The choice of historical data product dominates uncertainty in projections of climate impacts in a 2-degree world

Kevin Schwarzwald[1,2*], Nathan Lenssen[3,4], Radley Horton[1,2], Alia Bonanno[5], Gernot Wagner[5]

[1*]Climate School, Columbia University, Hogan Hall, 2910 Broadway, New York, 10025, NY, USA.
[2]Lamont-Doherty Earth Observatory, Columbia University, 61 Route 9W, Palisades, 10964, NY, USA.
[3] Colorado School of Mines, 1500 Illinois St., Golden, 80401, CO, USA.
[4]National Center for Atmospheric Research, orgaddressPO Box 3000, Boulder, 80307, CO, USA.
[5]Columbia Business School, 645 W 130th St, New York, 10027, NY, USA.

*Corresponding author(s). E-mail(s): kevin.schwarzwald@columbia.edu;

Abstract

Estimates of the risk of climate change on society rely on historical estimates of true weather conditions and future projections from global climate models (GCMs), which are typically bias-corrected and downscaled before use. Future projections of climate impacts are affected by uncertainty in the underlying climate data through multiple pathways, only some of which are regularly accounted for in the literature. We investigate the importance of the choice of gridded historical data product used to fit impact models and bias-correct and downscale GCMs, a decision that is often either *ad hoc* in climate impact studies or made for reasons orthogonal to a given product's performance for metrics and regions of interest, on the spread in projections of climate impacts. We re-estimate three climate impact models from the literature, relating exposure to daily mean or max temperature to annual GDP per capita growth, mortality, and payroll, using four different reanalysis products. We then project damages for each dose-response function using a novel ensemble of GCM projections that accounts for all sources of climate uncertainty, bias-corrected and downscaled to the same four reanalyses to estimate this "observational" uncertainty, and incorporating all runs from

multiple "Large Ensembles" of GCMs to estimate model uncertainty and internal variability. We find that the choice of gridded historical data product dominates the spread in future projections of GDP per capita growth, mortality, and payroll at a given Global Warming Level for most parts of the globe, particularly in the mid-latitudes. We thus recommend that users of climate data test the sensitivity of their results to the choice of historical data product and use products that have been evaluated for the metrics and regions of interest whenever possible, and call for more research into constraining uncertainties about past estimates of the climate.

Keywords: climate impacts, climate uncertainty, reanalysis, GCMs

1 Introduction

Projections of the socioeconomic impacts of climate change are increasingly used to understand and prepare for climate risks, however the *climate* uncertainties in these studies are often underestimated. Statistical projections of the impacts of climate change on society typically fit an econometric model relating a historical weather variable and a socioeconomic outcome, creating a 'dose-response' function, and then project damages into the future using expected changes in that weather variable under increased greenhouse gases [1]. Uncertainties in our knowledge of the past or future climate can thus enter these projections in two ways: through the future climate projections used to project damages, which typically come from ensembles of Global Climate Models (GCMs), and through the historical weather data used to fit the statistical model and post-process GCMs to make them usable in econometric contexts by adjusting their output to better fit the observed climate and provide data at higher resolutions (a process known as bias-correction and downscaling).

The uncertainties in the former are well-documented (if not always well-constrained) in the climate science literature, and increasingly accounted for in the climate impacts literature. They come from three primary sources (cf. Hawkins and Sutton [2]): (1) the uncertainty in future emissions and concentrations of 'forcings', i.e., greenhouse gases, aerosols, land use, and other human impacts on the earth systems. This uncertainty is generally termed 'scenario uncertainty' and accounted for by testing results against output from GCMs run on different forcing 'scenarios', such as the Shared Socioeconomic Pathways [3] used in the latest generation of climate models. (2) the uncertainty in the response of the earth system to these forcings, generally estimated as the spread of results across different GCMs and thus termed 'model uncertainty' and accounted for by testing results against an ensemble of GCMs. Though historically rarely considered in projections of climate impacts [4], model uncertainty is now commonly incorporated into climate impacts studies, also thanks to the availability of ready-to-use sets of bias-corrected and downscaled GCM ensembles such as NASA's NEX GDDP product [5]; (3) the uncertainty arising from the chaotic nature of the climate system, which supports a distribution of climate states even without

changing external or anthropogenic forcings. This uncertainty is termed 'internal variability' and is accounted for by testing results against many different runs of GCMs with the same inputs, but is rarely considered in projections of climate impacts [6], partially due to the lack of available ready-to-use datasets that sample it.

However, the impact of uncertainty in our understanding of past weather and climate patterns has yet to be systematically studied in the context of climate impacts projections, despite known issues with the historical observational network: ground-based weather observations networks are denser in wealthier countries and everywhere tend to be located near populated areas, along major roads, and away from regions of complex topography [7–10]. Even existing weather stations may have only restricted access to data for economic or legal reasons [1] or may suffer from temporal gaps in their records due to quality control issues, funding lapses, or conflict [10, 11].

To overcome these limitations, researchers use gridded historical observational products, which combine different observational sources with a quality control model and a statistical or dynamical (i.e., physics-based) 'assimilation' procedure to produce a spatiotemporally complete set of estimates of true historical conditions on a uniform grid across the globe. These products can differ drastically in their complexity, ranging from simple statistical interpolations of weather station observations to complex "reanalyses" that may use data from satellites, airplanes, and weather radars in addition to weather stations to constrain a physics-based model of the global climate. Each product may differ in which data it assimilates, how it deals with missing data, and the assimilation model itself – i.e., the interpolation function used in statistical products or the sub-grid parameterizations used to resolve processes below the model's resolution in reanalyses. Thus, different gridded historical observational products will produce different estimates of historical conditions at the same locations, and thus will result in different estimates of the relationship between climate and society if used to fit dose-response functions.

Importantly, however, the choice of which gridded historical data product to use in projections of climate impacts tends to either be *ad hoc* or is based on criteria orthogonal to a product's suitability as a historical "ground truth". For example, many studies (e.g., Carleton et al. [12], Rode et al. [13], Hultgren et al. [14]) use GMFD, an older generation reanalysis product (based on the NCEP / NOAA reanalysis, which uses an "antiquated" – cf. National Center for Atmospheric Research Staff [15] – data assimilation model) with few evaluations and no data past 2010, primarily because it is the dataset used to bias-correct and downscale NASA's NEX-GDDP product [5], a commonly-used set of future GCM projections. Many studies do not justify their choice of historical gridded data product at all, or do so using absolute, not comparative metrics, showing only that the chosen product appears reasonable for their purpose, without evaluating whether a different one would be better suited.

