A bias-corrected & downscaled massive ensemble to diagnose uncertainty in climate impact projections

Kevin Schwarzwald *[1,2], Nathan Lenssen[3,4], Radley M. Horton[1,2], and Gernot Wagner[5]

[1]Columbia Climate School, New York, NY, USA

[2]Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY, USA

[3]Colorado School of Mines, Golden, CO, USA

[4]NSF National Center for Atmospheric Research, Boulder, CO, USA

[5]Columbia Business School, New York, NY, USA

Abstract

Projections of climate change and climate impacts requires bias-corrected, downscaled output from ensembles of earth system models (ESMs). Potential impacts are uncertain due to modeling differences between ESMs, internal variability stemming from the chaos of the earth system, and differences in the historical reference datasets used to bias-corrected and downscale ESM output. Here, we introduce the Bias-Corrected and Downscaled Massive Ensemble (BCD-ME), a multi-model large ensemble of over 1,400 projections of daily mean and maximum temperature. The BCD-ME samples model and internal uncertainty with up to 86 runs from 12 Large Ensembles and uncertainty in the reference dataset by using 4 different reanalysis products to bias-correct and downscale output. Output is organized by Global Warming Levels (GWL), accounting for differences between forcing scenarios and ESM climate sensitivities. The ensemble contains 20-year daily time series for each GWL on a uniform 1° grid, bias-corrected using Quantile Delta Mapping, and statistics of 20-year time series for each GWL on a uniform 0.25° grid, downscaled using Quantile-Preserving Localized Analog Downscaling. The BCD-ME is stored in Analysis-Ready, Cloud-Optimized (ARCO) Zarr stores, enabling efficient computation of the 97 TB (uncompressed) dataset.

*Correspondence to ks3753@columbia.edu

Background & Summary

Climate impact assessments help predict changes in human and natural systems due to climate change. Generally, these studies seek to find a statistical or process-based relationship between the climate and an impact over some observed, historical period and then use this relationship to predict changes in the impact given some change to the climate. Estimates of future climates are critical in this process, generally referred to as "climate projections". Climate projections are most commonly derived from simulations made with dynamical Earth System Models (ESMs) as part of the Coupled Model Intercomparison Project (CMIP) [1].

There are three major knowledge barriers to properly using CMIP-class ESMs' climate projections in climate impacts studies. First, uncertainty in future climate must be included and properly propagated to the estimate of future impacts, stemming from differences between forcing scenarios, differences between ESMs, and the internal variability due to the chaotic nature of the earth system [2–6]. The climate impacts community is increasingly accounting for climate uncertainty due to possible future scenarios and differences in climate models [7]. Including uncertainty due to internal climate variability has been shown to be important as well, but has lagged, since proper treatment of internal uncertainty requires applying analyses to hundreds of climate model projections [6].

Second, raw values from ESM output are biased relative to the true Earth's climate in both the mean and higher moments of key climate variables. While ESMs simulate the climate with underlying physics, chemistry, and biology as currently understood, due to modeling limitations they do not exactly replicate the observed climate, and are run with no requirement that the output has equal statistics to the true Earth system [4;8]; ESM output must, thus, be bias-corrected before it can be used in impacts contexts [4].

Finally, CMIP-class ESMs generally utilize atmospheric models with spatial resolutions of $0.7 - 2.8°$ (approx. 100-300 km at the equator) and are therefore unable to provide local estimates of key variables at impact-relevant spatial scales, particularly in regions with complex geography [9;10]. The Multi-Model Large Ensemble Archive (MMLEA) provides a comparably cleaned and organized projection dataset at 2.5° resolution [11]. ESM output is, thus, often downscaled before use in impacts contexts, particularly in contexts where multi-model and/or large ensembles must be used.

Bias-correction and statistical downscaling methods require a spatially and temporally complete product of observed climate. Climate reanalyses or other gridded historical data products are typically used for this purpose as direct weather observations have gaps in spatial and temporal coverage, particularly prior to the satellite era. However, reanalyses provide imperfect estimates of climate variables with uncertainty arising due to limitations in the dynamical model, the assimilation scheme, and uncertainty in the data assimilated into the reanalysis, adding an additional source of uncertainty in climate projections.

To properly overcome all three of these knowledge barriers, a research group must therefore have sufficient computational resources and technical expertise to download and store hundreds of climate model projections, implement or develop bias-correction and downscaling methods, and apply these to each climate projection. Together, these steps are a massive investment in computing, disk storage, and software development that ought to be supported by expertise from climate and statistical experts. The scale of this investment is a major barrier to entry into studying climate impacts. Existing publicly-available products aimed at impacts research, such as NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP)[12], the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP)[13], or the Global Downscaled Projections for Climate Impact Research (GDPCIR)[14] provide global bias-corrected and downscaled climate projections but do not sample internal variability (only one ensemble member per ESM is provided). A notable exception are the recent GARD-LENS[15] and LOCA2[16] products, which provide output from several large ensembles at very high resolution over North America, or Semenov et al. (2025)[17], which does so at locations in the United Kingdom; however, these products are limited to particular regions. Furthermore, none of these products samples the uncertainty in datasets used to post-process ESM output (only one reanalysis or other gridded historical reference dataset is used in each product).

Summary of the BCD-ME

To fill this gap, in this data descriptor we present the Bias-Corrected & Downscaled Massive Ensemble (BCD-ME), a near-surface air temperature projection dataset designed to make it as straightforward as possible to include climate uncertainty in analyses. Here, we use "Massive" to denote that the ensemble is a multi-model large ensemble; a collection of large ensembles of projections from various ESMs. The BCD-ME addresses the knowledge barriers discussed above, including uncertainty due to the reanalysis product used. The key features of this dataset include:

- **Daily average and maximum temperature with near-global land coverage** to enable global and regional analyses of temperature projections and climate impact projections.

- **Indexed by global warming levels (GWLs)** to facilitate direct comparisons of ESMs with different climate sensitivities.

- **Bias-Corrected** time series on a uniform 1° grid using Quantile Delta Mapping (QDM)[18].

- **Downscaled** statistics of bias-corrected time series on a 0.25° grid using Quantile-Preserving Localized Analog Downscaling (QPLAD)[14]. Full time series at 0.25° generated with the same methods will be made public in future versions of the BCD-ME.

- **Sampling climate model uncertainty** using 12 different ESMs
- **Sampling internal variability** using between 5 and 86 ensemble members from each ESM.
- **Sampling "ground-truth" uncertainty** using 4 reanalyses of historical climate to bias-correct and downscale ESM output.
- **Provided in Zarr stores** and on the cloud for efficient computation.

We use four products as possible observational ground truths in the bias-correction and downscaling. (1) The European Centre for Medium-Range Weather Forecasting Reanalysis v5 (ERA5)[19] (2) The NASA Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA2)[20] (3) The Japanese Meteorological Agency Japanese Reanalysis for Three Quarters of a Century (JRA-3Q)[21] (4) The Global Meteorological Forcing Dataset (GMFD)[22]. The first three are state-of-the-art operational reanalyses. GMFD is included as it is used to bias-correct the NASA NEX-GDDP[12] and is therefore commonly used in climate impact research[23–25].

The BCD-ME contains two collections of data: (1) bias-corrected time series on a uniform $1^\circ \times 1^\circ$ grid and (2) selected metrics commonly used in climate impacts studies on a uniform $0.25^\circ \times 0.25^\circ$ (see Table 3 for descriptions). We provide statistics on the downscaled 0.25° grid, rather than the full daily time series, to increase the usability of this very large dataset. The BCD-ME time series collection provides bias-corrected time series 2m daily temperature (`tas`) and 2m maximum daily temperature (`tasmax`) for all locations from 56°S to 86°N (land-only for GMFD-based ensemble members), spanning all permanently inhabited regions. The 0.25° downscaled product provides bias-corrected and downscaled metrics for all land locations over the same spatial extent.

The presented BCD-ME dataset contains on the order of 100 times more impact assessment-ready temperature projections than the commonly used NASA NEX-GDDP dataset[12]. The organization of the BCD-ME enables users to easily account for all sources of climate uncertainty by indexing across uncertainty-relevant dimensions. The computational costs of scaling an impacts analysis by a factor of 100 is reduced by providing the BCD-ME as I/O-optimized Zarr stores that enable code bases to be easily scaled and parallelized on modest systems.

In the remainder of this paper, we outline the source data and methods used to construct the BCD-ME, provide a validation of the dataset, and provide suggestions for best practices when working with this data. We expect that the BCD-ME will be useful for users of climate projection information by enabling the robust quantification of uncertainty in future climate.

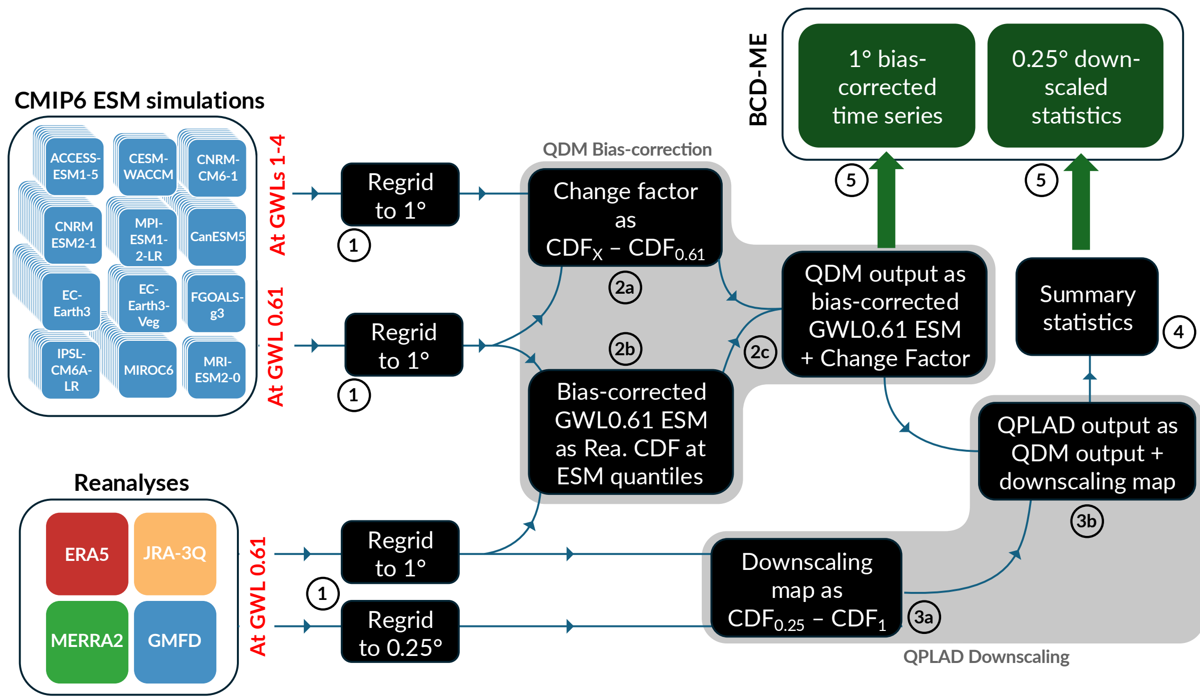

Figure 1: Diagram of bias-correcting and downscaling workflow. Reanalysis data for the 20 years around GWL0.61 (1982-2001) is regridded to uniform 1° and 0.25° grids; data from ESM runs for the 20 years around GWLs 0.61 and 0.5°C steps from GWL1 to 4 (or the highest GWL step reached) are regridded to a uniform 1° grid (Step **1**). Processing is done separately for each combination of an ESM run and a reanalysis 'ground truth'. 1-degree ESM and reanalysis data is used to create bias-corrected time series through the QDM methodology (Step **2**). That bias-corrected output is combined with a downscaling map from reanalysis between 1° and 0.25° to create downscaled time series through the QPLAD methodology (Step **3**). Finally, summary statistics (detailed in Table 3) of the 0.25°-degree time series are calculated (Step **4**). For each ESM run and reanalysis, one set of bias-corrected time series on a 1° grid and one set of statistics of the bias-corrected and downscaled time series on a 0.25° grid are thus saved, forming the BCD-ME (Step **5**).

