Geophysical Research Letters

RESEARCH LETTER
10.1029/2025GL116492

"Which Projections Do I Use?" Strategies for Climate Model Ensemble Subset Selection Based on Regional Stakeholder Needs

A. M. Wootten[1], E. C. Massoud[2], and C. Raymond[3]

[1]South Central Climate Adaptation Science Center, University of Oklahoma, Norman, OK, USA, [2]Integrated Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA, [3]Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, Los Angeles, CA, USA

Key Points:

- Within an ensemble, projected changes from individual downscaled models can differ dramatically pre and post downscaling
- Reliable interpretations require maximum retention of ensemble information and conservative uncertainty assumptions
- Where subsetting and downscaling are applied, model fitness-for-purpose and essential caveats must be clearly described

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. M. Wootten,
amwootte@ou.edu

Citation:

Wootten, A. M., Massoud, E. C., & Raymond, C. (2025). "Which projections do I use?" strategies for climate model ensemble subset selection based on regional stakeholder needs. *Geophysical Research Letters*, *52*, e2025GL116492. https://doi.org/10.1029/2025GL116492

Received 11 APR 2025
Accepted 11 JUN 2025

Author Contributions:

Conceptualization: A. M. Wootten
Formal analysis: A. M. Wootten
Methodology: A. M. Wootten, E. C. Massoud
Software: A. M. Wootten
Visualization: A. M. Wootten
Writing – original draft: A. M. Wootten, E. C. Massoud, C. Raymond
Writing – review & editing: A. M. Wootten, E. C. Massoud, C. Raymond

Abstract

Climate model (or earth system model) projections are increasingly used for climate adaptation planning and impact assessments. As part of this process, many end-users evaluate a subset of downscaled climate projections without being aware of the implications of downscaling methodology for statistics or event outcomes. Approaches for determining a subset of global climate models to use often focus on values from the raw models, rather than from their downscaled counterparts, in other words assuming that the statistical distribution of the multi-model ensemble does not change post downscaling. This study demonstrates that a downscaled ensemble will typically retain the change distribution as a raw ensemble, but individual models can differ dramatically post-downscaling. We recommend that subset-selection methods account for this possibility and that decision-relevant downscaled climate projections provide proper descriptions of fitness-for-purpose and essential caveats, so that non-specialists can interpret the results with an appropriate level of confidence.

Plain Language Summary

Many end-users are now making use of downscaled climate projections for climate adaptation planning and decision-making, oftentimes using a subset of climate projections for their activities. Approaches for selecting subsets of projections often do so based on evaluations of raw climate model data with little regard to the effect of downscaling. This study demonstrates that individual models' outputs can substantially differ pre and post statistical downscaling, which has profound implications for selecting climate projections for assessments and planning. We recommend that decision-relevant downscaled climate projections provide proper descriptions of fitness-for-purpose and all essential caveats to help end-users and boundary organizations interpret climate projection information with appropriate levels of confidence.

1. Introduction

The increasing interest in climate adaptation planning in recent years has heightened the need for future climate projections suitable for impact assessments. Numerous studies have used climate projections to assess climate impacts on multiple sectors, including ecosystems (Pillet et al., 2022), hydrology (Chen et al., 2021; Lewis et al., 2023; Swain et al., 2022), agriculture (Liu et al., 2022; Samimi et al., 2022), groundwater (Massoud et al., 2018) and infrastructure (Caprario et al., 2022; Lai et al., 2022; Zhang et al., 2022). Many such evaluations of possible climate futures necessitate high spatial or temporal resolution, thus requiring the use of downscaled global climate models [GCMs] or other publicly available climate projections. Furthermore, GCM analysis is notoriously data-intensive and historical GCM performance on metrics of interest varies widely, both factors which encourage analysis of only a subset of the available data (Liu et al., 2022) by either the downscaled-dataset producing entity or the end user. It is far from clear, however, that the uncertainties and implications of these choices are broadly understood and accounted for.

The challenge of selecting projections for use in applications is known as the "practitioner's dilemma" (Barsugli et al., 2013) and it includes the process of selecting a subset of projections from (or instead of) a larger ensemble. Many procedures to determine which "climate scenarios" or ensemble subsets to use in adaptation planning have arisen as a result. Some studies propose subset selection by assessing the ability of models to capture large scale physical processes in a target region (Di Virgilio et al., 2022; Eyring et al., 2019) to simulate equilibrium climate sensitivity (ECS) (Hausfather et al., 2022; Schlund et al., 2020), or to capture sensitivities in systems of interest (Miller et al., 2022; Snover et al., 2013). Other studies have put forward analytic (or algorithmic) solutions for selecting GCMs to form an ensemble subset (Herger et al., 2018, 2019; Lawrence et al., 2021; McSweeney &

10.1029/2025GL116492

Jones, 2016; McSweeney et al., 2015; Parding et al., 2020; A. M. Wootten et al., 2024). However, these studies tend to focus on variables related to simple climate model evaluations (e.g., long-term average temperature and precipitation); on the raw GCM output, when downscaled GCM output is often used in adaptation planning; and on continental scales, rather than regional ones. One important exception is Mahony et al. (2022), who considered ways of selecting ensembles specifically for statistical downscaling and regional interests.

Although numerous applications use subsetted and statistically downscaled projections, the required choices involved have not been backed up by robust literature quantifying the influence of the antecedent statistical downscaling process on subset selection, and thus on the ultimate implications for decision-making.