Studies have indeed shown one reanalysis to outperform others in comparison to other observational products (e.g., ERA5 in surface temperature in the North China Plain [16], West Africa [17], and in correlations with temperature and convective parameters in Europe and North America [18, 19], MERRA-2 in several indices of warm extremes over Europe [18] or wind potential in Morocco [20]; JRA-55 in detecting extremely cold years globally [21] and dry spells in North America [18]) to give a

few examples of a rich literature). However, these studies do not cover all areas of the global of interest to climate impacts researchers, and face several fundamental challenges common to all evaluations between gridded products and observations: gridded products are designed to produce an estimate of average conditions over a grid cell rather than exact conditions at a reference location, many gridded products assimilate in part or completely the publicly available station databases that could conceivably be used to provide independent verifications, and the aforementioned scarcity of observations outside of the wealthiest countries make verification challenging in the regions likely to experience the worst impacts of climate change (e.g., Carleton et al. [12]). Thus, even when evaluations are available, choosing a clear 'best' product to use in climate impacts research is difficult.

In short, current practices in statistical projections of climate impacts are likely to underestimate the uncertainty in their future estimates due to relying on single estimates of historical climate conditions for fitting relationships between climate and society and for 'ground-truthing' future GCM projections through bias-correction and downscaling.

To test the importance of this uncertainty in choice of historical data, we conduct a sensitivity analysis of climate impacts projections from the literature to multiple estimates of historical conditions. We reestimate three dose-response functions from the statistical econometric climate impacts literature relating exposure to daily temperature to annual all-cause mortality [12], national GDP per capita [22], and payroll [23], using four different reanalysis products to estimate the impact of the choice of historical 'ground truth' on estimates of the historical relationship between climate and society. We then create a novel "Bias-Corrected and Downscaled Massive Ensemble" (BCD-ME; Schwarzwald et al., *in prep*) of daily mean and maximum temperature time series on a 0.25-degree global grid that accounts for scenario uncertainty by normalizing changes by Global Warming Level (GWL), accounts for model and internal uncertainty by using 300 runs from 12 different GCMs, and accounts for observational uncertainty by bias-correcting and downscaling GCM output using the same four reanalysis products. We use the BCD-ME to project future changes in mortality, GDP per capita, and payroll due to climate change. Finally, we partition uncertainty in these estimates into contributions from model, internal, and historical sources.

We show that over much of the globe at a given GWL, the future spread in estimated climate impacts is dominated by the choice of historical gridded data product, primarily through its influence on the fitting of the original dose-response function, with a smaller role played by differences in 'ground truths' used to bias-correct and downscale future GCM output. Though the estimates of the sign of mean change in climate impacts is largely unaffected by the addition of estimates based on multiple historical gridded datasets, the total uncertainty in future damages is increased substantially. This increased spread implies higher values of the Social Cost of Carbon [24, 25].

2 Results

2.1 Fitting dose-response functions with multiple observational products

We first refit historical relationships between mean daily temperature exposure and GDP per capita growth [22], mean daily temperature exposure and annual all-cause mortality [12], and maximum daily temperature exposure and payroll [23] using four reanalysis products as historical 'ground truth': ERA5, MERRA2, JRA-3Q, and either GMFD (GDP, mortality) or PRISM (payroll) (Figure 1). For Burke et al. [22] and Behrer et al. [23], we replicate their primary regression results; for Carleton et al. [12] we use their 'pooled' specification (see Section 3.1 for details). For GDP per capita growth and payroll, we additionally fit dose-response functions using UDel and PRISM, respectively, to match the original papers' data. Note that for payroll, we do not use GMFD, since that product's maximum temperature data ends in 2008, which would necessitate a substantial decrease in sample size.

For the GDP per capita growth model of Burke et al. [22], two clusters of results emerge, with ERA5 and JRA-3Q producing nearly identical fits to the UDel fit used in the original paper, while GMFD and MERRA2 have significantly different fits, particularly for the effect of hotter temperatures (Figure 1**A**). For the latter products, the 'ideal' temperature that maximizes GDP per capita growth, i.e., the inflection point in the quadratic model, is nearly 5°C warmer than for ERA5, JRA-3Q, and UDel, and extreme heat is associated with lower damages.

GMFD produces a fit more in line with ERA5 and JRA-3Q for changes in all-cause mortality, though for extremely hot or cold days, the same exposure is estimated to result in fewer deaths for GMFD than for the latter products (Figure 1**B**). MERRA2 is a large outlier, once again estimating an 'ideal' temperature (that minimizes mortality) substantially higher than the other data products.

Finally, for payroll in the continental United States, fit on daily maximum instead of mean temperature, ERA5 and MERRA2 produce similar results to the PRISM data used in the original paper, while JRA-3Q suggests much milder (and statistically insignificant) damages for exposure to days above 90°F (Figure 1**C**).

In short, the choice of observational dataset has a substantial impact on the climate impact model fit. Though the relative ordering of fits by gridded historical data product differs between impact models, some patterns emerge. For dose-response functions based on mean daily temperature, MERRA2 consistently predicts a substantially higher inflection point for even-polynomial models than the other products, while ERA5 and JRA-3Q produce similar results. Whether GMFD is an outlier depends on the impact studied. This stylized behavior does not hold for the payroll model, which uses maximum daily temperature instead of mean daily temperature, and for which JRA-3Q is an outlier. Thus, the relationship between choice of data product and the resultant climate impact model is dependent on both the functional form of the impact model and the climate metric it is based on.

