Data Adoption Justification Memo (for California's Fifth Climate Change Assessment):
Targeted Extremes Database

Dan Cayan, Josh Mumford, Julie Kalansky Scripps Institution of Oceanography, UCSD.
CEC-funded agreement: EPC-20-006 Development of Climate Projections for California and Identification of General Use Projections
May 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. Looking through this data to identify extremes in the projections can be challenging. To help identify extremes, we have created a database (excel files) that identify a handful of different extremes from the entire LOCA2-Hybrid data set.

Prior to selecting the extremes, we worked with the CEC, other CEC supported research projects using the projections and Investor-Owned Utilities (IOUs) to identify the type of extremes that would support their research efforts and vulnerability assessments. Through this process, we learned that the various stakeholders had different uses and definitions of *extremes*; for almost all extremes, the duration was balanced by the intensity. In addition, users have different regions of interest. Due to this, the extremes were identified in three regions covering the northern, central, and southern parts of California (Figure 1) over separate historical and future periods (1950-2014, 2015-2040, 2041 -2070, 2071-2100; the latter two are mid-century and end of century for the CA 5th Climate Change Assessment respectively).

The extremes were calculated from area average precipitation and area average temperature for each region, identified for different durations covering few to several day periods where events often are more acute, one to several year periods which have lower intensities than the shorter period extremes but may still have high impacts. These durations were determined based on the typical impacts from the selected extremes; for example, floods range from 1-30 days, whereas dry and wet periods were provided over range from 1 to 20 years, to address issues of different longer duration extremes that are needed by varied users. The extremes and the associated durations are listed in Table 1. In addition to values in mm or °C units, the extremes are presented as fractions (%, precipitation) or difference [°C] of/from the historical (1950-2014) downscaled GCM modeled climatological values.

Table 1. Extremes that were identified and the durations

Variable	Extreme	Duration of Events Identified
Precipitation	Drought; driest	Water years*: 1, 2, 3, 5, 10 and 20 years
Precipitation	Wet Years; wettest	Water years: 1, 2, 3, 5, 10 and 20 years
Minimum Daily Temp (Tmin)	Cold Snaps; lowest Tmin	Day Average: 1, 2, 3, 7, 14, 30 days in Dec – Feb
Maximum Daily Temp (Tmax)	Heatwaves; highest Tmax	Day Average: 1, 2, 3, 7, 14, 30 days in Jun-Aug
Precipitation	Flood (inferred) highest precipitation	Accumulated Precipitation: 1, 2, 3, 7, 14, 30 day

*the water years begin on October 1st and go through September 30th the following year

Not included in the table is a description of the occurrence of compound extremes. However, the suite of these same and other variables, also provided in the LOCA2-Hybrid dataset (available on the Cal-Adapt Analytics Engine) allows for such a survey. For example, the LOCA2 precipitation and temperature projections can be used to estimate the co-occurrence of drought and heatwaves, where a 2-year drought was identified along with the highest 1-day temperature.

To facilitate using the dynamically downscaled data in extreme case studies, the table identifies which models were dynamically downscaled by WRF in the databases. The resulting list of extreme cases are represented in three excel spread sheets that together make the Targeted Extreme Database.

QA/QC & Uncertainty

As demonstrated by the results in the excel files (the Targeted Extreme Database), there is substantial variation between models, model variants (ensemble members), SSPs, and time periods. This variation expresses the uncertainty that exists in the occurrence of extremes, and the envelope of extremes that is presented across the entire set of model runs may provide a better representation of the range of outcomes that could occur than can be obtained from any single model run.

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

The targeted extreme databases are meant to support investigation of user-defined types of extremes because neither the duration, location nor period are pre-determined. Prior to examination of the database, users need to determine the characteristics of the extremes of interest, with particular focus on region, duration and intensity based on the options in the database.

Extreme cases can be used in scenario development or stress cases to ensure infrastructure can adapt to potential future extremes (e.g. Dumas et al. 2019). The structure of the database enables a relatively quick examination of the selected extreme amongst the LOCA2-Hybrid downscaled data (which is available on the Cal-Adapt Analytics Engine). This facilitates an assessment of how the selected extreme compares between the various GCMs and ensembles providing a range of extremes that could possibly occur during the early-, mid-, and late-21st Century future, and provides a comparison with extremes within the historical period of the respective model runs.

