deviation generated by the report function of the climQMBC package are the same (Fig. 5). The first row of the summary tables shows the ratio or difference in the mean value (or standard deviation) between the QM bias corrected and observed historical series. A perfect match between the corrected and observed value would result in a value of 1 for multiplicative change variables and 0 for the additive change variables. In the example shown in Fig. 5a (Mean Table) for precipitation, the ratio between the annual mean values of the QM method and the modeled series in the historical period is 1.000, showing a perfect match between the mean of the observations and corrected series in the reference period.

Mean Table:

Ratio between the future and historical period

Ratio between QM and Observed

QM performance in historical period: 1.000			
	Future	2030	2060
Modeled	0.920	0.975	0.892
QM	0.947	0.993	0.922
DQM	0.933	0.992	0.903
QDM	0.933	0.990	0.905
UQM	0.931	0.991	0.900
SDM	0.923	0.991	0.889

Modeled (GCM)

Quantile mapping methods

Ratio between the projected periods and the historical period

Standard deviation Table:

QM performance in historical period: 1.004			
	Future	2030	2060
Modeled	0.916	0.951	0.888
QM	1.046	1.045	1.038
DQM	1.077	1.042	1.081
QDM	1.063	1.042	1.044
UQM	1.081	1.048	1.083
SDM	1.117	1.053	1.128

Figure 5: Example of the summary tables generated by the report function of the climQMBC package, applied to a precipitation case. The Mean and Standard Deviation Tables show the summary report for the annual mean and standard deviation, respectively. First row presents the performance based on the ratio between the annual mean of the bias corrected and modeled series in the historical period. Second row presents headers for different periods. Third row and below presents the ratio between the periods and the historical period for each series presented in the first column. Values under the header 'Future, shows the ratio between the complete future period and the historical period, the third and following columns show the ratio between the projected period, centered in the year of the corresponding header (defined by the user), and the historical period.

The second row of the summary table shows the header of the periods that are compared with the historical period. This comparison considers information exclusively of each series (e.g. projected period and historical information of the UQM corrected series). The third row and below shows the ratio or difference between the future period, based on the header, and the historical period. In these rows the first column indicates the series considered (modeled series, QM corrected series, DQM corrected series, …), the second column, under the header 'Future, shows the ratio or difference between the complete future period and the historical period, the third and following columns show the ratio or difference between the projected period, centered in the year of the corresponding header (defined by the user), and the historical period. For a bias correction model to be good, its value (ratio or difference) should be close to the value of the "Modeled" row in each period, both for the mean and standard deviation. For the example shown in Fig. 5b (Standard deviation Table), the model ratio between the future centered in 2060 and historical is 0.888 (a change of -11.2 %) whereas the corresponding values for QM, DQM, QDM, UQM and SDM are 1.038, 1.081, 1.044, 1.083 and 1.128, respectively. Hence, all the methods show an increase in the standard deviation, rather than a decrease, being the SDM method with the greater discrepancy.. However, for the same future period, the ratio corresponding to changes in the mean for the SDM method is the closest to the modeled series (0.889 compared to 0.892, Mean Table).

The report function also provides a second report with three different figures illustrated in Fig. 6 and Fig. 7. The first figure (Fig. 6) shows the cumulative distribution of the modeled and corrected series in the historical and future periods, and the cumulative distribution of the observed series (left side). This figure also shows the observed, modeled, and corrected time series (right side). The second figure (Fig. 7a) shows the seasonal monthly mean of each series in the historical and future projected periods (the same projected periods as in the summary tables). Finally, the third figure (Fig. 7b) shows the same monthly seasonality, but for the standard deviation. In the second and third figures, the objective series is the observed monthly mean or standard deviation, scaled by the raw projected change (scaling factor based on the projected and the historical period) of the uncorrected model. Note that, ideally, the corrected value should capture the projected changes; hence, match the objective value.

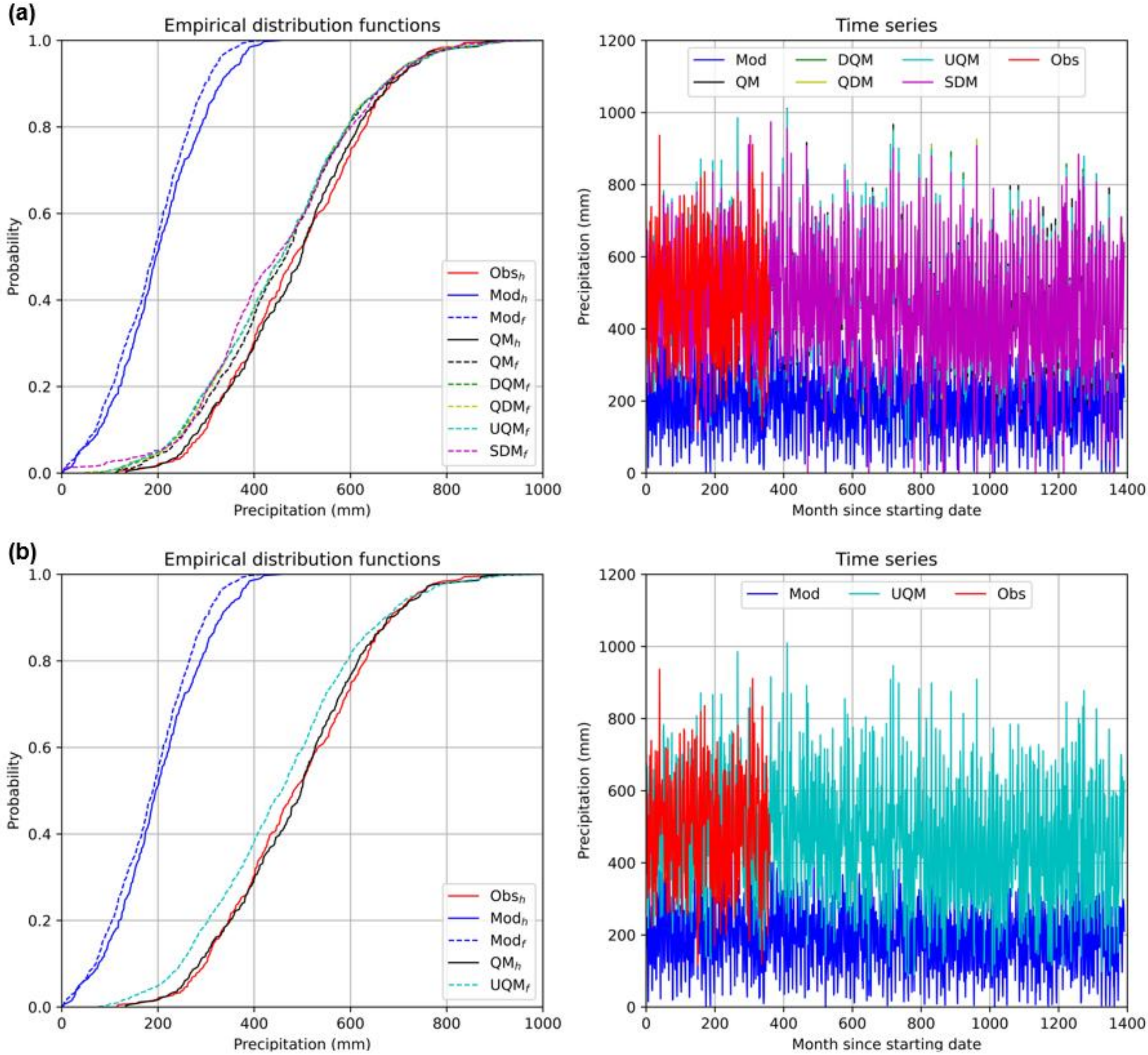


Figure 6: Example of the summary figures generated by the report function of the climQMBC package: cumulative distribution functions in the historical and future periods (left side) and time series (right side). (a) Default report function option displaying all methods at once. (b) Custom view using the optional argument fun=['UQM'] to display only the UQM method.

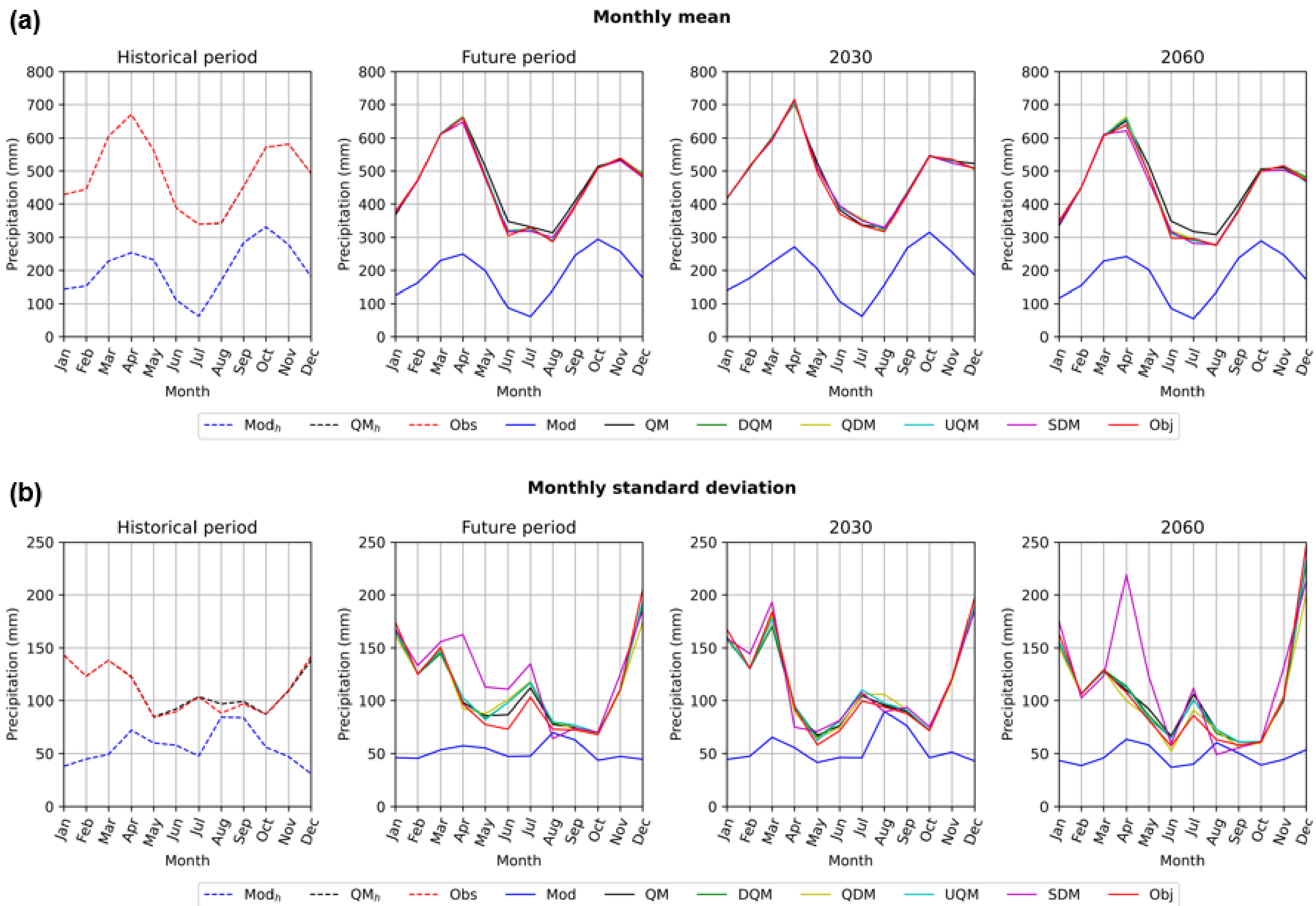


Figure 7: Example of the summary figures generated by the report function of the climQMBC package: (a) summary plots of the seasonal monthly mean, (b) summary plots of the seasonal monthly standard deviation. To display the user defined periods, the optional arguments y_init=1985 and y_wind=[2030, 2060] were set to define 1985 as initial year and to display the values in the windows centered in 2030 and 2060.

4 Application examples with synthetic series

In this section we assess the capabilities of the climQMBC package to systematically bias correct climatic variables of GCMs or RCMs. We compared the performance of the implemented bias correction methods in both the historical and projected periods based on the synthetic tester that is available as a function in the climQMBC package. A controlled bias was introduced into a synthetic time series, enabling a clear evaluation of each method-s ability to correct it. This example is designed to demonstrate the performance of bias correction methods under varying conditions, offering insights into their optimal use. Nevertheless, determining the most suitable method for each specific case lies beyond the scope of this study. The example highlights the importance of having an easy-to-use tool, such as the climQMBC package, to implement bias correction methods and discern the most adequate one to a specific case.

4.1 Description of the synthetic example and evaluation

Extending the synthetic example presented by Maurer and Pierce (2014), Cannon et al. (2015), and Chadwick et al. (2023), gamma distributions were used to generate multiple synthetic precipitation time series. The gamma probability density function with shape parameter k and scale parameter θ is given by:

$$f(x;k,\theta)=\frac{x^{k-1}\exp(-x/\theta)}{\theta^{k}\Gamma(k)},\quad x>0,\quad k>0,\quad \theta>0, \tag{19}$$

where $\Gamma(\cdot)$ is the gamma function. Using the method of moments the mean (μ) and standard deviation (σ) can be used to estimate the parameters with $\mu = k\theta$ and $\sigma = \sqrt{k}\theta$.

The example considers four observed historical series generated with 10,000 values, mean of μ_h=1,000 and coefficients of variation (CV) ranging from 0.25 to 1.00 (Table 2). The future period is generated using a gamma distribution with 49 possible changes (i.e., the combination of $\Delta\mu$ and $\Delta\sigma$ in Table 2) from the historical series. Note that the projected or future mean (μ^*) and future standard deviation (σ^*) are obtained from $\mu^* = \mu(1+\Delta\mu)$ and $\sigma^* = \sigma(1+\Delta\sigma)$, respectively. Then, bias in the mean and standard deviation were added ($\Delta\mu_{bias}$ and $\Delta\sigma_{bias}$, respectively), following Eq. (20). These range from -50 % to +50 % (Table 2) for both the observed historical and future objective series, resulting in 49 combinations of bias. The same bias is added in the historical and in the future periods.

To obtain the bias values (x_i^b), the generated historical or future series (x_i^o) goes through Eq. (20). Here, one can choose the specific bias being added in the mean ($\Delta\mu_{bias}$) and standard deviation ($\Delta\sigma_{bias}$) independently, while preserving the original sequence of events.

$$x_i^b = \mu^o(1+\Delta\mu_{bias}) + (x_i^o - \mu^o)(1+\Delta\sigma_{bias})\,, \tag{20}$$

where μ^o is the mean of the observed historical or future objective series. As such, in this synthetic exercise the bias corrected series, if perfectly done, should be the same as the synthetic observed historical or future objective series.

Table 2: Coefficient of variation, projected changes in the mean and standard deviation, and bias in the mean and standard deviation considered in the synthetic example.

Variable	Values
Coefficient of Variation (CV)	0.25, 0.50, 0.75, 1.00
Projected change in mean ($\Delta\mu$)	-50 %, -25 %, -10 %, 0 %, 10 %, 25 %, 50 %
Projected change in standard deviation ($\Delta\sigma$)	-50 %, -25 %, -10 %, 0 %, 10 %, 25 %, 50 %
Bias in mean ($\Delta\mu_{bias}$)	-50 %, -25 %, -10 %, 0 %, 10 %, 25 %, 50 %
Bias in standard deviation ($\Delta\sigma_{bias}$)	-50 %, -25 %, -10 %, 0 %, 10 %, 25 %, 50 %

For the bias correction process, we evaluated all the implemented bias correction methods in the climQMBC package. The bias correction uses three of the four set of series, both modeled series (projected and historical) and the historical observed, while the projected observed is kept for verification purposes, given that it is the "true series". In terms of implementation, as the climQMBC package considers moving windows and this example does not consider progressive changes in the variables, the biased series considered a duplicated future period so that the correction of the second future period neglects the information of the historical period. This means that the continuous series are the concatenation of the historical period, the future period (first copy) and the future period (second copy), finally evaluating the historical and second future period.

To assess the performance of each method, we considered (1) performance metrics based on statistical properties, measured as the ratio between statistics of the bias corrected series and those of the expected series, for which a value of 1 means a perfect agreement, and 2) goodness-of-fit metrics (i.e. Nash-Sutcliffe efficiency (NSE) proposed by Nash and Sutcliffe (1970), and Kling-Gupta efficiency (KGE) proposed by Gupta et al. (2009)) to evaluate the point-by-point agreement between the bias corrected and expected series. Values of 1 for NSE and KGE indicate perfect agreement. These metrics, their range and possible values are summarized in Table 3. Note that for this application, where the biased series share the exact sequence of events as its original series, the NSE and KGE metrics are appropriate, while this usually does not happen in real comparisons between GCM and observations. As such, these metrics are not proposed as metrics to evaluate bias correction performance in general, but only when the proposed synthetic tester is used.

Table 3: Performance metrics considered.

Variable	Range	Values
Ratio between the corrected and expected mean, standard deviation and percentiles[1]	[0, ∞)	<1: Underestimation =1: Perfect agreement >1: Overestimation
Ratio between the corrected and expected skewness[2]	(-∞, ∞)	<1: Underestimation =1: Perfect agreement >1: Overestimation
Nash-Sutcliffe efficiency (NSE; Nash and Sutcliffe, 1970)	(-∞, 1]	<1: Bias among values =1: Perfect agreement among values <0: Worse than the mean
Kling-Gupta efficiency (KGE; Gupta et al., 2009)	(-∞, 1]	<1: Bias among values, mean or standard deviation =1: Perfect correlation between values and perfect fit among mean and standard deviation <1-$\sqrt{2}$: Worse than the mean (Knoben et al, 2019)

1 Percentiles 5 %, 10 %, 50 %, 90 %, and 95 % are considered

2 Negative values indicate change in the direction of the distribution asymmetry

4.2 Bias correction comparison in the historical period

Regarding the performance of the bias correction methods in the historical period, Fig. 8 shows the NSE values for each combination of coefficient of variation (CV), bias in mean ($\Delta\mu_{bias}$) and bias in standard deviation ($\Delta\sigma_{bias}$). Despite the synthetic series were generated with a strictly positive distribution function, the procedure to add bias result in some negative values. These values were replaced with random values below 0.01, according to the implementation of the climQMBC package. The numbers in the grid of Fig. 8 indicate the percentage of negative values in the biased synthetic time series resulting from the procedure to add bias. Figure 8 shows only values for the QM and SDM methods because the DQM, QDM and UQM methods are equivalent to the QM method in the historical period (Chadwick et al., 2023).

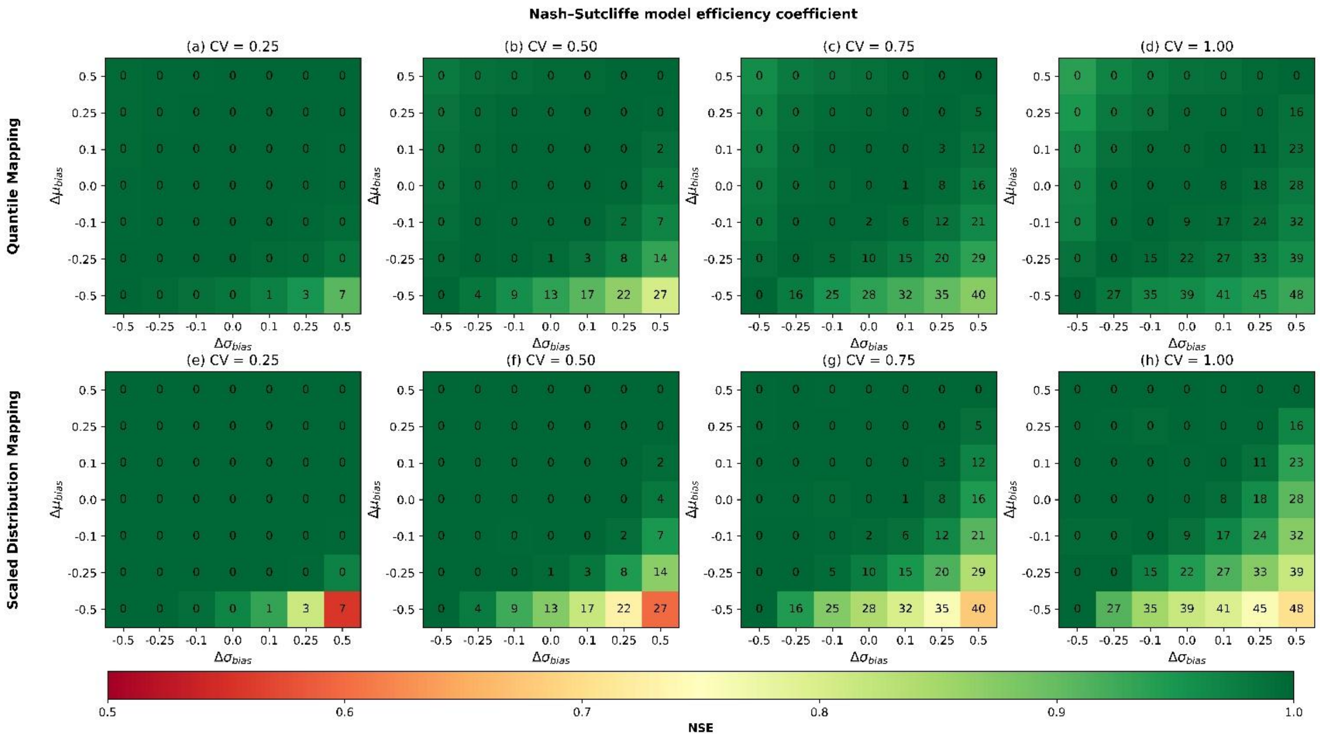


Figure 8: NSE values of the QM and SDM methods in the historical period for different combinations of the observations' CV, $\Delta\mu_{bias}$ and $\Delta\sigma_{bias}$. Green color indicates a perfect fit between the corrected and observed values (NSE ~ 1), and red tones indicate a lower quality fit (NSE < 1). The numbers in the grid indicate the percentage of negative values in the biased synthetic time series resulting from the procedure to add bias (these negative values were replaced with random values below 0.01).

Based on the NSE values (Fig. 8), the QM and SDM methods may fit differently to the observed values under different circumstances, with the QM method being the one with better fit when the bias in the standard deviation increases and the bias in the mean decreases to negative values (bottom right corner of each grid). On the other hand, when the bias in the mean increases and the bias in the standard deviation decreases to negative values (top left corner of each grid), NSE values for the

QM method tend to slightly decrease for high values of CV, tendency not observed for the SDM. Both methods perform well in terms of NSE values when the bias in the mean and standard deviation are similar. When performance is analyzed based on the mean, standard deviation, skewness, and percentiles, or based on the KGE values, results showed similar patterns (results not shown).

Figure 9 presents boxplots of the performance metrics shown in Table 3 for the ratio between the corrected and observed mean, standard deviation, skewness, percentiles 5th, 10th, 50th, 90th and 95th, alongside the NSE and KGE values in the historical period. The boxplots consider the results after adopting all the CV and bias conditions depicted in Table 2. For all plots in Fig. 9, a value of 1 indicates a perfect agreement between the corrected and observed statistics. Figure 9 shows only values for the QM and SDM methods because the DQM, QDM and UQM methods are equivalent to the QM method in the historical period (Chadwick et al., 2023).

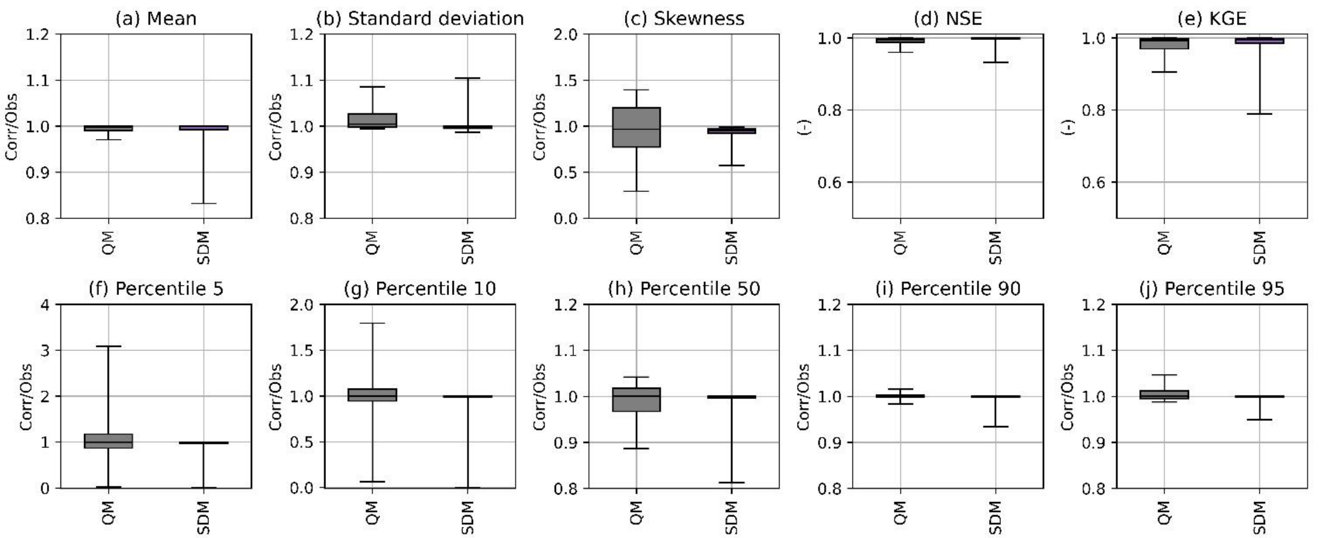


Figure 9: Performance comparison of QM and SDM in the historical period using the ratio between the bias corrected and observed mean (a), standard deviation (b), skewness (c), and percentiles (f to j), in addition to the NSE (d) and KGE (e) values. For all metrics, a value of 1.0 indicates a perfect agreement between the corrected and observed values, as depicted in Table 3.

In all cases, both the QM and SDM have boxes centered in a value of 1.0. Although more stable in terms of a more constrain box (representing the middle 50 % of the data) than the QM method, the SDM method resulted also in more outliers. For the mean, standard deviation, NSE, KGE, percentiles 50 %, 90 % and 95 %, the SDM shows a smaller box than the QM, but the SDM have larger whiskers. For the skewness, the SDM shows a smaller box and whiskers than the QM, and for percentiles 5 % and 10 %, the SDM shows a smaller box than the QM, and the QM shows whiskers for over and underestimation, while SDM only for underestimation. Hence, if the results of the SDM method are not properly analyzed, they may result in larger biases than QM. Having a tool to easily compute and compare among bias correction methods allows the analysis and facilitate the selection of the most adequate method depending on the case study under analysis.

4.3 Projected scenarios

Regarding the performance of the bias correction methods in the future, Fig. 10 shows the cumulative probability of the NSE and KGE values, resulting from the synthetic exercise considering all the combinations of CV, added biases and projected changes shown in Table 2. The desired result, indicating that almost all cases have a perfect agreement with the expected values, is to have cumulative probabilities as close as possible to the value of 1.0; hence curves whose mass is concentrated near NSE = 1 (that is, shifted toward the right edge of the plot) have better performance.

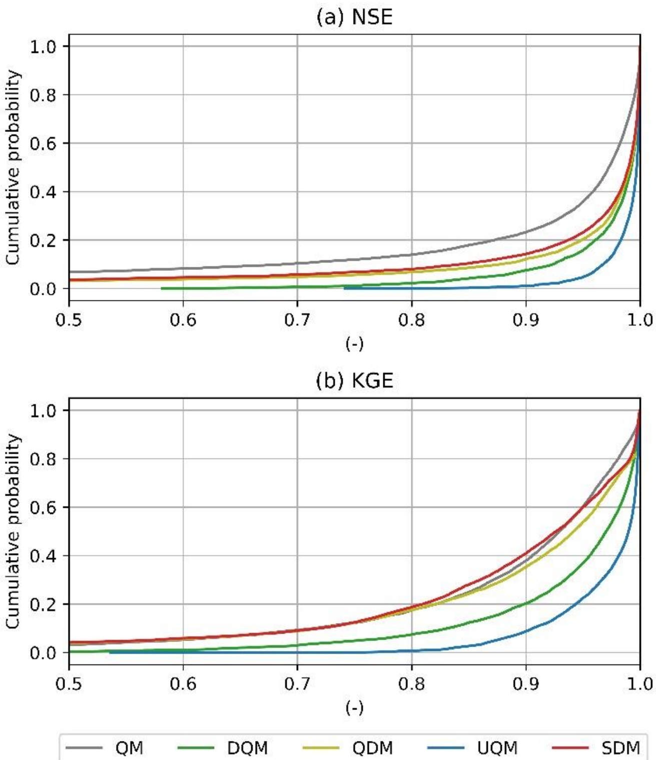


Figure 10: Cumulative probability of the NSE (a), and KGE (b) values of the synthetic example for each implemented method in the climQMBC package.

