Data Adoption Justification Memo (for California's Fifth Climate Change Assessment):
General Use Projections

Julie Kalansky, Peter Yao, Dan Cayan, David Pierce - Scripps Institution of Oceanography, UCSD.
CEC-funded agreement: EPC-20-006 Development of Climate Projections for California and Identification of General Use Projections
August 2024

Methods and Prior Relevant Work:

The climate data and projections developed for California as part of the project funded by the California Energy Commission (CEC) through its Electric Program Investment Charge (EPIC) Program, EPC-20-006, consist of upwards of 20 TB of data, which can be overwhelming if one is not accustomed to working with such large data sets. To lower this data barrier and to reduce the number of individual projections that might initially be considered, a subset of the downscaled CMIP6 Global Climate Models (GCMs) and ensemble members was selected as "general use projections". These general use projections are meant as an entry point to working with this larger full suite of LOCA2-Hybrid climate data and to begin to explore the range of possible outcomes. Given the broad range of variability and longer term outcomes that are represented across the LOCA2-Hybrid suite, the general use projections are a minimum set of GCMs that could be used in analysis of climate change impacts and assessments. They can also serve as a smaller, common set of projections for comparison between cross-sector, cross-jurisdiction, and/or cross-agency work if not all agencies are using the same larger data set of GCMs and ensemble members. The general use projections should not be considered a preferred set of GCMs or ensemble members, but rather an entry point into climate data utilization that mitigates challenges from using the larger data set.

The full suite of LOCA2-Hybrid climate data is a result of the hybrid-statistical downscaling method.[1] This downscaling method created data for 15 GCMs with a total of 199 ensemble runs. The process used to select the general use projections from the full suite included several criteria, listed below with associated rationales.

1. Ensured the models selected capture much of the range (see below, figure 1) of projected future changes in the GCMs and ensemble members for selected metrics, scenarios and time-periods. This selection aimed to help mitigate uneven sampling that might arise in using a smaller set of GCMs (e.g. not capturing possible extremes, biasing the results towards a certain projection) rather than the full suite of LOCA2-Hybrid GCMs.
2. Prioritized LOCA2-Hybrid GCMs and ensemble members that have also been dynamically downscaled by WRF with a priori bias correction. These WRF projections are considered to be superior to WRF runs with no a priori bias correction[2], and furthermore, this selected set enables the use of both dynamically downscaled and hybrid downscaled data sets if needed.
3. Prioritized ensemble members for which the full set of LOCA2-Hybrid downscaled variables were available (some GCM providers did not save all of the requisite data for LOCA2 downscaling) in order to provide a robust general use projection dataset.

[1] Please see https://www.energy.ca.gov/sites/default/files/2024-03/04_HybridDownscaling_DataJustificationMemo_Pierce_Adopted_ada.pdf

[2] Please see the Data Justification Memo "01 - Bias Correction in the WRF and LOCA2 Projections" https://www.energy.ca.gov/media/9326

4. Assessed future changes in California in both means and extremes (hot days, floods, and droughts) to ascertain that the general use subset of the LOCA2-Hybrid data suite covered a plausible range of extremes. A limited set of metrics was employed in this evaluation because having too many uncorrelated variables and metrics reduces the ability of the metrics to distinguish between models. Model range was assessed using temperature, precipitation, and wind.
5. Avoided metrics that are derivatives of several variables for the same reason as in 4. above.
6. Focused the assessment on the range of model outcomes for SSP370 with a 2045-2074 mid-century period based on input from the energy sector. This time period was used because it is far enough in the future to capture climate change signals, and energy users indicated that it is a high priority time horizon for planning and adaptation actions. SSP370 was considered because it is a mid-range climate forcing and climate change scenario that exhibits greenhouse gas emissions that are consistent with those that Earth is currently experiencing, and additionally is the only SSP for which WRF data that is part of EPC-20-006 is available.
7. Excluded GCMs that have a climate sensitivity that is unrealistically high compared to historical observations, based on Tokarska et al., 2020 and Hausfather et al., 2022. These models produce a rate of future warming that is unlikely to come to pass. This choice does not prevent any stakeholder from analyzing the data in terms of warming levels but avoids the requirement that the data from the general use projections must be analyzed in warming levels.

The variables and metrics for the general use projections selection process were based on input from the energy stakeholders and scientific considerations. Based on this we selected to use precipitation, maximum temperature, and wind. These are the primary variables that were of interest from the energy sector in terms of extremes such as heat waves, droughts, and fire weather. Primary climate variables (seasonal or annual mean temperature and precipitation) were used rather than the extremes because the extremes can be defined numerous ways, however, we find there is a good correspondence between projected changes in the means and various extremes of interest to the energy community. The relationship between the climate variable and the associated extremes are described in the full General Use Projection Memo (Kalansky et al., 2024). Further, using a statewide metric was necessary for the general use projections. Temperature and precipitation trends are more homogenous throughout California than some other metrics like solar radiation.

We determined that the following 5 models and ensemble members best met the 7 prioritized criteria stated above.

