Downscaling and bias-correction contribute considerable uncertainty to local climate projections in CMIP6

David C Lafferty[1] and Ryan L. Sriver[1]

[1]Affiliation not available

April 30, 2023

Abstract

Efforts to diagnose the risks of a changing climate often rely on downscaled and bias-corrected climate information, making it important to understand the uncertainties and potential biases of this approach. Here, we perform a variance decomposition to partition uncertainty in global climate projections and quantify the relative importance of downscaling and bias-correction. We analyze simple climate metrics such as annual temperature and precipitation averages, as well as several indices of climate extremes. We find that downscaling and bias-correction often contribute substantial uncertainty to local decision-relevant climate outcomes, though our results are strongly heterogeneous across space, time, and climate metrics. Our results can provide guidance to impact modelers and decision-makers regarding the uncertainties associated with downscaling and bias-correction when performing local-scale analyses, as neglecting to account for these uncertainties may risk overconfidence relative to the full range of possible climate futures.

Supporting Information

Downscaling and bias-correction contribute considerable uncertainty to local climate projections in CMIP6

David C. Lafferty[a]* & Ryan L. Sriver[a]

[a]Department of Atmospheric Sciences, University of Illinois Urbana-Champaign

*Corresponding author email: davidcl2@illinois.edu

Abstract

Efforts to diagnose the risks of a changing climate often rely on downscaled and bias-corrected climate information, making it important to understand the uncertainties and potential biases of this approach. Here, we perform a variance decomposition to partition uncertainty in global climate projections and quantify the relative importance of downscaling and bias-correction. We analyze simple climate metrics such as annual temperature and precipitation averages, as well as several indices of climate extremes. We find that downscaling and bias-correction often contribute substantial uncertainty to local decision-relevant climate outcomes, though our results are strongly heterogeneous across space, time, and climate metrics. Our results can provide guidance to impact modelers and decision-makers regarding the uncertainties associated with downscaling and bias-correction when performing local-scale analyses, as neglecting to account for these uncertainties may risk overconfidence relative to the full range of possible climate futures.

Main

Climate change is a global phenomenon that manifests on regional to local scales [1]. Managing the risks of a changing climate thus requires accurate, high-resolution climate projections as well as an understanding of the associated uncertainties. One of our primary sources of information about future climate change is ensembles of coupled general circulation models (GCMs) run under various greenhouse gas emissions scenarios [2]. However, GCM projections of future climate are highly uncertain, owing to three primary factors: model uncertainty, arising from differences in the structures and parameters of GCMs and thus their responses to the same radiative forcing input; scenario uncertainty, arising from the range of possible future greenhouse gas emissions trajectories; and internal variability, arising from the chaotic nature of the Earth system.

Understanding the relative importance of each of these sources of uncertainty can help guide research agendas and inform the modeling choices of end-users. Several previous studies have made important progress towards this goal for a variety of both climate and socioeconomic outcomes [3–8]. Hawkins and Sutton [3] (hereafter, HS09) use model outputs from the Coupled Model Intercomparison Project Phase 3 (CMIP3) to partition uncertainty in global and regional temperature projections, later extending their analysis to precipitation [5]. More recently, Lehner et al. [6] (hereafter, L20) leverage single model initial condition large ensembles (SMILEs) alongside CMIP6 outputs to better characterize internal variability, particularly at regional to local scales where its influence can be dominant. Using a similar SMILE-based approach, Blanusa et al. [7] (hereafter, B23) highlight the importance of internal variability in driving daily temperature and precipitation extremes.

While these works have led to many useful insights, they primarily rely on GCM outputs that are typically viewed as unsuitable for downstream analyses owing to their coarse spatial resolutions and systematic biases [9]. GCM outputs often need to be downscaled (to increase the spatial resolution) and bias-corrected (to remove systematic biases) before being considered suitable for the wide variety of end-uses in which they might be employed, including impact assessments [10, 11], adaptation planning [12], infrastructure design [13], and financial risk disclosures [14]. However, constructing a downscaled and bias-corrected ensemble requires making several methodological choices [15, 16], including the configuration of the downscaling and bias-correction algorithms, the selection and temporal slicing of the underlying observational dataset, and the sampling of parent GCMs and greenhouse gas emissions scenarios. These choices can combine to produce considerable differences in the projections from various ensembles such that users who rely on different datasets may attain meaningfully different results [17–22]. This, in turn, has motivated a separate body of work aimed at quantifying the importance of downscaling and bias-correction relative to other sources of uncertainty, but these studies often report mixed conclusions [23–28]. For example, Chegwidden et al. [25] analyze hydrologic variables in the Pacific Northwest region of North America and find that the choice of downscaling algorithm does not contribute meaningfully to projection spread compared to the influence of scenarios, GCMs, and hydrologic models. In contrast, Wootten et al. [27] focus on meteorological variables in the southeastern United States and conclude that impact assessments using only a single set of downscaled and bias-corrected GCMs may suffer from overconfidence. Many of the conflicting results in this literature can be explained by different studies focusing on distinct and often small geographic regions, or on varying sets of meteorological or hydrological variables. Each study also relies on a unique sampling of GCMs, scenarios, and downscaling and bias-correction algorithms, which can lead to different uncertainty decompositions.

In this work, we aim to address the above literature gaps by quantifying the contribution of downscaling and bias-correction to projection uncertainty for a variety of climate metrics at global scale. Following the simple variance decomposition approach of previous works [27], we account for scenario uncertainty, model uncertainty, downscaling and bias-correction uncertainty, and interannual variability. Our approach involves calculating the variance along each axis of uncertainty to obtain estimates of the time-evolving relative contribution of each source to the total projection spread (see Methods section). We focus on statistically downscaled and bias-corrected ensembles and include, to our knowledge, all global, publicly available datasets with parent GCMs taken from the CMIP6 repository [29]. This leads to a meta-ensemble comprising approximately 200 downscaled and bias-corrected model outputs across 4 emissions scenarios, 22 parent CMIP6 models, and 5 downscaling and bias-correction algorithms (Supplementary Table 1). Owing to data availability, we are restricted to analyzing metrics of climate change derived from daily maximum or minimum temperature and daily precipitation. Our selection of indicators includes annual temperature and precipitation averages as well as several indices of climate extremes due to their potential for large impacts on a broad variety of human-environment systems [30].

Our uncertainty partitioning results are strongly heterogeneous across space, time, and climate metrics. However, in general, we find that downscaling and bias-correction contribute a non-negligible fraction of the total projection spread and in many cases can represent the primary source of uncertainty. Downscaling and bias-correction are particularly important over the near term (early-to-mid 21st century), in projections of precipitation, in projections of extremes, in regions of complex terrain, and in regions where historical observations disagree.

Our results indicate that in many instances, relying on a single set of downscaled and bias-corrected outputs may risk overconfidence. For stakeholders or impact modelers who lack the computational capacity to extensively sample across all four sources of uncertainty, our results may also assist in deciding which factors to prioritize.

Results

Hereafter, to improve readability, we use the terms "downscaled" or "downscaling" to encompass the outputs or methods of downscaled and bias-corrected ensembles, unless the distinction between downscaling and bias-correction is important.

Variance decomposition of climate averages

We begin by analyzing indicators of long-term climatic change, namely annual average temperature and annual total precipitation. Before moving to the global picture, we focus on three example locations: New Delhi, India; Seattle, USA; and Lagos, Nigeria. In addition to being populous and economically important cities with distinct climates, these locations allow a comparison to previous works (L20, B23). The variance decomposition results for each city, as well as each individual downscaled projection, is shown in Figure 1. There is broad agreement on the sign of change for both temperature and precipitation, with average temperatures generally increasing in all locations (Fig. 1a-c) and total precipitation slightly increasing in New Delhi and Seattle (Fig. 1g-h) while remaining approximately constant in Lagos (Fig. 1i). However, there is considerable projection spread for all metrics and locations, and the resulting variance decompositions lead to different interpretations as to the driving factors. For temperature projections (Fig. 1d-f), the contribution of scenario uncertainty is similar in all three locations, starting small and only becoming non-negligible after around 2050. The reverse is true for interannual variability, which is more important in the first half of the century and declines over time. Similarly, the relative contribution of downscaling is largest over the near term and declines over time. However, there are considerable differences in magnitude across the three cities: temperature projections in New Delhi show little dependence on the choice of downscaled ensemble (Fig. 1d) whereas downscaling is the dominant uncertainty in Lagos long into the 21st century (Fig. 1f). For precipitation projections, a qualitatively different uncertainty decomposition emerges (Fig. 1j-l). Interannual variability is much more important in all locations, while the contribution of scenario uncertainty virtually disappears. In Seattle, downscaling is responsible for a substantial fraction of the variance of precipitation projections (Fig. 1k), model uncertainty contributes a small but perceptible fraction, and the overall decomposition changes little over time. This contrasts with New Delhi (Fig. 1j) and Lagos (Fig. 1l), where model uncertainty is relatively more important and grows over time.