Based on the NSE and KGE values, the QM, DQM, QDM, UQM and SDM methods result in different series which may differ in their performance. Results using the synthetic tester example, which consist of adding a bias in the two first statistical moments (i.e. the mean and standard deviation), show that the UQM method performs best in eliminating the bias from the modeled series, followed by the DQM method, being the QM method the one with the lowest performance quality. The SDM and QDM methods perform similarly in terms of the NSE values, but the SDM performs worse when evaluated with the KGE. The performance of a method may differ based on the performance metric and the objectives of the user. Figure 11 presents boxplots of the performance metrics depicted in Table 3 representing the ratio between the corrected and expected mean, standard deviation, skewness, and percentiles 5, 10, 50, 90 and 95, and as well as the NSE and KGE values in the future scenario. Again, a value of 1 indicates a perfect fit between the corrected and expected statistics.

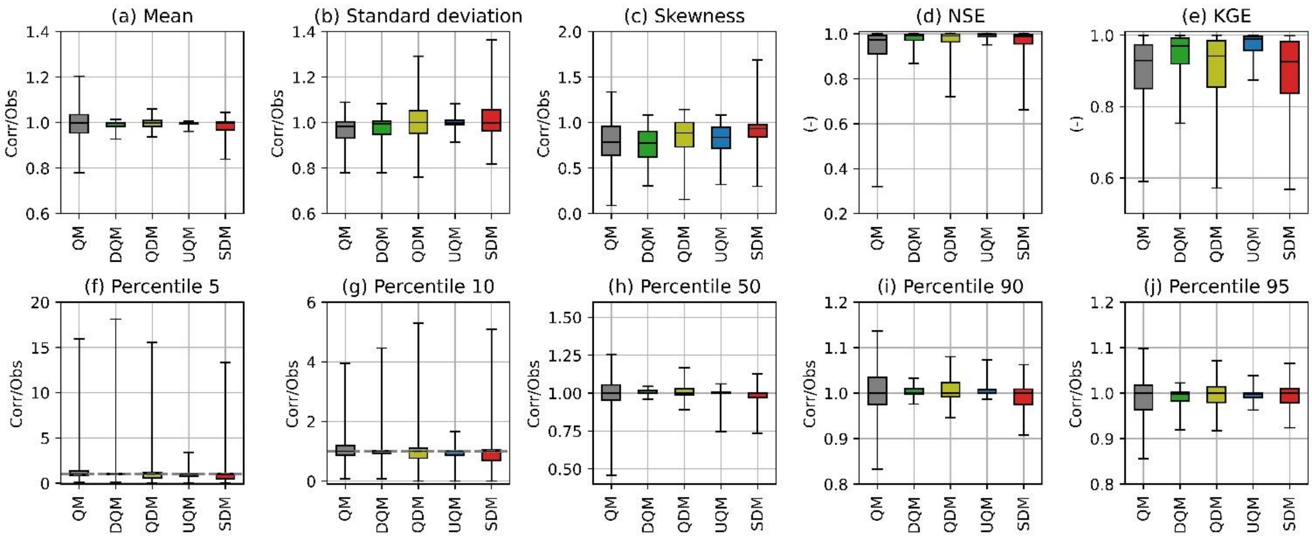


Figure 11: Performance comparison of the QM, DQM, QDM, UQM and SDM in the future period evaluated by means of the ratio between the bias corrected and objective future mean (a), standard deviation (b), skewness (c), and percentiles (f to j), in addition to the NSE (d) and KGE (e) fitness metrics. For all metrics, a value of 1.0 indicates a perfect agreement between the corrected and expected values, as depicted in Table 3.

Figure 11 shows that all the implemented methods might relate to the expected series based on different statistics; however, some methods are more stable in some statistics than others (e.g., UQM and DQM in the mean and percentiles, UQM in the standard deviation or the SDM, QDM and UQM in the skewness). Overall, in this exercise where bias was added in the first two statistics moments, the UQM, which aims at removing the bias in these moments, tends to be more stable than the other methods. Among the implemented methods in the climQMBC package, for this exercise, there is always a trend preserving method that is more stable than the QM method. This does not mean that the QM method should not be used. In fact, as mentioned above, all methods might perform well under certain circumstances, including the QM method. This is why an evaluation of the method for each specific situation is encouraged.

5 Conclusions

Quantile Mapping (QM) is one of the most widely used bias correction methods for climate change impacts, although it might distort raw GCM changes when future climate goes out of the range of the historical climate. Trend-preserving methods such as DQM, QDM, UQM, and SDM aim to address this limitation, each with its own advantages and disadvantages, although the QM method is still widely used because of its simplicity and ease of implementation. This paper presents the climQMBC package, a coding package written in Python, MATLAB and R that makes available all these point-to-point bias correction methods for daily, monthly and annual data. As such, the package enables incorporating several bias correction methods as a standard component of climate change impact assessment, as well as the identification of the most adequate method for a specific case. Being available in several programming languages avoids users having to change between languages and include the bias correction analysis in their existing workflow.

It is worth noting that several climate change impacts assessments are based on the analysis of compound events, which require methods that capture the inter-dependence of several variables or different points in space. The methods available in version 1.1.0 of the climQMBC package are strictly univariate. As such, results from the climQMBC package must be evaluated carefully when analyzing compound events. To date, existing multi-variable methods succeed in capturing these inter-dependences, but sometimes struggle on representing local climate and capturing raw GCM trends, leading to the user to consider a trade-off between the climate characteristics that are the most important for their analysis. Unfortunately, few works have focused on comparing methods or reporting these limitations (Aedo-Quililongo and Chong, 2024) and work is needed both in terms of method development and tools that provide these complex methods for easy and fair comparisons.

This work also proposes a synthetic tester for bias correction and an example in which it is tested. Here, bias is added in the first two statistical moments (i.e. mean and standard deviation) to evaluate the performance of the methods implemented in the climQMBC package. The example is based on a set of synthetic observed historical series, future objective series (assumed as “true series” as if we knew already the future) and biased historical and future series. The future es built as an independent series based on assumptions of the projected change in the mean and standard deviation. The bias is then added to each series considering a methodology that explicitly shifts both the mean and variability of the series while preserving the sequence of events. As such, the bias corrected series should be the same as the synthetic observed historical or future objective series. The performance is measured using goodness-of-fit metrics like the NSE and KGE, and the capability of replicating other statistical properties. It is important to note that in this example, the NSE and KGE are valid because point-by-point agreement between the bias corrected and expected series is desired. However, such metrics are not suitable for real case applications where the GCM sequence of events is not consistent with observations. This example is also implemented in the package so that users can make a brief theoretical comparison of the performance of each method for their specific case.

Based on the synthetic example, we demonstrate that all the methods implemented in the climQMBC package can successfully remove these biases in both the historical period and future scenarios; however, this performance cannot be guaranteed as a general rule. In general, among the methods, there is always one whose performance is more robust than the QM method. The UQM and DQM (methods that target statistical moments) resulted as the most stable methods capturing statistical moments, percentiles and recovering the original series in the synthetic example. Overall, we emphasize the need to assess the performance of the bias correction methods to avoid unwanted bias in the corrected series. However, there are cases where the best option among the available methods will not perform perfectly, and thus a tool like the climQMBC package will be useful to identify the most adequate method for each specific case.

Despite all trend-preserving methods outperform the QM method, in some cases the QM method should not be discarded because it is simple, computationally faster than the trend preserving methods, and might perform well enough under certain conditions. In terms of execution times, trend-preserving methods can take up to around 10, 30 and 40 times the mean execution time of the QM method for annual, monthly and daily frequencies, respectively. This may also be considered in the trade-off between methods as sometimes execution time can be restrictive. However, as machine learning approaches and applications are increasing over the years, these, based on synthetic results could develop methods with comparable results to the trend-preserving methods, but with lower execution cost.

Appendix A: Window length considerations to capture historical conditions and model trends

The climQMBC package considers that the length of the projected periods (backwards moving windows) is the same as the user defined length of the historical period for consistency on the characterization of each period (see Section 3.1.1). While the World Meteorological Organization (WMO) recommends 30 years to characterize Climate Normals (WMO, 2017), there are trade-offs between choosing larger periods to better capture statistical moments and shorter periods to better capture trends and avoid pseudo-stationarity bias. Appendices A1, A2 and A3 present three sensitivity analysis to make visible these trade-offs based on synthetic realizations using a gamma probability density function (eq. 19 in Section 4.1) and linear trend future scenarios, following Chadwick et al. (2023).

It is worth noting that the linear trend example is an extreme condition that allows exploring different issues arising from bias correction decisions. However, real life examples usually have combinations of oscillating or gradual patterns, where it is challenging to identify or address the issues.

Based on the results presented in Appendices A1, A2 and A3, there is no perfect condition in terms of window length to avoid biases in characterizing statistical moments while capturing the modeled trends and avoiding pseudo stationary bias. However, the importance of these potential biases are only relevant for time series with high coefficient of variations and strong trends. It is worth noting that the distribution of the data, which depends on the climate variable, is also an important factor that should be considered. We recommend using periods of lengths between 20 to 40 years aims balancing between the WMO recommendation of 30 years, properly characterizing the statistical moments, adequately capturing the modeled trend and reducing the pseudo-stationarity bias.

Appendix A1: Sample size (years) impacts on mean and standard deviation representativity

This analysis aims to show the impact of the sample size on the mean and standard deviation sample estimation. The analysis considers four conditions with objective mean of μ_{obj}=1,000 and standard deviation σ_{obj} of 250, 500, 750 and 1,000, resulting in coefficients of variation (CV) of 0.25, 0.50, 0.75 and 1.00, respectively. For each condition, the gamma function was used to generate cases with widow lengths from 10 to 60 values or years. Each case was repeated 10,000 times (realizations).

Figure A1 shows the comparison of the sample mean and standard deviation of each realization with the objective value (μ_{obj} and σ_{obj}) in terms of the ratio. The ratio is between each realization and the objective mean and standard deviation (Fig. A1a and Fig. A1c) and the variability among realizations (Fig. A1b and Fig. A1d).

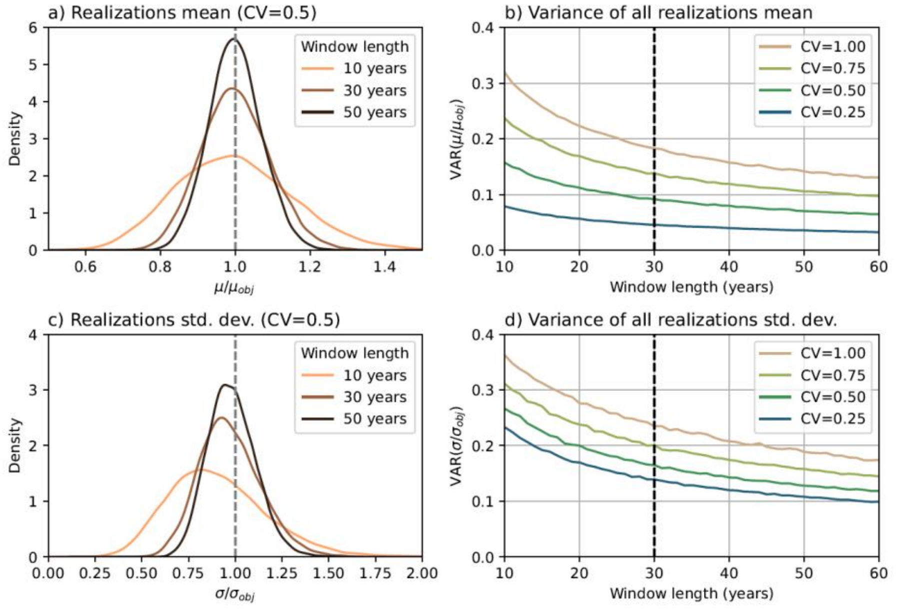


Figure A1: Density distribution of the ratio between each realization and the objective mean (a) and standard deviation (c), for a CV value of 0.5 and different window lengths. Standard deviation of the ratio between each realization and the objective mean (b) and standard deviation (d), for different CV values. Dashed lines in (b) and (d) represent the 30 year window recommended by WMO (2017).

Results show that, for a CV value of 0.5 and a window length of 10 years, the realizations mean is centered around the objective value, but some realizations can have an error of almost 40% (light colored lines in Fig. A1a). On the other hand, the realizations standard deviation is not centered around the objective, but some realizations present errors in standard deviation of almost 75% (light colored lines in Fig. A1c). When increasing the window length, this sample errors diminishes for both moments, and the standard deviation tends to be better centered around the objective standard deviation (dark colored lines in Fig. A1a and Fig. A1c).

The reduction of the variability with increasing window length is represented in Fig. A1b and Fig. A1d in terms of the standard deviation of the realizations mean and standard deviation for each CV condition. While all CV conditions follow the decreasing pattern, high CV values have consistently a higher dispersion. As such, longer window lengths reduce the potential bias on characterizing the sample statistical moments; however, conditions with higher CV result in less representative values.

Note that for centered distributions like normal distributions, the mean and standard deviation tend to be centered around the objective values with shorter window lengths than for skewed distributions like the gamma distribution. However, the potential errors or dispersion issue is not solved directly by changing the distribution.

Appendix A2: Backwards moving window length impacts on capturing the trend in the mean

This analysis aims to show the potential biases in capturing model trends of the mean when considering backwards moving windows (e.g. for a 30 years moving window, the value for 2029 represents the mean of the period 2000-2029). Backwards moving windows are have been used in previous studies (Chadwick et al., 2018; Chadwick et al., 2023) as are more suitable to capture GCM trends and generate continuous time series over forward and centered moving windows, as the latter result in discontinuities between the historical and future periods, along to not being able to compute values for the last years of the projections.

The analysis considers annual time series from 1955 to 2100 generated with the gamma function considering a mean of μ_h=1,000 and a CV of 0.1. Four scenarios were imposed on the time series where the historical period (1955-2014) has no trend and the future period (2015-2100) has a linear trend with projected changes of -50, -25, +25 and +50% between the historical period and year 2100. For each of these four scenarios, 10,000 realizations were generated. Figure A2a shows the trends in the annual mean.

Each of the generated realizations were averaged considering backward moving windows of lengths from 10 to 60 years. Figure A2b presents the average of the 10, 30 and 50 years moving mean realizations for the scenario with 50% increase for 2100, showing two potential biases that can arise from considering moving windows: trend lag and terminal error. Overall, the moving windows underestimate the magnitude of the change of the real linear trend; however, the difference between the real values and the moving mean increases in time until it remains constant. The trend lag is the lag between the real trend and the moving window trend, which is equal to half of the moving window length, while the terminal error is the maximum difference. Both the trend lag and the terminal error increase with larger window lengths, but the terminal error also increases with the magnitude of the trend, as shown in Fig. A2c.

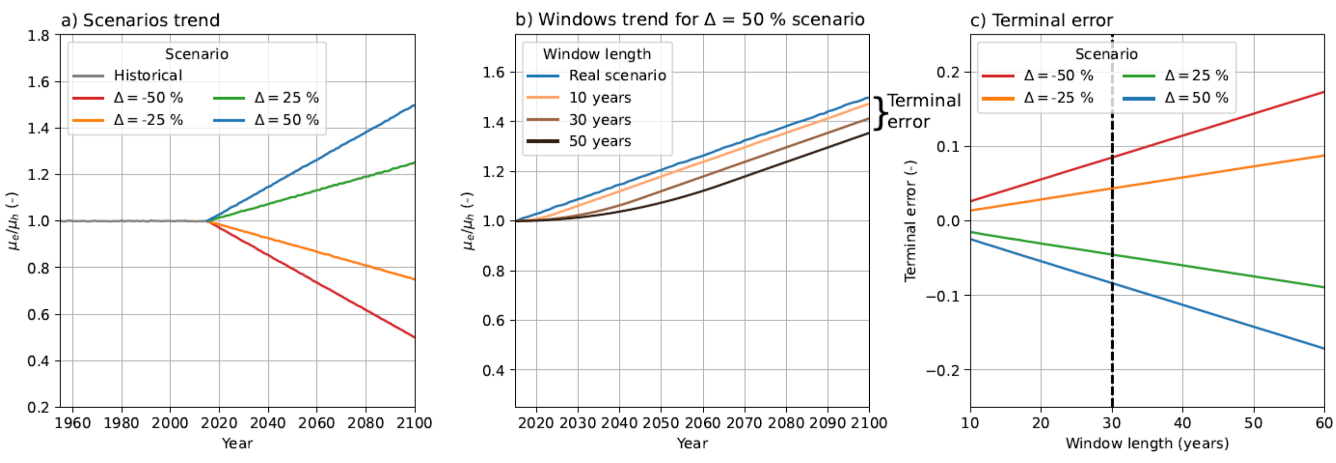


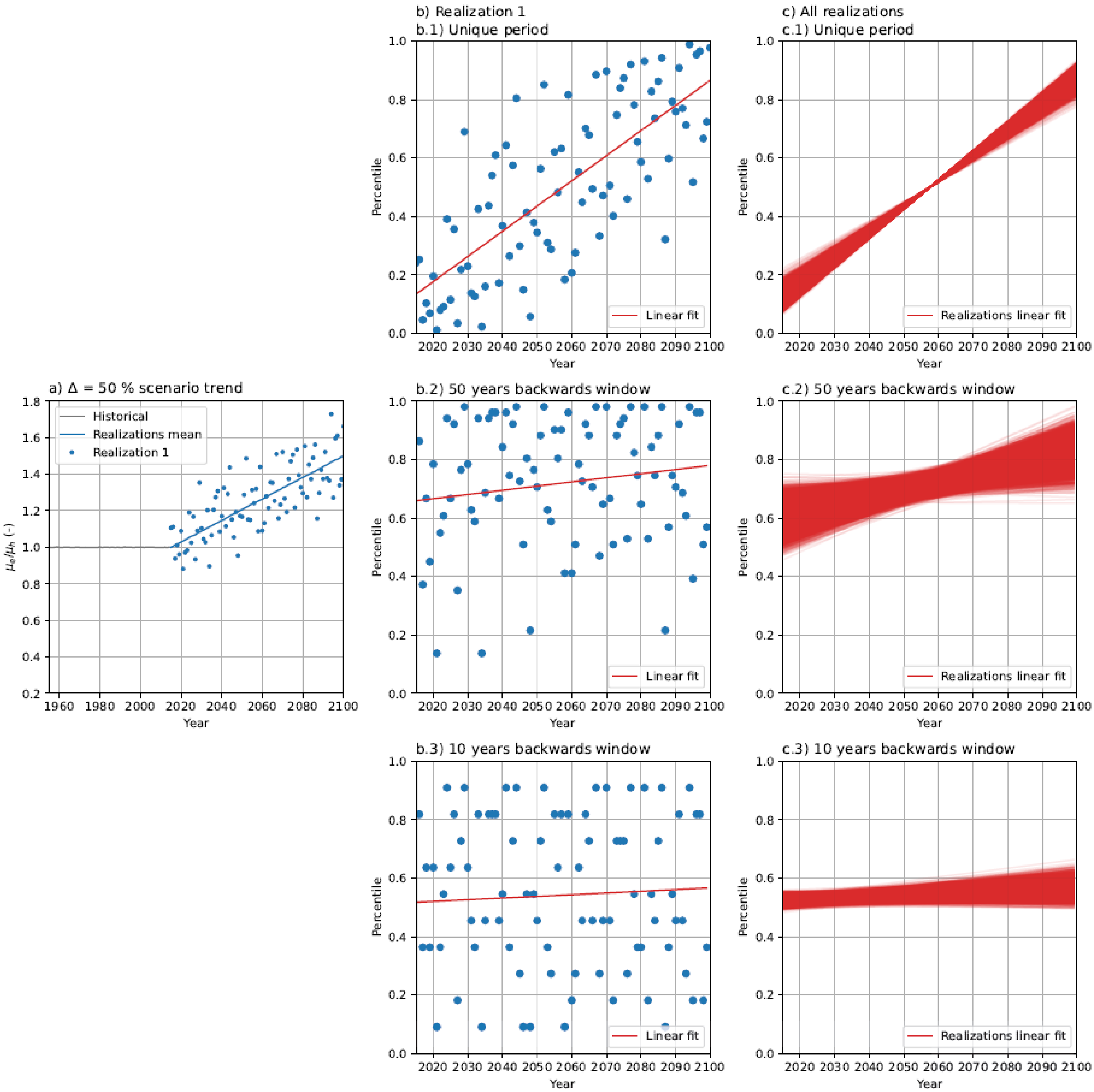
Figure A2: a) Mean of the linear trend realizations, divided by the historical mean; b) Comparison of real trend and the moving window with different window lengths; and c) Terminal error of for different window length for the four linear trend scenarios. Dashed lines in (c)represent the 30 year window recommended by WMO (2017).

Appendix A3: Moving windows and pseudo-stationarity bias

Pseudo stationary bias arises when the model biases are assumed constant over an analyzed period where climate is changing (Chadwick et al., 2023). This condition can be observed on increasing temperatures scenarios where the values at the end of the period are considerably higher than at the beginning of the period, resulting on a probability distribution where the values at the end of the period are on one of the distribution tails, while the initial values are in the other tail. Considering moving windows for each year analyzed allows reducing the pseudo stationary bias.

This analysis presents how using moving windows reduces the pseudo-stationarity bias. The analysis is based on the scenario with increase in 50% in 2100 generated on Appendix A2. Figure A3a shows the mean of the realizations and the future values of realization 1. Figure A3b (middle column of Fig. A3) shows the values in terms of percentiles of realization 1 and the percentiles linear trend, while Figure A3c (right column of Fig. A3) shows the linear trends of all realizations. Figure A3b1 and A3c1 (first row of Fig. A3) considers the period 2015 – 2100 as a unique period (i.e., no moving window is used here, this considers the traditional fix window approach). Figure A3b2 and A3c2 (second row of Fig. A3) considers a backwards moving window of 50 years for each year to define the corresponding percentile. Figure A3b3 and A3c3 (third row of Fig. A3) shows the same information as Fig. A3b2 and A3c2, but for moving windows of 10 years. Figure A4 shows the mean values and the slope of the percentile time series for different moving window values, along to the case where a unique period is considered. Results show that considering a unique period in a linear trend scenario result in percentiles where the extreme values are located at the beginning and at the end of the period, with a clear increasing trend (Fig. A3b1). On the other hand, when moving windows are considered, the slope of the trend in the probability domain decreases with the moving window length (Fig. A3b2, Fig. A3b3 and Fig. A4b). On the other hand, when a unique period is considered, the mean is centered in a percentile of 0.5, while for moving windows the mean is above 0.5 with a decreasing trend with decreasing window lengths (Fig. A3c2, Fig. A3c3 and Fig. A4b).

This analysis shows a trade-off in terms of the pseudo-stationary bias, where reducing the window length results in a reduction in this bias. On the other hand, considering the complete period may capture correctly the mean but have a tendency to overestimate percentiles at the end of the period and underestimate them at the beginning of the period. While these results are related to trends with monotonic increase, each analyzed area should explore this kind of conditions.



705 **Figure A3: a) Realization 1 and mean of the realizations for the scenario with linear increase of 50%, normalized by the historical mean. b) Realization 1 values in terms of percentiles and linear trend of the percentile probability values: b1) considering a unique period to build the distribution; b2) and b3 considering moving windows of 50 and 10 years for each year, respectively, to build the distribution. c) Same as b), but only the linear trend for all realizations.**

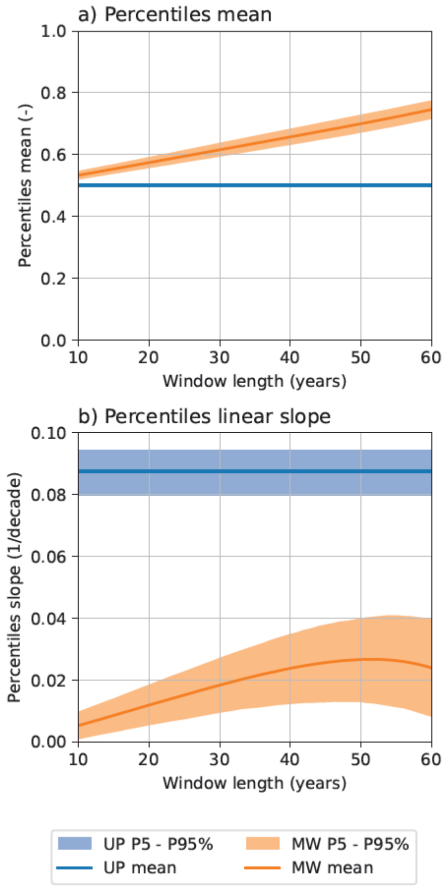


Figure A4: a) Mean of the percentiles of the realizations for the scenario with linear increase of 50% when a unique period is considered (UP) and when moving windows (MW) of different lengths are considered. b) Slope of the percentiles linear trend.

Appendix B: climQMBC package methods execution time comparison

To compare the different execution times of the bias correction methods available in the climQMBC package, this appendix presents a synthetic example using a gamma probability density function (eq. 19 in Section 4.1). The example is based on 10,000 realizations considering a fixed historical mean of 1,000, a historical period of 30 years and a future period of 60 years. Each realization considers a different standard deviation with values of 250, 500, 750 and 1,000, sampled randomly, a random bias in the mean and standard deviation ($\Delta\mu_{bias}$ and $\Delta\sigma_{bias}$, respectively) ranging from -50 % to +50 %, and a random projected change in the mean and standard deviation ($\Delta\mu$ and $\Delta\sigma$, respectively) ranging from -50 % to +50 %. The same bias was added in the historical and in the future periods, following the same procedure as in the synthetic example presented in Section 5 (eq. 20). The example was replicated for annual, monthly and daily values, following the same procedure, and considering multiplicative values and strictly non-negative series.

This example aims for a general reference of the execution time rather than a detailed performance comparison. As such, the execution times are analyzed relative to the mean execution time of the QM method (Table B1) and considering only the package implementation in Python.

Results show that the QM method is the fastest among the implemented in the climQMBC package. The second fastest method is the DQM, with execution times close to 2.1 times the QM method for annual and monthly frequency, and 6.8 times for daily frequency, explained by the additional calculation of the projected changes in the mean and its application to detrend and reimpose the trend to the corrected series. The slowest methods are the QDM, UQM and SDM, with execution times between 7 and 10 times the QM method for annual frequency, 22 to 27 times for monthly frequency and 32 to 39 times for daily frequency. While QDM, UQM and SDM are the slowest methods, QDM tends to be slightly faster than the UQM method. While both methods consider a new transform function for each value to be corrected, the QDM computes only the projected change of the quantile of the value to be corrected, and the UQM computes the projected change of the mean, standard deviation and skewness (if applicable) of the associated moving windows, resulting in a longer execution time. On the other hand, the SDM method considers several detrending, projected changes and interpolation steps that demand additional times.

While the three temporal frequencies consider the random near-zero replacement and moving windows, the daily frequency also considers the daily moving windows (to capture seasonality) and the calculation of the threshold to match the modeled number of rain days in the historical period to the observed number of rain days. These differences explain the additional execution times difference of the daily frequency compared to the annual and monthly frequencies.

Table B1: Execution times for the bias correction methods available in the climQMBC package, relative to the mean execution time of the QM method.