This study asks: when statistical downscaling is applied to a multi-model ensemble, what happens to the ensemble distribution (both the raw values and simulated change signal)? What happens to the individual members and their positions in the ensemble distribution? The remainder of this paper demonstrates the influence of statistical downscaling on a multi-model ensemble, via a case study of precipitation in the US Southern Great Plains region, and concludes with a discussion of how the results matter for the production of robust decision-relevant climate projections and end-user applications.

2. Materials and Methods

We use 29 GCMs from the Coupled Model Intercomparison Project Phase 5 (CMIP5, Andrews et al., 2012). The two ensembles in this study comprise the same 29 GCMs, where one ensemble is the raw CMIP5 data and the other ensemble is drawn from downscaled projections created with the Localized Canonical Analogs (LOCA, Pierce et al., 2014) statistical downscaling method (Table S1 in Supporting Information S1). For this study, the historical period used for the evaluation is 1976–2005 and the future period used for analysis of projected changes is 2070–2099 under Representative Concentration Pathway (RCP) 8.5 (Riahi et al., 2007; van Vuuren et al., 2011). RCP8.5 is used to maximize the change signals for each ensemble member for statistical purposes. In addition, the subset selection makes use of the Livneh v 1.2 gridded observation data set (Livneh et al., 2013) to evaluate historical skill. Raw CMIP5 GCM outputs are interpolated to the Livneh grid using bi-linear interpolation and masked to the continental United States after calculating the climatology for each variable of interest. LOCA2 downscaled outputs for the CMIP6 model suite were not available during the initial analysis, but the effects of model generation are likely small compared to the other issues examined in this study (Cannon, 2020; Knutti et al., 2013). For this analysis, we focus on the Southern Great Plains (SGP) region from the United States National Climate Assessment [NCA] and specifically on precipitation [pr] as an example. We further consider a hypothetical case where a user selects only precipitation projections from three members of the ensemble for additional impact assessments. Results for precipitation in the other NCA regions, the entire continental United States, and individual point locations are available in Supporting Information S1.

3. Downscaling Effects on Ensemble Distributions

Prior studies have mostly considered an ensemble of climate projections as reflecting samples from a distribution of possible future climates. This formulation typically comes in two varieties, one that conceives of a multi-model ensemble as reflecting a random sample of possibilities centered around the true climate (Jun et al., 2008; Tebaldi & Knutti, 2007) and another that conceives of an ensemble as being statistically indistinguishable or exchangeable samples not necessarily centered on the truth (Annan & Hargreaves, 2010, 2011). However, the effects of model genealogy and other limitations in earth-system knowledge or its numerical implementation all tend to lead to undersampling of the true possible distributions (Knutti, 2010; Knutti et al., 2013, 2017; Masson & Knutti, 2011; Sanderson et al., 2015, 2017). For this reason, many studies have considered both skill and independence of climate models when developing multi-model ensemble averages used to project future climates (i.e., Massoud et al., 2019, 2020; Sanderson et al., 2015, 2017; A. M. Wootten, Dixon, et al., 2020; A. M. Wooten, Massoud, et al., 2020; A. M. Wootten et al., 2023). Often neglected, however, is the effect of statistical downscaling on the characteristics of the downscaled multi-model ensemble or its implications for ensemble subset selection. Many statistical downscaling techniques, including LOCA, are designed to retain the original change signal at the different quantiles of a variable from a single target GCM, and these effects have been well-addressed elsewhere (e.g., Lanzante et al., 2017, 2019, 2020, 2021). Here, we instead focus on implications for multi-model ensemble-subset contexts, an area growing in importance as multi-model climate impacts analyses become ever more common across sectors and geographies (Jägermeyr et al., 2021).