We use Figure 1 only as a demonstration of the complex and sometimes non-intuitive nature of the interplay among these four uncertainty sources at local scales. The results of each variance decomposition arise from a combination of factors unique to each location. For example, the importance of downscaling uncertainty for Seattle precipitation may be related to its positioning in a mountainous region [19], whereas the dominance of downscaling uncertainty in Lagos temperature projections may be driven by disagreements among the underlying observational datasets used to perform the downscaling (Figs. S1-S2). Fully explaining each uncertainty decomposition would require expertise regarding the many physical processes affecting each location's climate, an understanding of their representations in the CMIP6 GCMs, and knowledge of how the resulting temperature and precipitation outputs are affected by each downscaling methodology.

We now apply our variance decomposition globally, continuing to focus on climate averages. These results are shown in Figure 2, where uncertainty sources are shown along each column and each row shows a 20-year period representing the early, mid, and late 21st century. The global results are largely in keeping with those of the three example cities. For annual average temperature, across almost all regions of the globe, there is a marked increase in the contribution of scenario uncertainty over time and a corresponding decrease in downscaling uncertainty and interannual variability. This matches the behavior of each of the locations shown in Figure 1, even if the magnitudes differ. For example, Lagos can be seen as an outlier in terms of the importance of downscaling uncertainty—by the late 21st century, downscaling still contributes around 25% of the total variance of Lagos temperature projections (Fig. 1f), almost double the global average. Figure 2a also shows that in many locations, model uncertainty grows to become the most important driver of variance by mid-century and continues to contribute a substantial fraction by late-century, though scenario uncertainty typically becomes larger. For annual total precipitation (Fig. 2b), interannual variability remains the dominant contributor, usually followed by downscaling uncertainty and model uncertainty while scenario uncertainty is almost always negligible. As in Figure 1, the precipitation decomposition changes little over

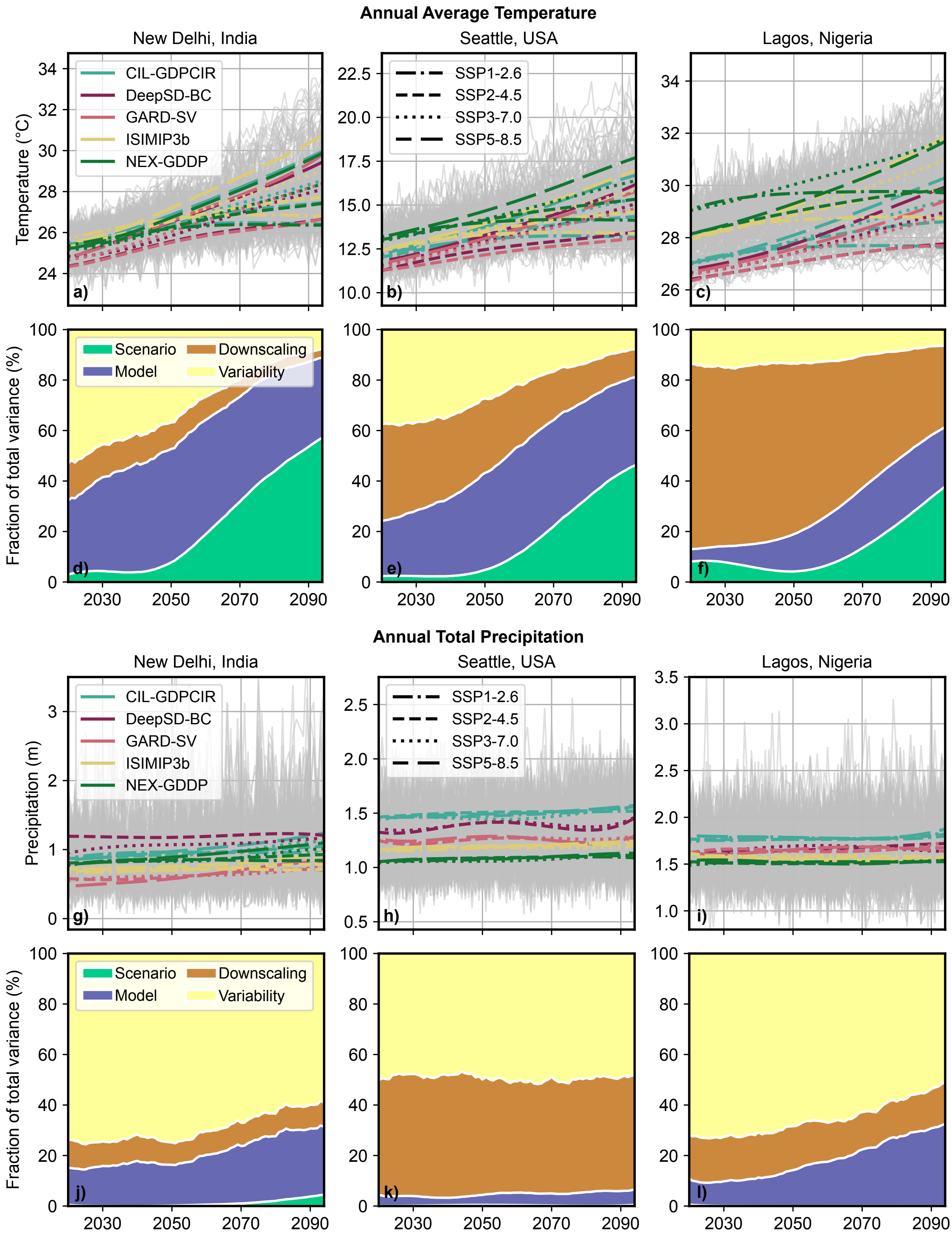

Figure 1: **Projections and variance decomposition of climate averages.** **a-c** Timeseries of annual average temperature from each downscaled model output. Gray lines show individual model outputs and colored lines of different styles show associated ensemble-scenario means. Outputs for each city are taken from the single grid point encompassing their respective locations. **d-f** Variance decomposition of annual average temperatures corresponding to the timeseries plots in a-c. The contribution of each uncertainty source is expressed as a percentage of the total variance. **g-i** Timeseries of annual total precipitation, similar to a-c. **j-l** Variance decomposition of annual total precipitation, similar to d-f.