- ACCESS-CM2 r1i1p1f1 (no WRF dynamical downscaled 3 km run)
- MPI-ESM1-2-HR r3i1p1f1
- EC-Earth r1i1p1f1
- FGOALS-g3 r1i1p1f1 (was not *a priori* bias corrected)
- MIROC6 r1i1p1f1

Guidance or Caveats on Best Practices for Use of Data Products:

Although using the general use projections as a minimum mitigates some pitfalls of using an arbitrarily selected subset of the full data set, several limitations remain when using a subset rather than the full suite of GCMs. Subsets do not represent the full range of all the possible climate projections, and in

particular may not fully characterize projected changes in extremes in a specific region of interest. For this reason, we always recommend using a broad set of models with as many of the GCMs as possible. The Cal-Adapt Analytics Engine enables exploration and analysis of the full suite of climate data without having to download and process the data independently, though it does require some training and experience with climate data. More information about the implication of model choice and sampling is available in the guidance on how to use climate data that is available on the Cal-Adapt: Analytics Engine website, https://analytics.cal-adapt.org/guidance/.

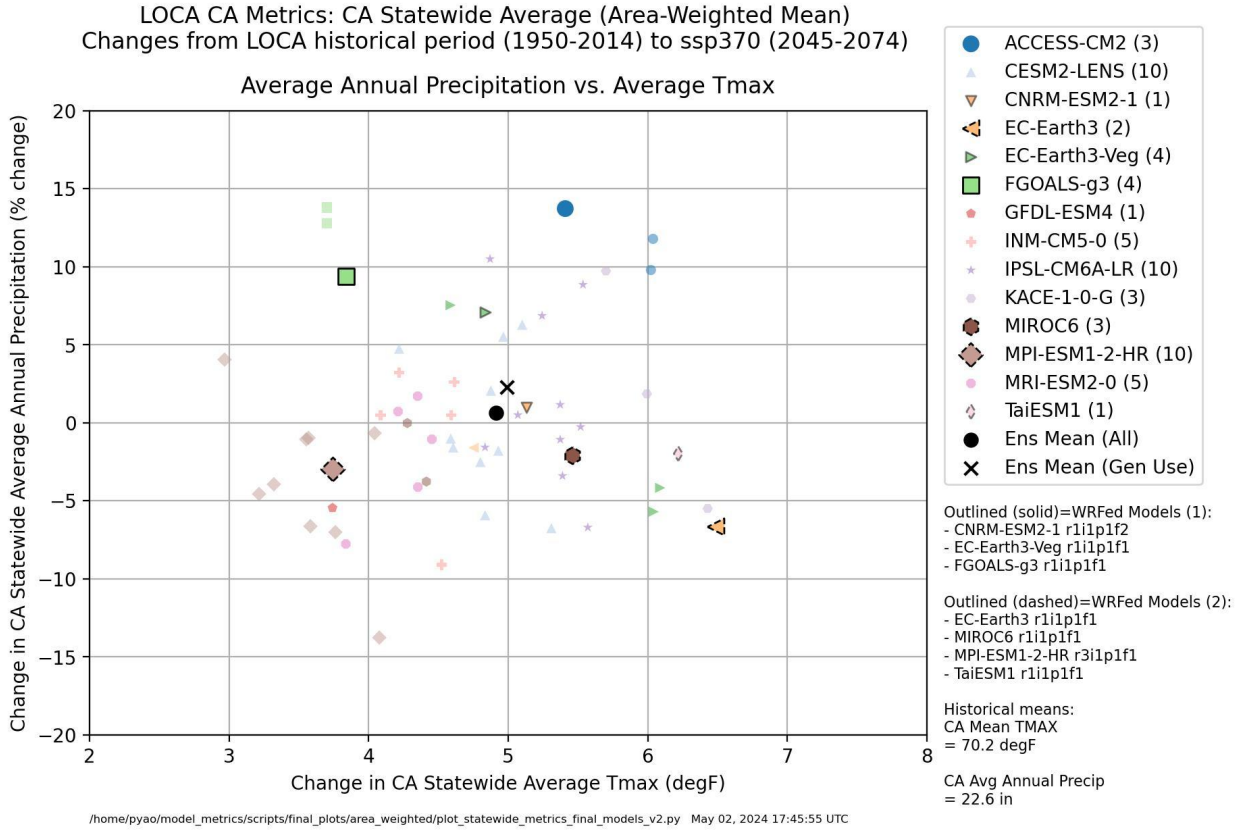

Figure 1. The change in CA statewide average Tmax (in degrees Farenheit) compared to change in statewide average annual precipitation (as a percent change) for the SSP370 scenario during mid-century (2045-2074) relative to the historical period (1950-2014), with the symbol indicating the GCM. Many of the GCMs have multiple ensemble members, indicated in parentheses after each GCM in the legend. The shapes outlined in black are the GCM ensemble members that have been dynamically downscaled with WRF. Symbols for the 5 model ensemble members selected for the general use projections are larger and more opaque than the remaining models. The filled black circle is the multi-model ensemble mean of all the models while the black x is the mean of the general use projections.

References

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

Kalansky, J.F., P. Yao, D. W. Pierce, D.R. Cayan (2024). Memo on General Use Projections. Deliverable to the California Energy Commission, https://www.energy.ca.gov/media/10478

Tokarska, K. B., M. B. Stople, S. Sippel, E. M. Fischer, C. J. Smith, F. Lehner, R. Knutti (2020). Past warming trend constraints future warming in CMIP6 models. *Science* 6(12). DOI: 10.1126/sciadv.aaz9549