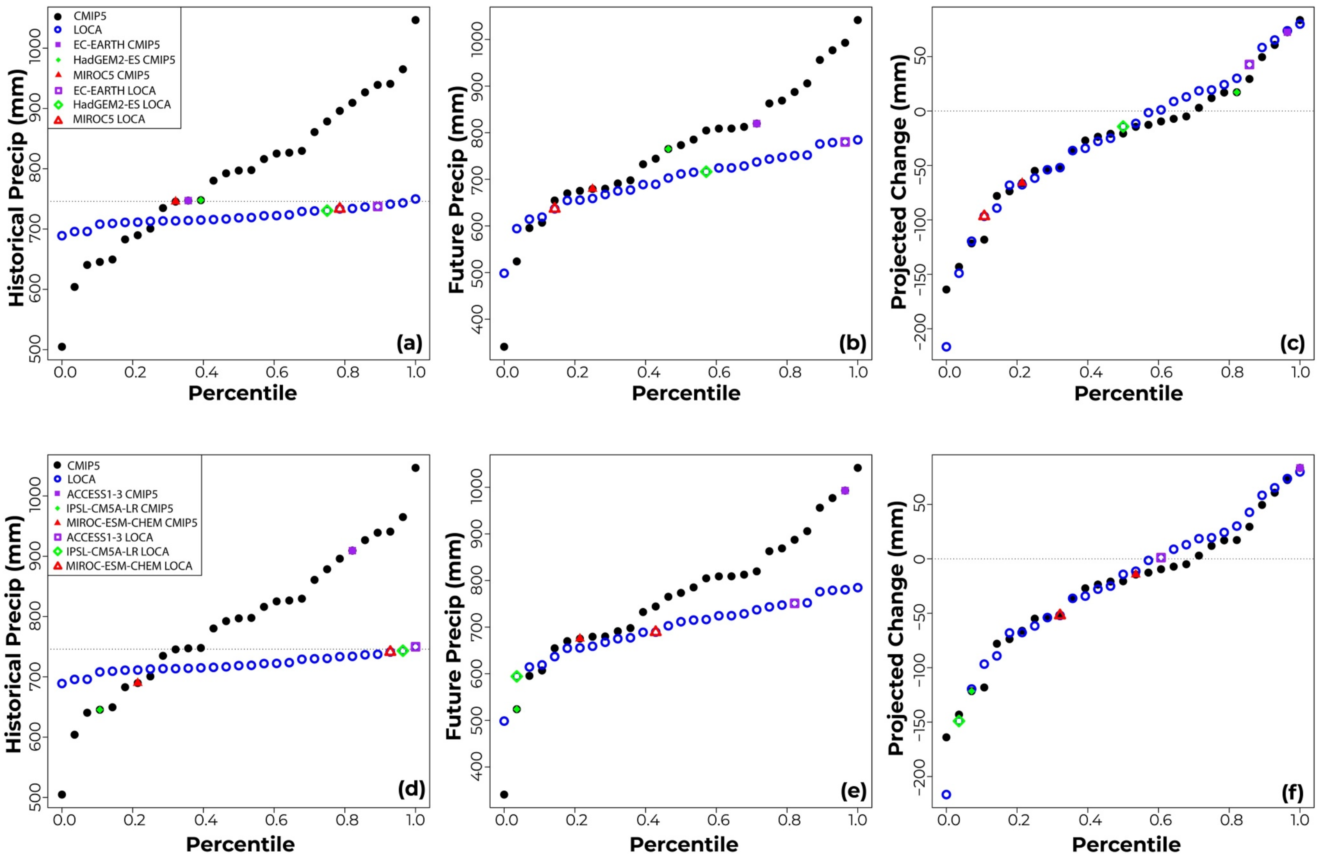

Figure 1. Effect of subsetting on projected future changes from a multi-model ensemble. Each plot shows a 29-member ensemble for annual precipitation climatology in the SGP. Panels (a–c) focus on a 3-member ensemble subset selected based on the bias of raw CMIP5 models; panels (d–f) focus on a 3-member ensemble subset selected based on the bias of LOCA downscaled CMIP5 models. CMIP5 models (pre-downscaling) are shown in black and the LOCA downscaled CMIP5 models are shown in blue. Panels (a and d) are the historical simulated precipitation (1976–2005) for the SGP arranged by percentile for both the CMIP5 and LOCA simulations, with a dashed line showing the historical observed value from Livneh v. 1.2. Panels (b and e) are the simulated future precipitation (RCP 8.5, 2070–2099) for the CMIP5 and LOCA simulations. Panels (c and f) are the simulated change signal for the future precipitation (RCP 8.5, 2070–2099).

Statistical downscaling generally corrects GCM biases and incorporates known physical influences on climatic variables on smaller scales (such as topographic effects on temperature and precipitation) that cannot be represented at the coarse resolution of a GCM (Schoof, 2013). Since GCMs vary in their resolutions and their fidelity in representing regionally important physical processes (Di Virgilio et al., 2022), statistical downscaling can alter the distribution of a multi-model ensemble of GCMs as well as the composition of a best-performing subset. Although the error of a given GCM is reduced following statistical downscaling (pushing all members in closer alignment with observations, and shrinking the variance of the multi-model ensemble), the biases are not identical across GCMs, nor are the corrections imposed by any statistical downscaling method. Therefore, a subtle aspect of statistical downscaling is that in addition to the entire multi-model ensemble distribution being altered, any individual GCM might not retain the same relative position in the ensemble pre and post downscaling. The future multi-model ensemble distribution is subject to the same caveats.

As an illustrative case, we consider the historical skill for the SGP domain mean annual precipitation of the CMIP5 and LOCA ensembles. In a ranking of the historical means (Figure 1a), the best-performing three-GCM subset—selected based on the raw CMIP5 model bias—would fall quite close to the historical observations for the region and in the 30–40th percentile of the CMIP5 multi-model distribution. The same three GCMs have higher error post-downscaling (pre-downscaling mean absolute error = 1.21 mm, post-downscaling mean absolute error = 12.17 mm), but remain close to the rest of the post-downscaled ensemble. The three selected GCMs exhibit substantial differences in their projected changes, which become slightly smaller (i.e., drier) after applying the LOCA method. Most notably, however, where each GCM falls within the multi-model distributions is very different between CMIP5 and LOCA, implying a substantial effect from the statistical downscaling approach in terms of precipitation change and the relative relationship among GCMs (Figures 1a–1c).

Repeating this example but starting with model skill assessed from the LOCA ensemble, we can observe again how downscaling interacts with the ensemble subset selection process. In Figure 1d, the three top GCMs chosen from the LOCA ensemble were, in the CMIP5 ensemble, at very different percentiles of the historical multi-model distribution, and had much more error (pre downscaling mean absolute error = 106.72 mm, post downscaling mean absolute error = 3.76 mm). With statistical downscaling, the ensemble distribution of precipitation and the relative placement of GCMs within the distribution are significantly altered (Figures 1d–1f).

The multi-model ensemble of projections is often considered a sample of the distribution of future climate with the true best estimate lying somewhere in the distribution (Massoud et al., 2019, 2023). For the SGP, statistical downscaling alters both the change signal of the individual GCMs and their position in the downscaled multi-model ensemble, complicating efforts to produce confident estimates at high resolutions. The significant inter-model variability in precipitation in the SGP region is likely driven by the inability of models to represent key physical processes to the region, such as the Great Plains Low Level Jet (GPLLJ) and teleconnections with tropical oceans (Chikamoto et al., 2024; Danco & Martin, 2018). Statistical downscaling attempts to correct errors in GCM output without addressing the underlying cause, thereby leading to concerns about consistency of interpretation of unphysical corrected data, especially for higher-order derived statistics (Chen & Zhang, 2021; Gutiérrez et al., 2013). The larger the GCM error, the larger the correction and the greater this concern—including about whether selected-subset models are 'right for the wrong reasons'.