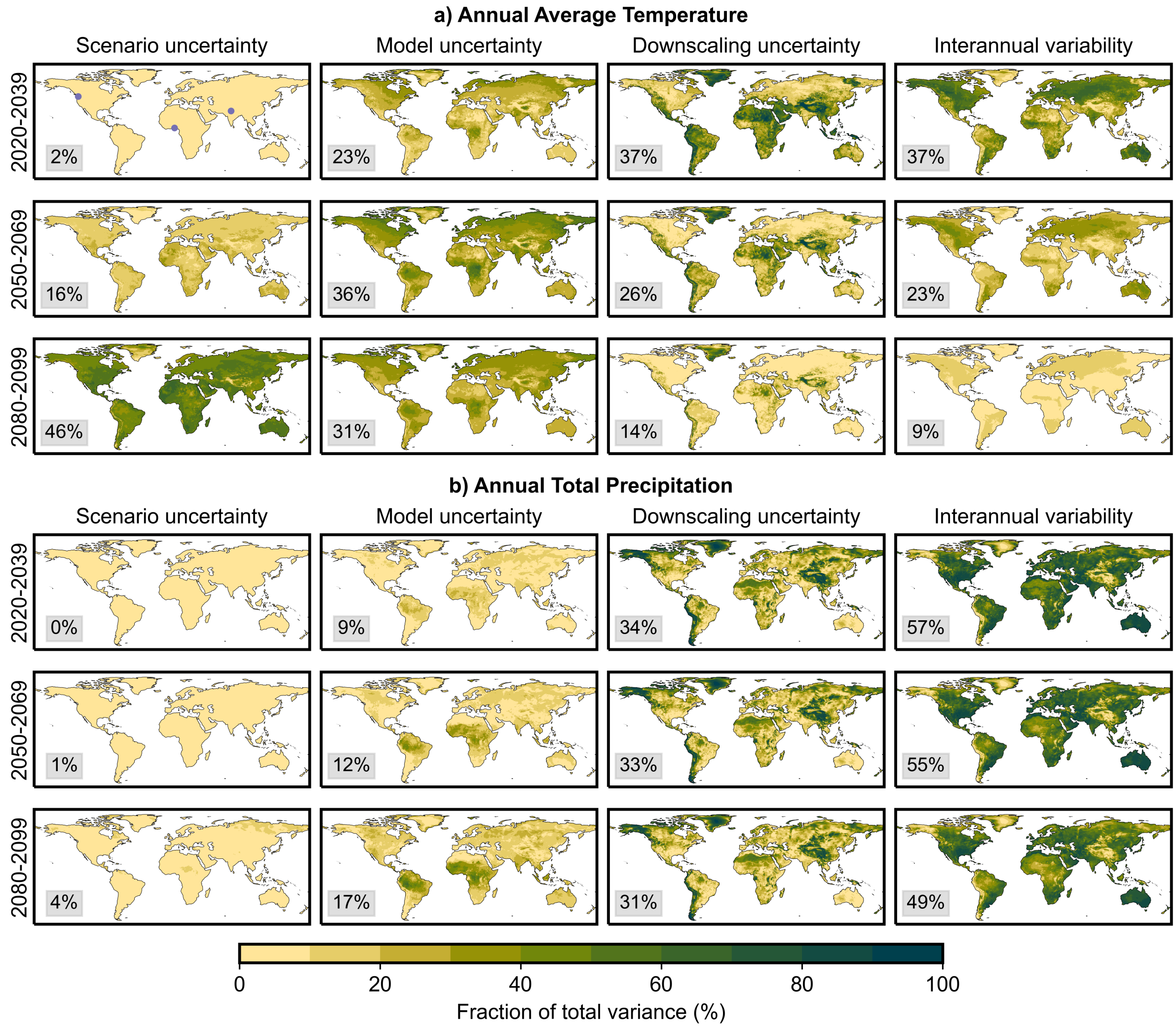

Figure 2: **Global variance decomposition of climate averages. a** Variance decomposition for annual average temperature. Each column shows the contribution from a different source of uncertainty, measured as the fraction of total variance. Each row depicts a 20-year period representing either the early, mid, or late 21st century. The purple dots in the upper left subplot show the locations of New Delhi, Seattle, and Lagos. **b** Variance decomposition for annual total precipitation, in the same layout as a. The gray boxes in the lower left of each subplot gives the global average.

time.

The global results shown in Figure 2 also reveal some important spatial patterns. Notably, regions of complex terrain are often associated with larger downscaling uncertainties. For both temperature and precipitation projections, major mountain ranges including the Rocky Mountains, the Andes, and the Himalayas exhibit comparatively large downscaling uncertainties with correspondingly lower contributions from other sources. This could be due to topographic influences on atmospheric dynamics that are not well represented in coarse-resolution GCMs, leading to methodological differences in the downscaling algorithms being amplified into a larger spread in outcomes [31]. However, the same regions also tend to show larger disagreements in the historical record (Fig. S5), which can drive differences in the projections [32, 33]. There are other regions, such as Greenland and parts of the Sahara desert, where large downscaling uncertainties are likely solely driven by observational disagreements.

Our global results broadly agree with HS09 and L20, and reproduce some aspects of the spatial patterns uncovered by L20. For example, we find for temperature projections that interannual variability is largest over the mid- and high-latitudes; for precipitation projections, we also find that model uncertainty is larger in the tropics compared to other regions. In our results, interannual variability remains considerably more important beyond the early 21st century, which arises because previous works apply decadal averages to each climate metric before performing the variance decomposition. In this study, we do not average any climate indices over time in order to ensure that our results remain sensitive to the entire distribution of possible outcomes in any given year.

Variance decomposition of climate extremes

While long-term averages are important indicators of climatic change, climate and weather extremes play an outsized role in driving environmental and socioeconomic impacts [34] and can be important in shaping public perceptions [35]. In this section, we therefore apply our variance decomposition approach to a suite of indices measuring climate extremes. We first focus on annual 1-day maxima for daily maximum temperature and daily precipitation. As with our discussion of climate averages, we focus initially on the three example cities before showing the global results. The 1-day maxima timeseries and variance decompositions for New Delhi, Seattle, and Lagos are shown in Figure 3, which adopts the same layout as Figure 1. There is strong agreement across models, scenarios, and downscaling methods that the magnitudes of temperature extremes are expected to increase in future (Figs. 3a-c). However, the associated uncertainty decompositions are qualitatively different. In New Delhi and Seattle, interannual variability plays an important role in driving the variance of 1-day temperature maxima (Figs. 3d-e), more so than for annual averages (Figs. 1d-e). In these locations, the temperature decompositions also show a similar temporal pattern to the annual average results, with a growing importance for scenario uncertainty and a declining contribution from interannual variability over time. The uncertainty partitioning for Lagos is qualitatively different (Fig. 3f). Here, the overwhelming contribution arises from downscaling uncertainty, which remains almost constant throughout the century. It is worthwhile emphasizing the tremendous differences in Lagos extreme temperature projections that arise due to downscaling—annual maximum temperatures throughout the century, from the same GCM and forcing scenario, can differ by 10°C depending on the downscaling algorithm applied (Fig. 3c). As we discuss in more detail in the Supporting Information, this is driven by sizeable (but localized) disagreements between observational datasets in coastal areas (Figs. S3, S12-S17).

The partitioning for 1-day precipitation maxima is qualitatively similar in each city and the relative contribution from each uncertainty source remains constant over time (Figs. 3j-i). Interannual variability is the primary driver of variance in New Delhi; in Seattle, interannual variability and downscaling contribute approximately equally; in Lagos, although interannual variability remains important, downscaling is the largest contributor, which again likely arises due to observational disagreements (Fig. S4). In New Delhi and Seattle, the decomposition for 1-day precipitation maxima (Figs. 3j-k) is fairly similar to the breakdown for annual total precipitation in those locations (Figs. 1j-k). In Lagos, downscaling plays a much more prominent role in driving the variance of precipitation extremes (Fig. 3l) relative to annual totals (Fig. 1l).

In Figure 4, we show global maps of the variance decompositions for annual 1-day maxima. The spatial patterns of these results share many commonalities with those of annual averages (Fig. 2). Specifically, regions of complex terrain and areas of relatively large observational disagreement (Fig. S6) are often associated with larger downscaling uncertainties. The temporal evolutions are also broadly similar—for both average metrics and 1-day maxima, the precipitation decomposition remains approximately constant over time and the temperature decomposition shows a similar pattern of increasing relative contributions from model and scenario uncertainty at the expense of downscaling uncertainty and interannual variability. In terms of the magnitude of the contribution from each source, the decomposition for 1-day precipitation maxima (Fig. 4b) is very similar to that for annual totals (Fig. 2b). One of the few differences is that interannual variability becomes slightly more important at the expense of model uncertainty, particularly in the tropics. For temperature projections, there are notable differences. Downscaling and interannual variability play a more important role at longer time horizons for annual maximum temperatures (Figs. 4a) compared to annual average temperatures (Fig. 2a). Recall that for annual average temperatures, scenario and model uncertainty account for most of the variance by the late 21st century (around 80%, globally averaged; Fig. 2a). The corresponding late-century breakdown for maximum temperatures is qualitatively different as each