Methods

There are four major components in the creation of the BCD-ME (see summary diagram in Figure 1). Section details the ESM and reanalysis data used as the raw inputs to the BCD-ME. Section discusses how GWLs are calculated for each climate projection, how the data is organized by these GWLs, and presents analysis justifying the use of GWLs. Section details the bias-correction methods and Section details the downscaling method.

Reanalysis	Resolution	Publisher	Citation
ERA5	0.25°	European Centre for Medium-Range Weather Forecasting (ECMWF)	19
MERRA2	0.5° x 0.625°	National Aeronautics and Space Administration (NASA)	27
JRA-3Q	0.375°	Japan Meteorological Agency (JMA)	21
GMFD	0.25 °	Princeton Hydrology Group	22

Table 1: Reanalysis products used as historical "ground truths" for bias-correction and downscaling

Input Data

CMIP6 Large Ensemble MIP

We use all simulations from CMIP6 models with more than 5 ensemble members available at time of processing on Pangeo's CMIP6 cloud storage (`https://console.cloud.google.com/marketplace/product/noaa-public/cmip6`) that have daily output for daily mean (`tas`) or max (`tasmax`) near-surface air temperature for a historical experiment starting at the latest on January 1, 1850, and a future ScenarioMIP experiment forced with any Shared Socioeconomic Pathway (SSP)[26], resulting in 12 ESMs with ensemble sizes ranging from 8 to 86. Currently, this includes 365 total members for `tas` and 190 for `tasmax`, with future versions of the dataset including more members as they are processed or become available on pangeo. These ensemble sizes include all projections run with the model, across all SSPs. A full description of the ESMs used is shown in Table 4. Note that ESM output is regridded bilinearly to a uniform 1-degree latitude/longitude grid before bias-correction (Step **1** in Figure 1).

Global Reanalysis

Bias-correction and downscaling both rely on knowing the Earth's true climate over some observed time period. However, there is non-negligible uncertainty in estimates of the true historical climate[28;29]. Despite this uncertainty, climate impact studies generally use a single reanalysis product, failing to sample both uncertainty in the raw data assimilated into the reanalysis as well as structural uncertainties in the underlying dynamical model. In the BCD-ME, we sample this reanalysis uncertainty by repeating the bias-correction and downscaling with four reanalyses (see Table 1): the European Centre for Medium-Range Weather Forecasting ERA5[19], NASA MERRA2[27], the Japan Meteorological Agency JRA-3Q[21], and GMFD[22]. ERA5, MERRA2, and JRA-3Q are latest-generation, continuously updating global reanalysis products, while GMFD is used as the "ground truth" for NASA's NEX GDDP[12] set of bias-corrected and downscaled climate projections and is therefore frequently used as a source of historical weather observations in projections of climate impacts[23–25]. While this is not a complete assessment of uncertainty in the observed climate, there are substantial differences between these four products, providing the ability to determine the sensitivity of climate impact projections to the choice of ground-truth observations.

ERA5[30], GMFD[31], and JRA-3Q[32] data were accessed through the NSF NCAR Geoscience Data Exchange (GDEX) and can be accessed through the replication code; MERRA-2[33] was downloaded from NASA's Goddard Earth Sciences Data and Information Services Center (GES DISC).

Reanalysis products are used at two resolutions (Step **1** in Figure 1). For bias-correction, output from reanalyses is bilinearily regridded to a uniform 1-degree latitude/longitude grid. For downscaling, output from reanalyses is bilinearily regridded to a uniform 0.25-degree latitude/longitude grid that exactly nests into the 1-degree grid such that each 1-degree grid cell contains 4 0.25-degree grid cells.

Note that 0.25 degree is a finer resolution than two of the reanalyses used (MERRA2 and JRA-3Q). Thus, the 0.25 degree projections from MERRA2 and JRA-3Q cannot contain information finer than the native grids of these reanalyses (See Table 1) and are constructed by interpolation from these native reanalysis grids. Nevertheless, we provide all four downscaled products at the 0.25 degree grid to facilitate comparison. Users should use caution when interpreting fine-scale spatial features in MERRA2 and JRA-3Q 0.25 degree projections.

Global Warming Levels

Rather than providing projections as continuous time series over the 21st century, we instead organize the BCD-ME by 20-year daily time series at given Global Warming Levels (GWLs). GWLs are degrees of global mean near-surface air temperature warming relative to the 1850-1900 baseline[34]. For instance, the '2°C world' envisioned by the Paris Agreement considers climate conditions and impacts at GWL2.

Future climate impacts are increasingly presented by GWL (including in the sixth IPCC assessment report), as are climate datasets aimed at impacts work[35;36], due to several key advantages over time-series-based projections. First, data presented at GWLs is agnostic to climate sensitivity, helping avoid the so called "hot model" problem from ESMs with possibly unrealistically high climate sensitivities in CMIP6[37]. Note that this removes intermodel differences in global climate sensitivity as a driver of model uncertainty; the intermodel spread will therefore be reduced compared to an analysis at a comparative set of calendar years.

Second, data presented at GWLs control for differences between climate scenarios, such as the Shared Socioeconomic Pathways (SSPs)[26] used in the CMIP6 ScenarioMIP set of experiments. These SSPs are socioeconomic storylines of future anthropogenic forcings, such as greenhouse gas and aerosol concentrations or land use patterns, and are not intended to provide a statistical sampling of plausible future conditions. Abstracting away from differences in warming rates between SSPs allows a user to assume their own preferred timeline of warming (including by examining scenarios such as the Paris Agreement or using GMSTs estimated from integrated assessment models).

Third, by pooling ensemble members from different SSPs at the same GWL, users of climate data are able to boost the ensemble size for each ESM. These larger ensemble sizes enable characterization of ESM internal variability for a larger set of ESMs[38].

GWL Identification and Data Organization

In the BCD-ME, we utilize a global warming level index `GWL` rather than indexing by a calendar year in the future, providing 20 years of daily simulations at 8 GWLs: 0.61, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, and 4.0 °C. The 0.61 °C equals the observed global mean temperature over the 20 year period of 1982-2001 in HadCRUT5[28] and was used as a reference time to bias-correct and downscale ESM output, and together with GWL1 (2001-2020 in HadCRUT5) may be useful for climate impact studies that investigate the changes in climate risk compared to a period close to present.

A GWL is defined by the global mean temperature anomaly relative to a 1850-1900 baseline following the IPCC definition for warming compared to the preindustrial. For each climate model ensemble member, we calculate the global mean surface annual temperature anomaly and compute a 20 year moving average. Then, for each GWL, we find the first year that this moving average crosses the given GWL, and define the 20 years around this year as the climate for this GWL[34;39].

For a given GWL, 20 years of daily temperature fields are then provided in the BCD-ME, indexed by `year` from `1:20` and `dayofyear` from `1:365`. As an example, in the r1i1p1f1 member of the MPI-ESM1-LR model run under SSP2.45, the 20-year rolling average of global mean temperature is more than 1 degree above its 1850-1900 mean for the first time in 2013. Thus, we define the 20-year time period of 2003-2022 as an ensemble member for this model for the 1.0 °C GWL. We additionally provide the calendar years from the ESM simulation as the auxiliary dimension `calendar_year`, but recommend users utilize the GWL setup of the data whenever possible. Note that many model projections will not reach every GWL, with only the warmest models under high emissions scenarios reaching the 4.0 °C GWL. The number of simulations available at each GWL for each model is presented in Table 2.

It is important to note that not all ESM-GWL combinations have sufficient ensemble sizes to provide robust estimates of internal variability. The exact ensemble size varies substantially by location and statistic, with required ensemble sizes for monthly mean temperature to be within an acceptable error 0.25 °C ranging from fewer than 10 over much of the global ocean and tropics to over 50 in the high northern latitudes[38]. However, we determined it was advantageous to provide a greater number of ESMs to enable better quantification of model uncertainty. Users should exercise caution when estimating internal variability with smaller ensemble sizes and confirm results against ensembles with larger ensemble sizes. This is of particular

Variable	GWL reached: Model	1.0	1.5	2.0	2.5	3.0	3.5	4.0
tas	ACCESS-ESM1-5	24	24	24	24	5	5	5
	CESM2-WACCM	8	8	8	7	7	4	3
	CNRM-CM6-1	18	18	18	18	15	12	6
	CNRM-ESM2-1	41	41	31	26	13	7	5
	CanESM5	86	86	86	86	86	86	79
	EC-Earth3	35	35	35	35	32	15	10
	EC-Earth3-Veg	13	13	13	13	13	9	6
	FGOALS-g3	10	10	10	7	3	0	0
	IPSL-CM6A-LR	28	28	28	28	28	18	17
	MIROC6	60	34	34	31	28	0	0
	MPI-ESM1-2-LR	30	30	30	20	16	6	0
	MRI-ESM2-0	12	8	8	8	7	2	0
	Total	365	335	325	303	253	164	131
tasmax	ACCESS-ESM1-5	6	6	6	6	6	6	6
	CESM2-WACCM	0	0	0	0	0	0	0
	CNRM-CM6-1	5	5	5	3	2	2	1
	CNRM-ESM2-1	5	5	3	3	2	1	1
	CanESM5	30	30	30	30	30	30	27
	EC-Earth3	22	22	22	22	19	8	3
	EC-Earth3-Veg	10	10	10	10	10	5	2
	FGOALS-g3	7	7	7	4	0	0	0
	IPSL-CM6A-LR	22	22	22	22	22	12	11
	MIROC6	37	25	25	22	19	0	0
	MPI-ESM1-2-LR	30	30	30	20	16	6	0
	MRI-ESM2-0	16	16	16	14	11	6	0
	Total	190	178	176	156	137	76	51

Table 2: Number of runs in the BCD-ME that (as of version 1) have data for at least a given GWL, by ESM, for near-surface mean daily air temperature (`tas`, top half) and near-surface maximum daily air temperature (`tasmax`, bottom half). Note that some ESM-GWL combinations are of insufficient size to fully capture internal variability[38].

importance at higher GWLs such as 3.5 and 4.0 where only a few members of a few ESMs reach these levels (Table 2).