Regardless of cause, we clearly observe the effect of statistical downscaling on the ensemble distribution and individual members. This is described for the SGP, but this demonstration is easily repeated for other regions, with similar results (summarized in Figure 2): when downscaling is applied, the ensemble distribution shifts, the “best” models change, and both the raw values and change signal can vary considerably (Figure 2, Figures S1–S12 in Supporting Information S1). Both quantitative and qualitative changes of this nature can be seen even for individual locations (Figure 2, Figures S13–S18 in Supporting Information S1, map for individual locations in Figure S19 in Supporting Information S1), of relevance to application-oriented studies that interpolate a target GCM to a high-resolution grid and/or evaluate against point validation data (A. M. Wootten, Dixon, et al., 2020; Yuan & Quiring, 2017).

When statistically downscaling a multi-model ensemble, the raw values of the distribution shift to be aligned with training data but with reduced variance (A. M. Wootten, Dixon, et al., 2020; A. M. Wootten et al., 2023). But what happens to the ensemble distribution of the change signal post-downscaling? In the analyzed case, the downscaled ensemble retains approximately the same mean and variance as the raw GCM ensemble. However, the individual models in an ensemble have different biases and the effect of bias correction from statistical downscaling will also differ between models. Otherwise stated, an ensemble of downscaled projections might retain the same distribution of projected changes as a raw ensemble (e.g., distributions in Figures 1c, 1f, and 2), but individual models might change dramatically post-downscaling (e.g., individual model points in Figures 1c, 1f, and 2), posing major concerns for subset-selection procedures that do not take these effects into account.

4. Relevance to Stakeholder Needs

We next turn our attention to what the above-mentioned effects of statistical downscaling may mean for stakeholder needs for downscaled climate projections. These effects are less of a concern if the capacity exists to analyze a complete ensemble of downscaled GCMs. However, for many applications, analyzing output from dozens of models is not feasible (Di Virgilio et al., 2022), nor do most end-users have the time, information or ability to sort through the implications of, for example, model independence. In these applications, judiciously selecting a subset of models is a requirement.

In this illustrative analysis we use model skill alone to determine the subset selection, but there are numerous other valid approaches, including those based on ECS, process representation, Bayesian model averaging, or global warming levels (Eyring et al., 2019; Hausfather et al., 2022; IPCC, 2022; Maraun & Widmann, 2018; Massoud et al., 2023; Tamoffo et al., 2019). We have already demonstrated that selecting based on skill alone and then downscaling often scrambles the change signal of the original models. Alternatively, one might select based upon the subset's ability to retain the future spread of the change signal from the larger ensemble. For example, some

10.1029/2025GL116492

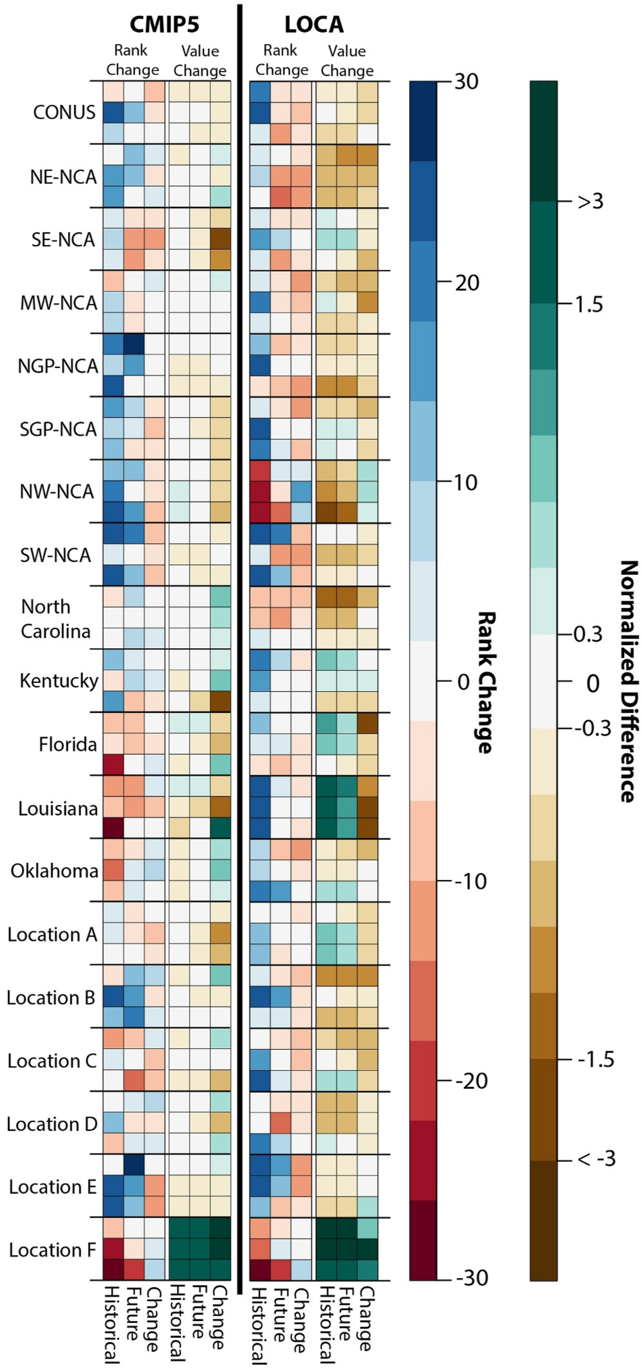

Figure 2.