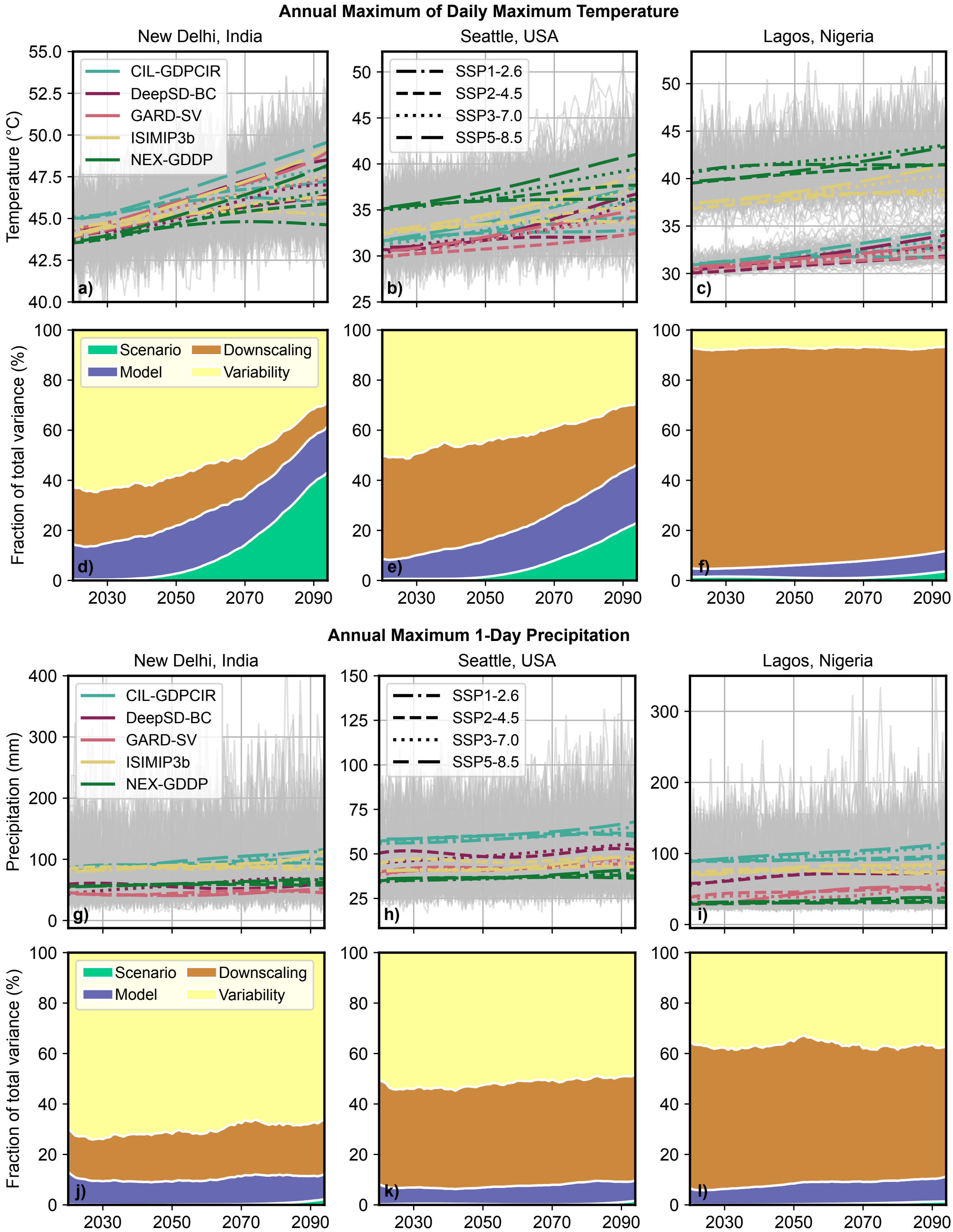

Figure 3: **Projections and variance decomposition of annual 1-day maxima. a-c** Timeseries of annual maximum temperature from each downscaled model output and associated ensemble-scenario means, in a similar format to Figure 1. **d-f** Variance decomposition of annual maximum temperatures corresponding to the timeseries plots in a-c. The contribution of each uncertainty source is expressed as a percentage of the total variance. **g-i** Timeseries of annual maximum 1-day precipitation, similar to a-c. **j-l** Variance decomposition of annual maximum 1-day precipitation, similar to d-f.

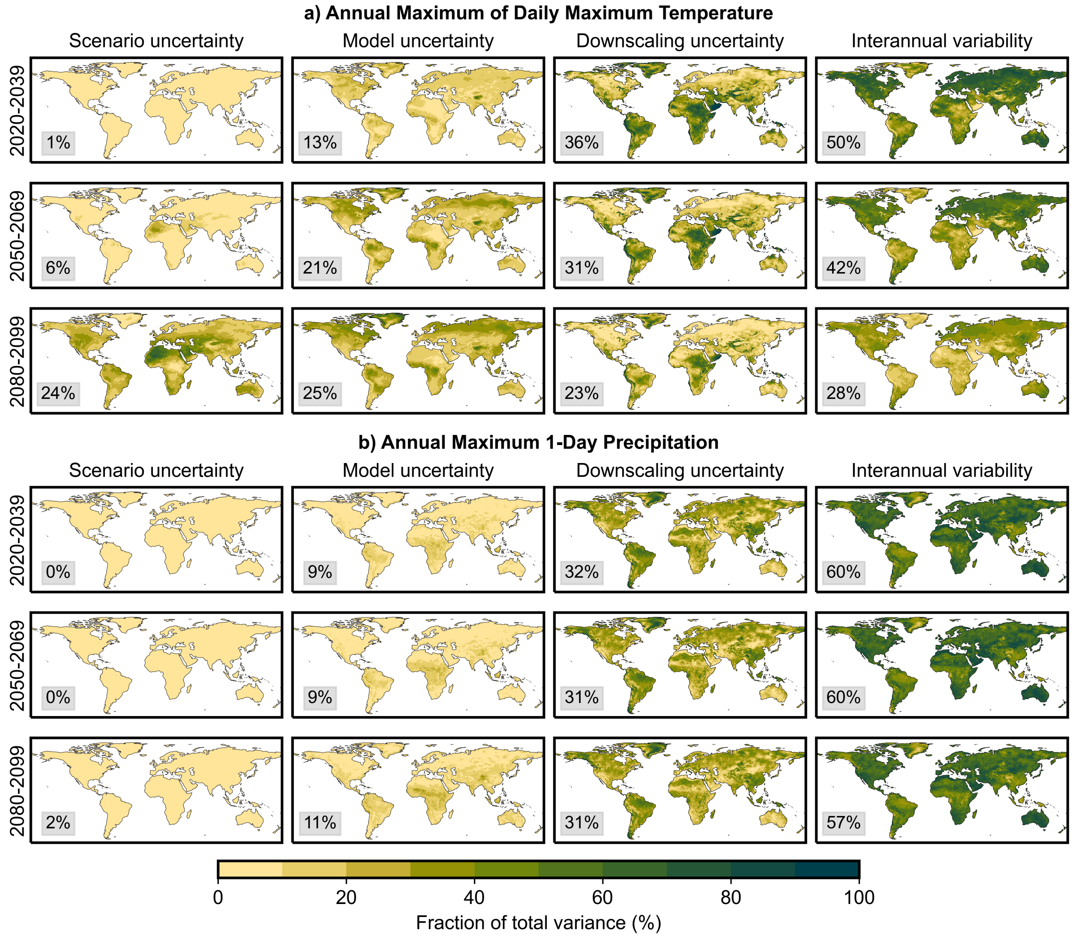

Figure 4: **Global variance decomposition of annual 1-day maxima.** **a** Variance decomposition for annual maximum temperature. As in Figure 2, columns delineate the contribution from each uncertainty source and rows demonstrate the temporal evolution. **b** Variance decomposition for annual maximum 1-day precipitation, in the same layout as a.

source contributes approximately equally (Fig. 4a).

We find qualitatively similar results for the annual maxima of daily average temperature and daily minimum temperature (Fig. S18), although downscaling is slightly less important in both cases. We also consider how the uncertainty partitioning changes for temporally compounding extremes by repeating the calculation for 5-day maxima (Fig. S19). This made very little difference for temperature projections; for precipitation, it led to a small decrease in the contribution from downscaling uncertainty and a corresponding increase in the importance of interannual variability.