Each store in the BCD-ME contains a coordinate called "`has_data`" (see Section), which provides a boolean flag for whether a particular ensemble member has data for a particular GWL, and the sample code linked below provides an example of how to use it to subset the ensemble to only models with more than a particular number of ensemble members for a particular GWL of interest.

Bias-Correction: Quantile Delta Mapping

Since the probability distributions of climate variables in an ESM differ from those in the true Earth, ESM output is generally bias-corrected in climate impacts context to make values directly usable in climate impact models trained on historical observations.

We use the Quantile Delta Mapping (QDM)[18] method (Step **2** in Figure 1). In addition to the future ESM time series that is to be bias-corrected, the QDM process requires a ground truth "reference" time series and a baseline time series from the same ESM simulation over a comparative historical period. In our implementation, the "reference" time series is taken from reanalysis. A common definition for this comparative historical period is to take the same years in the reference and baseline ESM time series; for reasons detailed above in Section , we instead match them by GWL.

For each variable, location, and day-of-year, QDM first calculates the empirical quantile of each future ESM timestep relative to the projection distribution within each ensemble member. That is, we calculate a time series of quantiles; if a given time step is the 0.6 quantile, it is warmer than 60% of all time steps for that location, day-of-year, and projection period.

Then, the values in the reference reanalysis time series are rearranged so that it has the same sequence of empirical quantiles and the future ESM time series. This preserves the actual values from the reanalysis, but arranges them to match the temporal sequence of the projection through the quantiles. This method produces what Cannon et al. (2015)[18] refer to as a 'detrended,' 'bias-corrected' future ESM time series.

Then, the difference between the projection and baseline ESM distributions is calculated at each quantile and day-of-year, determining the 'change factor' between the reference and future periods for the entire distribution. That is, if the 0.6 quantile is 20 °C in the baseline ESM time series, but 23 °C in the future ESM time series, the change factor would be +3 °C.

Finally, this change factor is added to the reanalysis values at the future ESM quantiles to reintroduce the climate change signal. Thus, each bias-corrected projection time series is the reanalysis series reordered to match the quantile series from the ESM projection series and adjusted at each quantile by the difference between the reference reanalysis period and the ESM projection.

More formally, the change factor is based on differences in empirical cumulative distribution functions (CDFs). Each day-of-year is processed separately, using a moving window of 31 days centered at that day-of-year to calculate CDFs. At a given location and each day of year, the changes between the empirical CDF of the ESM run at a baseline GWL (0.61, representing 1982-2001 in the real world) and the empirical CDF of a future GWL is found, using 100 equally-spaced quantiles to estimate the CDF (Step **2a**). Then, for each day in the future period, the same quantile is found in the reference reanalysis dataset (Step **2b**), and the change factor–the difference in the ESM CDF at that quantile and day of year between the reference and future GWL–is added to it (Step **2c**). Since, for each day-of-year, 100 equally-spaced quantiles are used to create the change factor but the model set contains 621 (20 years x 31 days) elements, linear interpolation is performed in quantile space to determine the change factor for each quantile. For ranks beyond the edges of the change factor distribution, the slope for the linear interpolation is calculated from the lowest and second-lowest quantiles (for values below the minimum) or between the highest and second-highest quantiles (for values above the maximum).

The difference between the future ESM and GWL 0.61 ESM CDFs thus provides the necessary change factor of the reanalysis data. That is, for $x_m(t)$, a ESM projection m at time t in the future, we calculate the adjusted $x^*_m(t)$ as:

$$x^*_m(t) = x_m(t) + \left(F_o^{-1}(q) - F_m^{-1}(q)\right) \tag{1}$$

where F_o^{-1} and F_m^{-1} are the empirical inverse CDFs of the observed historical climate and the ESM climate respectively. Finally, q is the quantile associated with the value of $x_m(t)$ in the ESM's simulation. We roughly follow the notation in Section 3.1 of Gergel et al. (2024)[14] which contains more background on this method.

As an example, if the 0.6 quantile of a ESM simulation at a location for March 1st at GWL1 is 14.5 °C and the 0.6 quantile of the same ESM run at this same location and day-of-year at GWL0.61 is 13.8 °C, the 0.6 quantile from the reanalysis for this day-of-year is projected by adding 0.7 °C.

Downscaling: Quantile-Preserving Localized-Analog Downscaling

The Large Ensemble ESMs simulation used here for temperature projections have native grids ranging in resolution from 0.7° to 2.8° (approx. 100-300 km at the equator). There is need in the climate impact community for information at as fine a spatial scale as possible, motivating the downscaling of these products. Here, we implement Quantile-Preserving Localized-Analog Downscaling (QPLAD) for statistically downscaling the ESM output to a finer resolution reanalysis grid, following Gergel et al. (2024)[14] (Step **3** in Figure 1).

Similar to the QDM bias-correction, QPLAD is a non-parametric quantile-based method that provides estimates of the temperature at the fine-scale. First, a downscaling map is created using the reference reanalysis datasets (Step **3a** in Figure 1). We use the 1982-2001 reanalysis data regridded to a 1-degree grid from above, and also bilinearily regrid the reanalysis data to a uniform 0.25-degree grid that perfectly slots into the 1-degree grid such that each low-resolution grid cell exactly contains 4 high-resolution grid cells. For each reanalysis, the downscaling map is the difference in empirical CDFs (using all empirical quantiles) between the 1-degree data and the 0.25-degree data. As before, separate CDFs are calculated for each day-of-year, using a 31-day rolling window across each of the 20 years centered at that day-of-year, resulting in 620 (31 x 20) empirical quantiles.

Then, for each day at every location and GWL in each 1-degree bias-corrected ESM simulation, the downscaling map at the relevant day-of-year and quantile (as before, calculated relative to a 31-day-of-year moving window across years) is added to its value to produce values at each of the 4 nested 0.25-degree grid cells within (Step **3b** in Figure 1). See Gergel et al. (2024)[14] for more details.

We note that, of the four reanalysis datasets used as downscaling "ground truths", JRA-3Q (0.375°) and MERRA2 (0.5° x 0.675°) are of lower resolution than the 0.25-degree resolution of the downscaled datasets. We use them to downscale to the same uniform 0.25-degree grid as reference, but note that values of the downscaling map at a higher resolution than their native resolution are statistically interpolated and do not provide more information from the model. Users should therefore treat those values with caution, and a corresponding flag (`resolution_flag`) is included with downscaled data indicating these reanalyses.

Downscaled statistics

Due to the size of the dataset, this initial version of the BCD-ME does not provide full time series at the 0.25-degree resolution of the downscaled data. Rather, we save summary statistics averaged across the 20 years for each GWL (Step **4** in Figure 1; see Table 3 for details). Our choice of statistics is motivated by the empirical climate impacts literature, which relates historical variations in climate (characterized by some weather variable) to socioeconomic outcomes through simple "dose-response functions," and then project damages by plugging in projected changes in this climate variable into the function[40;41]. Common inputs to these dose-response functions include the number of days per year in temperature bins[42–44] and annually summed polynomials of daily temperature[23–25].

For both mean and maximum daily temperature, we thus provide the number of days in 1 °C and 5 °F bins (`tasbinC`, `tasbinF`, `tasmaxbinC`, and `tasmaxbinF`). In addition to their common use in non-parametric statistical impact models, bin-days allow the analysis of exceedences of particular temperature thresholds, which are commonly used as triggers in public heat warning systems[45]. We also provide summed polynomials up to degree 4, allowing for flexible nonlinear dose-response functions, as `tassumpoly` and `tasmaxsumpoly`:

$$sumpoly = \sum_{d=1}^{365} T^k \text{ for } k \in \{1, 2, 3, 4\}$$

In other words, k = 2 shows the sum of squared temperature. The $k = 1$ polynomial can also be used to calculate annual average temperature, by dividing by 365.

This set of statistics is sufficient to create projections of most temperature-based impact functions summarized in the reviews of Carleton et al. (2016)[40] (see their Fig. 3) or Hsiang (2026)[41] (see their Fig. 1).

Future updates to the BCD-ME will include full 0.25-degree resolution time series.

Metric	Units	Description	Notes
Bias-corrected variables (1° grid)			
`tas`	K	Near-surface mean daily temperature	Time series
`tasmax`	K	Near-surface maximum daily temperature	Time series
Bias-corrected and downscaled metrics (0.25° grid)			
`tasbinC`	days / year	Mean days per year in Celsius bins	1-deg bins from -21.5 to 51.5°C
`tasbinF`	days / year	Mean days per year in Fahrenheit bins	5-deg bins from -5 to 130 F
`tassumpoly`	$^\circ\mathrm{C}^k$ for $k \in (1, 2, 3, 4)$	Mean annual sums of polynomial mean temperature	
`tasmaxbinC`	days / year	Mean days per year in Celsius bins	1-deg bins from -21.5 to 51.5°C
`tasmaxbinF`	days / year	Mean days per year in Fahrenheit bins	5-deg bins from -5 to 130 F
`tasmaxsumpoly`	$^\circ\mathrm{C}^k$ for $k \in (1, 2, 3, 4)$	Mean annual sums of polynomial max temperature	

Table 3: Description of saved variables in the 1° bias-corrected and 0.25° bias-corrected and downscaled BCD-ME products.

Data Records

The BCD-ME[46–48] is presented as an ensemble with multiple ESMs and multiple ensemble members from each of the ESMs. Rather than providing the full daily time series over the projection period of 2015-2100, we instead provide 20-year daily time series (for the 1-degree bias-corrected set) and statistics of 20-year daily time series (for the downscaled 0.25-degree set) from each ESM such that the ESM ensemble mean has a specific global mean temperature value or global warming level (GWL). There has been a recent shift to using GWLs instead of time periods in climate projection-related studies, primarily due to the fact that the CMIP6 generation contains models with climate sensitivities outside the range deemed to be plausible or the so-called "hot model" problem[37]). Thus, we provide 20-year series at 0.61 °C (equivalent to the mean observed global temperature from 1982-2001 in HadCRUT5[28]) of warming relative to 1850-1900 (historical) as well as GWLs of 1.0-4.0 °C in 0.5 °C increments. Including the four different reanalyses used in bias-correction and the GWLs, the dataset can be viewed as a 7-dimensional array

$$\texttt{BCD-ME}[\texttt{lon}, \texttt{lat}, \texttt{time}, \texttt{GWL}, \texttt{reanalysis}, \texttt{ESM}, \texttt{experiment/ensemble\_member}] \tag{2}$$

for the 1-degree time series and

$$\texttt{BCD-ME}[\texttt{lon}, \texttt{lat}, \texttt{D}, \texttt{GWL}, \texttt{reanalysis}, \texttt{ESM}, \texttt{experiment/ensemble\_member}] \tag{3}$$

for the 0.25-degree statistics, with the dimension D being unique to each particular variable (e.g., temperature bin).