	Annual		Monthly		Daily	
Method	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
QM	1.0	0.2	1.0	0.4	1.0	0.2
DQM	2.2	0.5	2.0	0.4	6.8	4.6
QDM	8.8	2.0	22.6	0.7	32.6	6.0
UQM	9.8	2.1	24.8	0.7	33.4	6.1
SDM	7.1	1.6	26.5	1.1	38.7	5.9

Code availability

The source code of the climQMBC package for Python, R and MATLAB are openly available at GitHub (https://github.com/saedoquililongo/climQMBC) and has been archived in Zenodo with the DOI https://doi.org/10.5281/zenodo.18392899 (Aedo-Quililongo et al., 2026). Instructions to import the library can be found at https://github.com/saedoquililongo/climQMBC/blob/main/Python/README.md for Python, https://github.com/saedoquililongo/climQMBC/blob/main/R/README.md for R and https://github.com/saedoquililongo/climQMBC/blob/main/Matlab/README.md for MATLAB. Examples for testing the package can be found at https://github.com/saedoquililongo/climQMBC/blob/main/Python/climQMBC_tester_Python.py for Python, https://github.com/saedoquililongo/climQMBC/blob/main/R/climQMBC_tester_R.R for R and https://github.com/saedoquililongo/climQMBC/blob/main/Matlab/climQMBC_tester_Matlab.m for MATLAB. Data for the tester codes can be found at https://github.com/saedoquililongo/climQMBC/tree/main/Sample_data. Both the observed and modeled series must be a column vector of daily, monthly or annual data, without considering leap days, starting on the same date, and be continuous until the ending date. If daily frequency is specified, the length of the column vector should be a multiple of 365 and for monthly frequency, it should be a multiple of 12.

Author contributions

Sebastián Aedo-Quililongo: Conceptualization, Formal Analysis, Software, Writing – original draft. Cristián Chadwick: Conceptualization, Formal Analysis, Funding Acquisition, Software, Supervision, Writing – original draft. Fernando González-Leiva: Software, Writing – review & editing. Jorge Gironás: Funding Acquisition, Writing – review & editing

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors want to acknowledge the valuable feedback from users of the climQMBC package that has led to continuous improvement of the package. We would also like to acknowledge the relevant feedback from the reviewers and editors; it has helped to strengthen the manuscript.

Financial support

This research was supported by funding from the National Agency for Research and Development, Chile, Grant ANID/FONDECYT/Regular 1200135 and 1241163. We would also like to acknowledge the Grant ANID/FONDECYT/Iniciación 11220952; the Grant Centro de Desarrollo Urbano Sustentable ANID CEDEUS CIN250009.

References

Aboitiz, C., Chadwick, C., Barichivich, J., and Perez-Quezada, J.: A framework to compare the performance of spatial disaggregation methods of statistical climate downscaling in South America, J. Hydrol.: Reg. Stud., 67, https://doi.org/10.1016/j.ejrh.2026.103834, 2026.

Aedo-Quililongo, S., Chadwick, C., González-Leiva, F., and Gironás, J.: saedoquililongo/climQMBC: Version 1.1.0, Zenodo [code], https://doi.org/10.5281/zenodo.18392899, 2026.

Aedo-Quililongo, S.: ¿Por qué mis proyecciones de cambio climático difieren de las de los Modelos de Circulación General?, in Proceedings of the IAHR XXXI Congreso Latinoamericano de Hidráulica, Medellín, Colombia, 1-4 October 2024, 2775-2782, ISBN: 978-90-834302-6-3, 2024.

Aedo-Quililongo, S., and Chong, P.: Propuesta Metodológica Para El Escalamiento Local A Nivel Diario De Modelos De Circulación General, in Proceedings of the IAHR XXXI Congreso Latinoamericano de Hidráulica, Medellín, Colombia, 1-4 October 2024, 2783-2793, ISBN: 978-90-834302-6-3, 2024.

Andres, N., Vegas Galdos, F., Lavado Casimiro, W. S., and Zappa, M.: Water resources and climate change impact modelling on a daily time scale in the Peruvian Andes, Hydrol. Sci. J., 59, 2043-2059, https://doi.org/10.1080/02626667.2013.862336, 2014.

Ballarin, A., Sone, J., Gesualdo, G., Schwamback, D., Reis, A., Almagro, A., and Wendland, E.: CLIMBra - Climate Change Dataset for Brazil, Sci Data, 10, 47, https://doi.org/10.1038/s41597-023-01956-z, 2023.

Bayissa, Y., Melesse, A., Bhat, M., Tadesse, T., and Shiferaw, A.: Evaluation of Regional Climate Models (RCMs) Using Precipitation and Temperature-Based Climatic Indices: A Case Study of Florida, USA, Water, 13, 2411, https://doi.org/10.3390/w13172411, 2021.

Bedia, J., Baño-Medina, J., Legasa, M. N., Iturbide, M., Manzanas, R., Herrera, S., Casanueva, A., San-Martín, D., Cofiño, A. S., and Gutiérrez, J. M.: Statistical downscaling with the downscaleR package (v3.1.0): contribution to the VALUE intercomparison experiment, Geosci. Model Dev., 13, 1711–1735, https://doi.org/10.5194/gmd-13-1711-2020, 2020.

Blin, N.: Understanding groundwater dynamics in a changing climate: Numerical modeling and remote sensing insights from the Salar del Huasco Basin, Ph.D. thesis, Pontificia Universidad Católica de Chile, https://repositorio.uc.cl/server/api/core/bitstreams/acf55587-dc5f-41fa-8d7e-e23cba55082d/content, 2023.

Bourgault, P., Huard, D., Smith, T., Logan, T., Aoun, A., Lavoie, J., Dupuis, É., Rondeau-Genesse, G., Alegre, R., Barnes, C., Laperrièr, A., Biner, S., Caron, D., Ehbrecht, C., Fyke, J., Keel, T., Labonté, M-P., Lierhammer, L.,Low, J-F., Quinn, J., Roy, P., Squire, D., Stephens, A., Tanguy, M., and Whelan, C.: xclim: xarray-based climate data analytics, J. Open Source Softw., 8, 5415, https://doi.org/10.21105/joss.05415, 2023.

Brekke, L., Thrasher, B., Maurer, E. P., and Pruitt, T.: Downscaled CMIP3 and CMIP5 climate projections: Release of downscaled CMIP5 climate projections, comparison with preceding information, and summary of user needs, US Department of Interior, Bureau of Reclamation, Denver, CO, USA, 104 pp., 2013.

Bürger, G., Sobie, S., Cannon, A. J., Werner, A., and Murdock, T.: Downscaling extremes: An intercomparison of multiple methods for future climate, J. Clim., 26, 3429–3449, https://doi.org/10.1175/JCLI-D-12-00249.1, 2013.

Cannon, A. J., Sobie, S., and Murdock, T.: Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes?, J. Clim., 28, 6938-6959, https://doi.org/ 10.1175/JCLI-D-14-00754.1, 2015.

Cannon, A. J.: Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables, Clim. Dyn., 50, 31–49, https://doi.org/10.1007/s00382-017-3580-6, 2017.

Chadwick, C., Gironás, J., Vicuña, S., Meza, F., and McPhee, J.: Using a Statistical Pre-Analysis Approach as an Ensemble Technique for the Unbiased Mapping of GCM Changes to Local Stations, J. Hydrometeorol., 19, 1447-1465, https://doi.org/10.1175/JHM-D-17-0198.1, 2018.

Chadwick, C., Gironás, J., González-Leiva, F., and Aedo, S.: Bias adjustment to preserve changes in variability: the unbiased mapping of GCM changes, Hydrol. Sci. J., 68, 1184-1201, https://doi.org/10.1080/02626667.2023.2201450, 2023.

Chen, D., Dai, A., and Hall, A.: The convective-to-total precipitation ratio and the "drizzling" bias in climate models, J. Geophys. Res. Atmos., 126, https://doi.org/10.1029/2020JD034198, 2021.

Chen, J., Brissette, F. P., Chaumont, D., and Braun, M.: Performance and uncertainty evaluation of empirical downscaling methods in quantifying the climate impacts on hydrology over two North American river basins. J. Hydrol., 479, 200–214, https://doi.org/10.1016/j.jhydrol.2012.11.062, 2013.

Christensen J. H., Boberg, F., Christensen, O. B., and Lucas-Picher, P.: On the need for bias correction of regional climate change projections of temperature and precipitation, Geophys. Res. Lett., 35, https://doi.org/10.1029/2008GL035694, 2008.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, L.: Hydrology and Earth System Sciences HESS Opinions "Should we apply bias correction to global and regional climate model data?", Hydrol. Earth Syst. Sci., 16, 3391–3404, https://doi.org/10.5194/hess-16-3391-2012, 2012.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fernandez-Palomino, C.A., Hattermann, F.F., Krysanova, V., Vega-Jácome, F., Menz, C., Gleixner, S., and Bronstert, A.: High-resolution climate projection dataset based on CMIP6 for Peru and Ecuador: BASD-CMIP6-PE, Sci Data, 11, 34, https://doi.org/10.1038/s41597-023-02863-z, 2024.

Fowler, H. J., Blenkinsop, S., and Tebaldi, C.: Linking climate change modelling to impacts studies: Recent advances in down-scaling techniques for hydrological modelling, Int. J. Climatol., 27, 1547–1578, https://doi.org/10.1002/joc.1556, 2007.

Fowler, H.J., Mearns, L.O., Wilby, R.L.: Downscaling Future Climate Projections: Compounding Uncertainty But Adding Value?, in: Uncertainty in Climate Change Research, Springer, edited by: Mearns, L.O., Forest, C.E., Fowler, H.J., Lempert, R., Wilby, R.L, https://doi.org/10.1007/978-3-031-85542-9_18, 2025.

Gergel, D. R., Malevich, S. B., McCusker, K. E., Tenezakis, E., Delgado, M. T., Fish, M. A., and Kopp, R. E.: Global Downscaled Projections for Climate Impacts Research (GDPCIR): preserving quantile trends for modeling future climate impacts, Geosci. Model Dev., 17, 191–227, https://doi.org/10.5194/gmd-17-191-2024, 2024.

González-Leiva, F., Gironás, J., Chadwick, C., Meza, F., and Vicuña, S.: When will meteorological droughts become the new normal climatic condition in Central-Southern Chile?, Hydrol. Sci. J, 71, 1068–1086. https://doi.org/10.1080/02626667.2026.2625979, 2026.

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol, 377, 80-91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Gutiérrez, J. M., San-Martín, D., Brands, S., Manzanas, R., and Herrera, S.: Reassessing Statistical Downscaling Techniques for Their Robust Application under Climate Change Conditions, J. Clim., 26, 171–188, https://doi.org/10.1175/JCLI-D-11-00687.1, 2013.

Guzmán, J., Gironás, J., Sandoval, C., Chadwick, C., Cienfuegos, R., Nash, S., Muñoz, S., and Escauriaza, C.: Evaluating climate-driven changes in peak flows through continuous hydrological simulation: the case of the Marga Marga basin in central Chile, J. Water. Clim. Change, 17, 870-887, https://doi.org/10.2166/wcc.2026.267, 2026.

Hagemann, S., Chen, C., Haerter, J. O., Heinke, J., Gerten, D., and Piani, C.: Impact of a Statistical Bias Correction on the Projected Hydrological Changes Obtained from Three GCMs and Two Hydrology Models, J. Hydrometeorol., 12, 556–578, https://doi.org/10.1175/2011JHM1336.1, 2011.

Hakala, K., Addor, N., Teutschbein, C., Vis, M., Dakhlaoui, H., and Seibert, J.: Hydrological Modeling of Climate Change Impacts, Encyclopedia of Water: Science, Technology, and Society, https://doi.org/10.1002/9781119300762.wsts0062, 2019.

Hanel, M., Kožín, R., Heřmanovský, M., and Roub, R.: An R package for assessment of statistical downscaling methods for hydrological climate change impact studies, Environ. Model. Softw., 95, 22-28, https://doi.org/10.1016/j.envsoft.2017.03.036, 2017.

Hnilica, J., Hanel, M., and Puš, V.: Multisite bias correction of precipitation data from regional climate models. Int. J. Climatol., 37, 2934–2946, https://doi.org/10.1002/joc.4890, 2017.

Iturbide, M., Bedia, J., Herrera, S., Baño-Medina, J., Fernández, J., Frías, M.D., Manzanas, R., San-Martín, D., Cimadevilla, E., Cofiño, A.S., and Gutiérrez, J.M.: The R-based climate4R open framework for reproducible climate data access and post-processing, Environ. Model. Softw., 111, 42-54, https://doi.org/10.1016/j.envsoft.2018.09.009, 2019.

Kassahun, D., Tilahun, M., Tesema, Z., Meseret, M., Lakew, E., Tesfa, A., Bitew, A., Lakew, M., Bimrow, T., Ayalew, M., Kebede, A., Hunegnaw, B., Adene, A., Dagnaw, D., Alemayehu, M., Tarekegn, G.: Climate-resilient genetic selection in cattle: Insights from resilience phenotypes of Fogera cattle under climate variabilities, Vet. Anim. Sci., 30, https://doi.org/10.1016/j.vas.2025.100538, 2025.

Kim, K. B., Bray, M., and Han, D.: Exploration of optimal time steps for daily precipitation bias correction: a case study using a single grid of RCM on the River Exe in southwest England, Hydrol. Sci. J., 61, 289-301, https://doi.org/10.1080/02626667.2015.1027207, 2016.

Knoben, W. J. M., Freer, J. E., and Woods, R. A.: Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. Hydrol. Earth Syst. Sci., 23, 4323-4331, https://doi.org/10.5194/hess-23-4323-2019, 2019.

Lafferty, D.C., and Sriver, R.L: Downscaling and bias-correction contribute considerable uncertainty to local climate projections in CMIP6, npj Clim Atmos Sci, 6, 158, https://doi.org/10.1038/s41612-023-00486-0, 2023.

Lange, S.: Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0), Geosci. Model Dev., 12, 3055–3070, https://doi.org/10.5194/gmd-12-3055-2019, 2019.

Lanzante, J. R., Dixon, K. W., Nath, M. J., Whitlock, C. E., and Adams-Smith, S.: Some Pitfalls in Statistical Downscaling of Future Climate, Bull. Amer. Meteor. Soc., 99, https://doi.org/10.1175/BAMS-D-17-0046.1, 2018.

Maraun, D., Wetterhall, E. F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H., Sauter, T., Themeßl, M., Venema, V., Chun, K., Goodess, C., Jones, R., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, Rev. Geophys, 48, https://doi.org/10.1029/2009RG000314, 2010.

Maraun, D.: Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums, Geophys. Res. Lett., 39, https://doi.org/10.1029/2012GL051210, 2012.

Maraun, D.: Bias Correcting Climate Change Simulations - a Critical Review, Curr. Clim. Change Rep., 2, 211–220, https://doi.org/10.1007/s40641-016-0050-x, 2016.

Martínez, D., Chadwick, C., and Plaza-Aguilar, A.: The Time of Emergence (ToE) of the Andean Mediterranean sclerophyllous forest of Quillaja saponaria (Mol.) and Lithraea caustica (mol.) Hox. & Arn, For. Ecol. Manag., 544, https://doi.org/10.1016/j.foreco.2023.121169, 2023.

Maurer, E. P., Das, T., and Cayan, D. R.: Errors in climate model daily precipitation and temperature output: Time invariance and implications for bias correction, Hydrol. Earth Syst. Sci., 17, 2147–2159, https://doi.org/10.5194/hess-17-2147-2013, 2013.

Maurer, E. P., and D. W Pierce.: Bias correction can modify climate model simulated precipitation changes without adverse effect on the ensemble mean, Hydrol. Earth Syst. Sci., 18, 915–925, https://doi.org/10.5194/hess-18-915-2014, 2014.

Mauritsen, T., Stevens, B., Roeckner, E., Crueger, T., Esch, M., Giorgetta, M., Haak, H., Jungclaus, J. H., Klocke, D., Matei, D., Mikolajewicz, U., Notz, D., Pincus, R., Schmidt, H., and Tomassini, L.: Tuning the climate of a global model, J. Adv. Model. Earth Syst., 4, https://doi.org/10.1029/2012MS000154, 2012.

Menapace, A., Dhawan, P., Dalla Torre, D., Kaffas, K., Crespi, A., Larcher, M., Righetti, M., and Cannon, A. J.: Review of bias correction methods for climate model outputs in hydrology, J. Hydrol., 660, 133213, https://doi.org/10.1016/j.jhydrol.2025.133213, 2025.

Meyer, J., Kohn, I., Stahl, K., Hakala, K., Seibert, J., and Cannon, A. J.: Effects of univariate and multivariate bias correction on hydrological impact projections in alpine catchments, Hydrol. Earth Syst. Sci., 23, 1339–1354, https://doi.org/10.5194/hess-23-1339-2019, 2019.

Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., Ilyina, T., Kleine, T., Kornblueh, L., Li, H., Modali, K., Notz, D., Pohlmann, H., Roeckner, E., Stemmler, I., Tian, F., and Marotzke, J.: A higher-resolution version of the Max Planck Institute Earth System Model (MPI-ESM 1.2-HR), J. Adv. Model. Earth Sy., 10, 1383–1413, https://doi.org/10.1029/2017MS001217, 2018.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Nash, J. E., and J. V. Sutcliffe.: River flow forecasting through conceptual models: Part 1. A discussion of principles, J. Hydrol., 10, 282-290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Nguyen, N., Anh, T., Nguyen, H., and Dang, D.: Quantile mapping technique for enhancing satellite-derived precipitation data in hydrological modelling: A case study of the Lam River Basin, Vietnam. J. Hydroinform, 26, 2026-2044, https://doi.org/10.2166/hydro.2024.225, 2024.

Petrovic, D., Fersch, B., and Kunstmann, H.: Heat wave characteristics: evaluation of regional climate model performances for Germany, Nat. Hazards Earth Syst. Sci., 24, 265–289, https://doi.org/10.5194/nhess-24-265-2024, 2024.

Pierce, D. W., Cayan, D. R., Maurer, E. P., Abatzoglou, J. T., and Hegewisch, K. C.: Improved bias correction techniques for hydrological simulations of climate change, J. Hydrometeorol., 16, 2421-2442, https://doi.org/10.1175/JHM-D-14-0236.1, 2015.

Portalanza, D., Ortega, C., Garzon, L., Bello, M., Zuluaga, C. F., Bresciani, C., Durigon, A., Ferraz, S.: Projected Heat Waves in Ecuador under Climate Change: Insights from HadGEM-RegCM4 Coupled Model, Earth, 5, 90-109, https://doi.org/10.3390/earth5010005, 2024.

Slater, L. J., Arnal, L., Boucher, M.-A., Chang, A. Y.-Y., Moulds, S., Murphy, C., Nearing, G., Shalev, G., Shen, C., Speight, L., Villarini, G., Wilby, R. L., Wood, A., and Zappa, M.: Hybrid forecasting: blending climate predictions with AI models, Hydrol. Earth Syst. Sci., 27, 1865–1889, https://doi.org/10.5194/hess-27-1865-2023, 2023.

Soares, P. M. M., Johannsen, F., Lima, D. C. A., Lemos, G., Bento, V. A., and Bushenkova, A.: High-resolution downscaling of CMIP6 Earth system and global climate models using deep learning for Iberia, Geosci. Model Dev., 17, 229–259, https://doi.org/10.5194/gmd-17-229-2024, 2024.

Switanek, B. M., Troch, P., Castro, L. C., Leuprecht, A., Chang, H. I., Mukherjee, R., and Demaria, M. C. E.: Scaled distribution mapping: A bias correction method that preserves raw climate model projected changes, Hydrol. Earth Syst. Sci., 21, 2649-2666, https://doi.org/10.5194/hess-21-2649-2017, 2017.

Teutschbein, C., and Seibert, J.: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, J. Hydrol., 456–457, 12–29, https://doi.org/10.1016/j.jhydrol.2012.05.052, 2012.

Themeßl, M., Gobiet, A., and Heinrich, G.: Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal, Clim. Change, 112, 449–468, https://doi.org/10.1007/s10584-011-0224-4, 2012.

Thrasher, B., Wang, W., Michaelis, A., Melton, F., Lee, T., and Nemani, R.: NASA Global Daily Downscaled Projections, CMIP6, Sci Data, 9, 262, https://doi.org/10.1038/s41597-022-01393-4, 2022.

Vicuña, S., Ávila, A., Gironás, J., Melo, Ó., Rivera, D., and Molinos, M.: Contingent Agreements to Reduce the Impacts of Droughts on Human Water Consumption, in: Droughts in Chile. Global Issues in Water Policy, Springer, edited by: Rivera, D., Donoso, G., https://doi.org/10.1007/978-3-031-85040-0_16, 2025.

Walton, D. B., Sun, F., Hall, A., and Capps, S.: A Hybrid Dynamical-Statistical Downscaling Technique. Part I: Development and Validation of the Technique, J. Clim., 28, 4597-4617. https://doi.org/10.1175/JCLI-D-14-00196.1, 2015.

Wang, A., Zhang, S., Jiang, H., and Wang, H.: Bias correction of seasonal precipitation prediction in Southern China: Pointwise and spatial corrections using deep learning methods, J. Hydrol. Reg. Stud., 61, https://doi.org/10.1016/j.ejrh.2025.102737, 2025.

WMO, W. M. O.: WMO Guidelines on the Calculation of Climate Normals, Technical, WMO, Geneva, https://library.wmo.int/idurl/4/55797 (last access: 13 August 2026), 2017.

Wood, A. W., Maurer, E. P., Kumar, A., and Lettenmaier, D. P.: Long-Range Experimental Hydrological Forecasting for the Eastern United States, J. Geophys. Res. Atmos., 107, https://doi.org/10.1029/2001JD000659, 2002.

Yuval, J. and O'Gorman, P.A.: Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions, Nat. Commun., 11, 3295, https://doi.org/10.1038/s41467-020-17142-3, 2020.

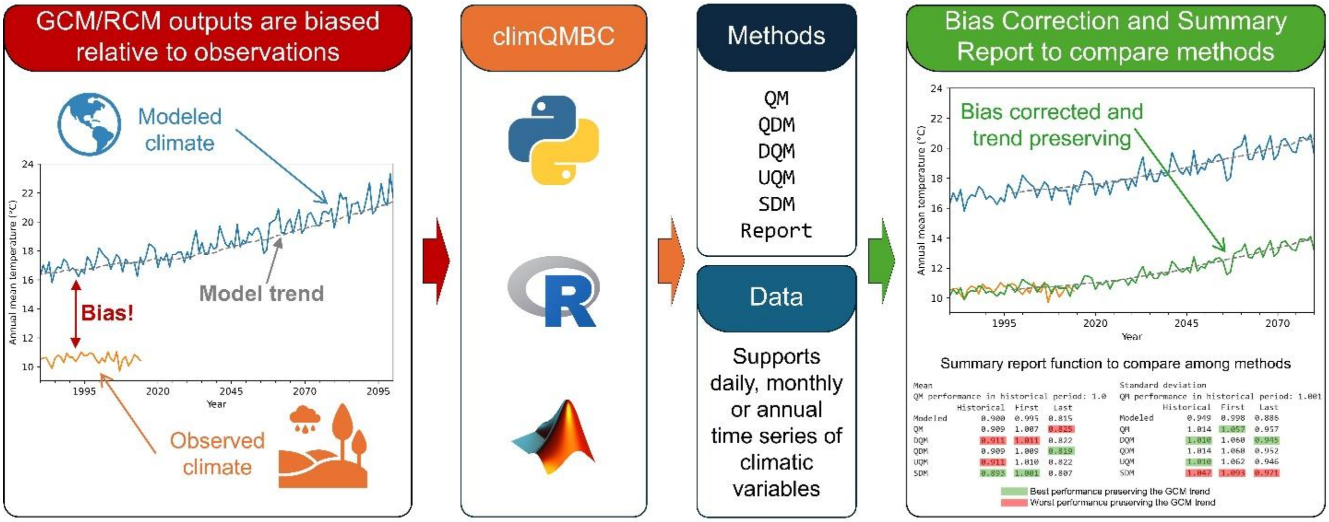
climQMBC: A package with multiple bias correction methods of GCM climatic variables at daily, monthly and annual scale, developed in Python, R and MATLAB

Sebastián Aedo-Quililongo[1], Cristián Chadwick[2], Fernando González-Leiva[3], Jorge Gironás[3,4,5]

[1]Stockholm Environment Institute, Latin America Centre, Bogotá, 110231, Colombia
[2]Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibánez, Santiago, 7941169, Chile
[3]Departamento de Ingeniería Hidráulica y Ambiental, Pontificia Universidad Católica de Chile, Santiago, 7820436, Chile
[4]Centro de Cambio Global UC, Pontificia Universidad Católica de Chile, Santiago, 7820436, Chile
[5]Centro de Desarrollo Urbano Sustentable ANID CEDEUS CIN250009, Santiago, 7820436, Chile

Correspondence to: Cristián Chadwick (cristian.chadwick@uai.cl)

Abstract. Climate change projections are studied using General Circulation Models (GCMs). GCMs ~~are models that~~ simulate climate on a broad scale, hence they cannot be directly used in local impact studies, such as, ~~for example,~~ hydrological studies. GCM~~s~~ results must be~~go through a process of~~ downscaled~~ing~~ and bias corrected, to adjust their results in terms of spatial scale and reduce their bias before being used ~~at the~~ locally~~ scale~~. However, downscaling and bias correction contributes ~~a non-negligible contribution~~ to climate change impacts assessment uncertainty. Quantile Mapping (QM) is one of the most widely used approaches ~~for~~ in bias correcting~~ GCM climate outputs~~. However, ~~in its conventional formulation~~~~classic~~standard QM assumes a time-invariant correction function, which ~~potentially results in~~might add~~itional~~ bias~~es~~. This has motivated the development of trend-preserving variations, which perform QM while ~~accounting for a non-stationary correction function and~~ aiming to preserve ~~the~~ raw GCM ~~signal~~changes. Unfortunately, choosing which variation to use is not straight-forward. We present the climQMBC package (https://github.com/saedoquililongo/climQMBC or https://doi.org/10.5281/zenodo.18392899) as an easy-to-use tool to compare point-to-point quantile mapping approaches. climQMBC is available in Python, R and MATLAB, and contains the ~~classic~~ standard QM method and four trend-preserving variations: Detrended Quantile Mapping (DQM), Quantile Delta Mapping (QDM), Unbiased Quantile Mapping (UQM) and Scaled Distribution Mapping (SDM). This package has a built-in summary report that allows comparing methods in terms of their capability of preserving raw GCM trends and a synthetic tester of bias correction methods. The synthetic tester develops times series to which bias is added allowing testing bias correction methods, given that one knows the objective time series ~~after~~before the bias correction. ~~analysis to complement the comparison based on the climate characteristics, where bias is added in the first two statistical moments (i.e. mean and standard deviation). An exercise based on synthetic exercise the synthetic analysis function~~The synthetic tester showed that the UQM and DQM methods ~~(methods that target statistical moments)~~ are the most reliable capturing statistical moments changes~~, percentiles and recovering the objective series~~.~~most reliable methods are the UQM and DQM.~~



1 Introduction

Results from General Circulation Models (GCMs) ~~or Regional Climate Models (RCMs)~~ are often used to assess climate change impacts on water resources, crop productivity, heat waves, among others climate risks (Fowler et al., 2007; Cannon, 2017; Portalanza et al., 2024). However, GCMs spatial resolution is not suitable for local-scale analysis as these models have biases associated with the parametrization of sub-processes, and their results correspond to mean values over large area (~ hundreds of km) (Mauritsen et al., 2012; Yuval and O'Gorman, 2020). ~~On the other hand, RCMs are climate models with higher spatial resolution, driven by GCMs results, allowing for a higher topography detail and sub-processes representation; however, RCMs results are also biased with respect to observations at the local scale (Christensen et al., 2008; Hakala et al., 2019; Bayissa et al., 2021; Petrovic et al., 2024).~~ To fulfill the limitations of climate models and allow the assessment of climate change impacts at the local scale a downscaling and bias correction (or bias adjustment) process is required, where downscaling is ~~the procedure transferring large-scale information from GCM to a local scale~~ most commonly associated with spatial disaggregation (Aboitiz et al., 2026) and other processes that allow using GCM information locally and bias correction is the process where climate model results are adjusted to be consistent with observations in terms of magnitude, variability, among other characteristics (Hakala et al., 2019).

Downscaling methods are generally classified into dynamical and statistical methods (Fowler et al., 2025). Dynamical downscaling commonly ~~consists on the~~ use ~~of~~ Regional Climate Models (RCMs), which are climate models with higher spatial resolution, better representing topography and physical processes ~~subprocesses~~ than GCMs in a limited area (Christensen et al., 2008; Hakala et al., 2019; Bayissa et al., 2021; Petrovic et al., 2024), driven by GCM results as boundary conditions. Statistical downscaling consists of~~n~~ building empirical relationships between GCM or RCM results and local climate variables (Hakala et al., 2019; Fowler et al., 2025), including transfer functions, weather typing, weather generators, among others. Bias correction can also be considered a statistical downscaling method as it, in some cases ~~inherently~~ closes the spatial scale gap. It is worth noting the emergence of hybrid approaches, combining dynamical and statistical downscaling methods, and machine learning approaches (Walton et al., 2015; Slater et al., 2023; Soares et al., 2024) for downscaling.