There are several possible measures of climate extremes beyond annual 1-day maxima. Different end-users may care about distinctive characteristics of a given extreme [36], including its magnitude and timing in relation to relevant human and/or environmental thresholds, its correlation structure across space and time, and whether it co-occurs with another hazard (i.e., a multivariate extreme) [37]. Although mindful that any set of indices will neglect many aspects of climate extremes that are important for specific sectors, we now define and analyze a suite of metrics that aim to be as broad as possible. We choose to analyze four threshold indices: the annual number of extremely hot days (defined as daily maximum temperature exceeding the local historical 99th percentile), the annual number of dry days (daily precipitation less than 1mm), and the annual number of extremely wet days (daily precipitation exceeding the local historical 99th percentile). The resulting uncertainty decompositions are shown at the global scale in Figure 5.

Several insights emerge from Figure 5. First, there continues to exist a clear qualitative difference between the precipitation- and temperature-based indices. The decomposition for dry days (Fig. 5b) and extremely wet days (Fig. 5c) is roughly constant over time and largely dominated by downscaling uncertainty and interannual variability while scenario uncertainty again contributes negligibly. In contrast, the results for extremely hot days (Fig. 5a) show a similar temporal pattern to previous temperature-derived metrics where model and scenario uncertainty play an increasingly important role at longer time horizons. Second, note that in many regions, model uncertainty is the most important factor by the late 21st century in projecting extremely hot days, which contrasts with our results for the non-threshold metric of temperature extremes, annual maxima (Fig. 2a). This is likely related to the large spread in CMIP6 climate sensitivities [38]. Since we define an extremely hot day in reference to a constant (but local) temperature threshold, higher-sensitivity GCMs will tend to cross that threshold earlier than lower-sensitivity GCMs, leading to a relative increase in model uncertainty. Third, for all metrics analyzed thus far, the annual number of dry days is markedly the most sensitive to the choice of downscaled ensemble. This may be related to observational disagreements regarding the historical frequency of dry days (Fig. S7) but could also be driven in part by methodological differences in whether and how the bias-correction algorithms adjust their outputs based on minimum precipitation thresholds. Finally, our results for extremely hot days and extremely wet days are in reasonable qualitative agreement with those of B23, notwithstanding some differences in the magnitudes that arise due to our inclusion of downscaling uncertainty and our decision not to apply decadal averaging. For extremely hot days (compare Fig. 5a with Fig. 5 of B23), we find very similar trends in the relative contributions from model uncertainty, scenario uncertainty, and interannual/internal variability. For extremely wet days (compare Fig. 5b with Fig. 6 of B23), both sets of results are dominated by interannual/internal variability. We also

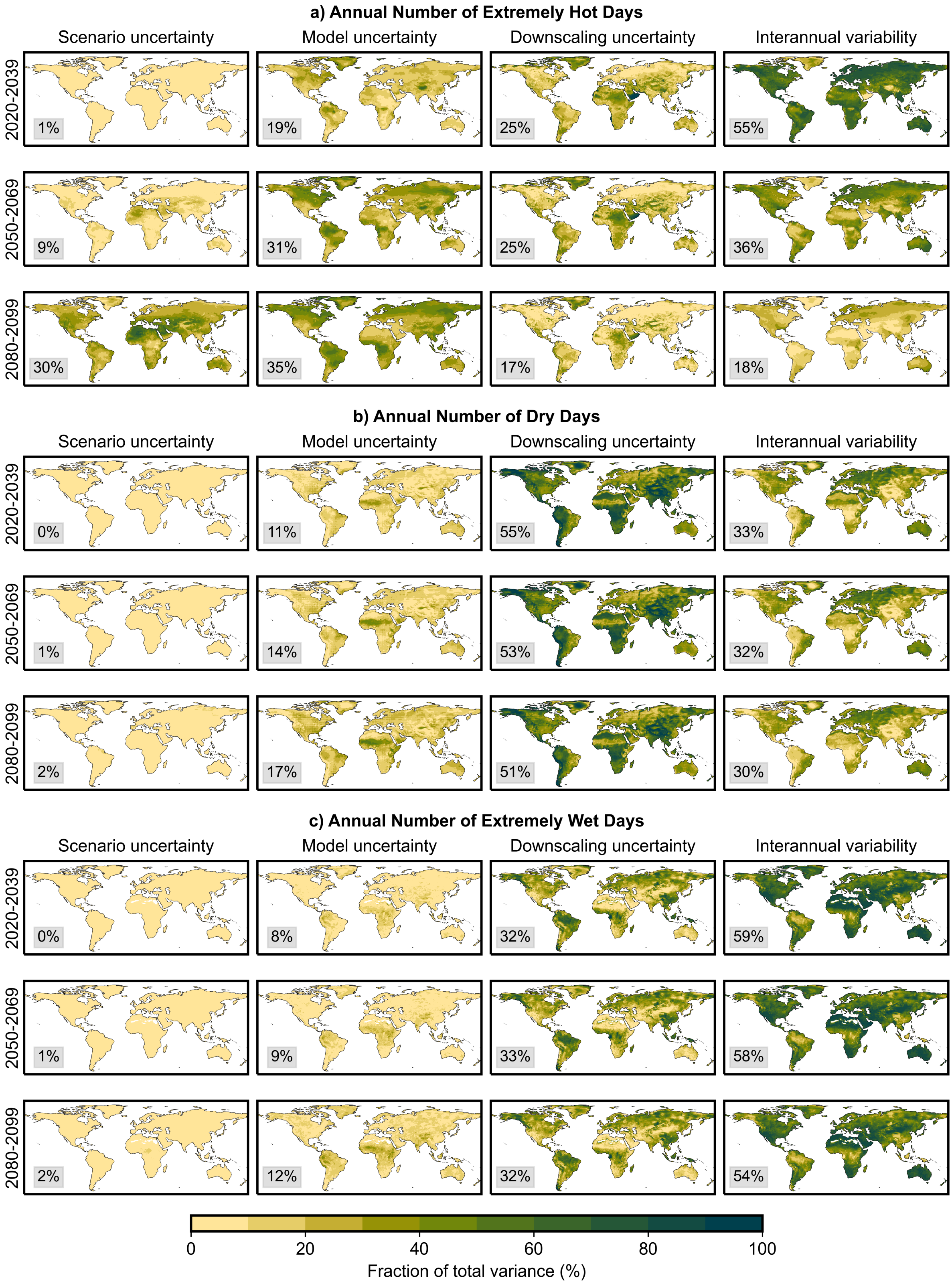

Figure 5: **Global variance decomposition of threshold indices of climate extremes.** Variance decomposition for: **a** annual number of extremely hot days, **b** annual number of dry days, and **c** annual number of extremely wet days. As in Figures 2 & 4, columns delineate the contribution from each uncertainty source and rows demonstrate the temporal evolution. Extremely hot days and extremely wet days are defined to occur when daily maximum temperature and daily precipitation exceed their local 99th percentiles, respectively, where percentiles are calculated over 1980-2014 (see Methods). Dry days are defined to occur when daily precipitation is less than 1mm.

find an increased role for model uncertainty in West Africa and the Amazon rainforest, although to a far lesser extent than B23 since downscaling also represents a significant contribution in those regions.

In the Supporting Information, we test the sensitivity of these results to several different threshold definitions (Figs. S20-S32). Broadly, we find that downscaling becomes less important if daily average or minimum temperatures are considered instead of daily maximums, and interannual variability becomes more important if more extreme thresholds are used. Calculating the historical quantiles from a separate observational dataset can lead to some differences in the contribution from downscaling uncertainty, but this does not change the qualitative results. We also include extensions to account for temporally compounding extremes by calculating the longest consecutive run of days crossing each threshold, the main effect of which is to increase the importance of interannual variability (Figs. S20-32). Lastly, we also investigate a simple multivariate metric, extremely hot and dry days (Figs. S33-S34), which shows a very similar decomposition to that for extremely hot days. This indicates that conditioning the occurrence of daily temperature extremes on concurrent low precipitation does little to alter the uncertainty decomposition, although it is unclear whether this result would hold over longer timescales.