An important consideration for scenario and/or stress case development is the downscaling method used to develop data used in these applications. Dynamically WRF downscaled projections are dynamically consistent throughout the event meaning that the atmospheric features and the physical processes that govern them are part of the downscaling methodology. For example. If a user is concerned about a large atmospheric river event for flooding, the winds and precipitation are linked through these processes in the dynamical downscaling. In contrast, each variable, which the except of Tmin and Tmax, are downscaled separately and do not maintain this same dynamical consistency as the dynamically WRF data. For extremes cases, this may be an important consideration. As such, the dynamically downscaled models are noted in the database. By comparing the dynamically downscaled extremes to the rest of the LOCA2-Hybrid models in the Targeted Extreme Database places the extremes in context of the other models as mentioned previously. As a reminder, as described in other associated data justification memos, the research team has recommended to use the a-priori bias corrected dynamically downscaled model runs (MIROC6, TaiESM1, EC-Earth3, MPI-ESM1-2-HR, EC-Earth2-Veg). Other non-a priori bias corrected models should only be employed on condition that the user is familiar with the associated errors that have been documented.

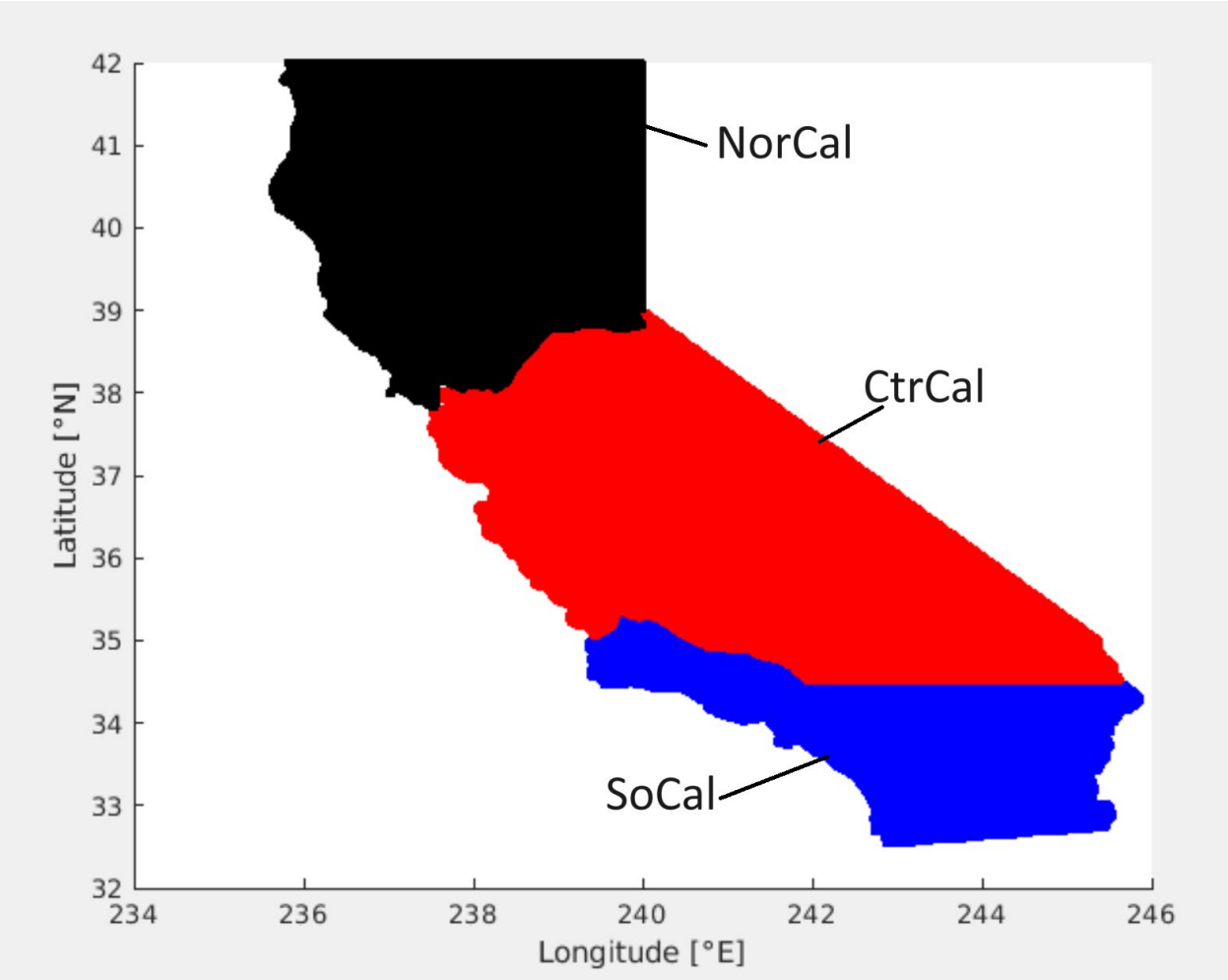

Figure 1. Northern, central and southern California regions over which extremes are identified.