Several bias correction methods of climatic variables have been developed to ~~(or bias adjustment) methods have been developed to fulfill the limitations of climate models and~~ allow the assessment of climate change impacts at the local scale (e.g. Maraun et al., 2010; Teutschbein and Seibert, 2012; Gutiérrez et al., 2013; Maraun, 2016, Bedia et al., 2020). Bias correction methods range from simply~~e~~ adjusting~~tments~~ ~~to~~ the mean to sophisticated ~~through~~ adjustments ~~to~~ of the distribution (aiming for several statistical moments), from univariate to multivariate, and from time-invariant to trend-preserving. One approach that stands out is that of quantile mapping, given its easy implementation, low need of calibration parameters, low computational cost, and applicability to several climate variables (e.g., temperature and precipitation) (Andres et al., 2014; Kim et al., 2016).

The ~~classical~~ standard Quantile Mapping method (QM; Wood et al., 2002) removes the climate model bias by matching the historical model distribution to that of the observed climate. Then the same correction is applied to the projected period. This implies a time-invariant assumption where the transfer function obtained with information from the historical or reference period is valid for the projected periods (Christensen et al., 2008; Maraun, 2012; Themeßl et al., 2012; Brekke et al., 2013; Chen et al., 2013). Several studies have raised concerns towards this stationarity assumption of the QM method and its capability to preserve the climate models signal (Hagemann~~Hageman~~, 2011; Ehret et al., 2012; Maurer et al. 2013; Maurer and Pierce, 2014; Cannon et al., 2015; Pierce et al., 2015; Hnilica et al., 2017; Chadwick et al., 2018; Chadwick et al, 2023). As a response to this issue, several authors have developed trend-preserving versions of the quantile mapping approach focusing on preserving the raw changes or signals of the climate models (e.g. Bürger et al., 2013; Cannon et al., 2015; Switanek et al., 2017; Chadwick et al., 2023).

Some of the trend-preserving methods are: the Detrended Quantile Mapping (DQM; Bürger et al., 2013), developed to preserve the projected changes in the mean; the Quantile Delta Mapping (QDM; Cannon et al., 2015), developed to preserve the projected changes in the quantiles; the Scaled Distribution Mapping (SDM; Switanek et al., 2017), developed to preserve the likelihood of events and changes in quantiles; and the Unbiased Quantile Mapping (UQM; Chadwick et al., 2023), developed to preserve the changes in statistical moments: mean and standard deviation. Several studies have shown that these modified versions outperform the ~~classical~~ standard QM method (e.g. Cannon et al., 2015; Switanek et al., 2017; Chadwick et al., 2023). Currently, the QM method and its trend-preserving variants are widely used in several areas of analysis (Fowler et al., 2007; Cannon, 2017; Portalanza et al., 2024).

Another relevant dimension for bias correction methods is the capability to capture local interrelations or correlations between variables or in space like the MBCn (Cannon, 2017) and the ISIMIP3BASD (Lange, 2019), which are based on trend-preserving variations of the QM, also incorporating the relationships between variables (or points in space). These approaches are relevant to account for compound events like hot-dry conditions, precipitation-temperature co-variations in snowpack modeling, interrelations that univariate methods do not (Meyer et al., 2019; Menapace et al., 2025). However, while successfully capturing local correlations between variables, multi-variable bias correction methods may fail to capture local conditions and raw GCM trends (Aedo-Quililongo and Chong, 2024), currently existing a trade-off in the representativity of these methods.

Despite the efforts over the last decades, it has been challenging to determine the most robust bias correction strategy, with no consensus regarding the most adequate for each specific condition, adding a layer of uncertainty to the climate change impacts assessment (Hakala et al., 2019; Lafferty and Sriver, 2023; Fowler et al., 2025). Other uncertainties in the climate change impacts assessment, before the assessment model (~~for example~~~~example~~e.g., a hydrological model), include: 1) GCM structure, spatio-temporal resolution and parametrization, leading to different responses to the same radiative forcing; 2) RCM boundary conditions, which may be inherited from GCM results; 3) scenarios, which define a broad range of potential drivers of GCMs; 4) internal variability, related to the chaotic nature of climate and represented by small perturbations of the GCM or RCM initial conditions; and 5) downscaling, specifically in the spatial disaggregation.~~the big number of methods to achieve~~ In many cases, downscaling and bias correction can be the main source of uncertainty, with heterogeneity~~ously~~ ~~distributed relevance~~ across space, time and climate metrics (Lafferty and Sriver, 2023).

Lafferty and Sriver (2023) ~~work~~ compares uncertainty contribution of GCM, scenario, internal variability and downscaling and bias correction at the global scale and suggest that downscaling and bias correction contributes around ~~a~~ 30% of total uncertainty for annual and extreme precipitations between 2020 and 2099, with higher contributions in areas with complex topographies. ~~where~~ The higher uncertainty in complex topography is caused by atmospheric dynamics ~~may not~~ being not w~~s~~ell represented ~~because of~~ by GCM due to coarse-resolution, while internal variability contributes around 50 to 60%. On the other hand, downscaling and bias correction contribution to uncertainty for temperatures contributes around 40% in the short term and decreases in time as GCM and scenario contribution increases. This means that choosing a bias correction method is not a simple decision and should be considered carefully to avoid leading to decisions that are not robust as different downscaled and bias correction methods can result in different products and, in consequence, different hazard characterizations (Hanel et al., 2017; Aedo-Quililongo, 2024; Aedo-Quililongo and Chong, 2024). ~~It is worth noting that there is no consensus regarding the quantile mapping method that is the most adequate for each specific condition. However, as Hakala et al. (2019) detail, the climate change impacts assessment is surrounded by uncertainties in each of the assessment steps, including the GCMs drivers and construction, the downscaling process (RCMs and/or bias correction), and the tool to assess the specific risk (e.g. hydrological models). Thus, choosing a bias correction method should not be a simple decision as it could add additional uncertainty or bias to the related assessment (Hanel et al., 2017; Aedo-Quililongo, 2024; Aedo-Quililongo and Chong, 2024).~~ Due to the complexity of these methods and application strategies, there is a need for development of benchmarks and comparison of different methods to ~~Hence, there is the need for tools to implement several quantile mapping methods to compare their results and~~ discern the most appropriate for each specific condition (~~Polasky et al., 2023~~Fowler et al., 2025).

Freely available global and regional downscaled datasets prevent ~~avoid~~ users or stakeholders ~~to~~from performing a downscaling and bias correction process in their analysis. Examples of global downscaled products are the Global Downscaled Projections for Climate Impacts Research (GDPCIR; Gergel et al., 2024) or the NASA Globaly Daily Downscaled Projections CMIP6 (NEX-GDDP-CMIP6; Thrasher et al., 2022), and examples of regional downscaled products are the Climate Change Dataset for Brazil (CLIMBra; Ballarin et al., 2023) or the High-resolution climate projection dataset based on CMIP6 for Perú and Ecuador (BASD-CMIP6-PE; Fernandez-Palomino et al., 2024). These products consider global or regional products as surrogate of observations and consider both a specific reference period and strategy for the downscaling and bias correction process. Additionally, even if validated, at local scale, the extent of the validation may not be appropriate for all uses and need to be analyzed with caution (Lanzante et al., 2018; Aedo-Quililongo and Chong, 2024). As such, these downscaled products are a good alternative for a preliminary or regional analysis, while for local studies both the analysis of different methods and the application of downscaling and bias correction must be part of the workflow to reduce the aforementioned biases and uncertainties.

Several efforts have been made to provide users with different versions of the quantile mapping approach for their climate change impacts assessment workflow. The publicly available coding packages for bias correction of climatic variables based on the approach have the following characteristics: 1) the existing coding packages are available mainly in R (MBC, https://github.com/cran/MBC; downscaleR, https://doi.org/10.5281/zenodo.5070432; snowQM, https://doi.org/10.5281/zenodo.10257951), Python (pyCAT, https://github.com/wegener-center/pyCAT; xclim, https://doi.org/10.5281/zenodo.18352200; ibicus, https://doi.org/10.5281/zenodo.8101842) and MATLAB (GCD, https://github.com/thomasmosier/GCD), and a few in Julia (ClimateTools.jl, https://doi.org/10.5281/zenodo.8047681), NCAR programming language (QQmap, https://doi.org/10.5281/zenodo.4009101), Excel (ClimateScenarioAnalysisToolbox, https://github.com/mxgiuliani00/ClimateScenarioAnalysisToolbox), among others; 2) ~~there used to be a research lab with a package available in two programming languages, but one of these is no longer available; hence,~~ at this moment, none of the explored tools is available in more than one programming language with the exact structure and functionality to allow a unique framework for different user needs or capabilities; 3) the more sophisticated packages are usually frameworks, like the xclim (Bourgault et al., 2023) and downscaleR (Iturbide et al., 2019), that aim for a more complete tool that escapes from the scope of users that only require the downscaling process in their workflow; thus, a deep understanding of the workflow and more structured inputs are needed, which could lead to avoid using that package; and 4) except for the MBC package, which includes several multivariate versions of quantile mapping, but not several trend-preserving bias correction methods, to the best of our knowledge, no package includes more than three contrasting trend preserving bias correction methods to systematically compare their performance in a simple and user-friendly framework, comparison needed to reduce uncertainties regarding bias correction (Fowler et al., 2025). Thus, there is the need for comprehensive coding packages that implement and allow comparing a wide range of quantile mapping methods to facilitate the obtention of local climatic projections and discern the most appropriate one for specific applications.

This paper presents and assesses the climQMBC coding package (Aedo-Quililongo et al., 2026), an easy-to-use tool written in Python, R and MATLAB to perform point-to-point bias correction of climatic variables with the QM, DQM, QDM, UQM and SDM methods for daily, monthly and annual time series. This package aims to make available the QM and several trend-preserving QM methods to facilitate the performance assessment of these methods, making it easier to incorporate the bias correction process as a standard part of climate change impact assessment of final users of local climatic projections that might not be directly related to climate sciences. The climQMBC package also includes a summary report utility to easily compare the performance of the implemented methods at the monthly temporal scale, allowing the user to identify the most appropriate method for their specific situation, along ~~to~~with a synthetic tester of bias correction methods. The synthetic tester develops times series to which bias is added (i.e., in the mean and standard deviation) allowing testing bias correction methods, given that one knows the objective time series ~~after~~before the bias correction. ~~a synthetic analysis to complement the comparison based on the climate characteristics, where bias is added in the first two statistical moments (i.e. mean and standard deviation).~~ While application examples for all challenging climatic conditions for bias correction methods are extensive and considering that the climQMBC package has already been used in several studies (Blin, 2023; Martínez et al., 2023; Nguyen et al., 2024; Kassahun et al., 2025; Vicuña et al., 2025; Wang et al., 2025; González-Leiva et al., 2026; Guzmán et al., 2026), ~~F~~for demonstration purposes and as a proposed benchmark, we present an ~~synthetic example~~ analysis based on the developed synthetic ~~analysis~~ tester to show the capabilities of the climQMBC package to systematically bias correct climatic variables and to show the general performance of the implemented methods. Additionally, we show with a few lines of code how to bias correct real data with the climQMBC package and compare the performance of the implemented methods with the summary report utility.

2 Climate change impact assessment

Climate change assessment at local scales based on projections from climate models (GCMs or RCMs) requires a downscaling and/or bias adjustment/correction process to overcome the differences in resolution and biases of climate models and local climate. The bias corrected climatic variables are then used to assess climate change impacts on water availability, crop productivity, wildfire, or heat waves occurrence, among many others. Figure 1 presents a general workflow of a typical assessment study of climate change impacts, highlighting in grey where the climQMBC package is involved.

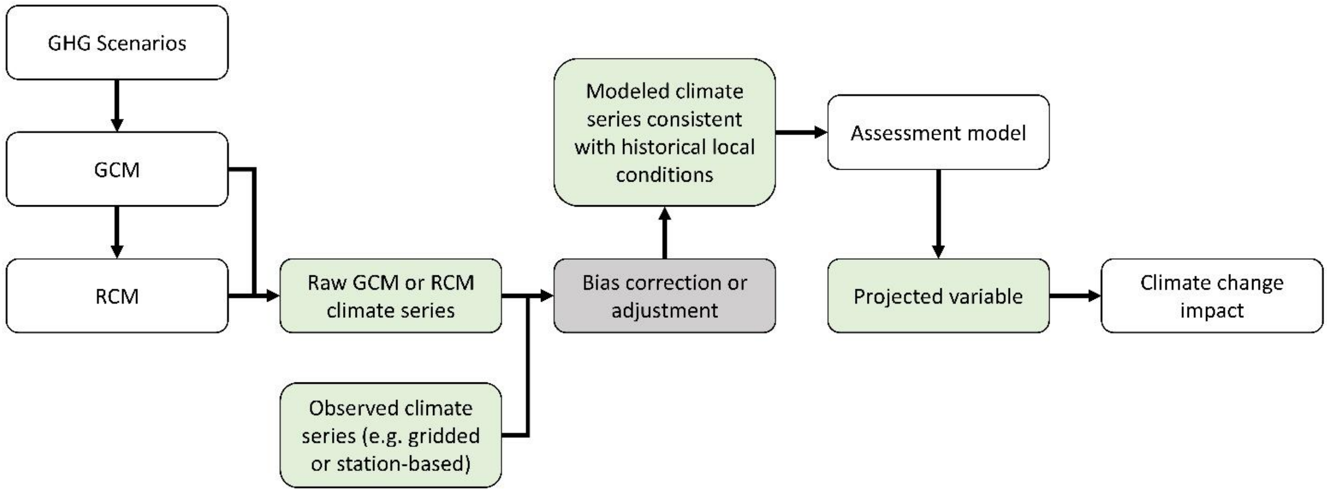


Figure 1: General workflow of a typical assessment study of climate change impacts. Green boxes indicate data or time series. Grey box shows where the climQMBC package is involved in the workflow.

3 climQMBC package

3.1 Package characteristics and functionalities

The climQMBC package is a coding package available for Python, R and MATLAB that contains five point-to-point bias correction methods: QM, DQM, QDM, UQM, and SDM. These methods were implemented for variables associated to both relative (or multiplicative) and absolute (or additive) changes (e.g., precipitations and temperature, respectively). The package supports point or station-based data at daily, monthly or annual time step. The bias correction process and a unified description of methods implemented in the climQMBC package are summarized in the following subsections.

3.1.1 Bias correction process in the climQMBC package

The climQMBC package takes as input (1) an observed and modeled series at one specific point, (2) a flag indicating if the variable allows negative values, (3) a flag indicating if the variable is related to relative or absolute changes (does not apply for the QM method), and (4) a flag indicating if the series frequency is daily, monthly or annual. In the SDM method, a single flag is used to indicate if the variable allows negative values and is related to absolute changes or if the variable does not allow negative values and is related to relative changes, as this method was built explicitly for temperature and precipitation. On the other hand, the modeled series is a single series that merges the historical and future projected periods starting on the same date as the observed series. In the climQMBC package, the trend-preserving bias correction methods capture the transient model changes by a continuum approach using moving windows, as recommended by Chadwick et al. (2018), to capture the non-stationary signals of the GCM. Hence, for each value x of the modeled future period, the climQMBC package considers moving windows with length equal to the number of observed years, ending in the value to be corrected (Fig. 2). The moving window strategy is an improvement implemented in Chadwick et al. (2023), as the original DQM, QDM and SDM consider future data as a single window.

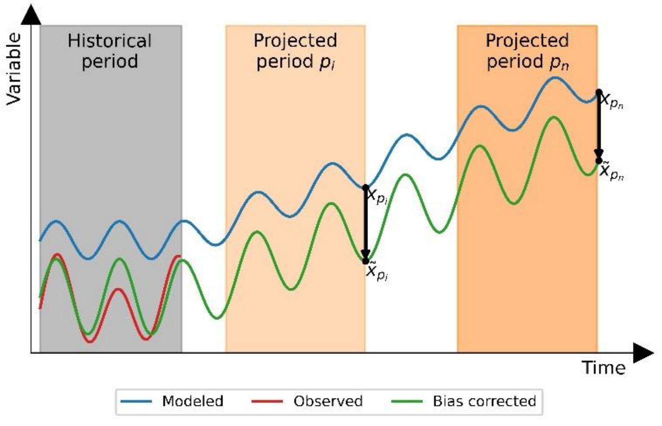


Figure 2: Schematic illustration of the trend-preserving bias correction (i.e., all but the QM) continuum approach considered in the climQMBC package. Each projected period p ends in the value x which is corrected to a value $\tilde{x}$, considering the change rates between the projected and historical period of the modeled series.

As an example of the moving window, if the observed period is 1985-2014, then the moving window will consider 30 years, and when correcting a future year, e.g., 2030, the moving window will cover 2001-2030. Each moving window is used to compute the change rates of the modeled series and the probability distribution function (if applies), and to modify the modeled value x to a corrected value $\tilde{x}$. Periods of 20 to 40 years are recommended (see Appendix A for a brief analysis of the trade-off of selecting larger or smaller windows). Shorter periods might not capture correctly the variable statistics, while longer ones might dampen the signal of the variable or induce a "pseudo-stationary bias" of the transient changes (Chadwick et al., 2023). Note that during the historical period, DQM, QDM and UQM methods become the standard QM method as the change rates are 0 for additive changes and 1 for relative changes (Chadwick et al., 2023).

When the frequency is set to monthly, the bias correction process is performed for each month of the year independently (e.g., each January is corrected based on the information of January in the historical and projected period). If the frequency is set to daily, the bias correction process is performed for each day of the year considering a centered moving window of length defined by the user (e.g., if the window is of 31 days, then the 16th of July uses the days between July 1st to July 31st to compute the required statistics for the bias correction).

The values of the historical period and each projected period are used to fit a probability distribution function. Furthermore, changes in the mean (δ_μ), standard deviation (δ_σ) or quantiles (Δ), depending on each method, between the projected and historical period of the modeled series are computed with the following equations:

$$\delta_\mu = \begin{cases} \dfrac{\overline{x_{m,p}}}{\overline{x_{m,h}}}, & \text{relative change variable} \\ \overline{x_{m,p}} - \overline{x_{m,h}}, & \text{absolute change variable} \end{cases}, \tag{1}$$

$$\delta_{\sigma} = \begin{cases} \frac{\widehat{x_{m,p}}}{\widehat{x_{m,h}}} & \text{, relative change variable} \\ \widehat{x_{m,p}} - \widehat{x_{m,h}}, & \text{absolute change variable} \end{cases}, \tag{2}$$

$$\Delta = \begin{cases} \frac{F_{m,p}^{-1}(\tau_p)}{F_{m,h}^{-1}(\tau_p)} & \text{, relative change variable} \\ F_{m,p}^{-1}(\tau_p) - F_{m,h}^{-1}(\tau_p) & \text{, absolute change variable (all but SDM),} \\ \frac{\widehat{x_{o,h}}}{\widehat{x_{m,h}}}\left(F_{m,p}^{-1}(\tau_p) - F_{m,h}^{-1}(\tau_p)\right), & \text{absolute change variable (SDM)} \end{cases} \tag{3}$$

with

$$\tau_p = F_{m,p}(x), \tag{4}$$

Subindexes o and m refer to the observed and modeled series, respectively, whereas subindexes h and p refer to the historical and projected period, respectively. Accents $\bar{}$ and $\tilde{}$ denote mean and standard deviation, respectively. F and F^{-1} are the cumulative distribution function (CDFs) and its inverse CDF (invCDFs), respectively; and τ is the non-exceedance probability. The climQMBC package assigns a probability distribution function for each day or month and period based on the method of moments and by default selects the distribution with the lowest Kolmogorov-Smirnov test statistic, which is more sensitive to differences in ~~near~~ the center of the distribution and less sensitive to differences in the tails~~by default chooses the distribution considering the one with the lowest error in the Kolmogorov-Smirnov test~~. The exception is the SDM method in which a gamma or normal distribution is fitted by the maximum likelihood method for variables associated with precipitation and temperature, respectively (Switanek et al., 2017). The available distribution functions in the climQMBC package are the Normal, Log-normal, Two-parameter Gamma, Three-parameter Gamma, Log-Gamma, Gumbel, and Exponential. If the Kolmogorov-Smirnov test fails for all the available distributions, the climQMBC package prints a warning and continues with the procedure. Additionally, the user can define the desired distribution function to be used, and the package will estimate the parameter with the method of moments, skipping the Kolmogorov-Smirnov test.

For variables that do not allow negative values, the climQMBC package internally filters those distributions that allow negative values (Normal, Three-parameter Gamma, Gumbel and Exponential distributions). On the other hand, if the series has negative values, the climQMBC package discards those distributions that are strictly positive (Log-normal, Two-parameter Gamma and Log-Gamma distributions).

To ensure stability when computing the invCDFs, τ values are constrained to 0.001<τ<0.999. Additionally, to improve the management of non-precipitation days, the climQMBC package follows the two parameters "random near zero replacement" presented in Chadwick et al. (2023). The first parameter is a threshold that represents physically zero precipitation values ~~values below a threshold representing physically zero precipitation value~~ (1 mm as default, independently of the temporal scale. Hence, 1 mm/day, 1 mm/month and 1 mm/year for daily, monthly and annual precipitation, respectively) so that all values below this threshold are replaced with random numbers that are barely positive. The second parameter is a scaling factor (1/100 as default) of the physically zero precipitation values that defines the magnitude of the "barely positive" replacement values. After performing the bias correction methods, the corrected values below the physically zero precipitation values are replaced with zeros.

As an example of the “random near zero replacement” considering the default values in the climQMBC package, all values bellow 1 mm/day (for daily values) would be replaced with random values between 0 and 1/100 mm/day. If as a result of the trend-preserving procedures one of these values is scaled by a factor of 5, the corrected value would be ~~of~~ 5/100 mm/day and then would be replaced by a zero value at the end of the process. The two parameters scheme ensure that even for unrealistic changes with factors of 10, the replaced values will still be removed after the correction, avoiding these values to create false rain-day values. On the other hand, if all values are below the physically zero precipitation value (or only zero values), the correction would result in no precipitation values.

As such, ~~This procedure, referred to as “random near zero replacement” (Chadwick et al., 2023)~~ this procedure allows fitting distributions and performing bias correction to avoid numerical problems in periods with large number of physically zero precipitation values, or periods with only zero values. Nevertheless, the threshold and the magnitude of the random values should be analyzed to enhance the methods performance depending on the specific conditions. For daily data, the number of wet days is corrected as most models have a drizzling bias that generates rainy days but with intensity lower than observed (Chen et al., 2021). To correct the number of rainy days, in all methods but SDM, an internal threshold representing physically zero precipitation value for the modeled series is fitted to match the number of rainy days of the observed series in the reference period. Then, only rainy days are corrected (values above the physically zero precipitation threshold), but, if the number of rainy days to fit the distributions is less than 30, some of the near-zero replaced values are considered to complete 30 values (these would later be removed as physically zero precipitation values). The SDM has its own method of correcting rainy days, as described in Switanek et al. (2017).

3.1.2 Description of the bias correction methods implemented in the climQMBC package

The bias correction methods available in the climQMBC package are now briefly described using a unified mathematical notation. Their main objectives and corresponding considerations are summarized in Table 1, while Fig. 3 shows a schematic illustration of the procedure performed in the projected periods and the CDF involved, highlighting the main differences among methods. In terms of execution times, considering the implementation in the climQMBC package, the QM method is the fastest, followed by the DQM (twice the QM execution time for annual and monthly data, and 7 times for daily data), being the QDM, SDM and UQM the slowest (around 7 to 10 times the QM execution time for annual data, 22 to 27 times for monthly data and 32 to 39 times for daily data). Appendix B presents further details of the execution time comparison.

Table 1: Summary of the bias correction methods available in the climQMBC package

Method	Abbreviation	Summary	Reference
Quantile Mapping	QM	Considers a transfer function built with information from the observed and modeled historical period and applies it to the modeled series. Corrected	Wood et al. (2002)

		series values will match its CDF with the observed series, maintaining the model occurrence of the events and limiting future scenarios exceeding this range.	
Detrended Quantile Mapping	DQM	Considers the same transfer function as in the QM method but applies it to a detrended modeled series, to then reimpose the trends. Corrected series tends to preserve changes in the mean of the modeled series.	Bürger et al. (2013)
Quantile Delta Mapping	QDM	The QM method is applied to the modeled future to detrend the modeled values. Then, the change between the future value and its respective quantile in the modeled historical period is used to rescale the series. Corrected series tends to preserve changes in the quantiles of the modeled series.	Cannon et al. (2015)
Unbiased Quantile Mapping	UQM	The QM method is applied to the modeled future, and the changes of the moments (e.g. mean and standard deviation) are imposed in the observed CFD to scale the transfer function. Corrected series tends to preserve changes in the moments of the modeled series.	Chadwick et al. (2023)
Scaled Distribution Mapping	SDM	Scales the observed CDF by the ratio between the future and historical modeled CDFs to correct the recurrence interval. A scaling factor like in the QDM method is then applied to the scaled probability distribution. Finally, the corrected values are allocated accordingly to likelihood of events. For precipitation, the SDM method additionally adjusts rain-day frequency. Corrected series tends to preserve changes in quantiles (as in the QDM method) and likelihood of events.	Switanek et al. (2017)

Note: In the DQM, QDM, UQM and SDM methods the future period is divided in moving windows as an improvement implemented in Chadwick et al. (2023) to prevent pseudo-stationary biases.

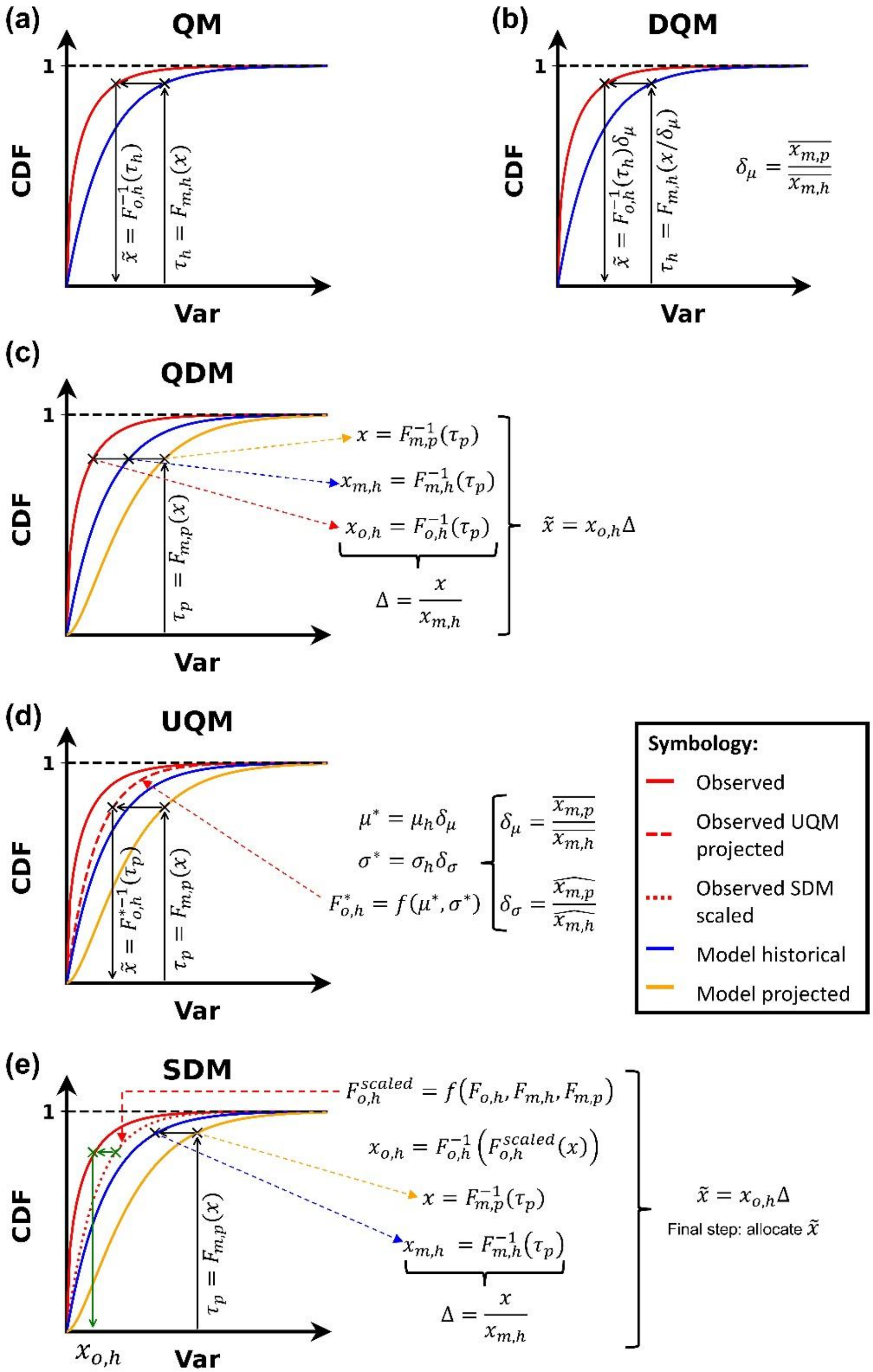


Figure 3: Schematic illustration of the main processes computed in the (a) QM, (b) DQM, (c) QDM, (d) UQM and (e) SDM methods available in the climQMBC package, at each projected period for variables related to relative changes (e.g. precipitation). For variables related to absolute changes (e.g. temperatures) the deltas (*δ* and *Δ*) are computed and applied additively, rather than multiplicatively.