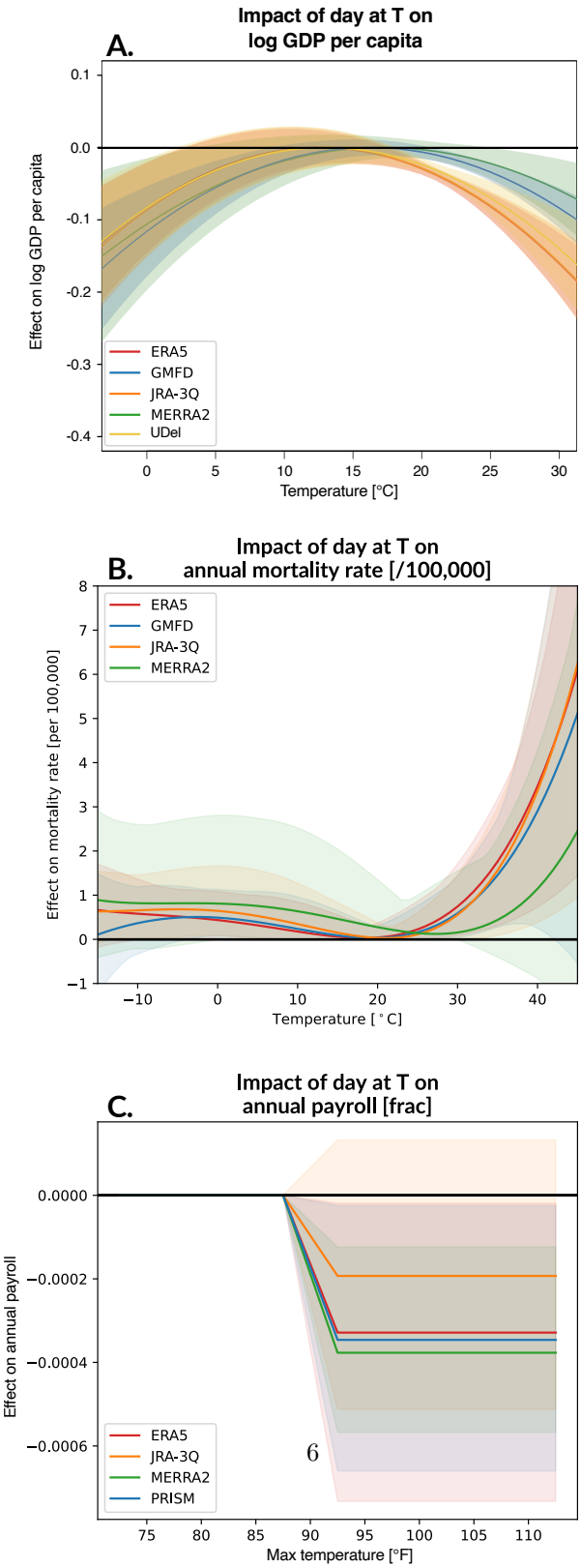


Fig. 1 Dose-response functions fit using multiple gridded historical datasets. Shading shows the maximum and minimum spread of Monte Carlo draws, thick lines show the mean fit. UDel (in panel **A**) and PRISM (**C**) are included for reference; these were the original products used by each respective paper.

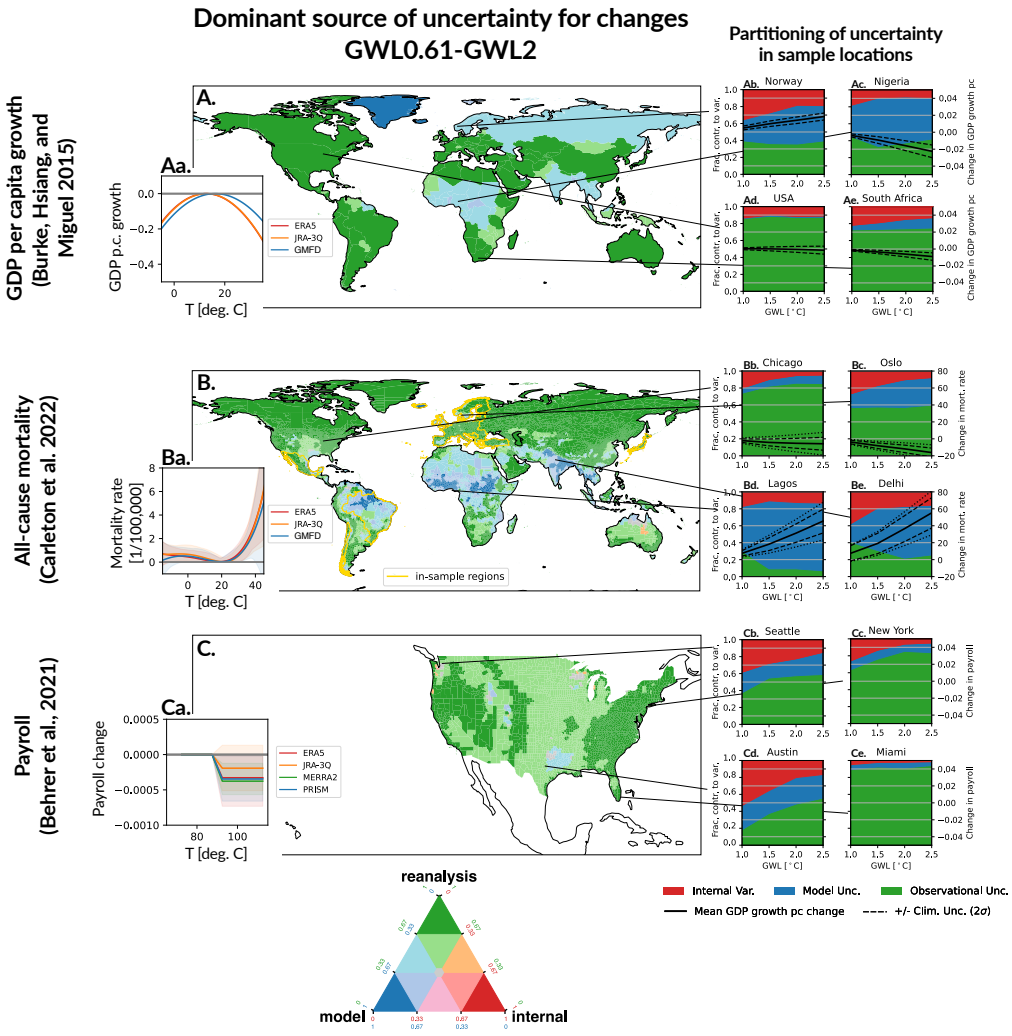


Fig. 2 Partitioning of uncertainty for projections of the impact of climate change on GDP per capita growth (subplots **A**-**Ae**), mortality (**B**-**Be**), and payroll (**C**-**Ce**). Uncertainty is calculated as the variance across changes between GWL 0.61 (1982-2001) and GWL 2 for each metric. Observational uncertainty in panels **A** and **B** is calculated without data from MERRA2. Maps (**A**, **B**, **C**) are colored by the dominant source of climate uncertainty, with solid colors representing areas where reanalysis (green), model (blue), or internal (red) uncertainty makes up more than 2/3 of the spread and lighter colors representing areas with a mix of uncertainty sources (see triangular color legend at bottom). The panels on the right-hand side show the evolution of the partitioning of uncertainty over different GWLs for select representative locations in addition to the evolution over time of mean damages across the ensemble (solid black line) and 2σ of spread across sources of climate uncertainty (dashed lines). Panels **Aa**, **Ba**, and **Ca** show the damage functions from Figure 1 for reference, excluding MERRA2 in panels **Aa** and **Ba** as it is not used in the map of partitioned uncertainty. In Panel **B**, yellow borders show regions with mortality data used to fit the dose-response function.

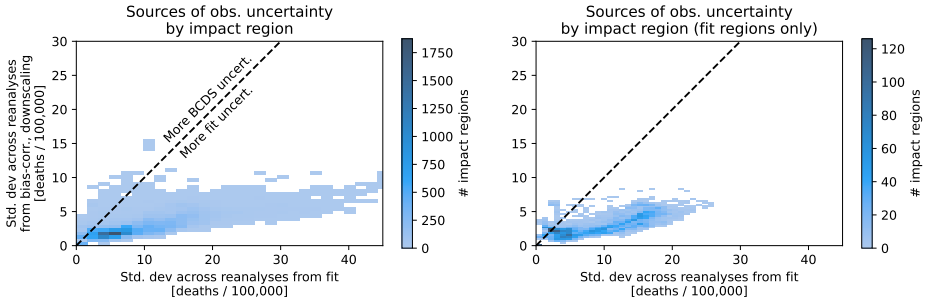


Fig. 3 Histograms of regions showing whether observational uncertainty in projections of changes in mortality rate (GWL0.61 to GWL2, as in Figure 2**B**) is primarily driven by the choice of historical product used to bias-correct and downscale future GCM output (y-axis) or used to fit the dose-response function (x-axis). The left panel includes all regions in the globe, the right panel only those that were in-sample for the original model fit (yellow-bordered regions in Figure 2).