Implications for risk assessment

Current impact analyses often rely on a single set of downscaled climate model outputs. Our results so far suggest that this approach may lead to overconfidence by generating an artificially narrow probability distribution relative to the full range of plausible climate futures. To demonstrate this effect, we provide a stylized example around characterizing mid-century hot and wet extremes in Seattle, shown in Figure 6. Using four previously defined indices of extremes, Figure 6 illustrates the effects of only sampling from one downscaled ensemble by comparing the resulting probability distribution to that obtained by using the entire meta-ensemble. For every metric examined, key distributional statistics such as the median and 95th percentile vary considerably among each downscaled ensemble as well as in relation to the full ensemble. For extremely hot days (Fig. 6a), extremely wet days (Fig. 6c), and maximum 1-day precipitation (Fig. 6d), neglecting to sample across downscaled ensembles can induce greater distributional changes than neglecting to sample across emissions scenarios. The distributions from different ensembles are most similar for annual maximum temperature (Fig. 6b), though there are still notable differences. Consider, for example, the extraordinary 2021 Pacific Northwest heatwave,

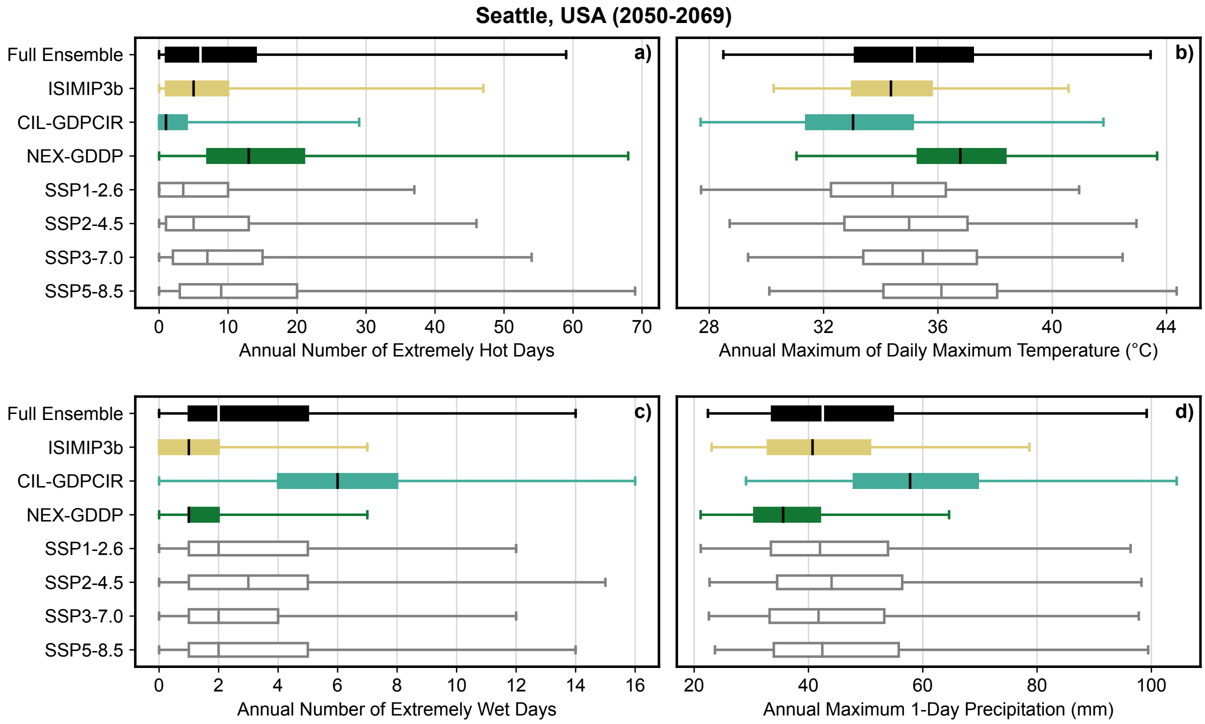

Figure 6: **Hazard characterization depends on modeling choices.** Comparison of the probability distribution generated by relying on the full meta-ensemble (all downscaled outputs and scenarios; black boxplot), any one downscaled ensemble (including all scenarios; colored boxplots), or any one scenario (including all ensembles; gray boxplots). Distributions are constructed for the grid point containing Seattle over 2050-2069 for different metrics: **a** annual number of extremely hot days, **b** annual maximum temperature, **c** annual number of extremely wet days, and **d** annual maximum 1-day precipitation. Boxplot whiskers span the 99% range. Details on each downscaled ensemble and the Shared Socioeconomic Pathway (SSP) scenarios can be found in the Methods section and Supporting Information. We neglect the carbonplan ensembles here since they contain a limited number of models.

which has been extensively studied after breaking several temperature records throughout the region [39–42], leading to widespread impacts across many sectors [43]. During this event, Seattle-Tacoma airport recorded a temperature of 42.2°C [44]. Figure 6b shows that estimates of the likelihood of surpassing this record by mid-century depend strongly on the choice of downscaled ensemble, as two from three ensembles project that this record is unlikely to be broken by mid-century even under extreme emissions scenarios.

Figure 6 presents a highly simplified example that neglects many of the challenges of implementing risk and decision analyses in a nonstationary climate [45, 46]. It nonetheless serves to illustrate how modeling choices surrounding downscaled data sources can induce substantively different hazard characterizations. The consequences of relying on a single downscaled ensemble may be more or less severe in other locations and for other hazards, but these results suggest that careful consideration should be given to the role of downscaling uncertainty within any broader risk assessment.

Discussion

Our main finding, that downscaling and bias-correction often contribute considerable uncertainty in local climate projections, is robust to a number of methodological checks that we outline in the Methods section and Supporting Information, though there are several possible avenues of future research. First, note that despite our simplified treatment of internal variability (see associated discussion in the Methods section), we nonetheless find that interannual variability is an important driver of uncertainty for many metrics. For several precipitation-based metrics and indices of extremes, the combined contribution of interannual variability and downscaling drive a large share of the variance. This would suggest that future work characterizing uncertainties around the role of internal variability at local scales would be highly valuable. The framework presented here could be extended to include downscaled initial condition ensembles [47], but to our knowledge such an ensemble does not yet exist at global scale. Independent estimates of internal variability at local scales, potentially derived from hybrid statistical techniques [48], could also be used to test for potential biases in the model-derived representation used here.

Second, we sample only a subset of the many different methods that can be used to downscale and bias-correct climate data. Many GCMs in our meta-ensemble are only downscaled in two different ways, and thus our estimate of the downscaling uncertainty (the variance across downscaling methods) likely suffers from biases associated with low sample size. We partially mitigate this bias by averaging each individual estimate across GCMs but expanding the meta-ensemble to include more downscaling algorithms should lead to more robust estimates. Most of the downscaling algorithms we consider are univariate approaches that do not adjust their outputs for spatial correlations (Supplementary Table 2), so expanding the meta-ensemble in a targeted manner that accounts for these aspects of the downscaling procedure could be particularly beneficial. We also do not include any dynamical downscaling approaches, which may provide some advantages over statistical methods [49]. In general, adding more ensembles to the uncertainty decomposition could result in an increase or decrease in the relative importance of downscaling, depending on whether the additional ensembles exhibit similar projections [27].

Third, we again highlight that our selection of climate metrics is necessarily limited. Since all of the indices we analyze are calculated annually, we are unable to probe extremes that manifest on longer timescales (for example, the magnitude of a 10-year return period event) and we aggregate over seasonal information that is important for many sectors. A useful extension to this work could test how these aspects of climate hazards alter the variance decompositions. Additionally, moving beyond standardized meteorological indices to analyze targeted metrics that are relevant for specific sectors may lead to qualitatively different results [50].

Finally, note that variance decomposition is only one of many possible approaches to characterize uncertainty. More formal sensitivity analysis techniques can be applied to understand specific aspects of the outcome space [51], including user-defined binary responses [52], and ensure that inferences are relevant for downstream decision analyses [53]. We also stress that for many analyses, projections of future climate represent only one source of uncertainty. Climate projections are often used to drive sectoral models that contain their own structural and parametric uncertainties [54–57]. Socioeconomic outcomes of interest may well be more sensitive to the representation of these environmental and/or human system dynamics, and sound risk management strategies should account for the uncertainty in each relevant system as well as their interactions [58].