To facilitate computation with this massive array, all of the data has been archived in the `Zarr` format, with a separate `Zarr` store for each ESM and data type. For example, for the ESM ACCESS-ESM1-5, one `Zarr` store contains all 1° bias-corrected time series in variables `tas` and `tasmax`, and one `Zarr` store contains all 0.25° downscaled statistics from both `tas`-based and `tasmax`-based variables. We recommend working with the data in Python with the `Xarray` and `Dask` libraries as these packages make parallelized I/O and computation relatively simple and efficient to implement[49].

The Zarr stores are chunked to facilitate regional analyses. In other words, a chunk will contain all days (for bias-corrected time series) or all bins or polynomial degrees (for downscaled statistics) for a sub-continental-scale region for one ensemble member, GWL, and reanalysis.

Since not every ensemble member reaches every GWL, and not every ensemble member has data for each variable, each Zarr store contains a metadata coordinate `flag_has_data`, a `variable x ensemble member x GWL` boolean array detailing which elements have data.

Data from different SSPs, but the same run identifier (e.g., "r1i1p1f1"), are assumed to branch off the historical run with that run identifier. If a GWL includes calendar years from before 2015 (the branching point between the "historical" and "SSP" experiments) and multiple SSPs are available with that same run identifier, then the pre-2015 data for those SSPs will be identical, and the effective sample size is lower. Each Zarr store therefore contains a metadata coordinate `flag_has_historical_duplicates`, an `ensemble member x GWL` boolean array that is `True` for a GWL for any ensemble member combination if at least one other ensemble member has exactly the same data for at least part of that GWL.

Finally, since some reanalyses used have native resolutions above 0.25°, Zarr stores for bias-corrected and downscaled 0.25° statistics have a metadata coordinate `flag_resolution`, a `reanalysis x 1` boolean array that is `True` for a reanalysis that is interpolated past its native data resolution before use as a downscaling ground truth, and values at individual locations should be interpreted with caution.

The full BCD-ME data is stored and hosted as `Zarr` stores in two repositories: one that will be frozen upon the publication of this article and one that will be updated. The dataset is archived in the National Science Foundation (NSF) National Center for Atmospheric Research (NCAR)'s Geoscience Data Exchange (GDEX) at `https://doi.org/10.5065/HX83-BX33`. The NSF NCAR GDEX provides open access to data through HTTPS and Globus, promotes FAIR principles including issuing a DOI, and enables users to work quickly with the data on the NCAR computer systems. On GDEX, 1° time series are stored under the store names `bcd_me_qdm_[ESM name].zarr` and 0.25° statistics are stored under the store names `bcd_me_qdm-qplad_[ESM name].zarr`.

The BCD-ME is also freely available via the Earthmover Data Marketplace in an Analysis-Ready, Cloud-Optimized format[49]. This version of the BCD-ME is a "cloud-native data repository"[49] that allows fast and efficient processing despite its size, with access occurring natively within `xarray`. The 1-degree° time series are stored at `https://doi.org/10.82924/9p8f-hd64` by ESM, and the 0.25-degree° statistics are stored at `https://doi.org/10.82924/k92n-qb44` by ESM. This version of the dataset will not be frozen, but any updates to the data will be fully tracked through version control, similar to how code revisions are documented through git repositories. Users accessing this repository will have the latest version of this data and improved I/O due to the optimizations of handling array data provided by Icechunk.

Time series and statistics bias-corrected and downscaled to ERA5, MERRA-2, and GMFD are provided under a Creative Commons Attribution 4.0 International License (CC-BY-4.0). Time series and statistics bias-corrected and downscaled to JRA-3Q are provided under a Creative Commons Attribution Non Commercial Share Alike 4.0 International License (CC-BY-NC-SA-4.0) license.

Technical Validation

GWL Assumptions and Validation

The primary assumption underlying presenting results by GWL is that local climate conditions are dependent only on the amount of warming (represented by GMST) rather than the pathway of warming taken up to that point. Two notable ways in which this could affect results across the ScenarioMIP archive is through the speed of warming–for instance, the Earth is projected to reach 2.0 °C GWL in 2046 under SSP 2-4.5, but an earlier 2038 under the higher emission SSP 5-8.5 (in the BCD-ME, 2036-2055 from the former simulation and 2028-2047 from the latter are presented as equivariant simulations at GWL2)–and the dependence on the exact mix of greenhouse gases, aerosols, and other forcings, which may be different at a certain GWL between SSPs.

To confirm that our use of GWL is justified, we show that GWLs are a very good predictor of regional climate distributions. The distributions in the difference between GWL 2 and the historical GWL 0.61 at the regional level are indistinguishable between SSPs with no notable bias in mean, variance, or other moments for a particular SSP (Figure 2). Even SSP3-7.0, which represents a different aerosol trajectory than the other SSPs[50], shows a similar distribution of temperature at each GWL to the other scenarios. This analysis was repeated with other GWLs, with similar results.

Physicality of Temperature

As a first-pass validation of the final BCD-ME data products, we check for unexpected missing values and confirm that all temperatures are physically possible. For downscaled statistics, we additionally verify that average bin-days per year sum up to 365.

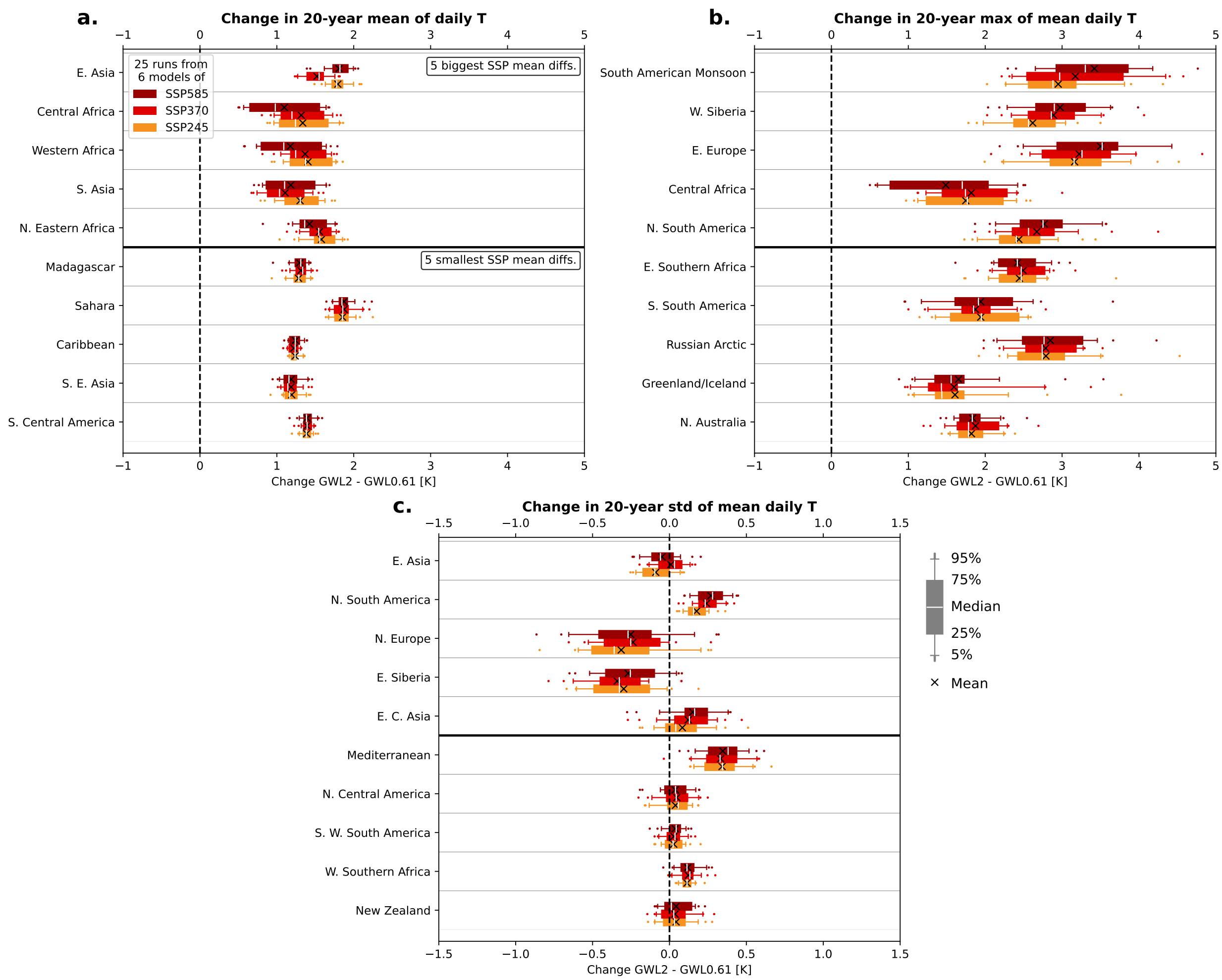

Figure 2: Changes in regionally aggregated statistics for (**a.**) mean daily temperature (**b.**) max of daily mean temperature and (**c.**) standard deviation of daily mean temperature from GWL0.61 to GWL2 by SSP. Each boxplot spans all ESMs and ensemble members. A subset of the BCD-ME is used to ensure the same number of ensemble members are in each sample to allow direct comparisons. All three panels are calculated using the 20 years surrounding each given GWL. The top half of each plot shows the 5 IPCC WG1 regions with the largest differences between SSPs at GWL2, the lower half the regions with the smallest differences.