2.2 Partitioning uncertainty in projected damages

We now compare the magnitude of the contribution to the spread of projected damages from historical product choice to the contribution from two other sources of climate uncertainty: model uncertainty and internal variability. Figure 2 shows the dominant source of uncertainty for each location and each climate impact model for changes in damages between GWL 0.61 (1982-2001 in the real world) and GWL 2. In areas in solid green in panels **A**, **B**, and **C**, the choice of observational product makes up more than 2/3 of the total variance in damages. Overall, this choice is the most important source of climate uncertainty, particularly in the mid-latitudes, with model uncertainty still making up the majority of the spread in large parts of the tropics, and internal variability playing a smaller role. Given MERRA2's outlier fits (see Discussion), we have chosen to not include it in the partitioned uncertainty calculation for Figure 2**A** and **B**; including it results in observational uncertainty dominating the spread of damages in nearly every region of the globe.

The partitioning of uncertainty may however vary by region and with lead times. The smaller panels on the right side of the figure show the evolution of the climate uncertainty mix for several representative locations for each impact model. As has been found elsewhere, internal variability tends to contribute a larger percentage of total variance at shorter lead times, with model uncertainty increasing its share of total uncertainty at higher GWLs. For each impact model, however, one can find regions with different combinations of internal, model, and observational uncertainty at most studied GWLs, showing the difficulty of ascertaining *a priori* which source of uncertainty dominates.

Finally, we want to understand if uncertainty from historical data products is primarily driven through the fitting dose-response functions or through bias-correcting and downscaling future GCM output. Figure 3 shows, for each region's projected mortality rate change, whether there is a greater spread across observational products in the dose-response function fit or across post-processing procedures (see Section 3.3.1 for details). For mortality projections in most parts of the world, most of the observational uncertainty comes from fitting the dose-response functions.

3 Methods

3.1 Re-fitting dose-response functions using multiple gridded historical data products

We use three dose-response functions from the statistical econometric climate impacts literature: Carleton et al. [12], relating exposure to daily mean temperature at the ADM2 level to annual all-cause mortality over much of the globe using a quartic polynomial, Burke et al. [22], relating exposure to daily mean temperature at the country level to annual growth in GDP per capita using a quadratic polynomial, and Behrer et al. [23], relating exposure to daily maximum temperature and annual payroll at the U.S. county level using a threshold-defined model. These three dose-response functions sample different levels of spatial aggregation, different metrics of temperature, and different functional forms for the dose-response function to provide a representative sampling of the climate econometric literature.

For each study, we use publicly-available replication code and data, merely replacing the historical weather data used with new estimates of the same metrics from four reanalysis products: the European Centre for Medium-Range Weather Forecasting's ERA5 [26], NASA's MERRA2 [27], the novel Japan Meteorological Agency's JRA-3Q [28], and GMFD (e.g., Thrasher et al. [5]). 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 set of bias-corrected and downscaled climate projections. ERA5 (e.g., [29–31]), MERRA2 (e.g., [32, 33]), and GMFD (e.g., [12–14, 34]) have been used extensively to fit climate impact models, as has JRA-3Q's predecessor, JRA-55 (e.g., [35–37]). Though relative evaluations of each reanalysis product exist for some variables in some regions, their relative performance away from weather stations and outside of data-heavy parts of the world is difficult to estimate. Impacts studies rarely use more than one product or justify the use of one over another, confirming the in practice ad hoc nature of the choice of observational data product.

As reference, we also use the datasets by Willmott and Matsuura [38] (commonly referred to as "UDel") and PRISM [39] to compare results to those by Burke et al. [22] and Behrer et al. [23] in their original papers, respectively.

We generally aim to use the primary or 'preferred' dose-response model specification for each study. For Carleton et al. [12], however, using public replication data leads to discrepancies with the original paper, since part of their replication data (mortality data from the USA and China) were not shared publicly due to restrictive access agreements with the data providers. The resultant sample size is strongly reduced, leading to highly unstable model behavior in some parts of the world for their

'preferred' specification, which contains interactions with income and long-run temperature (whose results are shown in e.g., their Figure IV). Thus, for our main text figures showing uncertainty partitioning for mortality, we use their simplified 'pooled' model, which fits one dose-response function for the entire sample. Though this limits the external validity of the model, it also reduces the impact of observational uncertainty by averaging over many regional discrepancies; we thus believe that our primary conclusions are unaffected by this choice.

3.2 Building a Bias-Corrected and Downscaled Massive Ensemble (BCD-ME)

To sample all primary sources of climate uncertainty in projections of impacts, we require a set of future bias-corrected, downscaled projections from multiple runs of multiple GCMs, using multiple gridded historical data products as 'ground truths'. Existing publicly-available products aimed at impacts research, such as NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP; Thrasher et al. [5]) or the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP; Frieler et al. [40]) provide bias-corrected and downscaled climate projections, but do not sample internal variability (only one ensemble member per GCM is provided) or the uncertainty in datasets used to post-process GCM output (only one reanalysis is used in each product).

We thus create a new dataset of downscaled, bias-corrected GCM simulations of daily mean and maximum temperature, the Bias-Corrected and Downscaled Massive Ensemble (BCD-ME). This ensemble is based on "Large Ensembles" of GCMs from the 6th Phase of the Coupled Model Intercomparison Project (CMIP6; the current state-of-the-art set of GCMs), which sample internal variability by providing multiple model runs for each greenhouse gas forcing scenario, with initial conditions slightly varied between ensemble members [41]. Due to the chaotic nature of the climate system, these runs quickly diverge to span a broad range of possible climate states for a given set of forcings.

We take every simulation of daily mean and maximum temperature in the CMIP6 ScenarioMIP set of experiments, available on the Pangeo online climate data store[1]), from GCMs having 8 or more published runs, and containing data starting in 1850, resulting in 300 simulations of mean temperature and 139 of maximum temperature from 12 different GCMs (see Table B for information on models). We string together historical and future ensemble members with the same run identifier[2]. The result is a dataset of 1200 future global, 0.25-degree projections of near-surface air temperature. This product is the first of its kind to sample internal variability and observational uncertainty at this scale.

To focus on aspects of uncertainty in climate data alone, we correct for scenario uncertainty by using climate data in 20-year increments organized by Global Warming Level (GWL) rather than by calendar year. In other words, we take data from a GCM

[1] https://pangeo-data.github.io/pangeo-cmip6-cloud/index.html

[2] We note that not all ScenarioMIP runs with a particular ensemble identifier are branched from the historical run with the same identifier, however we do not see meaningful discontinuities in 2014-2015 at the transition between the historical and ScenarioMIP experiment inputs.

run from any Shared Socioeconomic Pathway (SSP, Riahi et al. [3]) and convert its calendar year axis into global mean surface temperature anomalies vs. its 1850-1900 average. Following [42, 43], 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. We thus produce 20-year time series for GWLs 0.61 (following [42], representing 1982-2001 in the real world as a reference time period) and in 0.5-degree increments for as many GWLs as are crossed in a given simulation. All impacts projections are calculated as the average annual impact across these 20 years.