studies use metrics like fractional range coverage (FRC) as a measure of the spread of the full ensemble captured by the subset (McSweeney & Jones, 2016; Parding et al., 2020). However, selected GCMs are unlikely to retain the same FRC value pre- and post-downscaling (compare Figures 1c and 1f for example). Indeed, with any subset-selection method, the skill and spread of the individual models selected will change to varying degrees post downscaling. Weighting via Bayesian methods is often a useful tool to account for model skill and independence, but this procedure also is sensitive to downscaling, both in individual models and in an ensemble sense (A. M. Wootten, Dixon, et al., 2020; A. M. Wootten et al., 2023). Additionally, many subset selection methods have been developed for use with raw GCM output rather than downscaled data, and approaches based on ECS may not be useful in regions with complex region-specific dynamics (Open-File Report, 2022; Schneider et al., 2022).

Statistical downscaling techniques remain common in many applications due to their conceptual and practical simplicity, and it is therefore important to understand how ensemble subsetting in these situations can be optimized. The computational capacity of some research centers now permits usage of an entire ensemble of GCMs in downscaling, but this remains limited to a few centers focusing on a handful of statistical downscaling techniques and decision-relevant variables. As a result, many users of climate projections for impacts assessments and additional modeling are limited to using a subset of GCMs. Therefore, straightforward and generalizable methods to determine an appropriate ensemble subset are relevant and needed. In addition, while LOCA is a prominent and sophisticated downscaling approach in the United States, it is not the only one. There is also no reason to presume that changes to the multi-model ensemble will be the same between downscaling approaches. It is therefore relevant to explore this issue further with other downscaled climate projections and downscaling approaches. This should include continued analysis with LOCA for aspects of the method that potentially lead to the differences observed here.

Based on the case studies considered here, we highlight several key concerns relevant to the challenge of navigating ensemble subset selection with statistical downscaling for stakeholder applications. First, in the case of an analytic or algorithmic approach to subset selection (as opposed to selecting based on physical features or ECS), it is probable that selecting from a raw versus downscaled ensemble will lead to a different subset of models. Second, while prior literature has argued that selecting based on the GCM ability to capture relevant processes can increase confidence, all models have biases

Figure 2. Changes in model rank and precipitation values due to downscaling for the top three GCMs in each domain. The figure compares subsets of GCMs selected as the top three performers (based on historical precipitation skill) either before downscaling (left panel) or after downscaling (right panel), for each domain. Within each panel, the left group of blocks shows the change in each GCM's rank (percentile position) within the 29-member ensemble after downscaling, where positive values indicate improved rank. The right group of blocks shows the corresponding change in precipitation values, where positive values indicate an increase in precipitation. Each group highlights three metrics: historical precipitation (left column), future precipitation (center column), and projected change in precipitation (right column). For example, for the SGP–NCA domain, the top-left blocks show that the GCM represented by the purple square in Figure 1a improved in rank for historical precipitation after downscaling, with minimal change in absolute value.

that can be affected (and in a physically inconsistent way) by downscaling, considerations that likely carry through to applications to stakeholders and potentially lead to mal-adaptive decisions and planning. However, the broad strengths and weaknesses of ensemble subset selection methods, let alone the additional effect of downscaling as discussed here, have rarely if at all been formally explored.

We therefore make the following recommendations: (a) Because the downscaling technique can significantly alter the multi-model ensemble distribution affecting both raw values and the change signal of individual models, and because there is typically little a priori reason to expect particular GCMs will be more accurate than others when evaluating across different variables, domains, or percentiles of a distribution, we recommend that applications using downscaled projections evaluate as much of the downscaled ensemble as possible to minimize the potential alteration of the raw values and change signal by the statistical downscaling approach. (b) Maximum retention of model information is beneficial to a well-rounded final assessment, requiring independent analyses for each question to avoid overlooking sources of skillful representation. (c) Finally, there is a need for scientists to explore comprehensively the strengths and weaknesses of various ensemble subset selection techniques across applications. Such analysis would contribute to developing standard approaches for creating decision-relevant downscaled climate projections for impacts assessments and climate adaptation planning.

Data Availability Statement

GCM data from CMIP5 models are available from the Earth System Grid Federation (ESGF) Repositories (ESGF User Support Team, 2019). LOCA downscaled CMIP5 GCM data are available through Lawrence Livermore National Laboratory (https://gdo-dcp.llnl.gov/downscaled_cmip_projections/dcpInterface.html). Finally, the R code and data files used for the analysis are available on Zenodo (A. Wootten, 2025).

Acknowledgments

We thank the anonymous reviewers for their comments and critiques to strengthen this article. Funding for this work was provided by the National Aeronautics and Space Administration through the Regional Climate Model Evaluation System (RCMES) project (Award Number: 106029 281945.02.03.03.67). This manuscript has been partially authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan. Financial support was provided by the University of Oklahoma Libraries' Open Access Fund.