Validation of QDM bias-correcting method

The QDM methodology uses the change in ESM temperature between GWLs and adds it to a reference bias-corrected model time series. The reference base temperature ESM series is determined using the CDF of a reanalysis dataset sampled at the quantiles of the ESM ensemble member using 31-day windows around each day-of-year to define quantiles. The CDF of this reference ESM time series should, thus, be expected to be nearly identical to the CDF of the reanalysis dataset. Differences between the two CDFs can emerge from two sources. First, we use 100 quantiles to estimate the true 621-element (31 days x 20 years) empirical CDF with quantiles between these 100 samples being linearly interpolated. Strong nonlinearities in the

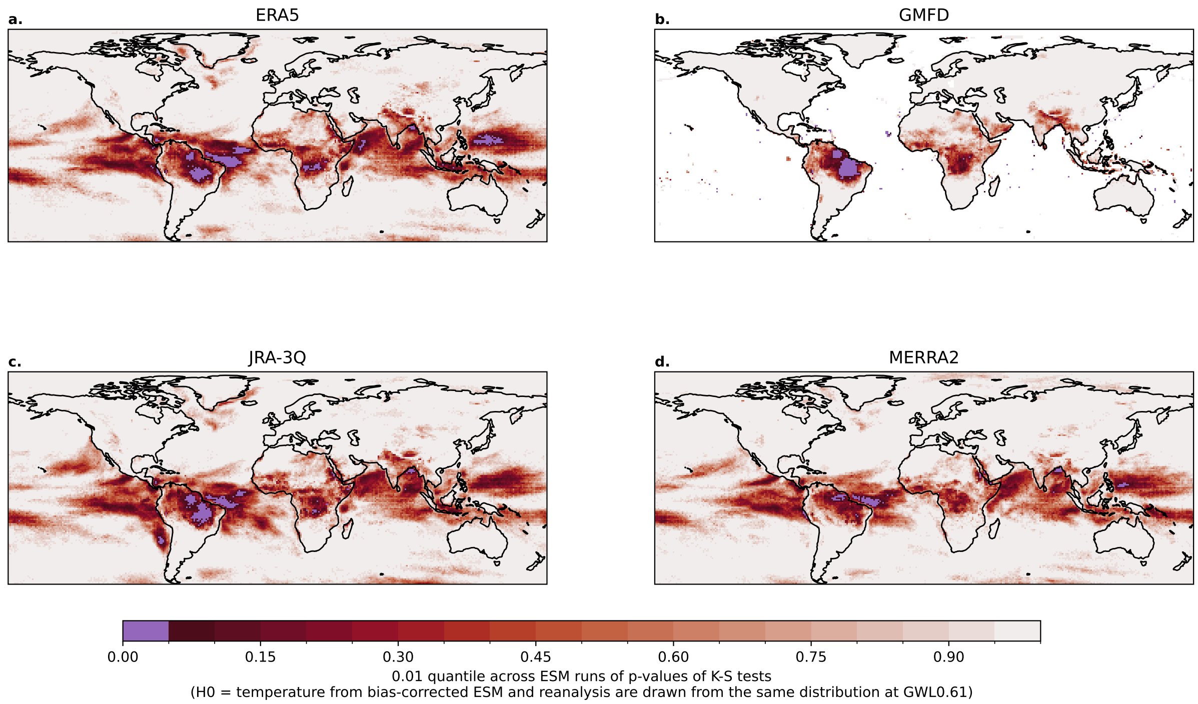

Figure 3: Maps of the 0.01 quantile across BCD-ME ensemble members of p-values of 2-sample Kolmogorov-Smirnov tests at each location against the GWL0.61 daily mean temperature distributions of each reanalysis product and each bias-corrected ensemble member of the BCD-ME being drawn from the same distribution.

CDF of either the reanalysis dataset or the ESM can, thus, cause discrepancies, as can the interpolation of extremes. Second, the 30-day rolling window used to estimate quantiles could artificially narrow the resultant distributions. For example, the rolling window around the hottest day of the year is composed of that day and 30 colder days.

We test the combined influence of these factors on the differences between CDFs of the bias-corrected ESM at GWL0.61 and the corresponding reanalysis between 1982-2001. We conduct 2-sample Kolmogorov-Smirnov tests at each grid cell, using a null hypothesis that both the bias-corrected ESM and the reanalysis CDFs are drawn from the same distribution (Figure 3). Despite the high sensitivity of the Kolmogorov-Smirnov test to small differences in CDF, low p-values are rare in our analysis. Fewer than 0.1% of grid cells in the BCD-ME ensemble member with the lowest average p-value have a p-value of less than 0.01, the benchmark significance value used in Cannon et al. (2015) [18]. These low p-values are spatially limited to locations over the tropics, particularly over the ocean.

In many of the locations with low p-values, the seasonal range of temperature is low and a 31-day rolling window captures a large fraction of the total annual temperature range. Thus, the narrowing effect of the

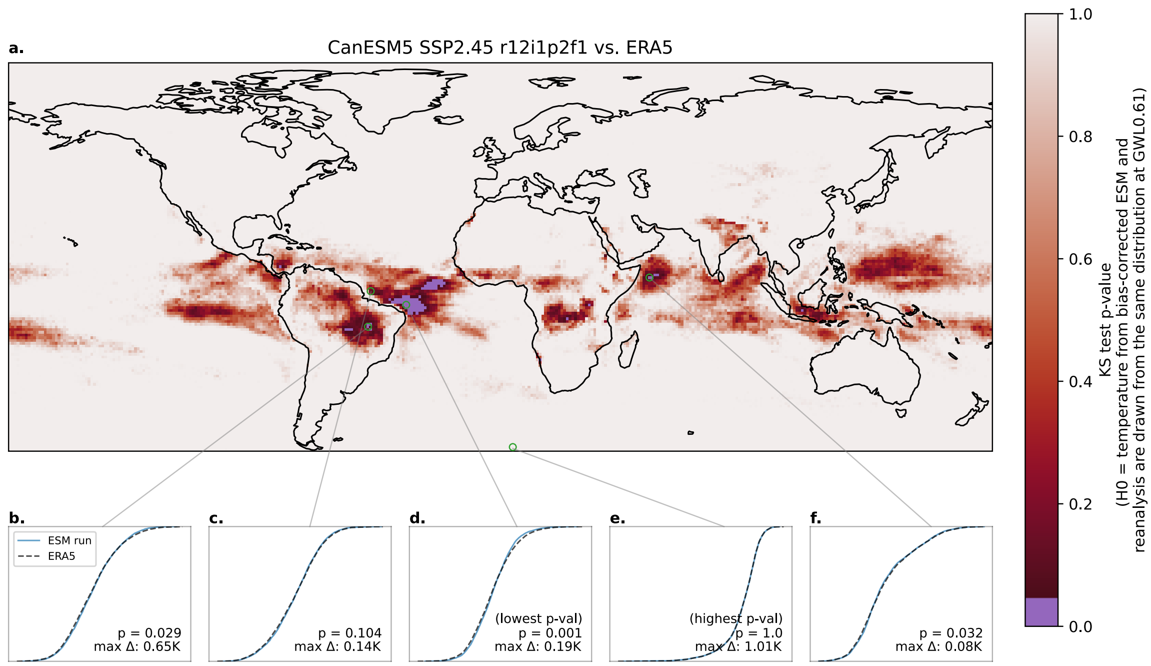

Figure 4: (**a.**) P-values of 2-sample Kolmogorov-Smirnov tests at each location comparing the ERA5 GWL0.61 daily mean temperature CDF with the CDF of a bias-corrected member from the BCD-ME. We show the r12i1p2f1 from CanESM5 run using SSP2.45, representing the BCD-ME ensemble member with the lowest average p-value, that is, the BCD-ME member for which the CDFs of the ERA5 reference distribution and the bias-corrected ESM are least similar on average. (**b.**–**f.**) CDFs of the ESM (blue line) and ERA5 (dashed black line) at various locations, including at the locations with the highest (**e.**) and lowest (**d.**) p-values. For each location, we provide the p-value and the maximum temperature difference between the two CDFs at the same quantile.

window would be expected to be most pronounced. Correspondingly, the most anomalous bias-corrected CDFs are slightly narrower than their reanalysis counterparts (e.g. Figure 4d). However, any differences in the CDF are negligible in practice; in the location with the most differing CDFs in the ensemble member with the lowest global average p-value, the largest difference between CDFs is only 0.19 °C (Figure 4**d.**).

Comparison with NASA NEX-GDDP

To verify projected values against previously published work, we compare the BCD-ME to corresponding members of the NASA NEX-GDDP dataset[12] which also provides bias-corrected and downscaled daily temperature data on a near-global 0.25° grid. The datasets are not exactly comparable – NEX-GDDP provides continuous time series instead of chunks organized by GWL, the bias-correction methodologies are slightly different, and NEX-GDDP uses 1960-2014 as a reference period instead of the 20 years around GWL0.61 used in the BCD-ME. Thus, mean values of temperature are not directly comparable as 1960-2014 spans a different set of GWLs for each ESM ensemble member.

We instead compare *changes* between GWL1.5 and GWL3 for a set of ESMs, experiments, and ensemble members common to both the BCD-ME and NEX-GDDP to mitigate these differences in methods (Figure 5). To best match with NEX-GDDP, we use the the BCD-ME as bias-corrected with GMFD. We determine the GWL1.5 and GWL3 20-year periods in NEX-GDDP following our methods in Section . We expect the changes in temperature between GWL1.5 and GWL3 to be similar in NEX-GDDP and the BCD-ME, though not identical, for the reasons mentioned above.

Differences between mean temperatures at each location are on the order of 0.1 °C with a global mean of about 0 in every model run tested (Figure 5). This close alignment of the two products, despite differences in methodology, provides evidence that the BCD-ME is an appropriate product for extending analyses conducted with NEX-GDDP to account for additional sources of uncertainty.