The use of GWLs implies that metrics of interest vary only with average warming and are not pathway-dependent. This is the case for temperature and across the ScenarioMIP experiments [42]. Thus, GWLs allow not only to control for scenario uncertainty, but provide a way to study future impacts that is agnostic to climate sensitivity, helping avoid the so called "hot model" problem from GCMs with unphysically high climate sensitivities in CMIP6 [44]. Furthermore, we are thus able to boost the ensemble size for each GCM by pooling runs based on different SSPs, allowing for better accounting of internal variability, since otherwise many large ensembles have data available for fewer ensemble members per SSP than are necessary to characterize internal variability (see, e.g., Milinski et al. [45]).

3.2.1 Quantile Mapping Bias Correction

GCM output must be bias-corrected before using the output as you would data from observations of the climate. That is, GCMs simulate the Earth and have different probability distributions of climate variables from the true earth. Rudimentary bias-correction ensures that the GCM has the same mean state as the true Earth system. However, these methods do not correct for differences in higher moments of the distribution.

We employ the Quantile Delta Mapping (QDM) method developed by Cannon et al. [46], a non-parametric bias-correction method that projects an estimate of the observed climate by changes in the full CDF of the GCM. The QDM procedure uses the "delta method", i.e., it adds changes from an GCM to a time series taken from a reference data product. Each location is processed separately. Each day-of-year is processed separately, using a moving window of 30 days for each day of year to calculate CDFs. At a given location and each day of year, the changes between the CDF of the GCM run at a reference GWL (0.61, representing 1982-2001 in the real world) and the CDF of a future GWL is found, using 100 samples of quantiles to estimate the CDF. Then, for each day in the future period, the same quantile is found in the reference reanalysis dataset, and the difference in the CDF at that quantile and day of year of the GWL is added to it. The difference between the future GCM and GWL 0.61 GCM CDFs thus provides the necessary correction factor of the reanalysis data. That is, for $x_m(t)$, an GCM 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 GCM climate respectively. Finally, q is the quantile associated with the value of $x_m(t)$ in the GCM's simulation. We roughly follow the notation in Section 3.1 of [47] which contains more background on this method.

As an example, if the 0.6 quantile of the GCM output at a location at GWL1 is 14.5 °C and the 0.6 quantile of the same GCM run at this same location at GWL0.61 is 13.8 °C, the 0.6 quantile from the reanalysis for this day of year is projected by adding 0.8°C. This method ensures that trends present in the raw GCM output are preserved, a property not guaranteed by traditional delta-method bias-correction approaches that adjust the first moment or two of the distribution [47].

3.2.2 Quantile-Preserving Localized-Analog Downscaling

The Large Ensemble GCM simulation used here for temperature projections have native grids around or above 100 km (1 degree). Impacts are often nonlinear at high resolutions, there is thus need 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 GCM output to a finer resolution 0.25-degree grid, following Gergel et al. [47].

Similar to the QDM bias-correction, QPLAD is a non-parametric quantile-based method that provides estimates of the temperature at the fine-scale. The downscaling is performed by first aggregating the higher-resolution reanalysis output to the lower-resolution uniform 1-degree grid of the bias-corrected GCMs. Then, the downscaling model is trained by calculating the change in the empirical cdf between grids [47]. This delta is then used applied to map the lower-resolution GCM projections to the higher-resolution reanalysis grid. As with the QDM bias-correction, QPLAD preserves trends present in the raw output [47].

3.3 Partitioning uncertainty

We estimate the relative contribution of sources of uncertainty to the spread in projections of climate impacts using the Law of Total Variance [48]. Model uncertainty is estimated as the variance across model mean projections, internal uncertainty as the mean variance of projections across ensemble members within models, and observational uncertainty as the mean variance across reanalysis products used for fitting the statistical model and bias-correcting and downscaling the climate projections. This partitioning in sources of uncertainty allows a diagnosis of where geographically and for which lead times a given source of uncertainty substantially contributes to the spread in future climate risk.

3.3.1 Differentiating between uncertainty from the statistical fit and uncertainty from GCM post-processing

Uncertainty due to the choice of gridded historical data product can enter projections of climate impacts both through the fitting of the statistical impact function and through the GCM post-processing (bias-correcting and downscaling) procedure. To isolate which of these pathways dominates, we create cross-estimates of future

impacts across reanalyses using all four dose-response functions and all four sets of bias-corrected and downscaled simulations. In other words, we take the dose-response function fit using the reanalysis products ERA5, MERRA2, JRA-3Q, and GMFD and use each to project damages using GCM output bias-corrected and downscaled to all of ERA5, MERRA2, JRA-3Q, and GMFD, creating 16 sets of future damage estimates.

We then calculate the average variance across bias-corrected and downscaled runs, holding dose-response estimation fixed, and vice-versa, calculating the variance across dose-response estimations while holding bias-correction and downscaling ground truth fixed. Comparing the ratio between these two variances for a location identifies whether the spread in future damages is affected more by differences in dose-response functions or by differences in bias-correction and downscaling algorithms.

4 Discussion

In climate impacts studies, the choice of gridded historical data product used to fit relationships between climate and society and to post-process future GCM output has often been either *ad hoc* or done for reasons orthogonal to a product's suitability to accurately represent the desired climate metrics in the regions of study. We show in this study that this choice thus introduces implicit uncertainty in projections of future climate damages of a magnitude greater than the spread across different GCMs in many parts of the world for changes to a particular GWL. In other words, the most important source of climate uncertainty in projections of climate impacts may be the choice of historical data product. Since this uncertainty is not commonly addressed, uncertainties in future damages are likely underestimated, resulting in underestimations of the Social Cost of Carbon [24, 25] and possible missing of plausible but extreme tail risks. We also show that the partitioning of uncertainty depends not only on region and lead time, but also on the variable studied and the functional form of the dose-response model, with the Behrer et al. [23] payroll model, based on threshold exceedances of maximum daily temperature, producing a different clustering of gridded historical data products than the polynomial, mean daily temperature-driven models of Carleton et al. [12] and Burke et al. [22].

In this paper, we primarily use differences between reanalysis datasets as a measure of "observational uncertainty", since these datasets span many of the commonly-used data products in climate impacts studies. Reanalyses differ in which data sources (satellites, weather stations, etc.) they assimilate, how they characterize sub-grid processes, and their resolutions. We find that resolution is unlikely to explain the differences in projected damages (see, e.g., Figure A1, showing how the mortality dose-response function fit using ERA5 does not substantially change when ERA5 is regridded to the grids of the other reanalysis products), leaving differences in the configuration of the reanalysis model as drivers of uncertainty. Previous work has begun to disentangle these differences. For example, MERRA2 has known issues with its soil moisture model, which results in too wet soils, affecting precipitation and temperature near the surface [49–51] and possibly explaining the substantially different 'ideal' temperatures for GDP growth and mortality predicted by the product.