Our results have important implications for many users of downscaled climate products. Across almost all locations, time horizons, and indices of climatic change that we analyze, downscaling rarely represents a negligible source of uncertainty. This would imply that a strategy of sampling from more than one downscaled ensemble is advisable during risk or impact analyses that are sensitive to low-probability climate hazards, as has been suggested elsewhere [27, 59]. Such a sampling may represent a substantial increase in data and computational requirements, so we emphasize that it may not be necessary in all cases. Our results can provide some initial heuristic guidance in this regard—they suggest that downscaling uncertainty is particularly important over the near term, in projections involving precipitation or climate extremes, and in regions of complex topography or observational disagreement. In general, we urge end-users to follow existing recommendations regarding the use of downscaled climate products [16, 60], including taking a process-informed approach and relying on expert knowledge of local weather and climate phenomena [61]. End-users may also consider whether downscaled projections are the most appropriate method of generating future climate information; other complementary approaches might include applying GCM-simulated changes to gridded historical data [62] or developing a statistical model based on pointwise observations [63].

This work also adds to a growing body of literature applying an increasingly diverse set of tools to characterize the uncertainties of a changing climate and the resulting environmental and socioeconomic impacts. Deliberate efforts to coordinate methodological comparisons would help build confidence in the insights derived from this line of research, which in turn will be necessary to guide best practices for the increasing number of both public and private actors who are incorporating climate projections into their decision-making processes.

Methods

Data sources

We leverage five ensembles of statistically downscaled and bias-corrected GCM outputs: NASA NEX-GDDP-CMIP6 [64] (which we refer to as NEX-GDDP), CIL-GDPCIR [65], ISIMIP3BASD [66, 67] (which we refer to as ISIMIP3b), and two ensembles from carbonplan [68]: GARD-SV [69] and DeepSD-BC [70]. Some details on the configurations of each approach can be found in Supplementary Table 2. Each ensemble is filtered to ensure: (1) parent GCMs are available in at least 2 ensembles, (2) downscaled outputs are available for at least 3 Shared Socioeconomic Pathways (SSPs) [71], (3) downscaled outputs are missing no more than one variable (from tasmax, tasmin, and pr), and (4) downscaling is performed on the same simulation member of the parent GCM. Satisfying these requirements results in dropping 13 of 35 NEX-GDDP parent models and 8 of 25 CIL-GDPCIR parent models. All ISIMIP3b outputs are used. Additional outputs from different downscaling techniques are available in the carbonplan dataset but do not satisfy the above requirements. After calculating each metric in each ensemble, all outputs are conservatively re-gridded to a common 0.25° grid.

For the threshold metrics that require comparing projection outputs to historical quantiles, we rely on two observational datasets: the Global Meteorological Forcing Dataset (GMFD) for Land Surface Modeling [72] and the ERA5 reanalysis from the European Centre for Medium-Range Weather Forecasts [73]. These products are chosen because they are available globally at 0.25° spatial resolution. GMFD is the training dataset for the NEX-GDDP ensemble, and ERA5 is the training dataset for the CIL-GDPCIR ensemble and both carbonplan ensembles, although with different temporal extents. The ISIMIP3b ensemble is trained on W5E5 v2.0 [74,75], which is only available at 0.5° spatial resolution. The quantiles are calculated from daily data over 1980-2014. We conservatively re-grid both observational datasets to the native grid of each downscaled ensemble before calculating the threshold metrics. Our definition of extremely hot days and extremely wet days in the main results is based on daily maximum temperature and daily total precipitation exceeding the local 99th percentile from GMFD, respectively. In the Supporting Information, we compare the GMFD-calculated quantiles to those obtained from ERA5 (Figs. S8-S11).

Uncertainty Partitioning

Following previous works, we employ a simple variance decomposition approach to calculate the relative uncertainty arising from four sources: scenario uncertainty, model/GCM uncertainty, downscaling uncertainty, and interannual variability. Additionally, in a similar manner to Wootten et al. [27], we employ a weighting strategy that accounts for data coverage. Our method is as follows: let $x(t, s, m, d)$ represent a given climate metric in some location at year t from scenario s, parent GCM m, and downscaling method d. We first estimate the forced response $\hat{x}(t, s, m, d)$ by fitting a 4th order polynomial over 2015-2100. Interannual variability is then estimated as the centered rolling 11-year variance of the residuals between the extracted forced response and the raw outputs. The assumption of constant interannual variability was highlighted as one shortcoming of HS09, so in this work we allow the magnitude of interannual variability to evolve over time. The contribution of each remaining uncertainty source is calculated based on the forced response. Scenario uncertainty is estimated as the variance over scenarios of the multi-model, multi-method mean,

$$U_s(t) = \mathrm{var}_s \left[\frac{1}{N(s)} \sum_{m,d} \hat{x}(t, s, m, d) \right], \tag{1}$$

where $N(s)$ is the total number of downscaled outputs available for scenario s. The above definition may underestimate the true scenario uncertainty when the multi-model, multi-method response is weak. Brekke and Barsugli [76] propose taking the variance over scenarios before averaging to circumvent this issue:

$$U_s^{bb13}(t) = \frac{1}{N_m N_d} \sum_{m,d} \mathrm{var}_s \left[\hat{x}(t, s, m, d) \right]. \tag{2}$$

Here, N_m and N_d are the number of distinct GCMs and downscaling methods in our meta-ensemble, respectively. Our main results are based on the former definition of scenario uncertainty, following much of the existing literature. In the Supporting Information we show that scenario uncertainty is indeed larger under the Brekke and Barsugli definition, although this does not change the qualitative results (Figs. S35-S41). Model uncertainty is estimated as the weighted mean of the variance across models,

$$U_m(t) = \frac{1}{N_s N_d} \sum_{s,d} w_{s,d} \mathrm{var}_m \left[\hat{x}(t, s, m, d)\right] \tag{3}$$

where N_s is the number of distinct scenarios in the meta-ensemble. The weights $w_{s,d}$ are chosen such that if more parent GCMs are available for a given downscaling method and scenario (i.e., if the variance is calculated across more GCMs), those methods and scenarios are weighted higher:

$$w_{s,d} = \frac{m(s,d)}{\sum_{s,d} m(s,d)}. \tag{4}$$

Here, $m(s,d)$ indicates the number of parent models that have been downscaled using method d for scenario s. Downscaling uncertainty is estimated as the weighted mean of the variance across methods:

$$U_d(t) = \frac{1}{N_s N_m} \sum_{s,m} w_{s,m} \mathrm{var}_d \left[\hat{x}(t, s, m, d)\right], \tag{5}$$

where the weights $w_{s,m}$ are chosen such that if more downscaled outputs are available for a given GCM and scenario, those GCMs and scenarios are weighted higher:

$$w_{s,m} = \frac{d(s,m)}{\sum_{s,m} d(s,m)}. \tag{6}$$

Here, $d(s,m)$ indicates the number of downscaled outputs available from parent GCM m and scenario s. The weighting strategy can be made more intuitive with an example: from Supplementary Table 1, there are 5 different downscaled outputs available from the CanESM5 parent GCM whereas only 2 different downscaled outputs are available from CMCC-ESM2 (neglecting SSP availability). The weighting strategy assumes that the estimated downscaling uncertainty from CanESM5 provides more information about the true uncertainty than the estimate from CMCC-ESM2. In this illustrative example, our estimate for the true downscaling uncertainty would be a weighted average of the two individual estimates, where the CanESM5 estimate is weighted higher by a factor of 5/2. In the Supporting Information, we recalculate our main results without performing any weighting and show that the qualitative interpretations are unchanged (Figs. S42-S48).

We assume that the total variance in each year is given by the sum of each individual variance estimate. Our main results show the relative contribution of each uncertainty source measured as a fraction of the total variance. Thus, while it is possible and indeed common for the absolute uncertainty of each source to grow over time, the relative importance of any one source can decline if the others grow faster (Figs. S55-S61).

In general, the assumption that all uncertainty sources are independent is false. Our assumed total uncertainty, as the sum of each individual term, can thus be larger or smaller than true total uncertainty, given by the variance across all outputs:

$$U_{total}^{true}(t) = \mathrm{var}_{s,m,d}\left[x(t,s,m,d)\right]. \tag{7}$$

In the Supporting Information we compare the true total uncertainty with our assumed total uncertainty for each metric to show that our assumption of independence generally leads to small errors, although the discrepancy can reach 20% at some locations for the extreme metrics (Figs. S52-S54).