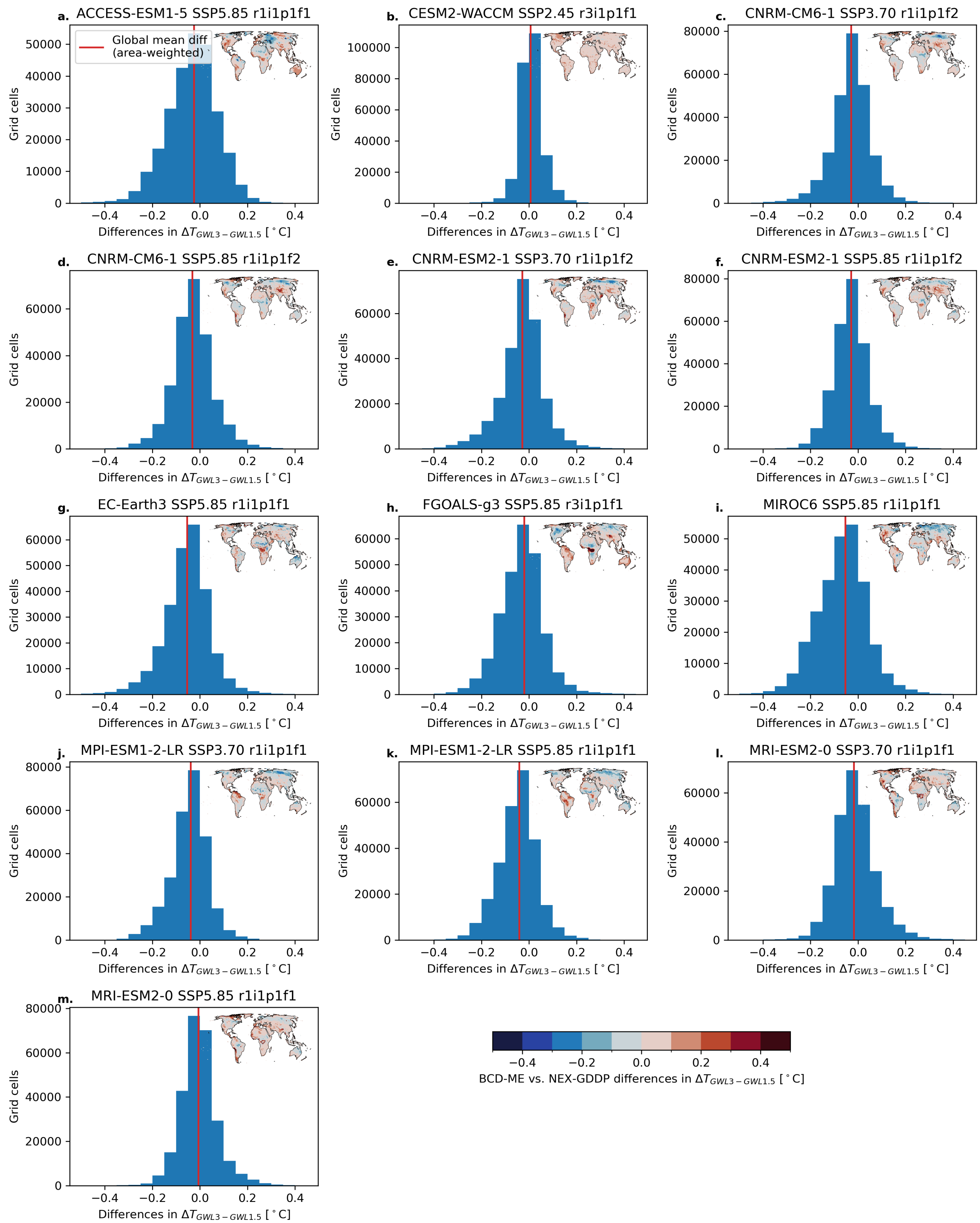

Figure 5: Histograms across grid cells of differences in mean daily temperature between changes from GWL1.5 to GWL3 between corresponding bias-corrected and downscaled ensemble members of the BCD-ME and NASA NEX-GDDP datasets. Inset maps show the spatial distribution of changes.

Uncertainty Partitioning Validation

The BCD-ME is specifically designed to test the sensitivity of climate impacts projections to three quantifiable sources of climate uncertainty: the irreducible internal variability arising from chaos in the climate system[6], the model uncertainty arising from the spread across ESMs outputs driven by the same inputs[7], and the observational uncertainty arising from differences between reanalysis estimates of the historical observed climates used to 'ground truth' projections[51]. When a researcher utilizes the BCD-ME to propagate these climate uncertainties to their projections of temperature and related impacts, understanding if the ultimate uncertainty is primarily due to irreducible uncertainty such as internal variability, or potential reducible uncertainty such as model or observational uncertainty can guide future efforts to better constrain projections.

We thus benchmark the use of the BCD-ME for such partitioning. As such, we calculate the uncertainty partition of various changes in temperature across bias-corrected and downscaled temperature data (Figure 6). We use the Law of Total Variance to partition the total BCD-ME ensemble variance into internal variability, model uncertainty, and reanalysis uncertainty. We note that large-scale patterns of uncertainty partitioning are very similar between the bias-corrected and downscaled projections.

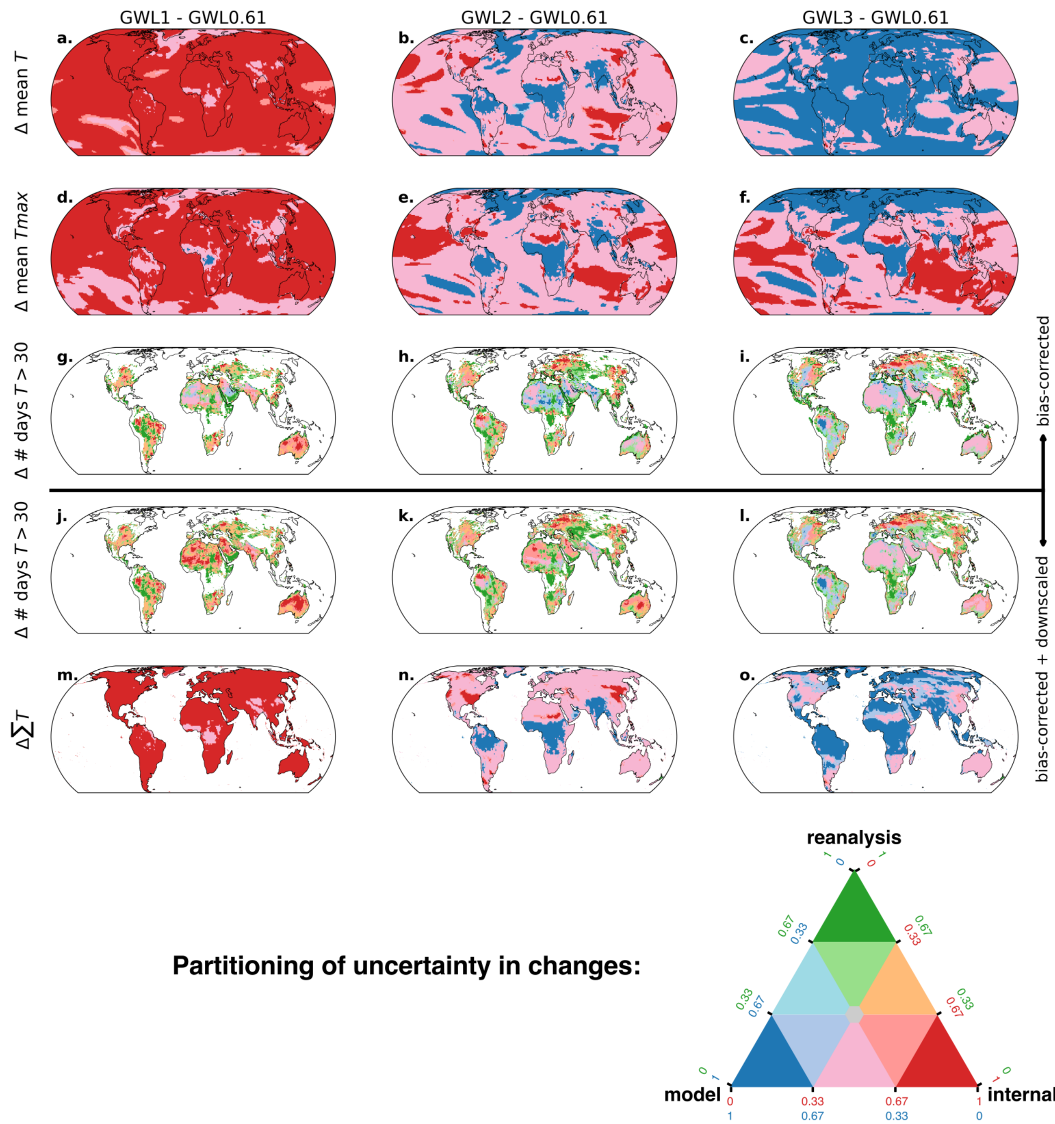

Figure 6: Sample partitioning of uncertainty using the BCD-ME for changes in mean daily temperature (panels **a. - c.**), maximum daily temperature (**d. - f.**), number of days / year with mean temperatures above 30 °C (**g. - l.**), and annual sum of daily temperatures (**m. - o.**), calculated with up to 12 runs from each ESM for each GWL. Partitioned uncertainty is shown for changes between GWL0.61 (equivalent to 1982-2001 in historical observations) and GWLs 1 (left panel), 2 (middle), and 3 (right panel). The first three rows show results for the bias-corrected set of ensembles, the last two rows show results for the bias-corrected and downscaled set of ensembles.

Usage Notes

The BCD-ME data structure is designed with the Python `Xarray` library and is best used in this environment. The currently plug-and-play compatibility of `Xarray` with `Zarr` for efficient and fast I/O and with `Dask` for parallel computing will provide the highest performance to software development ratio for most users. See the provided vignettes in the BCD-ME github repository for introductory examples. For context, the analysis used to create Figure 6 required 2.5 hours of computation (45 core-hours) of computation on NSF NCAR's Casper running on all 18 cores of a 2.3-GHz Intel Xeon Gold 6140 (Skylake) CPU with 10 GB RAM/core. As much of the computation is embarrassingly parallel, analysis can be expected to roughly scale with the number of CPU cores used, provided the system has sufficient I/O.

When using the BCD-ME, we recommend users cite both this study as well as the DOI of the BCD-ME data archive used to access the data. In addition, we recommend users cite the ESM simulation of any ESM used, following the citations provided in Table 4 (the replication code contains a way to automate the generation of these citations), in addition to the reanalyses used.

Model	Experiment
ACCESS-ESM1-5	SSP2-4.5[52]
	SSP5-8.5[53]
CESM2-WACCM	SSP2-4.5[54]
	SSP3-7.0[55]
	SSP5-8.5[56]
CNRM-CM6-1	SSP2-4.5[57]
	SSP3-7.0[58]
	SSP5-8.5[59]
CNRM-ESM2-1	SSP1-1.9[60]
	SSP1-2.6[61]
	SSP2-4.5[62]
	SSP3-7.0[63]
	SSP4-3.4[64]
	SSP4-6.0[65]
	SSP5-3.4-OS[66]

Continued on next page

Model	Experiment
	SSP5-8.5 [67]
CanESM5	SSP2-4.5 [68]
	SSP3-7.0 [69]
	SSP5-8.5 [70]
EC-Earth3	SSP2-4.5 [71]
	SSP3-7.0 [72]
	SSP5-8.5 [73]
EC-Earth3-Veg	SSP2-4.5 [74]
	SSP3-7.0 [75]
	SSP5-8.5 [76]
FGOALS-g3	SSP2-4.5 [77]
	SSP3-7.0 [78]
	SSP5-8.5 [79]
IPSL-CM6A-LR	SSP2-4.5 [80]
	SSP3-7.0 [81]
	SSP5-8.5 [82]
MIROC6	SSP2-4.5 [83]
	SSP3-7.0 [84]
	SSP5-8.5 [85]
MPI-ESM1-2-LR	SSP2-4.5 [86]
	SSP3-7.0 [87]
	SSP5-8.5 [88]
MRI-ESM2-0	SSP2-4.5 [89]
	SSP3-7.0 [90]
	SSP5-8.5 [91]

Table 4: ESM citation information by ESM and scenario, following best practices for citing CMIP simulations.