In line with Schwarzwald and Lenssen [6], we also show that internal variability can contribute a similar magnitude as the spread across GCMs to total climate uncertainty in projections of impacts in certain regions, such as the Pacific Northwest and Texas for payroll damages or Norway for GDP per capita growth changes (Figure 2).

Finally, though the BCD-ME allows us to test the sensitivity of climate impacts projections to model uncertainty, internal variability, and the choice of gridded historical data product – with the latter two rarely considered in these studies – further sources of climate uncertainty can impact damage projections beyond those studied here. Notably, the choice of post-processing methods used to downscale and bias-correct climate data can have a substantial impact on total uncertainty [52].

We conclude with two sets of recommendations, one for users of climate data when modeling future climate impacts, and one for climate scientists looking to inform climate impact research. Firstly, we urge data users to test the sensitivity of their results to the use of different latest-generation gridded historical data products, especially if no comparative evaluation of these products exists for variables of interests in the regions of study. Secondly, we call for research into better constraining historical observational uncertainty, though we recognize that this may not be feasible in all cases (the density of the historical observational network in the past is fixed, for example). An impacts-informed approach to studying observational uncertainty, using dose-response functions from the literature as done in this paper, can help highlight at what parts of the distribution of climate variables and in what regions of the world discrepancies between observational data products are most likely to affect projections of climate impacts, and thus help target improvements in these products.

Acknowledgements. The authors are grateful for insightful conversations with Tom Bearpark, Miriam Nielsen, Clara Deser, Amir Jina, Aandishah Samara, and Suzanna Camargo. N.L. is partially funded through NCAR which is sponsored by the National Science Foundation under Cooperative Agreement 1852977. R.H. and K.S. 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. A.B. is supported by Columbia Undergraduate Business Scholars (CUBS) Program and Columbia Business School Climate Knowledge Initiative (CKI).

Appendix A Additional figures

Appendix B Model data

References

[1] 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**(2), 181–198 (2013) https://doi.org/10.1093/reep/ret016

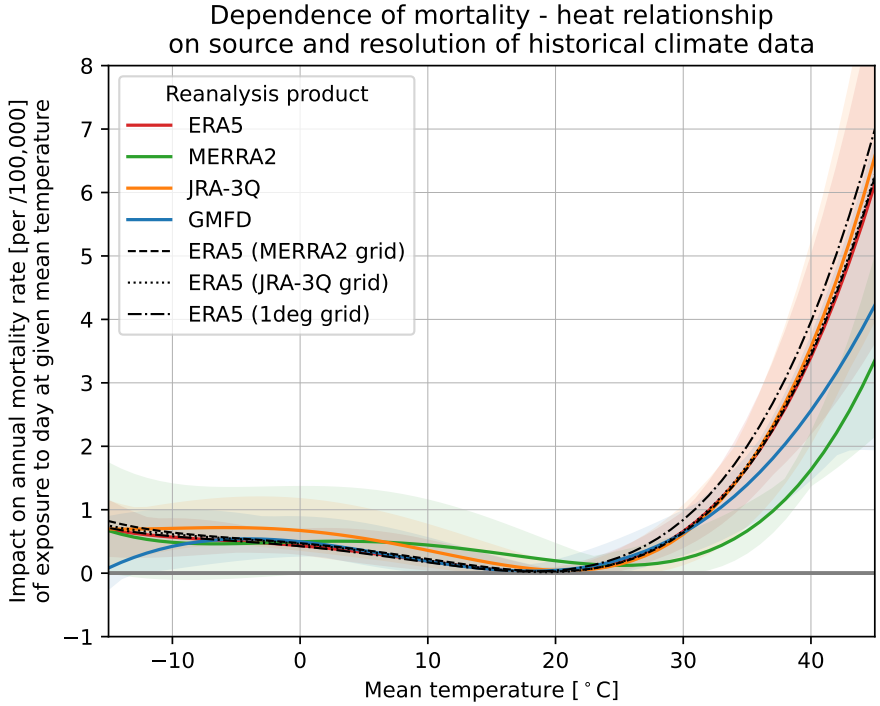


Fig. A1 Version of Figure 1**B** showing additionally the mean dose-response function fit for mortality using ERA5, regridded to the MERRA2 grid (dashed line), the JRA-3Q grid (dotted line) and an even lower resolution 1-degree grid (dot-dashed line). Dose-response functions using ERA5 at the resolutions of MERRA2 and JRA-3Q are almost identical to ERA5 at its native resolution, with only the 1-degree resolution version showing a slight change.

[2] Hawkins, E., Sutton, R.: The Potential to Narrow Uncertainty in Regional Climate Predictions. Bulletin of the American Meteorological Society **90**(8), 1095–1108 (2009) https://doi.org/10.1175/2009BAMS2607.1 . Chap. Bulletin of the American Meteorological Society

[3] Riahi, K., van Vuuren, D.P., Kriegler, E., Edmonds, J., O'Neill, B.C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J.C., Kc, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humpenöder, F., Da Silva, L.A., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J.C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A., Tavoni, M.: The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. Global Environmental Change **42**, 153–168 (2017) https://doi.org/10.1016/j.gloenvcha.2016.05.009

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	48	48	48	48	48	48	44
	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	26	26	26	26	26	16	15
	MIROC6	50	27	27	24	24	0	0
	MPI-ESM1-2-LR	15	15	15	15	13	6	0
	MRI-ESM2-0	12	8	8	8	7	2	0
	Total	300	273	263	251	206	124	94
tasmax	ACCESS-ESM1-5	6	6	6	6	6	6	6
	CNRM-CM6-1	5	5	5	3	2	2	1
	CNRM-ESM2-1	5	5	3	3	2	1	1
	CanESM5	24	24	24	24	24	24	23
	EC-Earth3	19	19	19	19	19	11	9
	EC-Earth3-Veg	15	15	15	15	15	10	7
	FGOALS-g3	10	10	10	7	3	0	0
	IPSL-CM6A-LR	16	16	16	16	16	10	10
	MIROC6	23	17	17	16	15	0	0
	MRI-ESM2-0	16	16	16	14	11	6	0
	Total	139	133	131	123	113	70	57

Table B1 Number of runs that have data for at least a given GWL, by GCM, for near-surface mean daily air temperature (`tas`, top half) and near-surface maximum daily air temperature (`tasmax`, bottom half).