References

Andrews, T., Gregory, J. M., Webb, M. J., & Taylor, K. E. (2012). Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models. *Geophysical Research Letters*, *39*(9), 1–7. https://doi.org/10.1029/2012GL051607

Annan, J. D., & Hargreaves, J. C. (2010). Reliability of the CMIP3 ensemble: Reliability of CMIP3. *Geophysical Research Letters*, *37*(2). https://doi.org/10.1029/2009GL041994

Annan, J. D., & Hargreaves, J. C. (2011). Understanding the CMIP3 multimodel ensemble. *Journal of Climate*, *24*(16), 4529–4538. https://doi.org/10.1175/2011JCLI3873.1

Barsugli, J. J., Guentchev, G., Horton, R. M., Wood, A., Mearns, L. O., Liang, X., et al. (2013). The practitioner's dilemma: How to assess the credibility of downscaled climate projections. *Eos, Transactions American Geophysical Union*, *94*(46), 424–425. https://doi.org/10.1002/2013EO460005

Cannon, A. J. (2020). Reductions in daily continental-scale atmospheric circulation biases between generations of global climate models: CMIP5 to CMIP6. *Environmental Research Letters*, *15*(6), 064006. https://doi.org/10.1088/1748-9326/ab7e4f

Caprario, J., Tasca, F. A., Santana, P. L., Azevedo, L. T. S., & Finotti, A. R. (2022). Framework for incorporating climate projections in the integrated planning and management of urban infrastructure. *Urban Climate*, *41*, 101060. https://doi.org/10.1016/j.uclim.2021.101060

Chen, J., & Zhang, X. J. (2021). Challenges and potential solutions in statistical downscaling of precipitation. *Climatic Change*, *165*(3–4), 63. https://doi.org/10.1007/s10584-021-03083-3

Chen, J., Arsenault, R., Brissette, F. P., & Zhang, S. (2021). Climate change impact studies: Should we bias correct climate model outputs or post-process impact model outputs? *Water Resources Research*, *57*(5), e2020WR028638. https://doi.org/10.1029/2020WR028638

Chikamoto, Y., Wang, S. S.-Y., Chang, H.-I., & Castro, C. L. (2024). Tropical interbasin interaction as effective predictors of late-spring precipitation variability in the Southern Great Plains. *Journal of Climate*, *37*(22), 6061–6080. https://doi.org/10.1175/JCLI-D-24-0013.1

Danco, J. F., & Martin, E. R. (2018). Understanding the influence of ENSO on the Great Plains low-level jet in CMIP5 models. *Climate Dynamics*, *51*(4), 1537–1558. https://doi.org/10.1007/s00382-017-3970-9

Di Virgilio, G., Ji, F., Tam, E., Nishant, N., Evans, J. P., Thomas, C., et al. (2022). Selecting CMIP6 GCMs for CORDEX dynamical downscaling: Model performance, independence, and climate change signals. *Earth's Future*, *10*(4), e2021EF002625. https://doi.org/10.1029/2021EF002625

ESGF User Support Working Team. (2019). Earth system gridfederation: Federated nodes. [website] Retrieved from https://aims2.llnl.gov/search/cmip6/

Eyring, V., Cox, P. M., Flato, G. M., Gleckler, P. J., Abramowitz, G., Caldwell, P., et al. (2019). Taking climate model evaluation to the next level. *Nature Climate Change*, *9*(2), 102–110. https://doi.org/10.1038/s41558-018-0355-y

Gutiérrez, J. M., San-Martín, D., Brands, S., Manzanas, R., & Herrera, S. (2013). Reassessing statistical downscaling techniques for their robust application under climate change conditions. *Journal of Climate*, *26*(1), 171–188. https://doi.org/10.1175/JCLI-D-11-00687.1

Hausfather, Z., Marvel, K., Schmidt, G. A., Nielsen-Gammon, J. W., & Zelinka, M. (2022). Climate simulations: Recognize the "hot model" problem. *Nature*, *605*(7908), 26–29. https://doi.org/10.1038/d41586-022-01192-2

Herger, N., Abramowitz, G., Knutti, R., Angelil, O., Lehmann, K., & Sanderson, B. M. (2018). Selecting a climate model subset to optimise key ensemble properties. *Earth System Dynamics*, *9*(1), 135–151. https://doi.org/10.5194/esd-9-135-2018

Herger, N., Abramowitz, G., Sherwood, S., Knutti, R., Angélil, O., & Sisson, S. A. (2019). Ensemble optimisation, multiple constraints and overconfidence: A case study with future Australian precipitation change. *Climate Dynamics*, *53*(3–4), 1581–1596. https://doi.org/10.1007/s00382-019-04690-8

IPCC. (2022). Global warming of 1.5°C: IPCC special report on impacts of global warming of 1.5°C above pre-industrial levels in context of strengthening response to climate change. In *Sustainable development, and efforts to eradicate poverty* (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781009157940

Jägermeyr, J., Müller, C., Ruane, A. C., Elliott, J., Balkovic, J., Castillo, O., et al. (2021). Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. *Nature Food*, *2*(11), 873–885. https://doi.org/10.1038/s43016-021-00400-y

Jun, M., Knutti, R., & Nychka, D. W. (2008). Spatial analysis to quantify numerical model bias and dependence. *Journal of the American Statistical Association*, *103*(483), 934–947. https://doi.org/10.1198/016214507000001265

Knutti, R. (2010). The end of model democracy? *Climatic Change*, *102*(3), 395–404. https://doi.org/10.1007/s10584-010-9800-2

Knutti, R., Masson, D., & Gettelman, A. (2013). Climate model genealogy: Generation CMIP5 and how we got there. *Geophysical Research Letters*, *40*(6), 1194–1199. https://doi.org/10.1002/grl.50256

Knutti, R., Sedláček, J., Sanderson, B. M., Lorenz, R., Fischer, E. M., & Eyring, V. (2017). A climate model projection weighting scheme accounting for performance and interdependence. *Geophysical Research Letters*, *44*(4), 1909–1918. https://doi.org/10.1002/2016GL072012

Lai, Y., Lopez-Cantu, T., Dzombak, D. A., & Samaras, C. (2022). Framing the use of climate model projections in infrastructure engineering: Practices, uncertainties, and recommendations. *Journal of Infrastructure Systems*, *28*(3), 04022020. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000685

