

Description of Data Sources



1. Climate Model Data

We used climate model output from the Coupled Model Intercomparison Project, Phase 6 (CMIP6). CMIP6 is a project where climate modeling centers from around the globe run their climate models for the same set of model experiments (i.e. use the same externally-defined inputs, such as greenhouse gas emissions) and provide their model output in a mostly consistent output format (e.g. use the same variable names). More information about CMIP6 is available at <https://www.wcrp-climate.org/wgcm-cmip>. For our tool, we use the output from (1) the historical experiment, which recreates the past to validate the models, and from (2) four experiments representing different possible future scenarios, which are described in more detail in [Section 1.2: Future Scenarios](#).

Due to the large size of climate model data, we use the google-cloud-based CMIP6 data archive and the ocean.pangeo.io analysis hub to perform analysis. CMIP6 data is also available for download via the [Earth System Grid](#). Currently, our tool only uses the subset of climate models that had been uploaded onto the google-cloud-based CMIP6 data archive as of November 30, 2019 (summarized in [Section 1.3, Data Availability](#)). We have structured our code so that this will be automatically updated as more model output becomes available on the google-cloud-based CMIP6 data archive, which is anticipated to occur on an ongoing basis over the next several months.

1.1 Data Structure

- **Type:** multiple netcdf files (1 file for each variable, model, scenario, and ensemble member)
- **Dimensions:** latitude/longitude/time
- **Size:** about 200 MB per variable/model/scenario/ensemble member combination, about 441 variable/model/scenario/ensemble members

1.2 Future Scenarios: Shared Socioeconomic Pathways

As part of CMIP6, climate modeling centers run various experiments that represent different possible pathways that the global community could take in the future, which are called shared socioeconomic pathways (SSPs). Different SSPs make different assumptions about how the world would develop (e.g. what if we continue to develop at current rates and fuel that development with our current fossil-fuel-heavy energy supply? what if we transition quickly to a more sustainable development trajectory oriented towards less material growth and using less

Description of Data Sources

carbon-intensive energy?). In our tool, we use the four “Tier 1” SSP model experiments, where “Tier 1” indicates that these are the highest-priority model experiments and most modeling centers will run them. These four SSPs are described in Table 1. More detail about the SSPs are provided by this [Carbon Brief summary article](#) (appropriate for general audiences) and this detailed technical summary by [Riahi et al. 2017](#) (best for audiences with some background in climate science and climate policy).

SSP Abbreviation	SSP Name
SSP1	Sustainability
SSP2	Middle-of-the-road
SSP3	Regional Rivalry
SSP5	Fossil Fueled Development

Table 1: Summary of Socioeconomic Pathways (SSPs) used in our analysis.

1.3 Data Availability

Table 2 summarizes the subset of climate models that had been uploaded onto the google-cloud-based CMIP6 data archive for each scenario as of November 30, 2019.

Scenario	Models with data available on google-cloud-based CMIP6 data archive
Historical	17
SSP1: Sustainability	10
SSP2: Middle-of-the-road	10
SSP3: Regional Rivalry	8
SSP5: Fossil Fueled Development	9

Table 2: Summary of data availability as of November 30, 2019

Description of Data Sources



2. Historical Observation Dataset

We use historical temperature observations from the Berkeley Earth Surface Temperature (BEST) dataset. The BEST dataset consists of monthly means of land surface air temperature observations that have been structured onto a $1^\circ \times 1^\circ$ latitude-longitude grid, although observations may not be available for every point on the grid at all time steps. The BEST dataset, and additional information about its development, is available at <http://berkeleyearth.org/>. A copy of the version of the dataset that we downloaded on November 22, 2019 is available [here](#).

2.1 Data Structure

- **Type:** netcdf
- **Dimensions:** latitude/longitude/time
- **Size:** 202 MB