[4] 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**(2), 461–471 (2015) https://doi.org/10.1162/REST_a_00478

[5] Thrasher, B., Wang, W., Michaelis, A., Melton, F., Lee, T., Nemani, R.: NASA Global Daily Downscaled Projections, CMIP6. Scientific Data **9**(1), 262 (2022) https://doi.org/10.1038/s41597-022-01393-4

[6] Schwarzwald, K., Lenssen, N.: The importance of internal climate variability in climate impact projections. Proceedings of the National Academy of Sciences **119**(42), 2208095119 (2022) https://doi.org/10.1073/pnas.2208095119

[7] Heft-Neal, S., Lobell, D.B., Burke, M.: Using remotely sensed temperature to estimate climate response functions. Environmental Research Letters **12**(1), 014013 (2017) https://doi.org/10.1088/1748-9326/aa5463

[8] Funk, C., Michaelsen, J., Verdin, J., Artan, G., Husak, G., Senay, G., Gadain, H., Magadazire, T.: The collaborative historical African rainfall model: Description

and evaluation. International Journal of Climatology **23**(1), 47–66 (2003) https://doi.org/10.1002/joc.866

[9] Dinku, T.: Overcoming challenges in the availability and use of climate data in Africa. ICT Update (2018)

[10] Dinku, T.: Challenges with availability and quality of climate data in Africa. In: Melesse, A.M., Abtew, W., Senay, G. (eds.) Extreme Hydrology and Climate Variability, pp. 71–80. Elsevier, ??? (2019). https://doi.org/10.1016/B978-0-12-815998-9.00007-5

[11] Muhire, I., Ahmed, F.: Spatiotemporal trends in mean temperatures and aridity index over Rwanda. Theoretical and Applied Climatology **123**(1), 399–414 (2016) https://doi.org/10.1007/s00704-014-1353-2

[12] Carleton, T., Jina, A., Delgado, M., Greenstone, M., Houser, T., Hsiang, S., Hultgren, A., Kopp, R.E., McCusker, K.E., Nath, I., Rising, J., Rode, A., Seo, H.K., Viaene, A., Yuan, J., Zhang, A.T.: Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits*. The Quarterly Journal of Economics **137**(4), 2037–2105 (2022) https://doi.org/10.1093/qje/qjac020

[13] Rode, A., Carleton, T., Delgado, M., Greenstone, M., Houser, T., Hsiang, S., Hultgren, A., Jina, A., Kopp, R.E., McCusker, K.E., Nath, I., Rising, J., Yuan, J.: Estimating a social cost of carbon for global energy consumption. Nature **598**(7880), 308–314 (2021) https://doi.org/10.1038/s41586-021-03883-8

[14] Hultgren, A., Carleton, T., Delgado, M., Gergel, D.R., Greenstone, M., Houser, T., Hsiang, S., Jina, A., Kopp, R.E., Malevich, S.B., McCusker, K.E., Mayer, T., Nath, I., Rising, J., Rode, A., Yuan, J.: Impacts of climate change on global agriculture accounting for adaptation. Nature **642**(8068), 644–652 (2025) https://doi.org/10.1038/s41586-025-09085-w

[15] National Center for Atmospheric Research Staff: The Climate Data Guide: NCEP-NCAR (R1): An Overview. Technical report, UCAR (December 2025)

[16] Huang, L., Fang, X., Zhang, T., Wang, H., Cui, L., Liu, L.: Evaluation of surface temperature and pressure derived from MERRA-2 and ERA5 reanalysis datasets and their applications in hourly GNSS precipitable water vapor retrieval over China. Geodesy and Geodynamics **14**(2), 111–120 (2023) https://doi.org/10.1016/j.geog.2022.08.006

[17] Bodjrenou, R., Cohard, J.-M., Hector, B., Lawin, E.A., Chagnaud, G., Danso, D.K., M'po, Y.N., Badou, F., Ahamide, B.: Evaluation of Reanalysis Estimates of Precipitation, Radiation, and Temperature over Benin (West Africa). Journal of Applied Meteorology and Climatology **62**(8), 1005–1022 (2023) https://doi.org/10.1175/JAMC-D-21-0222.1 . Chap. Journal of Applied Meteorology and

Climatology

[18] Keller, J.D., Wahl, S.: Representation of Climate in Reanalyses: An Intercomparison for Europe and North America. Journal of Climate **34**(5), 1667–1684 (2021) https://doi.org/10.1175/JCLI-D-20-0609.1 . Chap. Journal of Climate

[19] Taszarek, M., Pilguj, N., Allen, J.T., Gensini, V., Brooks, H.E., Szuster, P.: Comparison of Convective Parameters Derived from ERA5 and MERRA-2 with Rawinsonde Data over Europe and North America. Journal of Climate **34**(8), 3211–3237 (2021) https://doi.org/10.1175/JCLI-D-20-0484.1 . Chap. Journal of Climate

[20] El Kihel, B., El Kadri Elyamani, N.E., Chillali, A.: Evaluation of wind potential and economic prospects in the southern regions of Morocco with MERRA-2 and ERA5. Wind Engineering **49**(2), 475–498 (2025) https://doi.org/10.1177/0309524X241285864

[21] Yang, Y., Li, Q., Song, Z., Sun, W., Dong, W.: A comparison of global surface temperature variability, extremes and warming trend using reanalysis datasets and CMST-Interim. International Journal of Climatology **42**(11), 5609–5628 (2022) https://doi.org/10.1002/joc.7551

[22] Burke, M., Hsiang, S.M., Miguel, E.: Global non-linear effect of temperature on economic production. Nature **527**(7577), 235–239 (2015) https://doi.org/10.1038/nature15725

[23] 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**(9), 095001 (2021) https://doi.org/10.1088/2515-7620/abffa3

[24] Proistosescu, C., Wagner, G.: Uncertainties in Climate and Weather Extremes Increase the Cost of Carbon. One Earth **2**(6), 515–517 (2020) https://doi.org/10.1016/j.oneear.2020.06.002

[25] Lemoine, D.: The Climate Risk Premium: How Uncertainty Affects the Social Cost of Carbon. Journal of the Association of Environmental and Resource Economists (2021) https://doi.org/10.1086/710667

[26] Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., De Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.-N.: The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society **146**(730), 1999–2049 (2020) https://doi.org/10.1002/qj.3803

[27] Bosilovich, M.G., Akella, S., Coy, L., Cullather, R., Draper, C., Gelaro, R., Kovach, R., Liu, Q., Molod, A., Norris, P., Wargan, K., Chao, W., Reichle, R., Takasc, L., Vikhliaev, Y., Bloom, S., Collow, A., Firth, S., Labow, G., Partyka, G., Pawson, S., Reale, O., Schubert, S.D., Suarez, M.: MERRA-2: Initial Evaluation of the Climate. Technical Report NASA/TM-2015-104606/Vol.43, NASA (September 2015)

[28] Kosaka, Y., Kobayashi, S., Harada, Y., Kobayashi, C., Naoe, H., Yoshimoto, K., Harada, M., Goto, N., Chiba, J., Miyaoka, K., Sekiguchi, R., Deushi, M., Kamahori, H., Nakaegawa, T., Tanaka, T.Y., Tokuhiro, T., Sato, Y., Matsushita, Y., Onogi, K.: The JRA-3Q Reanalysis. Journal of the Meteorological Society of Japan. Ser. II **102**(1), 49–109 (2024) https://doi.org/10.2151/jmsj.2024-004

[29] Wang, J., Nikolaou, N., an der Heiden, M., Irrgang, C.: High-resolution modeling and projection of heat-related mortality in Germany under climate change. Communications Medicine **4**(1), 206 (2024) https://doi.org/10.1038/s43856-024-00643-3

[30] Callahan, C.W., Mankin, J.S.: Globally unequal effect of extreme heat on economic growth. Science Advances **8**(43), 3726 (2022) https://doi.org/10.1126/sciadv.add3726