Quantile Mapping (QM):

The QM method (Fig. 3a) bias corrects the modeled series by evaluating the modeled values (x) in the CDF of the modeled historical period ($F_{m,h}$) to get the associated probabilities of each value and then evaluates them in the invCDF of the observed historical period ($F_{o,h}^{-1}$) to obtain a vector of corrected values ($\tilde{x}$). This method aims to correct each value of the modeled series by matching its quantile computed on its historical period with the quantile of the observed values, and is formulated as:

$$\tilde{x} = F_{o,h}^{-1}\big(F_{m,h}(x)\big), \tag{5}$$

The QM method assumes that the correction process is time-invariant and in some cases may result in distortion of the raw model projected change rates. In most cases distortions are caused by projected climate being outside the range of historical modeled climate.

Detrended Quantile Mapping (DQM):

The DQM method (Fig. 3b) detrends the modeled series by computing the projected change in the mean of the modeled series for each rolling window (δ_μ, Eq. 1) and scaling the last value of each rolling window (x) to a range within the historical values. Then, the same transfer function described in the QM method is applied to the detrended values. Finally, the bias corrected values are rescaled multiplicatively or additively to obtain the corrected value ($\tilde{x}$). The procedure for each projected period is formulated as follows:

$$\tilde{x} = \begin{cases} F_{o,h}^{-1}\left(F_{m,h}(x/\delta_\mu)\right)\delta_\mu\,, & \text{relative change variable} \\ F_{o,h}^{-1}\left(F_{m,h}(x-\delta_\mu)\right)+\delta_\mu\,, & \text{absolute change variable} \end{cases}, \tag{6}$$

Quantile Delta Mapping (QDM):

In the QDM method (Fig. 3c) the non-exceedance probability τ_p (Eq. 4) of each value x to be corrected is calculated based on its associated moving window. Then, τ_p is evaluated in the observed invCDF to obtain the modeled projected values after correction, $x_{m,p\to o,h}$, as follows:

$$x_{m,p\to o,h} = F_{o,h}^{-1}(\tau_p), \tag{7}$$

given that τ_p depends on a moving window, $x_{m,p\to o,h}$ does not have the model trend. Finally, the percentile change Δ (Eq. 3) of the projections is estimated as the ratio between the projected climate and its respective quantile in the historical modeled period. Then, Δ is applied to $x_{m,p\to o,h}$ to obtain the corrected value $\tilde{x}$ as follows:

$$\tilde{x} = \begin{cases} x_{m,p\to o,h}\Delta\,, & \text{relative change variable} \\ x_{m,p\to o,h}+\Delta\,, & \text{absolute change variable'} \end{cases} \tag{8}$$

When Δ is estimated for relative changes, it can get indetermined as a result of a division by a near-zero value. To avoid this, the climQMBC package considers an upper limit for Δ when dividing by a physically no-precipitation value. This strategy mimics the implementation of the QDM method in the MBC package (Multivariate Bias Correction of Climate Model Outputs; Cannon, 2017), available in CRAN.

Unbiased Quantile Mapping (UQM):

In the UQM method (Fig. 3d) the projected change in the moments (i.e., mean and standard deviation) of the modeled series is first computed for each moving window (δ_μ and δ_σ, respectively; Eq. 2 and Eq. 3, respectively). Then, these change rates are used to project the observed historical mean (μ_h) and standard deviation (σ_h) into the future to obtain future values μ^* and σ^* as follows:

$$\mu^* = \begin{cases} \mu_h \delta_\mu, & \text{relative change variable} \\ \mu_h + \delta_\mu, & \text{absolute change variable'} \end{cases} \tag{9}$$

$$\sigma^* = \begin{cases} \sigma_h \delta_\sigma, & \text{relative change variable} \\ \sigma_h + \delta_\sigma, & \text{absolute change variable'} \end{cases} \tag{10}$$

Subsequently, μ^* and σ^* are used to produce future projections of the selected distribution for the historical observed values, $F^*_{o,h}$, as follows:

$$F^*_{o,h} = f(\mu^*, \sigma^*), \tag{11}$$

Finally, the non-exceedance probability of the modeled values x in each moving window (τ_p, Eq. 4) are evaluated in the inverse projected distribution invCDF, $F^{*-1}_{o,h}$, to obtain the bias corrected value $\tilde{x}$ as follows:

$$\tilde{x} = F^{*-1}_{o,h}(\tau_p), \tag{12}$$

Scaled Distribution Mapping (SDM):

The SDM method (Fig. 3e) considers several steps (Switanek et al., 2017; Chadwick et al., 2023). Note that they are presented in a sequence different than the originally depicted by Switanek et al. (2017), with the intention ~~of~~ to better illustrate the implications and structure of the method.

For temperatures, the first step is to linearly detrend both the observed and modeled series for each period. For precipitation, the first step is to remove all non-precipitation values (1~~<~~ mm as default, independently of the temporal scale. Hence, 1 mm/day, 1 mm/month and 1 mm/year for daily, monthly and annual precipitation, respectively~~1 mm/day as default~~) to account only for the rain values in each period. Then, the expected number of rain events for the corrected series (RD^{scaled}) is calculated by scaling multiplicatively (linear scaling) the number of rain events of each moving window ($RD_{m,p}$) with the ratio between the frequency of rainy days in the observed historical period ($RD_{o,h}/TD_{o,h}$) and in the modeled historical period ($RD_{m,h\cancel{mh}}/TD_{m,h}$) as follows:

$$RD^{scaled} = RD_{m,p} \times \frac{RD_{o,h}/TD_{o,h}}{RD_{m,h\cancel{mh}}/TD_{m,h}}, \tag{13}$$

where TD represents the total number of days, including rain and non-rain days. Hereafter, the SDM method is applied to a vector (bold notation) of sorted detrended temperature or non-zero precipitation values, to which the normal and gamma probability distribution functions are fitted by the maximum likelihood method, respectively.

In the second step the sorted vectors of the observed historical ($\boldsymbol{x}_{o,h}$), modeled historical ($\boldsymbol{x}_{m,h}$) and modeled projected period or moving window ($\boldsymbol{x}_{m,p}$) series are evaluated in their respective fitted CDF to obtain their associated probability vectors ($F_{o,h}(\boldsymbol{x}_{o,h})$, $F_{m,h}(\boldsymbol{x}_{m,h})$, $F_{m,p}(\boldsymbol{x}_{m,p})$). In the climQMBC package, each CDF is fitted considering a gamma distribution for positive precipitation and normal distribution for detrended temperature, using the maximum likelihood, following the SDM implementation in the pyCAT package. These probability vectors are used to estimate each recurrence intervals $\boldsymbol{RI}_{i,j}$ as follows:

$$\boldsymbol{RI}_{i,j} = \begin{cases} \frac{1}{1-F_{i,j}(x_{i,j})} & \text{, for precipitation} \\ \frac{1}{0.5-|0.5-F_{i,j}(x_{i,j})|}, & \text{for temperature} \end{cases} \quad ; \quad i = \{m, o\}, j = \{h, p\}\,, \tag{14}$$

In the third step, the recurrence interval of the observed historical series is linearly scaled by the ratio of the recurrence intervals of the model projected period and the model historical period (i.e., capturing the changes in recurrence intervals in the model) to build a scaled recurrence interval $\boldsymbol{RI}^{scaled}$ as follows:

$$\boldsymbol{RI}^{scaled} = \max\left(1, \boldsymbol{RI}_{o,h} \circ \boldsymbol{RI}_{m,p} \circ \frac{1}{\boldsymbol{RI}_{m,h}}\right) \;, \tag{15}$$

$\boldsymbol{RI}^{scaled}$ is built using the Hadamard product for vectors (i.e., multiplying element by element in a vector and preserving their dimensions), which requires equally sized vectors. Hence, prior to computing the precipitation recurrence intervals, the probability vectors are linearly interpolated to match the number of rain day events of the modeled projected period. The linear interpolation stretches or contracts the series to match the number of values in the modeled vector, keeping the original range and distribution of values.

In the fourth step, $\boldsymbol{RI}^{scaled}$ is used to go back into a scaled probability vector $\boldsymbol{F}_{o,h}^{scaled}$, which represents the scaled distribution function vector of the observed historical period as follows:

$$\boldsymbol{F}_{o,h}^{scaled} = \begin{cases} 1 - \frac{1}{\boldsymbol{RI}^{scaled}} & \text{, for precipitation} \\ \left(1 - \frac{1}{\boldsymbol{RI}^{scaled}}\right) \times sgn\left(F_{o,h}(\boldsymbol{x}_{o,h}) - 0.5\right) & \text{, for temperature} \end{cases}, \tag{16}$$

In the fifth step, $\boldsymbol{F}_{o,h}^{scaled}$ is sorted in descending order for precipitation, then, for both precipitation and temperature, it is evaluated in the invCDF of the observed historical period to get the associated value in the historical observed series domain $\boldsymbol{x}_{scaled\to o,h}$ as follows:

$$\boldsymbol{x}_{scaled\to o,h} = F_{o,h}^{-1}\left(\boldsymbol{F}_{o,h}^{scaled}\right), \tag{17}$$

In the sixth step, the vector with the projected change in the quantiles $\boldsymbol{\Delta}$ (Eq. 3. for SDM; similar to that of the QDM) is computed for all the values of the projected series and it is used to scale the vector computed in the previous step as follows:

$$\widetilde{\boldsymbol{x}}_{temp} = \begin{cases} \boldsymbol{x}_{scaled\to o,h}\boldsymbol{\Delta}, & \text{for precipitation} \\ \boldsymbol{x}_{scaled\to o,h} + \boldsymbol{\Delta}, & \text{for temperature} \end{cases}, \tag{18}$$

For precipitation, if the number of rainy days of the projected period ($RD_{m,p}$) is greater than the expected number of rain events (RD^{scaled}), the frequency of rain events is adjusted by linearly interpolating $\widetilde{\boldsymbol{x}}_{temp}$ to a length equal to RD^{scaled}.
As a final step for precipitation, each value of the $\widetilde{\boldsymbol{x}}_{temp}$ is placed in the location of the rainy values of the modeled projected temporal series, from the highest to the lowest. In case where RD^{scaled} is greater than $RD_{m,p}$, the remaining values are replaced with zeros. For temperatures the final step is to place back the $\widetilde{\boldsymbol{x}}_{temp}$ values in the correct temporal location based on the modeled time series, process that is done in descending order. Then, the linear trend of the modeled projected temporal series is added back. Because the climQMBC package considers a continuum approach to capture transient changes from the models, the corrected value $\tilde{x}$ of the projected period corresponds to the value of the last year of the corrected series; then, the complete procedure is repeated for the following year.

3.2 climQMBC package example and report function

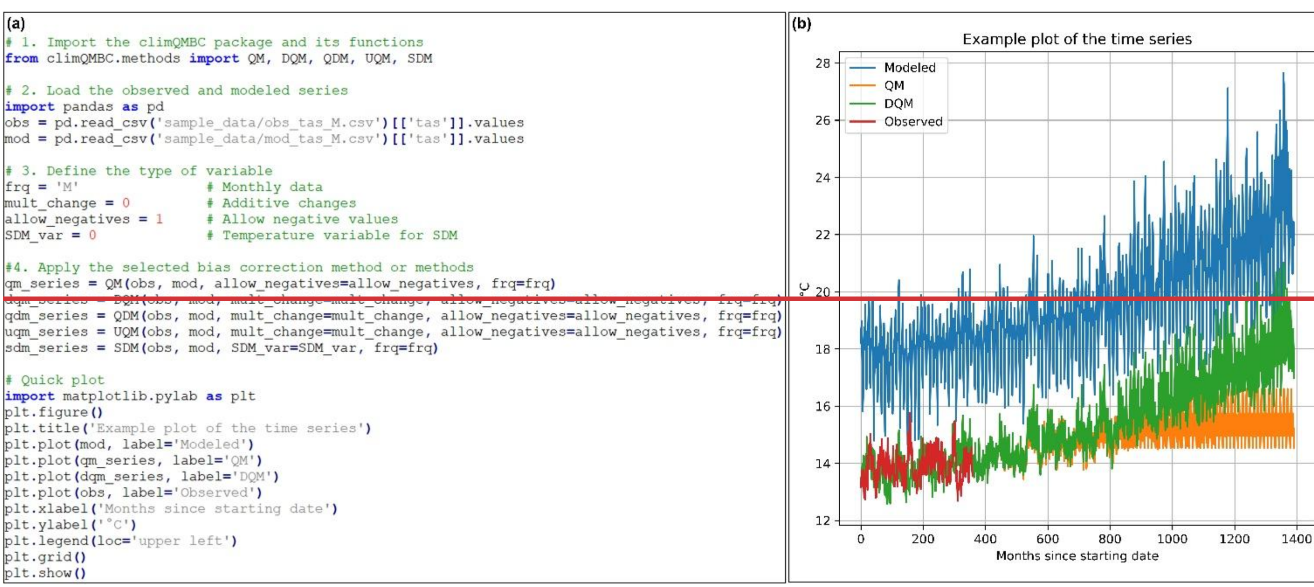
This section is based on an example that uses ~~the~~ sample data available in the package repository (https://github.com/saedoquililongo/climQMBC/tree/main/Sample_data). The sample data considers daily precipitation and mean temperature for the coordinate 4.64N and 75.48W, located near the Cocora Valley, Quindío, Colombia. The observational ERA5-Land reanalysis series was developed by ~~corresponds to the ERA5-Land reanalysis from~~ the European Centre for Medium-Range Weather Forecasts (Muñoz-Sabater et al., 2021) and for the modeled series considers the GCM MPI-ESM1-2-HR (Müller et al., 2018), ~~run by~~ under the SSP 5-8.5 scenario of the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (Eyring et al., 2016). ~~For both t~~The observational and modeled time series were extracted with~~,~~ the nearest point method ~~was used to extract the time series~~.

As an easy-to-use tool, the climQMBC package requires familiarity with importing and processing arrays, matrices, or vectors. Four steps are needed for its use: (1) Import the climQMBC package and its functions; (2) load the observed and modeled series (the package assumes that both series begin in the same date and that the reference period corresponds to the period of the observed series); (3) define the type of variable (i.e. negative or strictly non-negative, and variable associated with additive or multiplicative changes) and data frequency (daily, monthly or annually); and (4) apply the selected bias correction method or methods. These steps can be applied to multiple locations or grids cells to systematically downscale a region of interest (this example is included in the climQMBC tester scripts available in GitHub for Python, R and MATLAB at https://github.com/saedoquililongo/climQMBC/blob/main/Python/climQMBC_tester_Python.py, https://github.com/saedoquililongo/climQMBC/blob/main/R/climQMBC_tester_R.R and https://github.com/saedoquililongo/climQMBC/blob/main/Matlab/climQMBC_tester_Matlab.m, respectively). Also, different bias correction methods can be used and compared to identify the better ones in preserving the changes from raw GCM/RCMs for the specific conditions under study. Detailed demonstration of the capabilities of the package is provided in an example notebook available in its repository (https://github.com/saedoquililongo/climQMBC/blob/main/Example_notebook.ipynb).

Listing 1 illustrates a simple Python code to apply the five bias correction methods available in the climQMBC package to ~~the~~ temperature sample data available in the repository and then plot the QM and UQM results ~~for the results of the QM and UQM methods~~. Figure 4 presents the plots of Listing 1 and a close-up to the reference period and to ~~a~~ the period ~~where~~ when ~~the~~ QM and UQM ~~methods~~diverge, showing how the corrected series match the magnitude and variability of the observed series (Fig. 4b) and the potential differences in their results (Fig. 4c) which could lead to bias in a further climate change impacts assessment. More specifically, Fig. 4a shows the inability of the QM method to capture trends when the projected values are outside the range of the values of the reference period. ~~Figure 4 illustrates a simple Python code with an example, and the corresponding plotted results, based on the sample data available in the package repository (https://github.com/saedoquililongo/climQMBC/tree/main/Sample_data). Detailed demonstration of the capabilities of the package is provided in an example notebook available in its repository (https://github.com/saedoquililongo/climQMBC/blob/main/Example_notebook.ipynb).~~

```python
# 1. Import the climQMBC package and its functions
from climQMBC.methods import QM, DQM, QDM, UQM, SDM

# 2. Load the observed and modeled series
import pandas as pd
obs = pd.read_csv('sample_data/obs_tas_M.csv', index_col=0, parse_dates=True)
mod = pd.read_csv('sample_data/mod_tas_M.csv', index_col=0, parse_dates=True)

## Aggregate to annual scale and format to column vector
obs = obs[['tas']].resample('YE').mean().values
mod = mod[['tas']].resample('YE').mean().values

# 3. Define the type of variable (temperature)
frq = 'A'                   # Annual data
mult_change = 0             # Additive changes
allow_negatives = 1         # Allow negative values
SDM_var = 0                 # Temperature variable for SDM

# 4. Apply the selected bias correction method or methods
qm_series  = QM(obs, mod, allow_negatives=allow_negatives, frq=frq)
dqm_series = DQM(obs, mod, mult_change=mult_change, allow_negatives=allow_negatives, frq=frq)
qdm_series = QDM(obs, mod, mult_change=mult_change, allow_negatives=allow_negatives, frq=frq)
uqm_series = UQM(obs, mod, mult_change=mult_change, allow_negatives=allow_negatives, frq=frq)
sdm_series = SDM(obs, mod, SDM_var=SDM_var, frq=frq)

# 5. Plot time series
import matplotlib.pylab as plt
plt.figure()
plt.title('Example plot of the time series')
plt.plot(mod, label='Modeled')
plt.plot(qm_series, label='QM')
plt.plot(uqm_series, label='UQM')
plt.plot(obs, label='Observed')
plt.xlabel('Years since starting date')
plt.ylabel('°C')
plt.legend(loc='upper left')
plt.grid()
plt.show()
```

Listing 1: Python script showing the usage of the climQMBC package.



(a)

```python
# 1. Import the climQMBC package and its functions
from climQMBC.methods import QM, DQM, QDM, UQM, SDM

# 2. Load the observed and modeled series
import pandas as pd
obs = pd.read_csv('sample_data/obs_tas_M.csv')[['tas']].values
mod = pd.read_csv('sample_data/mod_tas_M.csv')[['tas']].values

# 3. Define the type of variable
frq = 'M'                # Monthly data
mult_change = 0          # Additive changes
allow_negatives = 1      # Allow negative values
SDM_var = 0              # Temperature variable for SDM

#4. Apply the selected bias correction method or methods
qm_series = QM(obs, mod, allow_negatives=allow_negatives, frq=frq)
dqm_series = DQM(obs, mod, mult_change=mult_change, allow_negatives=allow_negatives, frq=frq)
qdm_series = QDM(obs, mod, mult_change=mult_change, allow_negatives=allow_negatives, frq=frq)
uqm_series = UQM(obs, mod, mult_change=mult_change, allow_negatives=allow_negatives, frq=frq)
sdm_series = SDM(obs, mod, SDM_var=SDM_var, frq=frq)

# Quick plot
import matplotlib.pylab as plt
plt.figure()
plt.title('Example plot of the time series')
plt.plot(mod, label='Modeled')
plt.plot(qm_series, label='QM')
plt.plot(dqm_series, label='DQM')
plt.plot(obs, label='Observed')
plt.xlabel('Months since starting date')
plt.ylabel('°C')
plt.legend(loc='upper left')
plt.grid()
plt.show()
```

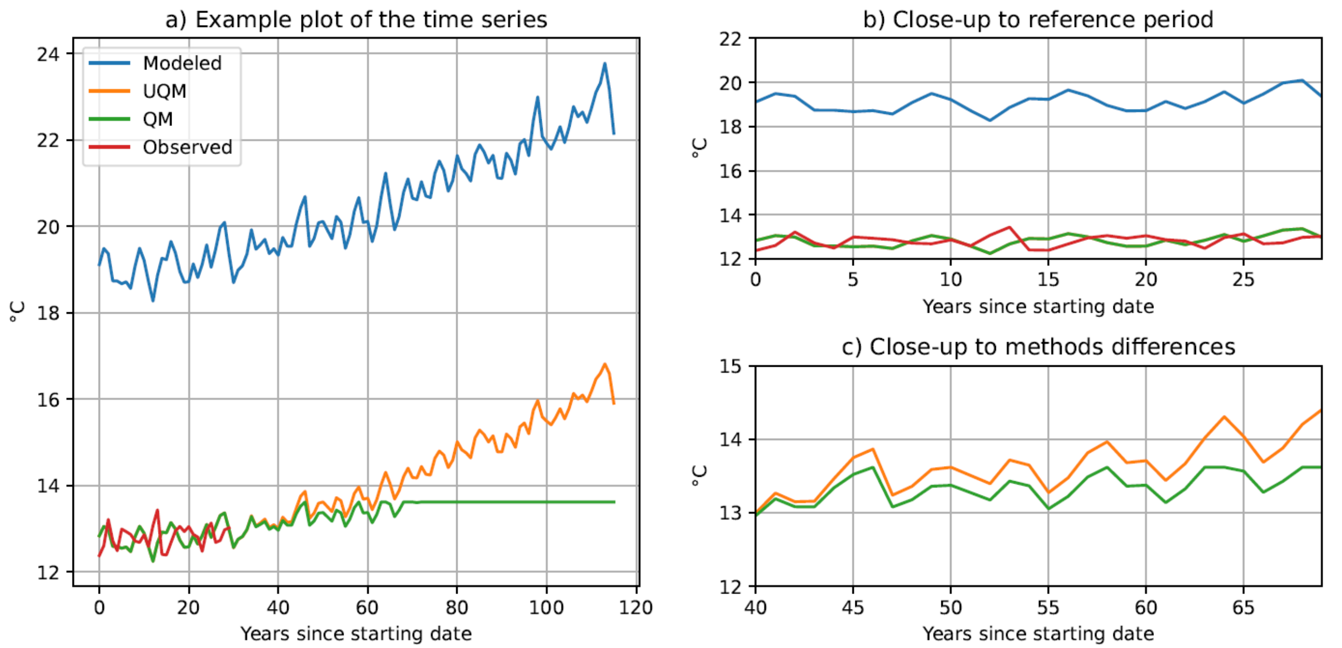


Figure 4: Example plot of the example Python code presented in Listing 1. a) Complete time series; b) close up to the reference period (~~it is worth noting~~ note that, in the reference period, ~~the~~ UQM ~~method~~ is the same as ~~the~~ QM ~~method~~); and c) close-up to ~~a~~ the period ~~where~~ when ~~the~~ UQM and QM ~~method show different results~~diverge.~~Python script showing the usage of the climQMBC package (a) and plotted results (b).~~

For monthly values, the climQMBC package has a summary report function to compare the performance or assess trade-offs between bias correction methods based on the preservation of the changes in the mean and standard deviation of the modeled series. This function outputs two reports: (1) a summary table with the overall performance of the bias correction methods in the historical period and in future periods for the mean and standard deviation (Fig. 5), and (2) a set of figures to visualize the main differences among methods in terms of monthly values and quantiles (Fig. 6 and Fig. 7). The report function requires the same inputs as the bias correction methods functions. Additional inputs of the report functions are the bias correction methods and future periods to be displayed in the summary tables and figures.

The structure of the summary tables for the mean and standard deviation generated by the report function of the climQMBC package are the same (Fig. 5). The first row of the summary tables shows the ratio or difference in the mean value (or standard deviation) between the QM bias corrected and observed historical series. A perfect match between the corrected and observed value would result in a value of 1 for multiplicative change variables and 0 for the additive change variables. In the example shown in Fig. 5a (Mean Table) for precipitation, the ratio between the annual mean values of the QM method and the modeled series in the historical period is 1.000, showing a perfect match between the mean of the observations and corrected series in the reference period. ~~1.002, showing an overestimation of the mean in 0.2 %.~~

Mean Table:

Ratio between the future and historical period

Ratio between QM and Observed

Modeled (GCM)

Quantile mapping methods

QM performance in historical period: 1.002				
	Future	2035	2065	2080
Modeled	0.761	0.893	0.749	0.657
QM	0.697	0.862	0.686	0.568
DQM	0.753	0.887	0.741	0.655
QDM	0.753	0.885	0.741	0.658
UQM	0.745	0.882	0.728	0.645
SDM	0.750	0.880	0.747	0.647

~~Ratio between the projected periods~~ and the historical period

Standard deviation Table:

QM performance in historical period: 0.989				
	Future	2035	2065	2080
Modeled	0.843	0.945	0.830	0.737
QM	0.836	0.948	0.841	0.708
DQM	0.871	0.965	0.874	0.775
QDM	0.858	0.957	0.857	0.756
UQM	0.866	0.960	0.865	0.770
SDM	0.878	0.975	0.887	0.773

Mean Table:

Ratio between the future and historical period

Ratio between QM and Observed

Modeled (GCM)

Quantile mapping methods

QM performance in historical period: 1.000			
	Future	2030	2060
Modeled	0.920	0.975	0.892
QM	0.947	0.993	0.922
DQM	0.933	0.992	0.903
QDM	0.933	0.990	0.905
UQM	0.931	0.991	0.900
SDM	0.923	0.991	0.889

Ratio between the projected periods and the historical period

Standard deviation Table:

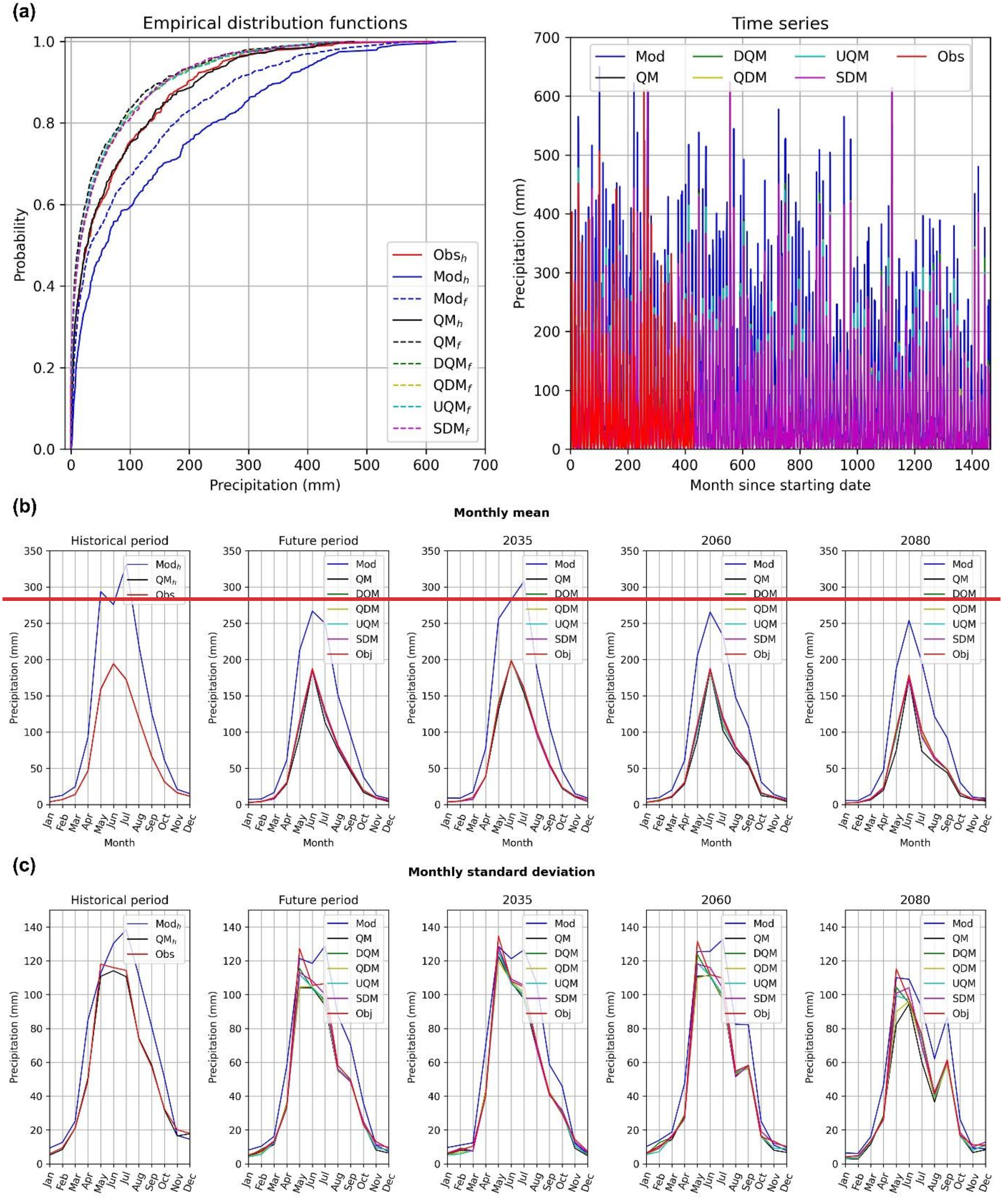
QM performance in historical period: 1.004			
	Future	2030	2060
Modeled	0.916	0.951	0.888
QM	1.046	1.045	1.038
DQM	1.077	1.042	1.081
QDM	1.063	1.042	1.044
UQM	1.081	1.048	1.083
SDM	1.117	1.053	1.128

Figure 5: Example of the summary tables generated by the report function of the climQMBC package, applied to a precipitation case. The Mean and Standard Deviation Tables show the summary report for the annual mean and standard deviation, respectively. First row presents the performance based on the ratio between the annual mean of the bias corrected and modeled series in the historical period. Second row presents headers for different periods. Third row and below presents the ratio between the periods and the historical period for each series presented in the first column. Values under the header 'Future, shows the ratio between the

complete future period and the historical period, the third and following columns show the ratio between the projected period, centered in the year of the corresponding header (defined by the user), and the historical period.