Data Availability

The BCD-ME[46–48] is available at https://doi.org/10.82924/9p8f-hd64 and https://doi.org/10.82924/k92n-qb44 at the Earthmover Data Marketplace, and at https://doi.org/10.5065/HX83-BX33 on the NSF NCAR Geoscience Data Exchange (GDEX).

Code Availability

All code required to replicate the BCD-ME and the analysis presented in this Data Descriptor is available on Zenodo[92] at https://doi.org/10.5281/zenodo.22311274 and on GitHub at https://github.com/ks905383/bcd_me/. The GitHub repository also contains sample code for accessing and working with the data. The GitHub repository will also remain available as a place to suggest improvements, note issues, and host discussions related to the project.

Author Contributions

Conceptualization: KS
Data Curation: KS
Formal Analysis: KS, NL
Funding Acquisition: KS, NL, RH, GW
Investigation: KS
Methodology: KS
Project Administration: KS
Resources: KS, NL
Software: KS
Supervision: KS, NL, RH, GW
Validation: KS, NL
Visualization: KS
Writing – Original Draft Preparation: NL
Writing – Review & Editing: KS, NL, RH, GW

Competing Interests

The authors have no known competing interests.

Acknowledgments

The authors thank Ryan Abernathy, Tom Bearpark, and Clara Deser for discussion that refined the BCD-ME analysis. We would like to acknowledge computing support from the Casper system (https://ncar.pub/casper) provided by the NSF National Center for Atmospheric Research (NCAR), sponsored by the National Science Foundation. We thank Thomas Cram, Adam Phillips and the NCAR Geoscience Data Exchange (GDEX) team for support with hosting the BCD-ME. We also thank Earthmover, including Ryan Abernathy and Brian Davis, for hosting the living repository of the BCD-ME.

Funding

K.S. and R.H. are supported by the National Oceanic and Atmospheric Administration's Regional Integrated Sciences and Assessments Program, grant NA21OAR4310313. K.S. is also supported by the Columbia Climate School Postdoctoral Fellowship Program. N.L. is partially funded through NCAR which is sponsored by the National Science Foundation under Cooperative Agreement 1852977. This work is also supported by the National Science Foundation as part of the Megalopolitan Coastal Transformation Hub (MACH) under NSF award ICER-2103754. This is MACH contribution number 113.

References

[1] Eyring, V. *et al.* Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development* **9**, 1937–1958 (2016). DOI 10.5194/gmd-9-1937-2016.

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

[3] Hawkins, E. & Sutton, R. The Potential to Narrow Uncertainty in Regional Climate Predictions. *Bulletin of the American Meteorological Society* **90**, 1095–1108 (2009). DOI 10.1175/2009BAMS2607.1.

[4] Auffhammer, M., Hsiang, S. M., Schlenker, W. & Sobel, A. Using Weather Data and Climate Model Output in Economic Analyses of Climate Change. *Review of Environmental Economics and Policy* **7**, 181–198 (2013). DOI 10.1093/reep/ret016.

[5] Rising, J., Tedesco, M., Piontek, F. & Stainforth, D. A. The missing risks of climate change. *Nature* **610**, 643–651 (2022). DOI 10.1038/s41586-022-05243-6.

[6] Schwarzwald, K. & Lenssen, N. The importance of internal climate variability in climate impact projections. *Proceedings of the National Academy of Sciences* **119**, e2208095119 (2022). DOI 10.1073/pnas.2208095119.

[7] Burke, M., Dykema, J., Lobell, D. B., Miguel, E. & Satyanath, S. Incorporating Climate Uncertainty into Estimates of Climate Change Impacts. *The Review of Economics and Statistics* **97**, 461–471 (2015). URL `https://doi.org/10.1162/REST_a_00478`.

[8] Maraun, D. Bias Correcting Climate Change Simulations - a Critical Review. *Current Climate Change Reports* **2**, 211–220 (2016). DOI 10.1007/s40641-016-0050-x.

[9] Schulze, R. Transcending scales of space and time in impact studies of climate and climate change on agrohydrological responses. *Agriculture, Ecosystems & Environment* **82**, 185–212 (2000). DOI 10.1016/S0167-8809(00)00226-7.

[10] Hsiang, S. *et al.* Estimating economic damage from climate change in the United States. *Science* **356**, 1362–1369 (2017). DOI 10.1126/science.aal4369.

[11] Maher, N. *et al.* The updated Multi-Model Large Ensemble Archive and the Climate Variability Diagnostics Package: New tools for the study of climate variability and change. *Geoscientific Model Development* **18**, 6341–6365 (2025). DOI 10.5194/gmd-18-6341-2025.

[12] Thrasher, B. *et al.* NASA Global Daily Downscaled Projections, CMIP6. *Scientific Data* **9**, 262 (2022). DOI 10.1038/s41597-022-01393-4.

[13] Frieler, K. *et al.* Scenario setup and forcing data for impact model evaluation and impact attribution within the third round of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP3a). *Geoscientific Model Development* **17**, 1–51 (2024). DOI 10.5194/gmd-17-1-2024.

[14] Gergel, D. R. *et al.* Global Downscaled Projections for Climate Impacts Research (GDPCIR): Preserving quantile trends for modeling future climate impacts. *Geoscientific Model Development* **17**, 191–227 (2024). DOI 10.5194/gmd-17-191-2024.

[15] Hartke, S. H. *et al.* GARD-LENS: A downscaled large ensemble dataset for understanding future climate and its uncertainties. *Scientific Data* **11**, 1374 (2024). DOI 10.1038/s41597-024-04205-z.

[16] Pierce, D. W., Cayan, D. R., Feldman, D. R. & Risser, M. D. Future Increases in North American Extreme Precipitation in CMIP6 Downscaled with LOCA. *Journal of Hydrometeorology* **24**, 951–975 (2023). DOI 10.1175/JHM-D-22-0194.1.

[17] Semenov, M. A., Senapati, N., Coleman, K. & Collins, A. L. A dataset of large ensemble of CMIP6-based transient climate scenarios for impact assessment in Great Britain. *Data in Brief* **61**, 111695 (2025). URL `https://linkinghub.elsevier.com/retrieve/pii/S2352340925004251`. DOI 10.1016/j.dib.2025.111695.

[18] Cannon, A. J., Sobie, S. R. & Murdock, T. Q. Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? *Journal of Climate* **28**, 6938–6959 (2015). DOI 10.1175/JCLI-D-14-00754.1.

[19] Hersbach, H. *et al.* The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* **146**, 1999–2049 (2020). DOI 10.1002/qj.3803.

[20] Gelaro, R. *et al.* The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate* **30**, 5419–5454 (2017). DOI 10.1175/JCLI-D-16-0758.1.

[21] Kosaka, Y. *et al.* The JRA-3Q Reanalysis. *Journal of the Meteorological Society of Japan. Ser. II* **102**, 49–109 (2024). DOI 10.2151/jmsj.2024-004.

[22] Sheffield, J., Goteti, G. & Wood, E. F. Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling. *Journal of Climate* **19**, 3088–3111 (2006). DOI 10.1175/JCLI3790.1.

[23] Carleton, T. *et al.* Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits*. *The Quarterly Journal of Economics* **137**, 2037–2105 (2022). DOI 10.1093/qje/qjac020.

[24] Hultgren, A. *et al.* Impacts of climate change on global agriculture accounting for adaptation. *Nature* **642**, 644–652 (2025). DOI 10.1038/s41586-025-09085-w.

[25] Rode, A. *et al.* Estimating a social cost of carbon for global energy consumption. *Nature* **598**, 308–314 (2021). DOI 10.1038/s41586-021-03883-8.

[26] Riahi, K. *et al.* The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change* **42**, 153–168 (2017). DOI 10.1016/j.gloenvcha.2016.05.009.

[27] Bosilovich, M. G. *et al.* MERRA-2: Initial Evaluation of the Climate. Tech. Rep. NASA/TM-2015-104606/Vol.43, NASA (2015).

[28] Morice, C. P. *et al.* An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set. *Journal of Geophysical Research: Atmospheres* **126**, e2019JD032361 (2021). DOI 10.1029/2019JD032361.

[29] Lenssen, N. *et al.* A NASA GISTEMPv4 Observational Uncertainty Ensemble. *Journal of Geophysical Research: Atmospheres* **129**, e2023JD040179 (2024). DOI 10.1029/2023JD040179.

[30] European Centre for Medium-Range Weather Forecasts. ERA5 Reanalysis (0.25 Degree Latitude-Longitude Grid) (2019, updated monthly). URL `https://doi.org/10.5065/BH6N-5N20`. DOI 10.5065/BH6N-5N20.

[31] Department of Civil and Environmental Engineering/Princeton University. Global Meteorological Forcing Dataset for Land Surface Modeling (2026). URL `https://doi.org/10.5065/JV89-AH11`. DOI 10.5065/JV89-AH11.

[32] Japan Meteorological Agency. Japanese Reanalysis for Three Quarters of a Century (JRA-3Q) (2023, updated monthly). URL `https://doi.org/10.5065/AVTZ-1H78`. DOI 10.5065/AVTZ-1H78.

[33] NASA Global Modeling and Assimilation Office (GMAO). MERRA2 statD_2d_slv_Nx: 2d, Daily, Aggregated Statistics, Single-Level, Assimilation, Single-Level Diagnostics (2026). URL `https://doi.org/10.5067/9SC1VNTWGWV3`. DOI 10.5067/9SC1VNTWGWV3.

[34] Seneviratne, S. I. & Hauser, M. Regional Climate Sensitivity of Climate Extremes in CMIP6 Versus CMIP5 Multimodel Ensembles. *Earth's Future* **8**, e2019EF001474 (2020). DOI 10.1029/2019EF001474.

[35] Adhikari, P. *et al.* Global Warming Induced Changes in Extreme Precipitation in the Western United States: Projections From Dynamically Downscaled CMIP6 GCMs. *Geophysical Research Letters* **52**, e2025GL116113 (2025). DOI 10.1029/2025GL116113.

[36] Gao, X., Sokolov, A. & Schlosser, C. A. A Large Ensemble Global Dataset for Climate Impact Assessments. *Scientific Data* **10**, 801 (2023). URL `https://www.nature.com/articles/s41597-023-02708-9`. DOI 10.1038/s41597-023-02708-9.

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

[38] Milinski, S., Maher, N. & Olonscheck, D. How large does a large ensemble need to be? *Earth System Dynamics* **11**, 885–901 (2020). DOI 10.5194/esd-11-885-2020.

[39] Batibeniz, F., Hauser, M. & Seneviratne, S. I. Countries most exposed to individual and concurrent extremes and near-permanent extreme conditions at different global warming levels. *Earth System Dynamics* **14**, 485–505 (2023). DOI 10.5194/esd-14-485-2023.

[40] Carleton, T. A. & Hsiang, S. M. Social and economic impacts of climate. *Science* (2016).