[31] Waidelich, P., Batibeniz, F., Rising, J., Kikstra, J.S., Seneviratne, S.I.: Climate damage projections beyond annual temperature. Nature Climate Change, 1–8 (2024) https://doi.org/10.1038/s41558-024-01990-8

[32] Amnuaylojaroen, T., Parasin, N., Limsakul, A.: Projections and patterns of heat-related mortality impacts from climate change in Southeast Asia. Environmental Research Communications **6**(3), 035019 (2024) https://doi.org/10.1088/2515-7620/ad3128

[33] Chen, W., Lu, X., Yuan, D., Chen, Y., Li, Z., Huang, Y., Fung, T., Sun, H., Fung, J.C.H.: Global PM2.5 Prediction and Associated Mortality to 2100 under Different Climate Change Scenarios. Environmental Science & Technology **57**(27), 10039–10052 (2023) https://doi.org/10.1021/acs.est.3c03804

[34] Ortiz-Bobea, A., Ault, T.R., Carrillo, C.M., Chambers, R.G., Lobell, D.B.: Anthropogenic climate change has slowed global agricultural productivity growth. Nature Climate Change **11**(4), 306–312 (2021) https://doi.org/10.1038/s41558-021-01000-1

[35] Tsujimoto, K., Kuriya, N., Ohta, T., Homma, K., Im, M.S.: Quantifying the GCM-related uncertainty for climate change impact assessment of rainfed rice production in Cambodia by a combined hydrologic - rice growth model. Ecological Modelling **464**, 109815 (2022) https://doi.org/10.1016/j.ecolmodel.2021.109815

[36] Poon, E.K.W., Kitsios, V., Pilcher, D., Bellomo, R., Raman, J.: Projecting Future

Climate Impact on National Australian Respiratory-Related Intensive Care Unit Demand. Heart, Lung and Circulation **32**(1), 95–104 (2023) https://doi.org/10.1016/j.hlc.2022.12.001

[37] Adeyeri, O.E., Zhou, W., Ndehedehe, C.E., Wang, X., Ishola, K.A., Laux, P.: Minimizing uncertainties in climate projections and water budget reveals the vulnerability of freshwater to climate change. One Earth **7**(1), 72–87 (2024) https://doi.org/10.1016/j.oneear.2023.12.013

[38] Willmott, C.J., Matsuura, K.: Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1950 - 1999), http://climate.geog.udel.edu/html_pages/README.ghcn_ts2.html (2001)

[39] Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., Pasteris, P.P.: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. International Journal of Climatology **28**(15), 2031–2064 (2008) https://doi.org/10.1002/joc.1688

[40] Frieler, K., Volkholz, J., Lange, S., Schewe, J., Mengel, M., Del Rocío Rivas López, M., Otto, C., Reyer, C.P.O., Karger, D.N., Malle, J.T., Treu, S., Menz, C., Blanchard, J.L., Harrison, C.S., Petrik, C.M., Eddy, T.D., Ortega-Cisneros, K., Novaglio, C., Rousseau, Y., Watson, R.A., Stock, C., Liu, X., Heneghan, R., Tittensor, D., Maury, O., Büchner, M., Vogt, T., Wang, T., Sun, F., Sauer, I.J., Koch, J., Vanderkelen, I., Jägermeyr, J., Müller, C., Rabin, S., Klar, J., Vega Del Valle, I.D., Lasslop, G., Chadburn, S., Burke, E., Gallego-Sala, A., Smith, N., Chang, J., Hantson, S., Burton, C., Gädeke, A., Li, F., Gosling, S.N., Müller Schmied, H., Hattermann, F., Wang, J., Yao, F., Hickler, T., Marcé, R., Pierson, D., Thiery, W., Mercado-Bettín, D., Ladwig, R., Ayala-Zamora, A.I., Forrest, M., Bechtold, M.: 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), 1–51 (2024) https://doi.org/10.5194/gmd-17-1-2024

[41] Deser, C., Lehner, F., Rodgers, K.B., Ault, T., Delworth, T.L., DiNezio, P.N., Fiore, A., Frankignoul, C., Fyfe, J.C., Horton, D.E., Kay, J.E., Knutti, R., Lovenduski, N.S., Marotzke, J., McKinnon, K.A., Minobe, S., Randerson, J., Screen, J.A., Simpson, I.R., Ting, M.: Insights from Earth system model initial-condition large ensembles and future prospects. Nature Climate Change **10**(4), 277–286 (2020) https://doi.org/10.1038/s41558-020-0731-2

[42] Seneviratne, S.I., Hauser, M.: Regional Climate Sensitivity of Climate Extremes in CMIP6 Versus CMIP5 Multimodel Ensembles. Earth's Future **8**(9), 2019–001474 (2020) https://doi.org/10.1029/2019EF001474

[43] 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**(2), 485–505 (2023) https://doi.org/10.5194/esd-14-485-2023

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

[45] Milinski, S., Maher, N., Olonscheck, D.: How large does a large ensemble need to be? Earth System Dynamics **11**(4), 885–901 (2020) https://doi.org/10.5194/esd-11-885-2020

[46] 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**(17), 6938–6959 (2015) https://doi.org/10.1175/JCLI-D-14-00754.1 . Chap. Journal of Climate

[47] Gergel, D.R., Malevich, S.B., McCusker, K.E., Tenezakis, E., Delgado, M.T., Fish, M.A., Kopp, R.E.: Global Downscaled Projections for Climate Impacts Research (GDPCIR): Preserving quantile trends for modeling future climate impacts. Geoscientific Model Development **17**(1), 191–227 (2024) https://doi.org/10.5194/gmd-17-191-2024

[48] Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E.M., Brunner, L., Knutti, R., Hawkins, E.: Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. Earth System Dynamics **11**(2), 491–508 (2020) https://doi.org/10.5194/esd-11-491-2020

[49] Hirschi, M., Stradiotti, P., Crezee, B., Dorigo, W., Seneviratne, S.I.: Potential of long-term satellite observations and reanalysis products for characterising soil drying: Trends and drought events. Hydrology and Earth System Sciences **29**(2), 397–425 (2025) https://doi.org/10.5194/hess-29-397-2025

[50] Tian, B., Fetzer, E.J., Teixeira, J.: Assessing the Tropospheric Temperature and Humidity Simulations in CMIP3/5/6 Models Using the AIRS Obs4MIPs V2.1 Data. Journal of Geophysical Research: Atmospheres **129**(15), 2023–040536 (2024) https://doi.org/10.1029/2023JD040536

[51] Draper, C., Reichle, R.H.: Assimilation of Satellite Soil Moisture for Improved Atmospheric Reanalyses. Monthly Weather Review **147**(6), 2163–2188 (2019) https://doi.org/10.1175/MWR-D-18-0393.1 . Chap. Monthly Weather Review

[52] Lafferty, D.C., Sriver, R.L.: Downscaling and bias-correction contribute considerable uncertainty to local climate projections in CMIP6. Preprint, Preprints (April 2023). https://doi.org/10.22541/essoar.168286894.44910061/v1