The second row of the summary table shows the header of the periods that are compared with the historical period. This comparison considers information exclusively of each series (e.g. projected period and historical information of the UQM corrected series). The third row and below shows the ratio or difference between the future period, based on the header, and the historical period. In these rows the first column indicates the series considered (modeled series, QM corrected series, DQM corrected series, …), the second column, under the header 'Future, shows the ratio or difference between the complete future period and the historical period, the third and following columns show the ratio or difference between the projected period, centered in the year of the corresponding header (defined by the user), and the historical period. For a bias correction model to be good, its value (ratio or difference) should be close to the value of the "Modeled" row in each period, both for the mean and standard deviation. For the example shown in Fig. 5b (Standard deviation Table), the model ratio between the future centered in 2060~~2065~~ and historical is 0.888 ~~0.830~~ (a change of -11.2 %~~-17 %~~) whereas the corresponding values for QM, DQM, QDM, UQM and SDM are 1.038, 1.081, 1.044, 1.083 and 1.128 ~~0.841, 0.874, 0.857, 0.865 and 0.887~~, respectively. Hence, all the methods show an increase in the standard deviation, rather than a decrease, being the SDM method with the greater discrepancy.~~underestimate the decrease in the standard deviation, being the QM method the closest one to the modeled series~~. However, for the same future period, the ratio corresponding to changes in the mean for the ~~QM~~ SDM method is the closest ~~largest as compared~~ to the modeled series (~~0.686~~0.889 compared to ~~0.749~~0.892, Mean Table).

The report function also provides a second report with three different figures illustrated in Fig. 6 and Fig. 7. The first figure (Fig. 6~~a~~) shows the cumulative distribution of the modeled and corrected series in the historical and future periods, and the cumulative distribution of the observed series (left side). This figure also shows the observed, modeled, and corrected time series (right side). The second figure (Fig. 7a~~6b~~) shows the seasonal monthly mean of each series in the historical and future projected periods (the same projected periods as in the summary tables). Finally, the third figure (Fig. 7b~~6c~~) shows the same monthly seasonality, but for the standard deviation. In the second and third figures, the objective series is the observed monthly mean or standard deviation, scaled by the raw projected change (scaling factor based on the projected and the historical period) of the uncorrected model. Note that, ideally, the corrected value should capture the projected changes; hence, match the objective value.


(a)
Empirical distribution functions
Time series
Probability
Precipitation (mm)
Month since starting date
Obs$_h$
Mod$_h$
Mod$_f$
QM$_h$
QM$_f$
DQM$_f$
QDM$_f$
UQM$_f$
SDM$_f$
Mod
QM
DQM
QDM
UQM
SDM
Obs
(b)
Monthly mean
Historical period
Future period
2035
2060
2080
Month
Obj
(c)
Monthly standard deviation

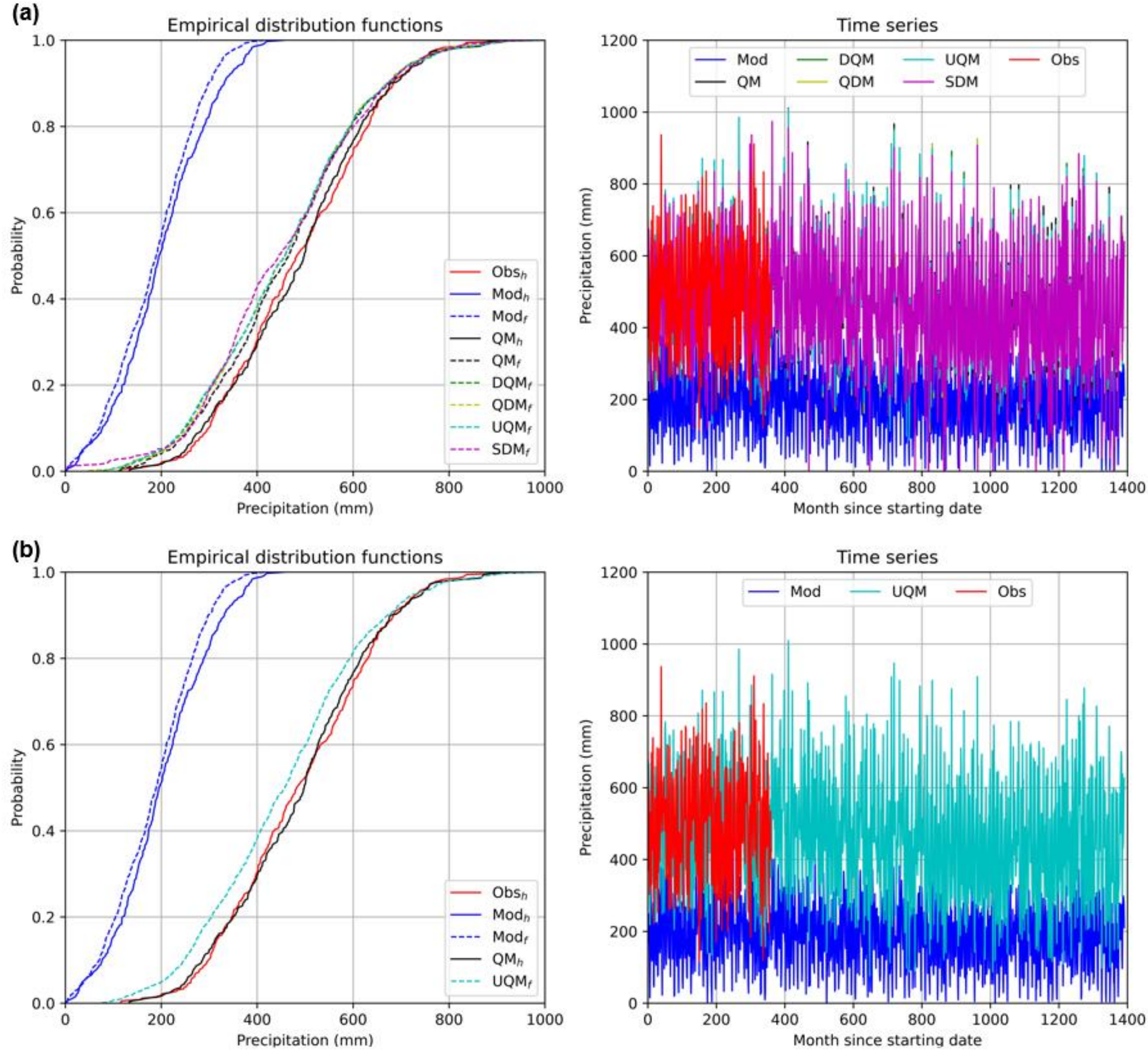


Figure 6: Example of the summary figures generated by the report function of the climQMBC package: cumulative distribution functions in the historical and future periods (left side) and time series (right side). (a) Default report function option displaying all methods at once. (b) Custom view using the optional argument fun=['UQM'] to display only the UQM method.~~. (a) cumulative distribution functions in the historical and future periods (left side), and time series (right side); (b) summary plots of the seasonal monthly mean, (c) summary plots of the seasonal monthly standard deviation.~~

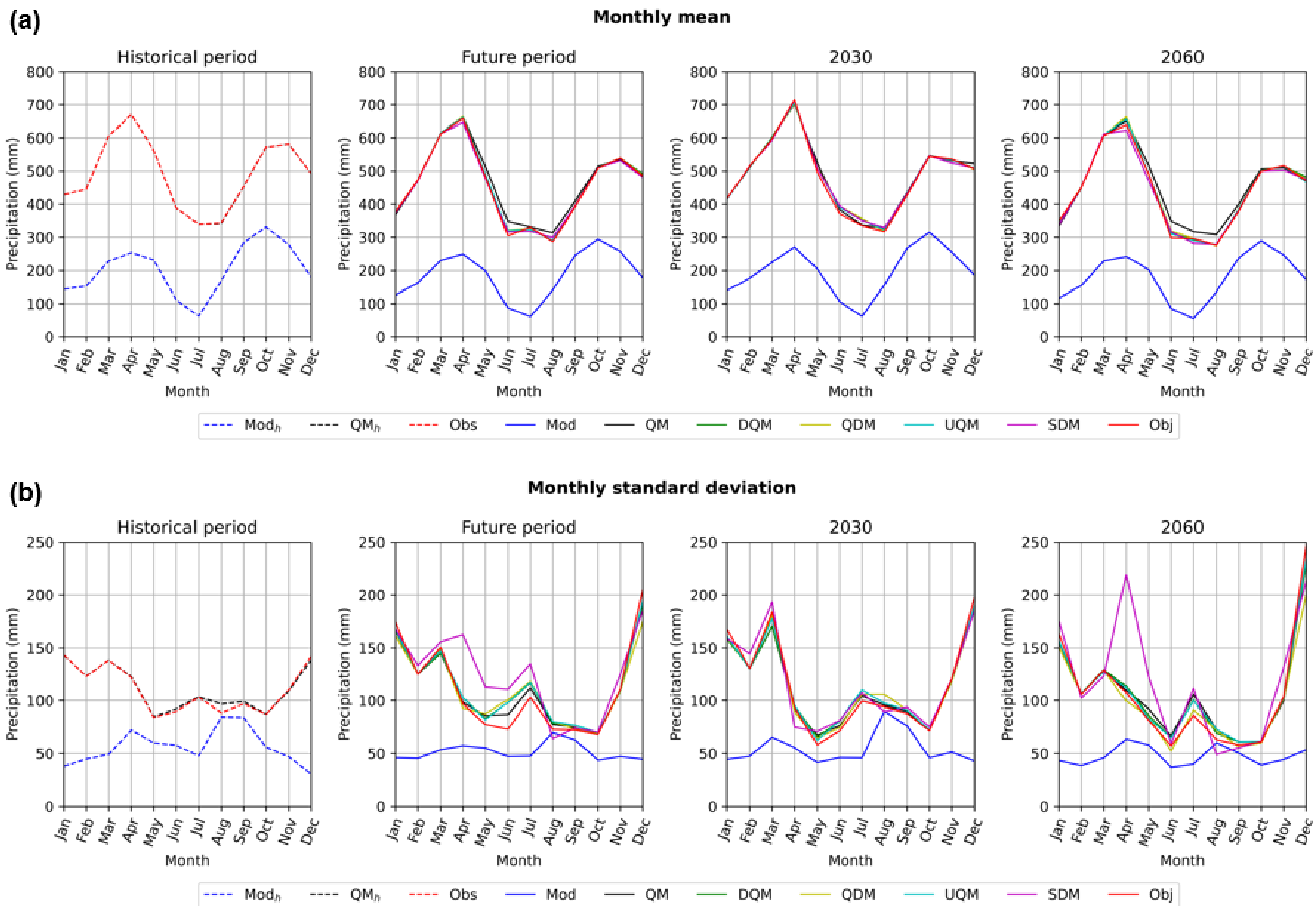


Figure 7: Example of the summary figures generated by the report function of the climQMBC package: (a) summary plots of the seasonal monthly mean, (b) summary plots of the seasonal monthly standard deviation. To display the user defined periods, the optional arguments y_init=1985 and y_wind=[2030, 2060] were set to define 1985 as initial year and to display the values in the windows centered in 2030 and 2060.

4 Application examples with synthetic series

In this section we assess the capabilities of the climQMBC package to systematically bias correct climatic variables of GCMs or RCMs. We compared the performance of the implemented bias correction methods in both the historical and projected periods based on ~~a synthetic example~~the synthetic tester that is available as a function in the climQMBC package. A controlled bias was introduced into a synthetic time series, enabling a clear evaluation of each method-s ability to correct it. This example is designed to demonstrate the performance of bias correction methods under varying conditions, offering insights into their optimal use. Nevertheless, determining the most suitable method for each specific case lies beyond the scope of this study. The example highlights the importance of having an easy-to-use tool, such as the climQMBC package, to implement bias correction methods and discern the most adequate one to a specific case.

4.1 Description of the synthetic example and evaluation

Extending the synthetic example presented by Maurer and Pierce (2014), Cannon et al. (2015), and Chadwick et al. (2023), gamma distributions were used to generate multiple synthetic precipitation time series. The gamma probability density function with shape parameter k and scale parameter θ is given by:

$$f(x;k,\theta) = \frac{x^{k-1}\exp(-x/\theta)}{\theta^k \Gamma(k)}, \quad x > 0, \quad k > 0, \quad \theta > 0, \tag{19}$$

where $\Gamma(\cdot)$ is the gamma function. Using the method of moments the mean (μ) and standard deviation (σ) can be used to estimate the parameters with $\mu = k\theta$ and $\sigma = \sqrt{k}\theta$.

The example considers four observed historical series generated with 10,000 values, mean of μ_h=1,000 and coefficients of variation (CV) ranging from 0.25 to 1.00 (Table 2). ~~To account for a known~~The future ~~series~~period~~, for each observed~~ is generated using a gamma distribution with 49 possible changes (i.e., the combination of $\Delta\mu$ and $\Delta\sigma$ in ~~table~~Table 2) from the historical series.~~, 49 future objective series were generated with the gamma function (independently from the observed historical series) considering combinations of changes in the mean and standard deviation ($\Delta\mu$ and $\Delta\sigma$, respectively) ranging from -50 % to +50 % (Table 2).~~ Note that the projected or future mean (μ^*) and future standard deviation (σ^*) are obtained from $\mu^* = \mu(1+\Delta\mu)$ and $\sigma^* = \sigma(1+\Delta\sigma)$, respectively. Then, bias in the mean and standard deviation ~~was~~were added ($\Delta\mu_{bias}$ and $\Delta\sigma_{bias}$, respectively), following Eq. (20). These range~~ing~~ from -50 % to +50 % (Table 2) for both the observed historical and future objective series, resulting in 49 combinations of bias~~for each observed historical – future objective combination~~. The same bias is added in the historical and in the future periods.

To obtain the bias values (x_i^b), the generated historical or future series (x_i^o) goes through Eq. (20). Here, one can choose the specific ~~add the desired~~ bias ~~with components both in the~~ being added in the mean ($\Delta\mu_{bias}$) and ~~in the~~ standard deviation ($\Delta\sigma_{bias}$) independently, while preserving the original sequence of events. ~~, the following procedure was applied to each value x_i^o of the generated observed historical or future objective series to get biased values x_i^b:~~

$$x_i^b = \mu^o(1+\Delta\mu_{bias}) + (x_i^o - \mu^o)(1+\Delta\sigma_{bias}), \tag{20}$$

where μ^o is the mean of the observed historical or future objective series. ~~This methodology ensures adding the desired bias by shifting both the mean and variability of the series while preserving the sequence of events.~~ As such, in this synthetic exercise the bias corrected series, if perfectly done, should be the same as the synthetic observed historical or future objective series.

~~Each generated series has an observed historical and future projected period. The future or projected period considers a multiplicative change compared to the historical period. Then we added bias to both period (historical and future or projected) to obtain a model synthetic series, which allows testing if the bias correction methods can remove the added bias. In this synthetic exercise the projected bias corrected climate should be the same as the synthetic observed projected climate.~~

~~We defined observed scenarios with a mean of μ_h=1,000 and coefficients of variation (CV) ranging from 0.25 to 1.00 (Table 2). For each observed scenario we generated several objective future scenarios with combinations of changes in the mean and standard deviation ($\Delta\mu$ and $\Delta\sigma$, respectively) ranging from -50 % to +50 % (Table 2). Note that the projected or future mean is obtained from $\mu^* = \mu\Delta\mu$, while an analogous process is used to obtain the future standard deviation. For each scenario 10,000 values were generated.~~

~~To each value x_i^o of the generated observed or objective future series, we added bias Δ_{bias}, with components both in the mean ($\Delta\mu_{bias}$) and in the standard deviation ($\Delta\sigma_{bias}$), to get biased values x_i^b by the following procedure, which shifts the mean and variability of the series:~~

$$\Delta_{bias}(x_i^o) = (1 + \Delta\mu_{bias}) + \frac{(x_i^o - \mu^o)(\Delta\sigma_{bias} - \Delta\mu_{bias})}{x_i^o}, \quad (20)$$

$$x_i^b = x_i^o \cdot \Delta_{bias}(x_i^o), \quad (21)$$

~~where μ^o is the mean of the observed or objective future series. This methodology ensures adding the desired bias in the mean ($\Delta\mu_{bias}$)) and standard deviation ($\Delta\sigma_{bias}$). We considered several combinations of bias in the mean and standard deviation with values ranging from -50 % to +50 % (Table 2).~~

Table 2: Coefficient of variation, projected changes in the mean and standard deviation, and bias in the mean and standard deviation considered in the synthetic example.

Variable	Values
Coefficient of Variation (CV)	0.25, 0.50, 0.75, 1.00
Projected change in mean ($\Delta\mu$)	-50 %, -25 %, -10 %, 0 %, 10 %, 25 %, 50 %
Projected change in standard deviation ($\Delta\sigma$)	-50 %, -25 %, -10 %, 0 %, 10 %, 25 %, 50 %
Bias in mean ($\Delta\mu_{bias}$)	-50 %, -25 %, -10 %, 0 %, 10 %, 25 %, 50 %
Bias in standard deviation ($\Delta\sigma_{bias}$)	-50 %, -25 %, -10 %, 0 %, 10 %, 25 %, 50 %

For the bias correction process, we evaluated all the implemented bias correction methods in the climQMBC package. ~~considering each original synthetic series as observed and those with added bias as the modeled series. There are four sets of synthetic series: historical observed, projected observed, historical modeled and projected modeled.~~ The bias correction uses three of the four set of series, both modeled series (projected and historical) and the historical observed, while the projected observed is kept for verification purposes, given that it is the "true series". In terms of implementation, as the climQMBC package considers moving windows and this example does not consider progressive changes in the variables, the biased series considered a duplicated future period so that the correction of the second future period neglects the information of the historical period. This means that the continuous series are the concatenation of the historical period, the future period (first copy) and the future period (second copy), finally evaluating the historical and second future period.

To assess the performance of each method, we considered (1) performance metrics based on statistical properties, measured as the ratio between statistics of the bias corrected series and those of the expected series, for which a value of 1 means a perfect agreement, and 2) goodness-of-fit metrics (i.e. Nash-Sutcliffe efficiency (NSE) proposed by Nash and Sutcliffe (1970), and Kling-Gupta efficiency (KGE) proposed by Gupta et al. (~~,~~2009)) to evaluate the point-by-point agreement between the bias corrected and expected series. Values of 1 for NSE and KGE indicate perfect agreement. These metrics, their range and possible values are summarized in Table 3. Note that for this application, where the biased series share the exact sequence of events as its original series, the NSE and KGE metrics are appropriate, while this usually does not happen in real comparisons between GCM and observations. As such, these metrics are not proposed as metrics to evaluate bias correction ~~methods~~ performance in general, but only when the proposed synthetic tester is used.

Table 3: Performance metrics considered.

Variable	Range	Values
Ratio between the corrected and expected mean, standard deviation and percentiles[1]	[0, ∞)	<1: Underestimation =1: Perfect agreement >1: Overestimation
Ratio between the corrected and expected skewness[2]	(-∞, ∞)	<1: Underestimation =1: Perfect agreement >1: Overestimation
Nash-Sutcliffe efficiency (NSE; Nash and Sutcliffe, 1970)	(-∞, 1]	<1: Bias among values =1: Perfect agreement among values <0: Worse than the mean
Kling-Gupta efficiency (KGE; Gupta et al., 2009)	(-∞, 1]	<1: Bias among values, mean or standard deviation =1: Perfect correlation between values and perfect fit among mean and standard deviation <1-$\sqrt{2}$: Worse than the mean (Knoben et al, 2019)

1 Percentiles 5 %, 10 %, 50 %, 90 %, and 95 % are considered

2 Negative values indicate change in the direction of the distribution asymmetry

4.2 Bias correction comparison in the historical period

Regarding the performance of the bias correction methods in the historical period, Fig. ~~7~~ 8 shows the NSE values for each combination of coefficient of variation (CV), bias in mean ($\Delta\mu_{bias}$) and bias in standard deviation ($\Delta\sigma_{bias}$). Despite the synthetic series were generated with a strictly positive distribution function, the procedure to add bias result in some negative values. These values were replaced with random values below 0.01, according to the implementation of the climQMBC package. The numbers in the grid of Fig. ~~7~~ 8 indicate the percentage of negative values in the biased synthetic time series resulting from the procedure to add bias~~that have some negative values and were replaced with with random values below 0.01, according to the implementation of the climQMBC package~~. Figure ~~7~~ 8 shows only values for the QM and SDM methods because the DQM, QDM and UQM methods are equivalent to the QM method in the historical period (Chadwick et al., 2023).

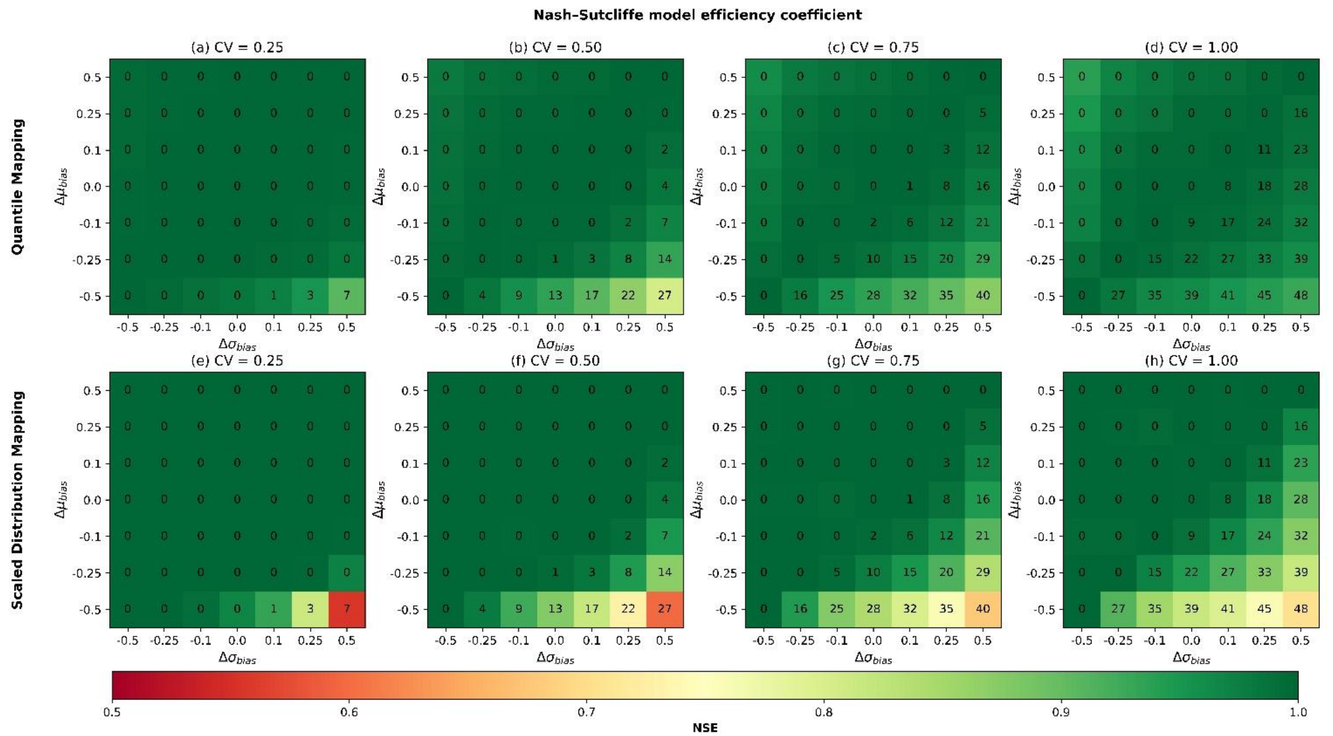


Figure 8~~7~~: NSE values of the QM and SDM methods in the historical period for different combinations of the observations' CV, $\Delta\mu_{bias}$ and $\Delta\sigma_{bias}$. Green color indicates a perfect fit between the corrected and observed values (NSE ~ 1), and red tones indicate a lower quality fit (NSE < 1). The numbers in the grid indicate the percentage of negative values in the biased synthetic time series resulting from the procedure to add bias~~percentage of synthetic time series that have some negative values~~ (~~that~~ these negative values were replaced with random values below 0.01).

Based on the NSE values (Fig. ~~7~~8), the QM and SDM methods may fit differently to the observed values under different circumstances, with the QM method being the one with better fit when the bias in the standard deviation increases and the bias in the mean decreases to negative values (bottom right corner of each grid). On the other hand, when the bias in the mean increases and the bias in the standard deviation decreases to negative values (top left corner of each grid), NSE values for the QM method tend to slightly decrease for high values of CV, tendency not observed for the SDM. Both methods perform well in terms of NSE values when the bias in the mean and standard deviation are similar. When performance is analyzed based on the mean, standard deviation, skewness, and percentiles, or based on the KGE values, results showed similar patterns (results not shown).

Figure ~~8~~ 9 presents boxplots of the performance metrics shown in Table 3 for the ratio between the corrected and observed mean, standard deviation, skewness, percentiles 5th, 10th, 50th, 90th and 95th, alongside the NSE and KGE values in the historical period. The boxplots consider the results after adopting all the CV and bias conditions depicted in Table 2. For all plots in Fig. ~~8~~9, a value of 1 indicates a perfect agreement between the corrected and observed statistics. Figure ~~8~~ 9 shows only values for the QM and SDM methods because the DQM, QDM and UQM methods are equivalent to the QM method in the historical period (Chadwick et al., 2023).