[41] Hsiang, S. The Global Economic Impact of Climate Change: An Empirical Perspective. *Annual Reviews* (2026). DOI 10.1146/annurev-economics-102224-021145.

[42] Deschênes, O. & Greenstone, M. Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US. *American Economic Journal: Applied Economics* **3**, 152–185 (2011). DOI 10.1257/app.3.4.152.

[43] Behrer, A. P., Park, R. J., Wagner, G., Golja, C. M. & Keith, D. W. Heat has larger impacts on labor in poorer areas*. *Environmental Research Communications* **3**, 095001 (2021). DOI 10.1088/2515-7620/abffa3.

[44] Schlenker, W. & Roberts, M. J. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proceedings of the National Academy of Sciences* **106**, 15594–15598 (2009). DOI 10.1073/pnas.0906865106.

[45] Casanueva, A. *et al.* Overview of Existing Heat-Health Warning Systems in Europe. *International Journal of Environmental Research and Public Health* **16**, 2657 (2019). DOI 10.3390/ijerph16152657.

[46] Schwarzwald, K., Lenssen, N., Horton, R. & Wagner, G. BCD-ME: Bias-Corrected and Downscaled Massive Ensemble. NSF National Center for Atmospheric Research (2026). URL `https://doi.org/10.5065/HX83-BX33`. DOI 10.5065/HX83-BX33.

[47] Schwarzwald, K., Lenssen, N., Horton, R. & Wagner, G. BCD-ME 1° bias-corrected (QDM) temperature time series. Earthmover Data Marketplace (2026). URL `https://doi.org/10.82924/9p8f-hd64`. DOI 10.82924/9p8f-hd64.

[48] Schwarzwald, K., Lenssen, N., Horton, R. & Wagner, G. BCD-ME 0.25° bias-corrected (QDM) and downscaled temperature statistics. Earthmover Data Marketplace (2026). URL `https://doi.org/10.82924/k92n-qb44`. DOI 10.82924/k92n-qb44.

[49] Abernathey, R. P. *et al.* Cloud-Native Repositories for Big Scientific Data. *Computing in Science & Engineering* **23**, 26–35 (2021). DOI 10.1109/MCSE.2021.3059437.

[50] Shiogama, H. *et al.* Important distinctiveness of SSP3–7.0 for use in impact assessments. *Nature Climate Change* **13**, 1276–1278 (2023). DOI 10.1038/s41558-023-01883-2.

[51] Casanueva, A. *et al.* Testing bias adjustment methods for regional climate change applications under observational uncertainty and resolution mismatch. *Atmospheric Science Letters* **21**, e978 (2020). DOI 10.1002/asl.978.

[52] Ziehn, T. *et al.* CSIRO ACCESS-ESM1.5 model output prepared for CMIP6 ScenarioMIP ssp245. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.4322`. DOI 10.22033/ESGF/CMIP6.4322.

[53] Ziehn, T. *et al.* CSIRO ACCESS-ESM1.5 model output prepared for CMIP6 ScenarioMIP ssp585. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.4333`. DOI 10.22033/ESGF/CMIP6.4333.

[54] Danabasoglu, G. NCAR CESM2-WACCM model output prepared for CMIP6 ScenarioMIP ssp245. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.10101`. DOI 10.22033/ESGF/CMIP6.10101.

[55] Danabasoglu, G. NCAR CESM2-WACCM model output prepared for CMIP6 ScenarioMIP ssp370. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.10102`. DOI 10.22033/ESGF/CMIP6.10102.

[56] Danabasoglu, G. NCAR CESM2-WACCM model output prepared for CMIP6 ScenarioMIP ssp585. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.10115`. DOI 10.22033/ESGF/CMIP6.10115.

[57] Voldoire, A. CNRM-CERFACS CNRM-CM6-1 model output prepared for CMIP6 ScenarioMIP ssp245. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.4189`. DOI 10.22033/ESGF/CMIP6.4189.

[58] Voldoire, A. CNRM-CERFACS CNRM-CM6-1 model output prepared for CMIP6 ScenarioMIP ssp370. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.4197`. DOI 10.22033/ESGF/CMIP6.4197.

[59] Voldoire, A. CNRM-CERFACS CNRM-CM6-1 model output prepared for CMIP6 ScenarioMIP ssp585. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.4224. DOI 10.22033/ESGF/CMIP6.4224.

[60] Voldoire, A. CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 ScenarioMIP ssp119. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.4182. DOI 10.22033/ESGF/CMIP6.4182.

[61] Voldoire, A. CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 ScenarioMIP ssp126. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.4186. DOI 10.22033/ESGF/CMIP6.4186.

[62] Voldoire, A. CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 ScenarioMIP ssp245. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.4191. DOI 10.22033/ESGF/CMIP6.4191.

[63] Voldoire, A. CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 ScenarioMIP ssp370. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.4199. DOI 10.22033/ESGF/CMIP6.4199.

[64] Voldoire, A. CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 ScenarioMIP ssp434. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.4213. DOI 10.22033/ESGF/CMIP6.4213.

[65] Voldoire, A. CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 ScenarioMIP ssp460. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.4217. DOI 10.22033/ESGF/CMIP6.4217.

[66] Voldoire, A. CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 ScenarioMIP ssp534-over. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.4221. DOI 10.22033/ESGF/CMIP6.4221.

[67] Voldoire, A. CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 ScenarioMIP ssp585. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.4226. DOI 10.22033/ESGF/CMIP6.4226.

[68] Swart, N. C. *et al.* CCCma CanESM5 model output prepared for CMIP6 ScenarioMIP ssp245. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.3685. DOI 10.22033/ESGF/CMIP6.3685.

[69] Swart, N. C. *et al.* CCCma CanESM5 model output prepared for CMIP6 ScenarioMIP ssp370. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.3690. DOI 10.22033/ESGF/CMIP6.3690.

[70] Swart, N. C. *et al.* CCCma CanESM5 model output prepared for CMIP6 ScenarioMIP ssp585. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.3696. DOI 10.22033/ESGF/CMIP6.3696.

[71] EC-Earth Consortium (EC-Earth). EC-Earth-Consortium EC-Earth3 model output prepared for CMIP6 ScenarioMIP ssp245. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.4880. DOI 10.22033/ESGF/CMIP6.4880.

[72] EC-Earth Consortium (EC-Earth). EC-Earth-Consortium EC-Earth3 model output prepared for CMIP6 ScenarioMIP ssp370. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.4884. DOI 10.22033/ESGF/CMIP6.4884.

[73] EC-Earth Consortium (EC-Earth). EC-Earth-Consortium EC-Earth3 model output prepared for CMIP6 ScenarioMIP ssp585. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.4912. DOI 10.22033/ESGF/CMIP6.4912.

[74] EC-Earth Consortium (EC-Earth). EC-Earth-Consortium EC-Earth3-Veg model output prepared for CMIP6 ScenarioMIP ssp245. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.4882. DOI 10.22033/ESGF/CMIP6.4882.

[75] EC-Earth Consortium (EC-Earth). EC-Earth-Consortium EC-Earth3-Veg model output prepared for CMIP6 ScenarioMIP ssp370. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.4886. DOI 10.22033/ESGF/CMIP6.4886.

[76] EC-Earth Consortium (EC-Earth). EC-Earth-Consortium EC-Earth3-Veg model output prepared for CMIP6 ScenarioMIP ssp585. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.4914. DOI 10.22033/ESGF/CMIP6.4914.

[77] Li, L. CAS FGOALS-g3 model output prepared for CMIP6 ScenarioMIP ssp245. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.3469. DOI 10.22033/ESGF/CMIP6.3469.

[78] Li, L. CAS FGOALS-g3 model output prepared for CMIP6 ScenarioMIP ssp370. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.3480. DOI 10.22033/ESGF/CMIP6.3480.

[79] Li, L. CAS FGOALS-g3 model output prepared for CMIP6 ScenarioMIP ssp585. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.3503`. DOI 10.22033/ESGF/CMIP6.3503.

[80] Boucher, O. *et al.* IPSL IPSL-CM6A-LR model output prepared for CMIP6 ScenarioMIP ssp245. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.5264`. DOI 10.22033/ESGF/CMIP6.5264.

[81] Boucher, O. *et al.* IPSL IPSL-CM6A-LR model output prepared for CMIP6 ScenarioMIP ssp370. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.5265`. DOI 10.22033/ESGF/CMIP6.5265.

[82] Boucher, O. *et al.* IPSL IPSL-CM6A-LR model output prepared for CMIP6 ScenarioMIP ssp585. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.5271`. DOI 10.22033/ESGF/CMIP6.5271.

[83] Shiogama, H., Abe, M. & Tatebe, H. MIROC MIROC6 model output prepared for CMIP6 ScenarioMIP ssp245. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.5746`. DOI 10.22033/ESGF/CMIP6.5746.

[84] Shiogama, H., Abe, M. & Tatebe, H. MIROC MIROC6 model output prepared for CMIP6 ScenarioMIP ssp370. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.5752`. DOI 10.22033/ESGF/CMIP6.5752.

[85] Shiogama, H., Abe, M. & Tatebe, H. MIROC MIROC6 model output prepared for CMIP6 ScenarioMIP ssp585. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.5771`. DOI 10.22033/ESGF/CMIP6.5771.

[86] Wieners, K.-H. *et al.* MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 ScenarioMIP ssp245. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.6693`. DOI 10.22033/ESGF/CMIP6.6693.

[87] Wieners, K.-H. *et al.* MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 ScenarioMIP ssp370. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.6695`. DOI 10.22033/ESGF/CMIP6.6695.

[88] Wieners, K.-H. *et al.* MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 ScenarioMIP ssp585. Earth System Grid Federation (2019). URL `https://doi.org/10.22033/ESGF/CMIP6.6705`. DOI 10.22033/ESGF/CMIP6.6705.

[89] Yukimoto, S. *et al.* MRI MRI-ESM2.0 model output prepared for CMIP6 ScenarioMIP ssp245. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.6910. DOI 10.22033/ESGF/CMIP6.6910.

[90] Yukimoto, S. *et al.* MRI MRI-ESM2.0 model output prepared for CMIP6 ScenarioMIP ssp370. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.6915. DOI 10.22033/ESGF/CMIP6.6915.

[91] Yukimoto, S. *et al.* MRI MRI-ESM2.0 model output prepared for CMIP6 ScenarioMIP ssp585. Earth System Grid Federation (2019). URL https://doi.org/10.22033/ESGF/CMIP6.6929. DOI 10.22033/ESGF/CMIP6.6929.

[92] Schwarzwald, K., Lenssen, N., Horton, R. & Wagner, G. Replication code for "A bias-corrected & downscaled massive ensemble to diagnose uncertainty in climate impact projections". Zenodo (2026). URL https://doi.org/10.5281/zenodo.22311274. DOI 10.5281/zenodo.22311274.