Lanzante, J. R., Adams-Smith, D., Dixon, K. W., Nath, M., & Whitlock, C. E. (2020). Evaluation of some distributional downscaling methods as applied to daily maximum temperature with emphasis on extremes. *International Journal of Climatology*, *40*(3), 1571–1585. https://doi.org/10.1002/joc.6288

Lanzante, J. R., Dixon, K. W., Adams-Smith, D., Nath, M. J., & Whitlock, C. E. (2021). Evaluation of some distributional downscaling methods as applied to daily precipitation with an eye towards extremes. *International Journal of Climatology*, *41*(5), 3186–3202. https://doi.org/10.1002/joc.7013

Lanzante, J. R., Dixon, K. W., Nath, M. J., Whitlock, C. E., & Adams-Smith, D. (2017). Some pitfalls in statistical downscaling of future climate. *Bulletin of the American Meteorological Society*, *99*(4), 791–803. https://doi.org/10.1175/BAMS-D-17-0046.1

Lanzante, J. R., Nath, M. J., Whitlock, C. E., Dixon, K. W., & Adams-Smith, D. (2019). Evaluation and improvement of tail behaviour in the cumulative distribution function transform downscaling method. *International Journal of Climatology*, *39*(4), 2449–2460. https://doi.org/10.1002/joc.5964

Lawrence, D. J., Runyon, A. N., Gross, J. E., Schuurman, G. W., & Miller, B. W. (2021). Divergent, plausible, and relevant climate futures for near- and long-term resource planning. *Climatic Change*, *167*(3–4), 38. https://doi.org/10.1007/s10584-021-03169-y

Lewis, J. W., Lytle, S. E., & Tavakoly, A. A. (2023). Climate change projections of continental-scale streamflow across the Mississippi River Basin. *Theoretical and Applied Climatology*, *151*(3–4), 1013–1034. https://doi.org/10.1007/s00704-022-04243-w

Liu, X., Liu, W., Tang, Q., Liu, B., Wada, Y., & Yang, H. (2022). Global agricultural water scarcity assessment incorporating blue and green water availability under future climate change. *Earth's Future*, *10*(4). https://doi.org/10.1029/2021EF002567

Livneh, B., Rosenberg, E. A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K. M., et al. (2013). A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions. *Journal of Climate*, *26*(23), 9384–9392. https://doi.org/10.1175/JCLI-D-12-00508.1

Mahony, C. R., Wang, T., Hamann, A., & Cannon, A. J. (2022). A global climate model ensemble for downscaled monthly climate normals over North America. *International Journal of Climatology*, *42*(11), 5871–5891. https://doi.org/10.1002/joc.7566

Maraun, D., & Widmann, M. (2018). Cross-validation of bias-corrected climate simulations is misleading. *Hydrology and Earth System Sciences*, *22*(9), 4867–4873. https://doi.org/10.5194/hess-22-4867-2018

Masson, D., & Knutti, R. (2011). Climate model genealogy: Climate model genealogy. *Geophysical Research Letters*, *38*(8). https://doi.org/10.1029/2011GL046864

Massoud, E. C., Espinoza, V., Guan, B., & Waliser, D. E. (2019). Global climate model ensemble approaches for future projections of atmospheric rivers. *Earth's Future*, *7*(10), 1136–1151. https://doi.org/10.1029/2019EF001249

Massoud, E. C., Lee, H. K., Terando, A., & Wehner, M. (2023). Bayesian weighting of climate models based on climate sensitivity. *Communications Earth and Environment*, *4*(1), 365. https://doi.org/10.1038/s43247-023-01009-8

Massoud, E. C., Lee, H., Gibson, P. B., Loikith, P., & Waliser, D. E. (2020). Bayesian model averaging of climate model projections constrained by precipitation observations over the contiguous United States. *Journal of Hydrometeorology*, *21*(10), 2401–2418. https://doi.org/10.1175/JHM-D-19-0258.1

Massoud, E. C., Purdy, A. J., Miro, M. E., & Famiglietti, J. S. (2018). Projecting groundwater storage changes in California's Central Valley. *Scientific Reports*, *8*(1), 12917. https://doi.org/10.1038/s41598-018-31210-1

McSweeney, C. F., & Jones, R. G. (2016). How representative is the spread of climate projections from the 5 CMIP5 GCMs used in ISI-MIP? *Climate Services*, *1*, 24–29. https://doi.org/10.1016/j.cliser.2016.02.001

McSweeney, C. F., Jones, R. G., Lee, R. W., & Rowell, D. P. (2015). Selecting CMIP5 GCMs for downscaling over multiple regions. *Climate Dynamics*, *44*(11–12), 3237–3260. https://doi.org/10.1007/s00382-014-2418-8

Miller, B. W., Schuurman, G. W., Symstad, A. J., Runyon, A. N., & Robb, B. C. (2022). Conservation under uncertainty: Innovations in participatory climate change scenario planning from U.S. national parks. *Conservation Science and Practice*, *4*(3). https://doi.org/10.1111/csp2.12633

Open-File Report, 2022 Open-File Report. (2022). (Open-File Report).