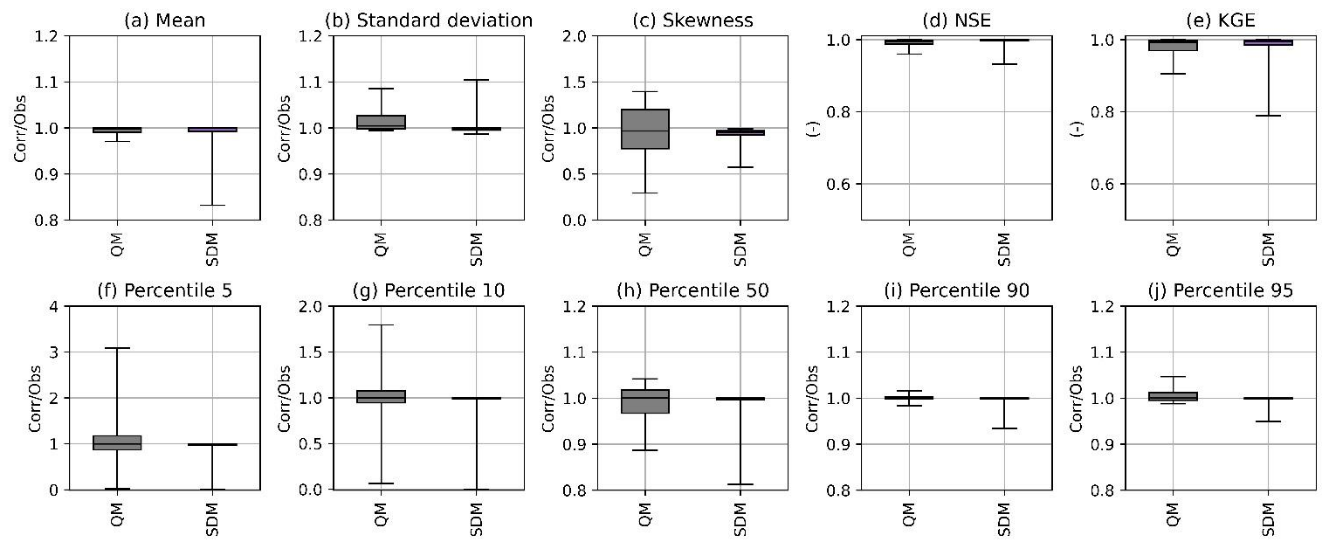


Figure 9~~8~~: Performance comparison of QM and SDM in the historical period using the ratio between the bias corrected and observed mean (a), standard deviation (b), skewness (c), and percentiles (f to j), in addition to the NSE (d) and KGE (e) values. For all metrics, a value of 1.0 indicates a perfect agreement between the corrected and observed values, as depicted in Table 3.

In all cases, both the QM and SDM have boxes centered in a value of 1.0. Although more stable in terms of a more constrain box (representing the middle 50 % of the data) than the QM method, the SDM method resulted also in more outliers. For the mean, standard deviation, NSE, KGE, percentiles 50 %, 90 % and 95 %, the SDM shows a smaller box than the QM, but the SDM have larger whiskers. For the skewness, the SDM shows a smaller box and whiskers than the QM, and for percentiles 5 % and 10 %, the SDM shows a smaller box than the QM, and the QM shows whiskers for over and underestimation, while SDM only for underestimation. Hence, if the results of the SDM method are not properly analyzed, they may result in larger biases than QM. Having a tool to easily compute and compare among bias correction methods allows the analysis and facilitate the selection of the most adequate method depending on the case study under analysis.

4.3 Projected scenarios

Regarding the performance of the bias correction methods in the future, Fig. ~~9~~ 10 shows the cumulative probability of the NSE and KGE values, resulting from the synthetic exercise considering all the combinations of CV, added biases and projected changes shown in Table 2. The desired result, indicating that almost all cases have a perfect agreement with the expected values, is to have cumulative probabilities as close as possible to the value of 1.0; hence curves whose mass is concentrated near NSE = 1 (that is, shifted toward the right edge of the plot) ~~indicate~~ have better performance.~~the closer to the right the curve is, the better.~~

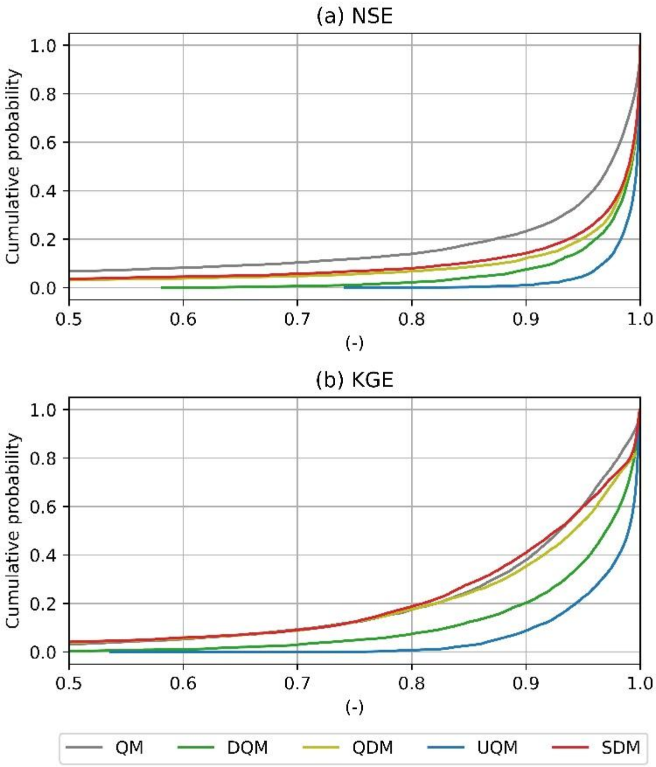


Figure 10~~9~~: Cumulative probability of the NSE (a), and KGE (b) values of the synthetic example for each implemented method in the climQMBC package.

Based on the NSE and KGE values, the QM, DQM, QDM, UQM and SDM methods result in different series which may differ in their performance. Results ~~of~~ using the synthetic tester example, which consist ~~on~~of adding a bias in the two first statistical moments (i.e. the mean and standard deviation), show that the UQM method performs best in eliminating the bias from the modeled series, followed by the DQM method, being the QM method the one with the lowest performance quality. The SDM and QDM methods perform similarly in terms of the NSE values, but the SDM performs worse when evaluated with the KGE. ~~The NSE and KGE indexes give an overall goodness-of-fit; nevertheless, other performance metrics commonly used to assess a bias correction method in terms of how certain statistics are preserved (e.g. mean, standard deviation, skewness, and percentiles) might be adopted.~~ The performance of a method may differ based on the performance metric and the objectives of the user. Figure ~~10~~ 11 presents boxplots of the performance metrics depicted in Table 3 representing the ratio between the corrected and expected mean, standard deviation, skewness, and percentiles 5, 10, 50, 90 and 95, and as well as the NSE and KGE values in the future scenario. Again, a value of 1 indicates a perfect fit between the corrected and expected statistics.

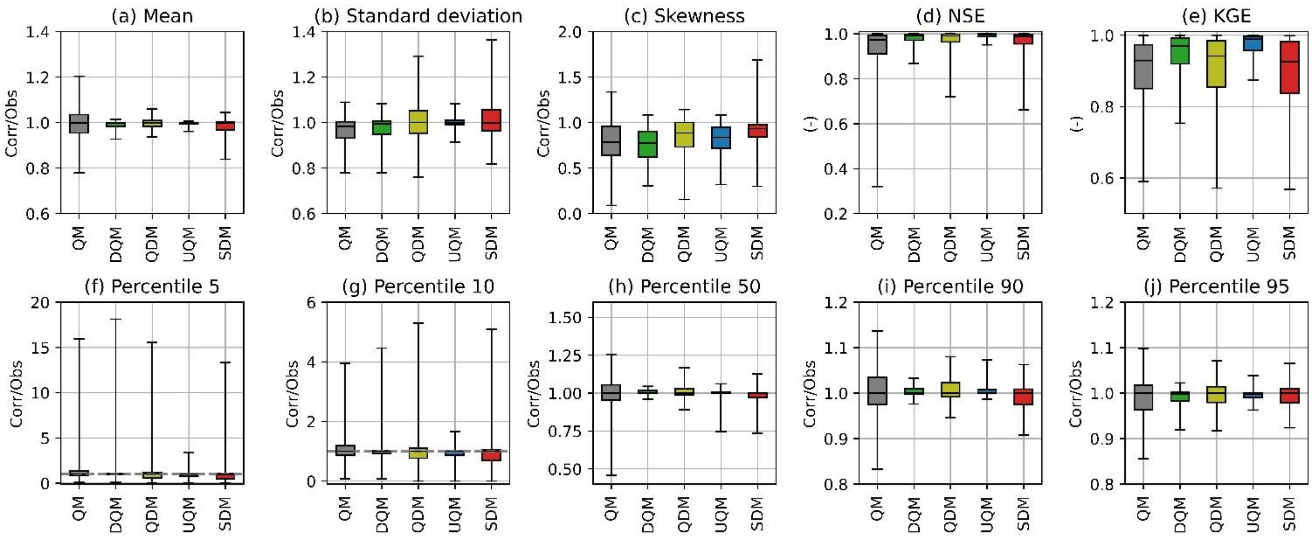


Figure 11~~0~~: Performance comparison of the QM, DQM, QDM, UQM and SDM in the future period evaluated by means of the ratio between the bias corrected and objective future mean (a), standard deviation (b), skewness (c), and percentiles (f to j), in addition to the NSE (d) and KGE (e) fitness metrics. For all metrics, a value of 1.0 indicates a perfect agreement between the corrected and expected values, as depicted in Table 3.

Figure ~~10~~ 11 shows that all the implemented methods might relate to the expected series based on different statistics; however, some methods are more stable in some statistics than others (e.g., UQM and DQM in the mean and percentiles, UQM in the standard deviation or the SDM, QDM and UQM in the skewness). Overall, in this exercise where bias was added in the first two statistics moments, the UQM, which aims at removing the bias in these moments, ~~UQM~~ tends to be more stable than the other methods. Among the implemented methods in the climQMBC package, for this exercise, there is always a trend preserving method that is more stable than the QM method. This does not mean that the QM method should not be used. In fact, as mentioned above, all methods might perform well under certain circumstances, including the QM method. This is why an evaluation of the method for each specific situation is encouraged.

5 Conclusions

Quantile Mapping (QM) is one of the most widely used bias correction methods for climate change impacts, although it might distort raw GCM changes when future climate goes out of the range of the historical climate. Trend-preserving methods such as DQM, QDM, UQM, and SDM aim to address this limitation, each with its own advantages and disadvantages, although the QM method is still widely used because of its simplicity and ease of implementation. This paper presents the climQMBC package, a coding package written in Python, MATLAB and R that makes available all these point-to-point bias correction methods for daily, monthly and annual data. As such, the package enables incorporating several bias correction methods as a standard component of climate change impact assessment, as well as the identification of the most adequate method for a specific case. Being available in several programming languages avoids users having to change between languages and include the bias correction analysis in their existing workflow.

It is worth noting that several climate change impacts assessments are based on the analysis of compound events, which require methods that capture the inter-dependence of several variables or different points in space. The methods available in version 1.1.0 of the climQMBC package are strictly univariate. As such, results from the climQMBC package must be evaluated carefully when analyzing compound events. To date, existing multi-variable methods succeed ~~o~~in capturing these inter-dependences, but sometimes struggle on representing local climate and capturing raw GCM trends, leading to the user to consider a trade-off between the climate characteristics that are the most important for their analysis.~~,~~ Unfortunately, few works have focused on comparing methods or reporting these limitations (Aedo-Quililongo and Chong, 2024) and work is needed both in terms of method development and tools that provide these complex methods for easy and fair comparisons.

This work also ~~presents~~proposes a synthetic tester for bias correction and an example in which it is tested. Here, ~~where~~ bias is added in the first two statistical moments (i.e. mean and standard deviation) to evaluate the performance of the methods implemented in the climQMBC package. The example is based on a set of synthetic observed historical series, future objective series (assumed as “true series” as if we knew already the future) and biased historical and future series. The future es built as an independent series based on assumptions of the projected change in the mean and standard deviation. The bias is then added to each series considering a methodology that explicitly shifts both the mean and variability of the series while preserving the sequence of events. As such, the bias corrected series should be the same as the synthetic observed historical or future objective series. ~~To address performance of the implemented method, performance metrics based~~The performance is measured using ~~on statistical properties and on~~ goodness-of-fit metrics like the NSE and KGE, and the capability of replicating other statistical properties ~~were considered~~. It is important to note that in this example, the NSE and KGE are valid because point-by-point agreement between the bias corrected and expected series is desired. However, such metrics are not suitable for real case applications where the GCM sequence of events is not consistent with observations. This example is also implemented in the package so that users can make a brief theoretical comparison of the performance of each method for their specific case.

~~Using a synthetic example~~Based on the synthetic example, we demonstrate that all the methods implemented in the climQMBC package can successfully ~~correct~~ remove these biases~~the modeled series~~ in both the historical period and future scenarios; however, this performance cannot be guaranteed as a general rule. In general, among the methods, there is always one whose performance is more robust than the QM method~~, being UQM and DQM the more stable ones~~. ~~Nonetheless, in some cases the QM method should not be discarded because it is simple, computationally faster than the trend preserving methods, and might perform well enough under certain conditions.~~ The UQM and DQM (methods that target statistical moments) resulted ~~in~~ as the most stable methods capturing statistical moments, percentiles and recovering the original series in the synthetic example ~~(example that targets statistical moments)~~. Overall, we emphasize the need to assess the performance of the bias correction methods to avoid unwanted bias in the corrected series. However, there are cases where the best option among the available methods will not perform perfectly, and thus a tool like the climQMBC package will be useful to identify the most adequate method for each specific case.

Despite all trend-preserving methods outperform the QM method, ~~Nonetheless,~~ in some cases the QM method should not be discarded because it is simple, computationally faster than the trend preserving methods, and might perform well enough ~~as the trend-preserving methods specific~~ under ~~under~~ certain conditions. In terms of execution times, trend-preserving methods can take up to around 10, 30 and 40 times the mean execution time of the QM method for annual, monthly and daily frequencies, respectively. This may also be considered in the trade-off between methods as some-times execution time can be restrictive. However, as machine learning approaches and applications are increasing over the years, these, based on synthetic results could develop methods with comparable results to the trend-preserving methods, but with lower execution cost.

Appendix A: Window length considerations to capture historical conditions and model trends

The climQMBC package considers that the length of the projected periods (backwards moving windows) is the same as the user defined length of the historical period for consistency on the characterization of each period (see Section 3.1.1). While the World Meteorological Organization (WMO) recommends 30 years to characterize Climate Normals (WMO, 2017), there are trade-offs between choosing larger periods to better capture statistical moments and shorter periods to better capture trends and avoid pseudo-stationarity bias. Appendices A1, A2 and A3 present three sensitivity analysis to make visible these trade-offs based on synthetic realizations using a gamma probability density function (eq. 19 in Section 4.1) and linear trend future scenarios, following Chadwick et al. (2023).

It is worth noting that the linear trend example is an extreme condition that allows exploring different issues arising from bias correction decisions. However, real life examples usually have combinations of oscillating or gradual patterns, where it is challenging to identify or address the issues.

Based on the results presented in Appendices A1, A2 and A3, there is no perfect condition in terms of window length to avoid biases in characterizing statistical moments while capturing the modeled trends and avoiding pseudo stationary bias. However, the importance of these potential biases ~~decreases~~are only relevant for time series with high~~ with the~~ coefficient of variations ~~of the data and the magnitude of the~~ and strong trends. It is worth noting that the distribution of the data, which depends on the ~~analyzed~~climate variable, is also an important factor that should be considered. ~~In this context, a~~We recommend~~ation of~~ using periods of lengths between 20 to 40 years aims ~~for a~~ balancing~~e~~ between the WMO recommendation of 30 years~~ and~~, properly characterizing the statistical moments, adequately captur~~e~~ing the modeled trend and reducing~~e~~ the pseudo-stationarity bias.

Appendix A1: Sample size (years) impacts on mean and standard deviation representativity

This analysis aims to show the impact of the sample size on the mean and standard deviation sample estimation. The analysis considers four conditions with objective mean of μ_{obj}=1,000 and standard deviation σ_{obj} of 250, 500, 750 and 1,000, resulting in coefficients of variation (CV) of 0.25, 0.50, 0.75 and 1.00, respectively. For each condition, the gamma function was used to generate cases with widow lengths from 10 to 60 values or years~~, representing the window length in years~~. Each case was repeated 10,000 times (realizations).

Figure A1 shows the comparison of the sample mean and standard deviation of each realization with the objective value (μ_{obj} and σ_{obj}) in terms of the ratio. The ratio is between each realization and the objective mean and standard deviation (Fig. A1a and Fig. A1c) and the variability among realizations (Fig. A1b and Fig. A1d).

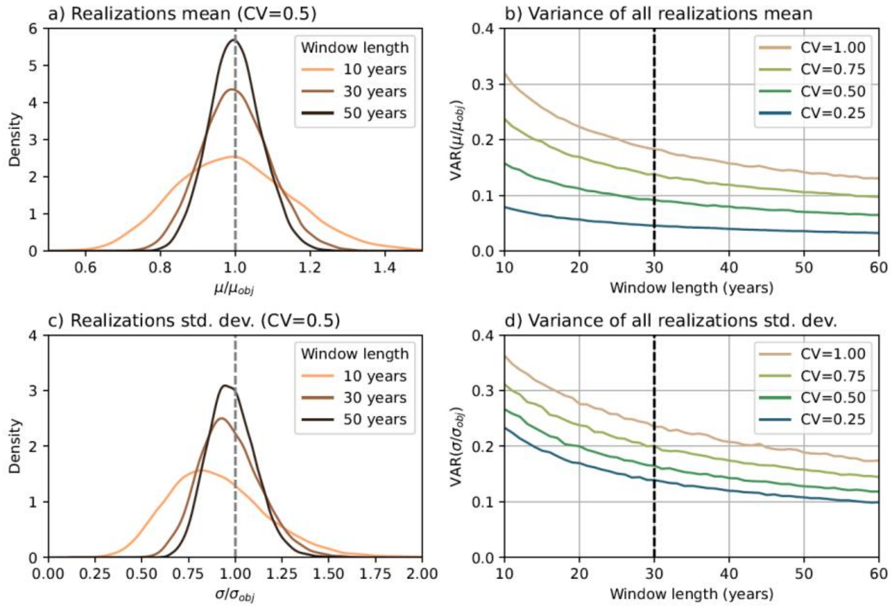


Figure A1: Density distribution of the ratio between each realization and the objective mean (a) and standard deviation (c), for a CV value of 0.5 and different window lengths. Standard deviation of the ratio between each realization and the objective ~~the~~ mean (b) and standard deviation (d), for different CV values. Dashed lines in (b) and (d) represent~~s~~ the 30 year window recommended by WMO (2017) ~~recommendation of windows length of 30 years~~.

Results show that, for a CV value of 0.5 and a window length of 10 years, the realizations mean is centered around the objective value, but some realizations can ~~be biased from~~ have an error of ~~the objective mean in~~ almost ~~a~~ 40% (light colored lines in Fig. A1a). On the other hand, the realizations standard deviation is not centered around the objective, ~~and~~but some realizations ~~are biased from the objective~~present errors in standard deviation ~~in~~of almost 75% (light colored lines in Fig. A1c). When increasing the window length, ~~the~~ this sample errors diminishes ~~biases decreas both for the realizations mean and standard deviation~~for both moments, and the ~~realizations~~ standard deviation ~~also~~ tends to be better centered around the objective standard deviation (dark colored lines in Fig. A1a and Fig. A1c).

The reduction of the variability with increasing window length is represented in Fig. A1b and Fig. A1d in terms of the standard deviation of the realizations mean and standard deviation for each CV condition. While all CV conditions follow the decreasing pattern, high CV values have consistently a higher dispersion. As such, longer window lengths reduce the potential bias on characterizing the sample statistical moments; however, conditions with higher CV result in less representative values.

Note that for centered distributions like ~~the~~ normal distributions, the mean and standard deviation tend to be ~~more~~ centered around the objective values with shorter window lengths than for skewed distributions like the gamma distribution. However, the potential ~~bias~~errors or dispersion issue is not solved directly by ~~a change~~changing ~~in~~ the distribution.

Appendix A2: Backwards moving window length impacts on capturing the trend in the mean

This analysis aims to show the potential biases in capturing model trends of the mean when considering backwards moving windows (e.g. for a 30 years moving window, the value for 2029 represents the mean of the period 2000-2029). Backwards moving windows are have been used in previous studies (Chadwick et al., 2018; Chadwick et al., 2023) as are more suitable to capture GCM trends and generate continuous time series over forward and centered moving windows, as the latter result in discontinuities between the historical and future periods, along to not being able to compute values for the last years of the projections.

The analysis considers annual time series from 1955 to 2100 generated with the gamma function considering a mean of μ_h=1,000 and a CV of 0.1. Four scenarios were imposed ~~to~~on the time series where the historical period (1955-2014) has no trend and the future period (2015-2100) has a linear trend with projected changes of -50, -25, +25 and +50% between the historical period and year 2100. For each of these four scenarios, 10,000 realizations were generated. Figure A2a shows the trends in the annual mean ~~of the realizations for each scenario~~.

Each of the generated realizations were averaged considering backward moving windows of lengths from 10 to 60 years. Figure A2b presents the ~~mean~~average of the 10, 30 and 50 years moving mean realizations for the scenario with 50% increase for 2100, showing two potential biases that can arise from considering moving windows: trend lag and terminal error. Overall, the moving windows underestimate the magnitude of the change of the real linear trend; however, the difference between the real values and the moving mean increases in time until it remains constant. The trend lag is the lag between the real trend and the moving window trend, which is equal to half of the moving window length, while the terminal error is the maximum difference. Both the trend lag and the terminal error increase with larger window lengths, but the terminal error also increases with the magnitude of the trend, as shown in Fig. A2c.

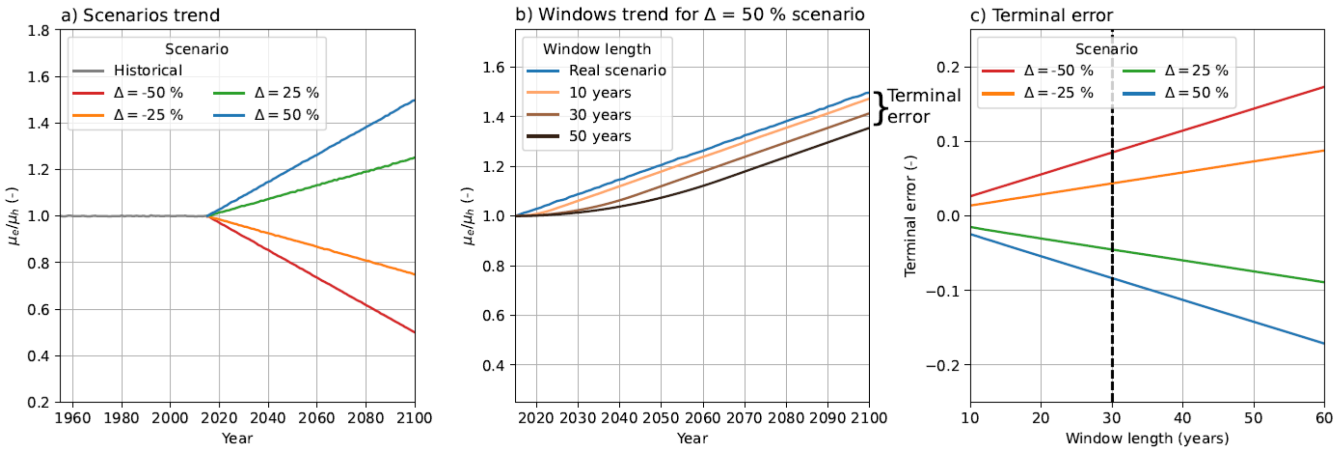


Figure A2: a) Mean of the linear trend realizations, ~~normalized~~divided by the historical mean; b) Comparison of real trend and the moving window with different window lengths; and c) Terminal error of for different window length for the four linear trend scenarios. Dashed lines in (c)represent the 30 year window recommended by WMO (2017) ~~recommendation of windows length of 30 years~~.

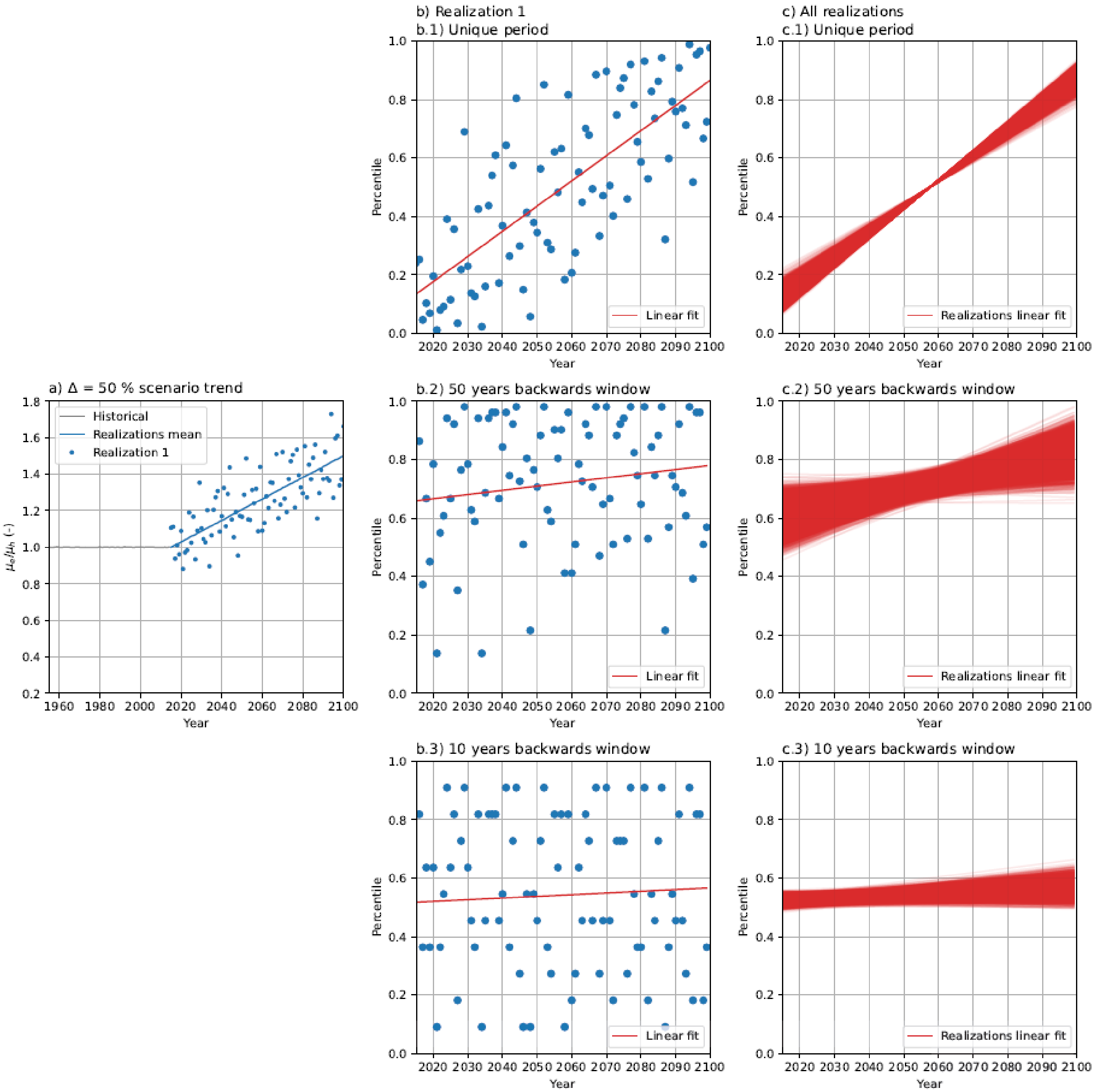
Appendix A3: Moving windows and pseudo-stationarity bias

Pseudo stationary bias arises when the model biases are assumed constant over an analyzed period where climate is changing (Chadwick et al., 2023). This condition can be observed on increasing temperatures scenarios where the values at the end of the period are considerably higher than at the beginning of the period, resulting on a probability distribution where the values at the end of the period are on one of the distribution tails, while the ~~values at the end of the period~~initial values are in the other tail. Considering moving windows for each year analyzed allows ~~detrending the probabilities,~~ reducing the pseudo stationary bias~~s~~.

This analysis ~~aims to show the trade-offs of~~presents how ~~recurring~~ using ~~to~~ moving windows ~~to~~ reduces the pseudo-stationarity bias. The analysis is based on the scenario with increase in 50% in 2100~~series~~ generated on Appendix A2 ~~and considering the scenario with increase in 50% in 2100~~. Figure A3a shows the mean of the realizations and the future values of realization 1. Figure A3b (middle column of Fig. A3) shows the values in terms of percentiles of realization 1 and the percentiles linear trend, while Figure A3c (right column of Fig. A3) shows the linear trends of all realizations. Figure A3b1 and A3c1 (first row of Fig. A3) considers the period 2015 – 2100 as a unique period (i.e., no moving window is used here, this considers the traditional fix window approach).~~,~~ ~~while~~ Figure A3b2 and A3c2 (second row of Fig. A3) considers a backwards moving window of 50 years for each year to define the corresponding percentile. Figure A3b3 and A3c3 (third row of Fig. A3) shows the same information as Fig. A3b2 and A3c2, but for moving windows of 10 years. Figure A4 shows the mean values and the slope of the percentile time series for different moving window values, along to the case where a unique period is considered. Results show that considering a unique period in a linear trend scenario result in percentiles where the extreme values are located at the beginning and at the end of the period, with a clear increasing trend (Fig. A3b1). On the other hand, when moving windows are considered, the slope of the trend in the probability domain decreases with the moving window length (Fig. A3b2,

Fig. A3b3 and Fig. A4b). On the other hand, when a unique period is considered, the mean is centered in a percentile of 0.5, while for moving windows the mean is above 0.5 with a decreasing trend with decreasing window lengths (Fig. A3c2, Fig. A3c3 and Fig. A4b).

This analysis shows a trade-off in terms of the pseudo-stationary bias, where reducing the window length results in a reduction in this bias. On the other hand, ~~and~~ considering the complete period may capture correctly the mean~~,~~ but have a tendency to overestimate percentiles at the end of the period and underestimate them at the beginning of the period. While these results are related to trends with monotonic increase, each analyzed area should explore this kind of conditions.~~g~~



785 **Figure A3: a) Realization 1 and mean of the realizations for the scenario with linear increase of 50%, normalized by the historical mean. b) Realization 1 values in terms of percentiles and linear trend of the percentile probability values: b1) considering a unique period to build the distribution; b2) and b3 considering moving windows of 50 and 10 years for each year, respectively, to build the distribution. c) Same as b), but only the linear trend for all realizations.**

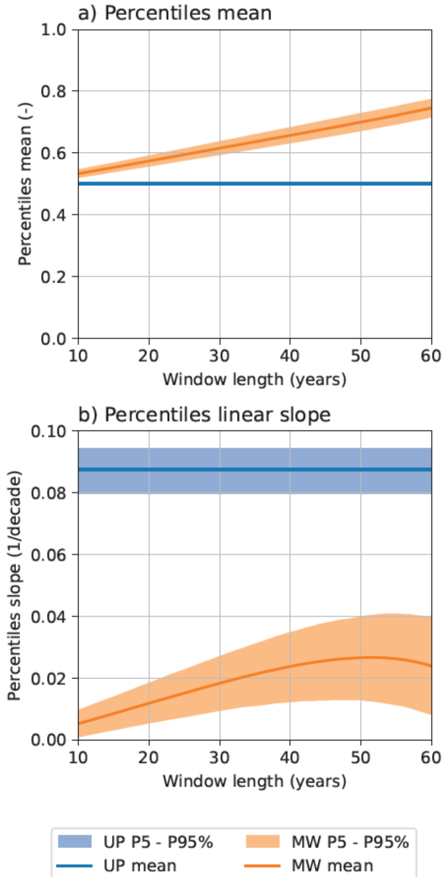


Figure A4: a) Mean of the percentiles of the realizations for the scenario with linear increase of 50% when a unique period is considered (UP) and when moving windows (MW) of different lengths are considered. b) Slope of the percentiles linear trend.

Appendix B: climQMBC package methods execution time comparison

To compare the different execution times of the bias correction methods available in the climQMBC package, this appendix presents a synthetic example using a gamma probability density function (eq. 19 in Section 4.1). The example is based on 10,000 realizations considering a fixed historical mean of 1,000, a historical period of 30 years and a future period of 60 years. Each realization considers a different standard deviation with values of 250, 500, 750 and 1,000, sampled randomly, a random bias in the mean and standard deviation ($\Delta\mu_{bias}$ and $\Delta\sigma_{bias}$, respectively) ranging from -50 % to +50 %, and a random projected change in the mean and standard deviation ($\Delta\mu$ and $\Delta\sigma$, respectively) ranging from -50 % to +50 %. The same bias was added in the historical and in the future periods, following the same procedure as in the synthetic example presented in Section 5 (eq.

20). The example was replicated for annual, monthly and daily values, following the same procedure, and considering multiplicative values and strictly non-negative series.

This example aims for a general reference of the execution time rather than a detailed performance comparison. As such, the execution times are analyzed relative to the mean execution time of the QM method (Table B1) and considering only the package implementation in Python.

Results show that the QM method is the fastest among the implemented in the climQMBC package. The second fastest method is the DQM, with execution times close to 2.1 times the QM method for annual and monthly frequency, and 6.8 times for daily frequency, explained by the additional calculation of the projected changes in the mean and its application to detrend and reimpose the trend to the corrected series. The slowest methods are the QDM, UQM and SDM, with execution times between 7 and 10 times the QM method for annual frequency, 22 to 27 times for monthly frequency and 32 to 39 times for daily frequency. While ~~the~~ QDM, UQM and SDM are the slowest methods, ~~the~~ QDM tends to be slightly faster than the UQM method. While both methods consider a new transform function for each value to be corrected, the QDM computes only the projected change of the quantile of the value to be corrected, and the UQM computes the projected change of the mean, standard deviation and skewness (if applicable) of the associated moving windows, resulting in a longer execution time ~~for the UQM method~~. On the other hand, the SDM method considers several detrending, projected changes~~,~~ and interpolation steps that demand additional times.

While the three temporal frequencies consider the random near-zero replacement and moving windows, the daily frequency also considers the daily moving windows (to capture seasonality) and the calculation of the threshold to match the modeled number of rain days in the historical period to the observed number of rain days. These differences explain the additional execution times difference of the daily frequency compared to the annual and monthly frequencies.

Table B1: Execution times for the bias correction methods available in the climQMBC package, relative to the mean execution time of the QM method.

	Annual		Monthly		Daily	
Method	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
QM	1.0	0.2	1.0	0.4	1.0	0.2
DQM	2.2	0.5	2.0	0.4	6.8	4.6
QDM	8.8	2.0	22.6	0.7	32.6	6.0
UQM	9.8	2.1	24.8	0.7	33.4	6.1
SDM	7.1	1.6	26.5	1.1	38.7	5.9

Code availability

The source code of the climQMBC package for Python, R and MATLAB are openly available at GitHub (https://github.com/saedoquililongo/climQMBC) and has been archived in Zenodo with the DOI https://doi.org/10.5281/zenodo.18392899~~https://doi.org/10.5281/zenodo.18392900~~ (Aedo-Quililongo et al., 2026). Instructions to import the library can be found at https://github.com/saedoquililongo/climQMBC/blob/main/Python/README.md for Python, https://github.com/saedoquililongo/climQMBC/blob/main/R/README.md for R and https://github.com/saedoquililongo/climQMBC/blob/main/Matlab/README.md for MATLAB. Examples for testing the package can be found at https://github.com/saedoquililongo/climQMBC/blob/main/Python/climQMBC_tester_Python.py for Python, https://github.com/saedoquililongo/climQMBC/blob/main/R/climQMBC_tester_R.R for R and https://github.com/saedoquililongo/climQMBC/blob/main/Matlab/climQMBC_tester_Matlab.m for MATLAB. Data for the tester codes can be found at https://github.com/saedoquililongo/climQMBC/tree/main/Sample_data. Both the observed and modeled series must be a column vector of daily, monthly or annual data, without considering leap days, starting ~~in~~on the same date, and be continuous until the ending date. If daily frequency is specified, the length of the column vector should ~~by~~be a multiple of 365 and for monthly frequency, it should be a multiple of 12.

Author contributions

Sebastián Aedo-Quililongo: Conceptualization, Formal Analysis, Software, Writing – original draft. Cristián Chadwick: Conceptualization, Formal Analysis, Funding Acquisition, Software, Supervision, Writing – original draft. Fernando González-Leiva: Software, Writing – review & editing. Jorge Gironás: Funding Acquisition, Writing – review & editing

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors want to acknowledge the valuable feedback from users of the climQMBC package that has~~ve~~ led to continuous improvement of the package. We would also like to acknowledge the relevant feedback from the reviewers and editors; it has helped to strengthen the manuscript.

Financial support

This research was supported by funding from the National Agency for Research and Development, Chile, Grant ANID/FONDECYT/Regular 1200135 and 1241163. We would also like to acknowledge the Grant ANID/FONDECYT/Iniciación 11220952; the Grant Centro de Desarrollo Urbano Sustentable ANID CEDEUS CIN250009.

References

Aboitiz, C., Chadwick, C., Barichivich, J., and Perez-Quezada, J.: A framework to compare the performance of spatial disaggregation methods of statistical climate downscaling in South America, J. Hydrol.: Reg. Stud., 67, https://doi.org/10.1016/j.ejrh.2026.103834, 2026.

Aedo-Quililongo, S., Chadwick, C., González-Leiva, F., and Gironás, J.: saedoquililongo/climQMBC: Version 1.1~~0~~.0, Zenodo [code], https://doi.org/10.5281/zenodo.18392899, 2026.

Aedo-Quililongo, S.: ¿Por qué mis proyecciones de cambio climático difieren de las de los Modelos de Circulación General?, in Proceedings of the IAHR XXXI Congreso Latinoamericano de Hidráulica, Medellín, Colombia, 1-4 October 2024, 2775-2782, ISBN: 978-90-834302-6-3, 2024.

Aedo-Quililongo, S., and Chong, P.: Propuesta Metodológica Para El Escalamiento Local A Nivel Diario De Modelos De Circulación General, in Proceedings of the IAHR XXXI Congreso Latinoamericano de Hidráulica, Medellín, Colombia, 1-4 October 2024, 2783-2793, ISBN: 978-90-834302-6-3, 2024.

Andres, N., Vegas Galdos, F., Lavado Casimiro, W. S., and Zappa, M.: Water resources and climate change impact modelling on a daily time scale in the Peruvian Andes, Hydrol. Sci. J., 59, 2043-2059, https://doi.org/10.1080/02626667.2013.862336, ~~2013~~2014.

Ballarin, A., Sone, J., Gesualdo, G., Schwamback, D., Reis, A., Almagro, A., and Wendland, E.: CLIMBra - Climate Change Dataset for Brazil, Sci Data, 10, 47, https://doi.org/10.1038/s41597-023-01956-z, 2023.

Bayissa, Y., Melesse, A., Bhat, M., Tadesse, T., and Shiferaw, A.: Evaluation of Regional Climate Models (RCMs) Using Precipitation and Temperature-Based Climatic Indices: A Case Study of Florida, USA, Water, 13, 2411, https://doi.org/10.3390/w13172411, 2021.

Bedia, J., Baño-Medina, J., Legasa, M. N., Iturbide, M., Manzanas, R., Herrera, S., Casanueva, A., San-Martín, D., Cofiño, A. S., and Gutiérrez, J. M.: Statistical downscaling with the downscaleR package (v3.1.0): contribution to the VALUE intercomparison experiment, Geosci. Model Dev., 13, 1711–1735, https://doi.org/10.5194/gmd-13-1711-2020, 2020.

Blin, N.: Understanding groundwater dynamics in a changing climate: Numerical modeling and remote sensing insights from the Salar del Huasco Basin, Ph.D. thesis, Pontificia Universidad Católica de Chile, https://repositorio.uc.cl/server/api/core/bitstreams/acf55587-dc5f-41fa-8d7e-e23cba55082d/content, 2023.

Bourgault, P., Huard, D., Smith, T., Logan, T., Aoun, A., Lavoie, J., Dupuis, É., Rondeau-Genesse, G., Alegre, R., Barnes, C., Laperrièr, A., Biner, S., Caron, D., Ehbrecht, C., Fyke, J., Keel, T., Labonté, M-P., Lierhammer, L.,Low, J-F., Quinn, J., Roy, P., Squire, D., Stephens, A., Tanguy, M., and Whelan, C.: xclim: xarray-based climate data analytics, J. Open Source Softw., 8, 5415, https://doi.org/10.21105/joss.05415, 2023.

Brekke, L., Thrasher, B., Maurer, E. P., and Pruitt, T.: Downscaled CMIP3 and CMIP5 climate projections: Release of downscaled CMIP5 climate projections, comparison with preceding information, and summary of user needs, US Department of Interior, Bureau of Reclamation, Denver, CO, USA, 104 pp., 2013.

Bürger, G., Sobie, S., Cannon, A. J., Werner, A., and Murdock, T.: Downscaling extremes: An intercomparison of multiple methods for future climate, J. Clim., 26, 3429–3449, https://doi.org/10.1175/JCLI-D-12-00249.1, 2013.

Cannon, A. J., Sobie, S., and Murdock, T.: Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes?, J. Clim., 28, 6938-6959, https://doi.org/ 10.1175/JCLI-D-14-00754.1, 2015.

Cannon, A. J.: Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables, Clim. Dyn., 50, 31–49, https://doi.org/10.1007/s00382-017-3580-6, 2017.

Chadwick, C., Gironás, J., Vicuña, S., Meza, F., and McPhee, J.: Using a Statistical Pre-Analysis Approach as an Ensemble Technique for the Unbiased Mapping of GCM Changes to Local Stations, J. Hydrometeorol., 19, 1447-1465, https://doi.org/10.1175/JHM-D-17-0198.1, 2018.

Chadwick, C., Gironás, J., González-Leiva, F., and Aedo, S.: Bias adjustment to preserve changes in variability: the unbiased mapping of GCM changes, Hydrol. Sci. J., 68, 1184-1201, https://doi.org/10.1080/02626667.2023.2201450, 2023.

Chen, D., Dai, A., and Hall, A.: The convective-to-total precipitation ratio and the "drizzling" bias in climate models, J. Geophys. Res. Atmos., 126, https://doi.org/10.1029/2020JD034198, 2021.

Chen, J., Brissette, F. P., Chaumont, D., and Braun, M.: Performance and uncertainty evaluation of empirical downscaling methods in quantifying the climate impacts on hydrology over two North American river basins. J. Hydrol., 479, 200–214, https://doi.org/10.1016/j.jhydrol.2012.11.062, 2013.

Christensen J. H., Boberg, F., Christensen, O. B., and Lucas-Picher, P.: On the need for bias correction of regional climate change projections of temperature and precipitation, Geophys. Res. Lett., 35, https://doi.org/10.1029/2008GL035694, 2008.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, L.: Hydrology and Earth System Sciences HESS Opinions "Should we apply bias correction to global and regional climate model data?", Hydrol. Earth Syst. Sci., 16, 3391–3404, https://doi.org/10.5194/hess-16-3391-2012, 2012.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fernandez-Palomino, C.A., Hattermann, F.F., Krysanova, V., Vega-Jácome, F., Menz, C., Gleixner, S., and Bronstert, A.: High-resolution climate projection dataset based on CMIP6 for Peru and Ecuador: BASD-CMIP6-PE, Sci Data, 11, 34, https://doi.org/10.1038/s41597-023-02863-z, 2024.

Fowler, H. J., Blenkinsop, S., and Tebaldi, C.: Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling, Int. J. Climatol., 27, 1547–1578, https://doi.org/10.1002/joc.1556, 2007.

Fowler, H.J., Mearns, L.O., Wilby, R.L.: Downscaling Future Climate Projections: Compounding Uncertainty But Adding Value?, in: Uncertainty in Climate Change Research, Springer, edited by: Mearns, L.O., Forest, C.E., Fowler, H.J., Lempert, R., Wilby, R.L, https://doi.org/10.1007/978-3-031-85542-9_18, 2025.

Gergel, D. R., Malevich, S. B., McCusker, K. E., Tenezakis, E., Delgado, M. T., Fish, M. A., and Kopp, R. E.: Global Downscaled Projections for Climate Impacts Research (GDPCIR): preserving quantile trends for modeling future climate impacts, Geosci. Model Dev., 17, 191–227, https://doi.org/10.5194/gmd-17-191-2024, 2024.

González-Leiva, F., Gironás, J., Chadwick, C., Meza, F., and Vicuña, S.: When will meteorological droughts become the new normal climatic condition in Central-Southern Chile?, Hydrol. Sci. J, 71, 1068–1086. https://doi.org/10.1080/02626667.2026.2625979, 2026.

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol, 377, 80-91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Gutiérrez, J. M., San-Martín, D., Brands, S., Manzanas, R., and Herrera, S.: Reassessing Statistical Downscaling Techniques for Their Robust Application under Climate Change Conditions, J. Clim., 26, 171–188, https://doi.org/10.1175/JCLI-D-11-00687.1, 2013.

Guzmán, J., Gironás, J., Sandoval, C., Chadwick, C., Cienfuegos, R., Nash, S., Muñoz, S., and Escauriaza, C.: Evaluating climate-driven changes in peak flows through continuous hydrological simulation: the case of the Marga Marga basin in central Chile, J. Water. Clim. Change, 17, 870-887, https://doi.org/10.2166/wcc.2026.267, 2026.

Hagemann, S., Chen, C., Haerter, J. O., Heinke, J., Gerten, D., and Piani, C.: Impact of a Statistical Bias Correction on the Projected Hydrological Changes Obtained from Three GCMs and Two Hydrology Models, J. Hydrometeorol., 12, 556–578, https://doi.org/10.1175/2011JHM1336.1, 2011.

Hakala, K., Addor, N., Teutschbein, C., Vis, M., Dakhlaoui, H., and Seibert, J.: Hydrological Modeling of Climate Change Impacts, Encyclopedia of Water: Science, Technology, and Society, https://doi.org/10.1002/9781119300762.wsts0062, 2019.

Hanel, M., Kožín, R., Heřmanovský, M., and Roub, R.: An R package for assessment of statistical downscaling methods for hydrological climate change impact studies, Environ. Model. Softw., 95, 22-28, https://doi.org/10.1016/j.envsoft.2017.03.036, 2017.

Hnilica, J., Hanel, M., and Puš, V.: Multisite bias correction of precipitation data from regional climate models. Int. J. Climatol., 37, 2934–2946, https://doi.org/10.1002/joc.4890, 2017.

Iturbide, M., Bedia, J., Herrera, S., Baño-Medina, J., Fernández, J., Frías, M.D., Manzanas, R., San-Martín, D., Cimadevilla, E., Cofiño, A.S., and Gutiérrez, J.M.: The R-based climate4R open framework for reproducible climate data access and post-processing, Environ. Model. Softw., 111, 42-54, https://doi.org/10.1016/j.envsoft.2018.09.009, 2019.

Kassahun, D., Tilahun, M., Tesema, Z., Meseret, M., Lakew, E., Tesfa, A., Bitew, A., Lakew, M., Bimrow, T., Ayalew, M., Kebede, A., Hunegnaw, B., Adene, A., Dagnaw, D., Alemayehu, M., Tarekegn, G.: Climate-resilient genetic selection in cattle: Insights from resilience phenotypes of Fogera cattle under climate variabilities, Vet. Anim. Sci., 30, https://doi.org/10.1016/j.vas.2025.100538, 2025.

Kim, K. B., Bray, M., and Han, D.: Exploration of optimal time steps for daily precipitation bias correction: a case study using a single grid of RCM on the River Exe in southwest England, Hydrol. Sci. J., 61, 289-301, https://doi.org/10.1080/02626667.2015.1027207, 2016.

Knoben, W. J. M., Freer, J. E., and Woods, R. A.: Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. Hydrol. Earth Syst. Sci., 23, 4323-4331, https://doi.org/10.5194/hess-23-4323-2019, 2019.

Lafferty, D.C., and Sriver, R.L: Downscaling and bias-correction contribute considerable uncertainty to local climate projections in CMIP6, npj Clim Atmos Sci, 6, 158, https://doi.org/10.1038/s41612-023-00486-0, 2023.

Lange, S.: Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0), Geosci. Model Dev., 12, 3055–3070, https://doi.org/10.5194/gmd-12-3055-2019, 2019.

Lanzante, J. R., Dixon, K. W., Nath, M. J., Whitlock, C. E., and Adams-Smith, S.: Some Pitfalls in Statistical Downscaling of Future Climate, Bull. Amer. Meteor. Soc., 99, https://doi.org/10.1175/BAMS-D-17-0046.1, 2018.

Maraun, D., Wetterhall, E. F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H., Sauter, T., Themeßl, M., Venema, V., Chun, K., Goodess, C., Jones, R., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, Rev. Geophys, 48, https://doi.org/10.1029/2009RG000314, 2010.

Maraun, D.: Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums, Geophys. Res. Lett., 39, https://doi.org/10.1029/2012GL051210, 2012.

Maraun, D.: Bias Correcting Climate Change Simulations - a Critical Review, Curr. Clim. Change Rep., 2, 211–220, https://doi.org/10.1007/s40641-016-0050-x, 2016.

Martínez, D., Chadwick, C., and Plaza-Aguilar, A.: The Time of Emergence (ToE) of the Andean Mediterranean sclerophyllous forest of Quillaja saponaria (Mol.) and Lithraea caustica (mol.) Hox. & Arn, For. Ecol. Manag., 544, https://doi.org/10.1016/j.foreco.2023.121169, 2023.

Maurer, E. P., Das, T., and Cayan, D. R.: Errors in climate model daily precipitation and temperature output: Time invariance and implications for bias correction, Hydrol. Earth Syst. Sci., 17, 2147–2159, https://doi.org/10.5194/hess-17-2147-2013, 2013.

Maurer, E. P., and D. W Pierce.: Bias correction can modify climate model simulated precipitation changes without adverse effect on the ensemble mean, Hydrol. Earth Syst. Sci., 18, 915–925, https://doi.org/10.5194/hess-18-915-2014, 2014.

Mauritsen, T., Stevens, B., Roeckner, E., Crueger, T., Esch, M., Giorgetta, M., Haak, H., Jungclaus, J. H., Klocke, D., Matei, D., Mikolajewicz, U., Notz, D., Pincus, R., Schmidt, H., and Tomassini, L.: Tuning the climate of a global model, J. Adv. Model. Earth Syst., 4, https://doi.org/10.1029/2012MS000154, 2012.

Menapace, A., Dhawan, P., Dalla Torre, D., Kaffas, K., Crespi, A., Larcher, M., Righetti, M., and Cannon, A. J.: Review of bias correction methods for climate model outputs in hydrology, J. Hydrol., 660, 133213, https://doi.org/10.1016/j.jhydrol.2025.133213, 2025.

Meyer, J., Kohn, I., Stahl, K., Hakala, K., Seibert, J., and Cannon, A. J.: Effects of univariate and multivariate bias correction on hydrological impact projections in alpine catchments, Hydrol. Earth Syst. Sci., 23, 1339–1354, https://doi.org/10.5194/hess-23-1339-2019, 2019.

Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., Ilyina, T., Kleine, T., Kornblueh, L., Li, H., Modali, K., Notz, D., Pohlmann, H., Roeckner, E., Stemmler, I., Tian, F., and Marotzke, J.: A higher-resolution version of the Max Planck Institute Earth System Model (MPI-ESM 1.2-HR), J. Adv. Model. Earth Sy., 10, 1383–1413, https://doi.org/10.1029/2017MS001217, 2018.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Nash, J. E., and J. V. Sutcliffe.: River flow forecasting through conceptual models: Part 1. A discussion of principles, J. Hydrol., 10, 282-290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Nguyen, N., Anh, T., Nguyen, H., and Dang, D.: Quantile mapping technique for enhancing satellite-derived precipitation data in hydrological modelling: A case study of the Lam River Basin, Vietnam. J. Hydroinform, 26, 2026-2044, https://doi.org/10.2166/hydro.2024.225, 2024.

Petrovic, D., Fersch, B., and Kunstmann, H.: Heat wave characteristics: evaluation of regional climate model performances for Germany, Nat. Hazards Earth Syst. Sci., 24, 265–289, https://doi.org/10.5194/nhess-24-265-2024, 2024.

Pierce, D. W., Cayan, D. R., Maurer, E. P., Abatzoglou, J. T., and Hegewisch, K. C.: Improved bias correction techniques for hydrological simulations of climate change, J. Hydrometeorol., 16, 2421-2442, https://doi.org/10.1175/JHM-D-14-0236.1, 2015.

~~Polasky, A., Evans, J., and Fuentes, J.: CCdownscaling: A Python package for multivariable statistical climate model downscaling, Environ. Model. Softw., 165, https://doi.org/10.1016/j.envsoft.2023.105712, 2023.~~

Portalanza, D., Ortega, C., Garzon, L., Bello, M., Zuluaga, C. F., Bresciani, C., Durigon, A., Ferraz, S.: Projected Heat Waves in Ecuador under Climate Change: Insights from HadGEM-RegCM4 Coupled Model, Earth, 5, 90-109, https://doi.org/10.3390/earth5010005, 2024.

Slater, L. J., Arnal, L., Boucher, M.-A., Chang, A. Y.-Y., Moulds, S., Murphy, C., Nearing, G., Shalev, G., Shen, C., Speight, L., Villarini, G., Wilby, R. L., Wood, A., and Zappa, M.: Hybrid forecasting: blending climate predictions with AI models, Hydrol. Earth Syst. Sci., 27, 1865–1889, https://doi.org/10.5194/hess-27-1865-2023, 2023.

Soares, P. M. M., Johannsen, F., Lima, D. C. A., Lemos, G., Bento, V. A., and Bushenkova, A.: High-resolution downscaling of CMIP6 Earth system and global climate models using deep learning for Iberia, Geosci. Model Dev., 17, 229–259, https://doi.org/10.5194/gmd-17-229-2024, 2024.

Switanek, B. M., Troch, P., Castro, L. C., Leuprecht, A., Chang, H. I., Mukherjee, R., and Demaria, M. C. E.: Scaled distribution mapping: A bias correction method that preserves raw climate model projected changes, Hydrol. Earth Syst. Sci., 21, 2649-2666, https://doi.org/10.5194/hess-21-2649-2017, 2017.

Teutschbein, C., and Seibert, J.: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, J. Hydrol., 456–457, 12–29, https://doi.org/10.1016/j.jhydrol.2012.05.052, 2012.

Themeßl, M., Gobiet, A., and Heinrich, G.: Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal, Clim. Change, 112, 449–468, https://doi.org/10.1007/s10584-011-0224-4, 2012.

Thrasher, B., Wang, W., Michaelis, A., Melton, F., Lee, T., and Nemani, R.: NASA Global Daily Downscaled Projections, CMIP6, Sci Data, 9, 262, https://doi.org/10.1038/s41597-022-01393-4, 2022.

Vicuña, S., Ávila, A., Gironás, J., Melo, Ó., Rivera, D., and Molinos, M.: Contingent Agreements to Reduce the Impacts of Droughts on Human Water Consumption, in: Droughts in Chile. Global Issues in Water Policy, Springer, edited by: Rivera, D., Donoso, G., https://doi.org/10.1007/978-3-031-85040-0_16, 2025.

Walton, D. B., Sun, F., Hall, A., and Capps, S.: A Hybrid Dynamical-Statistical Downscaling Technique. Part I: Development and Validation of the Technique, J. Clim., 28, 4597-4617. https://doi.org/10.1175/JCLI-D-14-00196.1, 2015.

Wang, A., Zhang, S., Jiang, H., and Wang, H.: Bias correction of seasonal precipitation prediction in Southern China: Pointwise and spatial corrections using deep learning methods, J. Hydrol. Reg. Stud., 61, https://doi.org/10.1016/j.ejrh.2025.102737, 2025.

WMO, W. M. O.: WMO Guidelines on the Calculation of Climate Normals, Technical, WMO, Geneva, https://library.wmo.int/idurl/4/55797 (last access: 13 August 2026), 2017.

Wood, A. W., Maurer, E. P., Kumar, A., and Lettenmaier, D. P.: Long-Range Experimental Hydrological Forecasting for the Eastern United States, J. Geophys. Res. Atmos., 107, https://doi.org/10.1029/2001JD000659, 2002.

Yuval, J. and O'Gorman, P.A.: Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions, Nat. Commun., 11, 3295, https://doi.org/10.1038/s41467-020-17142-3, 2020.