Parding, K. M., Dobler, A., McSweeney, C. F., Landgren, O. A., Benestad, R., Erlandsen, H. B., et al. (2020). GCMeval—An interactive tool for evaluation and selection of climate model ensembles. *Climate Services*, *18*, 100167. https://doi.org/10.1016/j.cliser.2020.100167

Pierce, D. W., Cayan, D. R., & Thrasher, B. L. (2014). Statistical downscaling using localized constructed analogs (LOCA). *Journal of Hydrometeorology*, *15*(6), 2558–2585. https://doi.org/10.1175/JHM-D-14-0082.1

Pillet, M., Goettsch, B., Merow, C., Maitner, B., Feng, X., Roehrdanz, P. R., & Enquist, B. J. (2022). Elevated extinction risk of cacti under climate change. *Nature Plants*, *8*(4), 366–372. https://doi.org/10.1038/s41477-022-01130-0

Riahi, K., Grübler, A., & Nakicenovic, N. (2007). Scenarios of long-term socio-economic and environmental development under climate stabilization. *Technological Forecasting and Social Change*, *74*(7), 887–935. https://doi.org/10.1016/j.techfore.2006.05.026

Samimi, M., Mirchi, A., Townsend, N., Gutzler, D., Daggubati, S., Ahn, S., et al. (2022). Climate change impacts on agricultural water availability in the middle Rio Grande basin. *JAWRA Journal of the American Water Resources Association*, *58*(2), 164–184. https://doi.org/10.1111/1752-1688.12988

Sanderson, B. M., Knutti, R., & Caldwell, P. (2015). A representative democracy to reduce interdependency in a multimodel ensemble. *Journal of Climate*, *28*(13), 5171–5194. https://doi.org/10.1175/JCLI-D-14-00362.1

10.1029/2025GL116492

Sanderson, B. M., Wehner, M., & Knutti, R. (2017). Skill and independence weighting for multi-model assessments. *Geoscientific Model Development*, *10*(6), 2379–2395. https://doi.org/10.5194/gmd-10-2379-2017

Schlund, M., Lauer, A., Gentine, P., Sherwood, S. C., & Eyring, V. (2020). Emergent constraints on equilibrium climate sensitivity in CMIP5: Do they hold for CMIP6? *Earth System Dynamics*, *11*(4), 1233–1258. https://doi.org/10.5194/esd-11-1233-2020

Schneider, D. P., Kay, J. E., & Hannay, C. (2022). Cloud and surface albedo feedbacks reshape 21st century warming in successive generations of an earth system model. *Geophysical Research Letters*, *49*(19). https://doi.org/10.1029/2022GL100653

Schoof, J. T. (2013). Statistical downscaling in climatology. *Geography Compass*, *7*, 249–265.

Snover, A. K., Mantua, N. J., Littell, J. S., Alexander, M. A., Mcclure, M. M., & Nye, J. (2013). Choosing and using climate-change scenarios for ecological-impact assessments and conservation decisions: Choosing and using climate-change scenarios. *Conservation Biology*, *27*(6), 1147–1157. https://doi.org/10.1111/cobi.12163

Swain, S., Taloor, A. K., Dhal, L., Sahoo, S., & Al-Ansari, N. (2022). Impact of climate change on groundwater hydrology: A comprehensive review and current status of the Indian hydrogeology. *Applied Water Science*, *12*(6), 120. https://doi.org/10.1007/s13201-022-01652-0

Tamoffo, A. T., Moufouma-Okia, W., Dosio, A., James, R., Pokam, W. M., Vondou, D. A., et al. (2019). Process-oriented assessment of RCA4 regional climate model projections over the Congo Basin under 1.5°C and 2°C global warming levels: Influence of regional moisture fluxes. *Climate Dynamics*, *53*(3–4), 1911–1935. https://doi.org/10.1007/s00382-019-04751-y

Tebaldi, C., & Knutti, R. (2007). The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *365*(1857), 2053–2075. https://doi.org/10.1098/rsta.2007.2076

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., et al. (2011). The representative concentration pathways: An overview. *Climatic Change*, *109*(1), 5–31. https://doi.org/10.1007/s10584-011-0148-z

Wootten, A. (2025). CMIP5 and LOCA climatologies for historical and RCP8.5 with R code and mask files [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.15192146

Wootten, A. M., Başağaoğlu, H., Bertetti, F. P., Chakraborty, D., Sharma, C., Samimi, M., & Mirchi, A. (2024). Customized statistically downscaled CMIP5 and CMIP6 projections: Application in the Edwards Aquifer region in south-central Texas. *Earth's Future*, *12*(10), e2024EF004716. https://doi.org/10.1029/2024EF004716

Wootten, A. M., Dixon, K. W., Adams-Smith, D. J., & McPherson, R. A. (2020). Statistically downscaled precipitation sensitivity to gridded observation data and downscaling technique. *International Journal of Climatology*, *41*(2), 1–22. https://doi.org/10.1002/joc.6716

Wootten, A. M., Massoud, E. C., Sengupta, A., Waliser, D. E., & Lee, H. (2020). The effect of statistical downscaling on the weighting of multi-model ensembles of precipitation. *Climate*, *8*(12), 138. https://doi.org/10.3390/cli8120138

Wootten, A. M., Massoud, E. C., Waliser, D. E., & Lee, H. (2023). Assessing sensitivities of climate model weighting to multiple methods, variables, and domains in the south-central United States. *Earth System Dynamics*, *14*(1), 121–145. https://doi.org/10.5194/esd-14-121-2023

Yuan, S., & Quiring, S. M. (2017). Evaluation of soil moisture in CMIP5 simulations over the contiguous United States using in situ and satellite observations. *Hydrology and Earth System Sciences*, *21*(4), 2203–2218. https://doi.org/10.5194/hess-21-2203-2017

Zhang, Y., Ayyub, B. M., & Fung, J. F. (2022). Projections of corrosion and deterioration of infrastructure in United States coasts under a changing climate. *Resilient Cities and Structures*, *1*(1), 98–109. https://doi.org/10.1016/j.rcns.2022.04.004