Methodological caveats

Here we outline some methodological caveats associated with our main results. First, the 4th order polynomial fit used to separate the forced response from interannual variability likely leads to an underestimate of the true extent of internal variability since the fit will interpret unforced fluctuations as being part of the forced response. L20 show that for coarse-resolution GCM outputs, this bias can be particularly acute at regional scales and for noisy output variables such as precipitation, reaching 50% of the total uncertainty in some cases. One approach to mitigate this bias is to average over large spatial scales but this would considerably reduce the influence of downscaling, which is our primary focus in this work. Alternatively, using a large number of model outputs may achieve a more robust averaged estimate. Our inclusion of 55 downscaled model outputs across 22 GCMs may be sufficient in many cases, but this is difficult to verify within the current framework. As noted in L20, more sophisticated methods of extracting the forced response could also be used (e.g., ref. [77]).

Second, our main results neglect interactions among uncertainty sources, which previous studies have shown to be significant in some instances [78]. To estimate the importance of interaction effects, we implement an ANOVA-based variance decomposition (described in the Supporting Information) for all metrics across our three example cities. We find that interactions are small for projections of climate averages (Fig. S50) but can sometimes be important for extremes (Fig. S51). B23 note that accounting for the interaction between model and scenario uncertainty may alter their results, but we find this effect to be small—the interaction between model and downscaling uncertainty is typically larger. Future research could investigate interaction effects in more detail and across more locations.

Finally, we do not evaluate model outputs against historical observations and instead make an implicit assumption that the outputs from each scenario, GCM, and downscaling method represent equally plausible realizations of future climate. There is an increasing number of GCM weighting techniques [79, 80] that account for historical performance while guarding against overfitting, some of which can induce significant changes in CMIP6 projections [81]. Future work might investigate how the application of such techniques alters the variance decompositions. However, note that what constitutes an appropriate weighting of downscaled outputs remains an area of active research [82]. Additionally, given the presence of large observational disagreements at local scales, particular care should be given to evaluating historical performance if different downscaling algorithms are trained on conflicting observational datasets [83].

Data availability

The NEX-GDDP ensemble is available from the NASA Center for Climate Simulation:

`https://ds.nccs.nasa.gov/thredds/catalog/AMES/NEX/GDDP-CMIP6/catalog.html`.

The CIL-GDPCIR ensemble is available on Microsoft Planetary Computer:

`https://planetarycomputer.microsoft.com/dataset/group/cil-gdpcir`.

The ISIMIP3b ensemble is available from the ISIMIP repository:

`https://data.isimip.org/`.

Both carbonplan ensembles are hosted on Microsoft Azure. Example code showing how to access the data can be found at the following GitHub repository:

`https://github.com/carbonplan/cmip6-downscaling/`.

The GMFD observational dataset is available from the National Center for Atmospheric Research:

`https://rda.ucar.edu/datasets/ds314.0/`.

The ERA5 reanalysis product is available from the Copernicus Climate Change Service: `https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5`.

Code availability

Code to reproduce this analysis is available at the following GitHub repository: `https://github.com/david0811/lafferty-sriver_inprep_tbd`.

Acknowledgements

This work was supported by the U.S. Department of Energy, Office of Science, Biological and Environmental Research Program, Earth and Environmental Systems Modeling, MultiSector Dynamics under Cooperative Agreement DE-SC0022141. Computations for this research were performed on the Pennsylvania State University's Institute for Computational and Data Sciences' Roar supercomputer. Computations for this research were performed on Microsoft Planetary Computer [84]. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies. The authors are grateful to participants of the Climate Discussion Group at the University of Illinois for useful feedback.

Author Contributions

D.C.L.: Conceptualization, methodology, data curation, formal analysis, visualization, writing—original draft, writing—review & editing. R.L.S.: Conceptualization, methodology, data curation, writing—review & editing, supervision, funding acquisition.

Competing Interests

The authors declare no competing interests.

References

[1] Doblas-Reyes, F. *et al.* Linking global to regional climate change. In Masson-Delmotte, V. *et al.* (eds.) *Climate Change 2021: The Physical Science Basis. Contribution of Work-*

ing Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, book section 10 (Cambridge University Press, Cambridge, UK and New York, NY, USA, 2021). URL https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Chapter10.pdf.

[2] Lee, J.-Y. *et al.* Future global climate: Scenario-based projections and near-term information. In Masson-Delmotte, V. *et al.* (eds.) *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, book section 4 (Cambridge University Press, Cambridge, UK and New York, NY, USA, 2021). URL https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Chapter04.pdf.

[3] Hawkins, E. & Sutton, R. The Potential to Narrow Uncertainty in Regional Climate Predictions. *Bulletin of the American Meteorological Society* **90**, 1095–1107 (2009).

[4] Bonan, D. B., Lehner, F. & Holland, M. M. Partitioning uncertainty in projections of Arctic sea ice. *Environmental Research Letters* **16**, 044002 (2021).

[5] Hawkins, E. & Sutton, R. The potential to narrow uncertainty in projections of regional precipitation change. *Climate Dynamics* **37**, 407–418 (2011).

[6] Lehner, F. *et al.* Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. *Earth System Dynamics* **11**, 491–508 (2020).

[7] Blanusa, M. L., López-Zurita, C. J. & Rasp, S. Internal variability plays a dominant role in global climate projections of temperature and precipitation extremes. *Climate Dynamics* 1–15 (2023).

[8] Schwarzwald, K. & Lenssen, N. The importance of internal climate variability in climate impact projections. *Proceedings of the National Academy of Sciences* **119** (2022).

[9] Rössler, O. *et al.* Challenges to link climate change data provision and user needs: Perspective from the COST-action VALUE. *International Journal of Climatology* **39**, 3704–3716 (2019).

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

[11] Rode, A. *et al.* Estimating a social cost of carbon for global energy consumption. *Nature* **598**, 308–314 (2021).

[12] Martinich, J. & Crimmins, A. Climate damages and adaptation potential across diverse sectors of the United States. *Nature Climate Change* **9**, 397–404 (2019). URL `https://doi.org/10.1038/s41558-019-0444-6`.

[13] Cook, L. M., Anderson, C. J. & Samaras, C. Framework for Incorporating Downscaled Climate Output into Existing Engineering Methods: Application to Precipitation Frequency Curves. *Journal of Infrastructure Systems* **23**, 04017027 (2017).

[14] Fiedler, T. *et al.* Business risk and the emergence of climate analytics. *Nature Climate Change* 1–8 (2021).

[15] Hewitson, B. C., Daron, J., Crane, R. G., Zermoglio, M. F. & Jack, C. Interrogating empirical-statistical downscaling. *Climatic Change* **122**, 539–554 (2014).

[16] Maraun, D. Bias Correcting Climate Change Simulations - a Critical Review. *Current Climate Change Reports* **2**, 211–220 (2016).

[17] Chen, J., Brissette, F. P., Chaumont, D. & Braun, M. Performance and uncertainty evaluation of empirical downscaling methods in quantifying the climate change impacts on hydrology over two North American river basins. *Journal of Hydrology* **479**, 200–214 (2013). URL `https://www.sciencedirect.com/science/article/pii/S0022169412010414`.

[18] Gutmann, E. *et al.* An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resources Research* **50**, 7167–7186 (2014).

[19] Jiang, Y. *et al.* Inter-comparison of multiple statistically downscaled climate datasets for the Pacific Northwest, USA. *Scientific Data* **5**, 180016 (2018).

[20] Lopez-Cantu, T., Prein, A. F. & Samaras, C. Uncertainties in Future U.S. Extreme Precipitation From Downscaled Climate Projections. *Geophysical Research Letters* **47** (2020).

[21] Malek, K. *et al.* Bias Correction of Hydrologic Projections Strongly Impacts Inferred Climate Vulnerabilities in Institutionally Complex Water Systems. *Journal of Water Resources Planning and Management* **148**, 04021095 (2022).

[22] Tabari, H., Paz, S. M., Buekenhout, D. & Willems, P. Comparison of statistical downscaling methods for climate change impact analysis on precipitation-driven drought. *Hydrology and Earth System Sciences* **25**, 3493–3517 (2021). URL https://hess.copernicus.org/articles/25/3493/2021/.

[23] Alder, J. R. & Hostetler, S. W. The Dependence of Hydroclimate Projections in Snow-Dominated Regions of the Western United States on the Choice of Statistically Downscaled Climate Data. *Water Resources Research* **55**, 2279–2300 (2019).

[24] Bürger, G., Sobie, S. R., Cannon, A. J., Werner, A. T. & Murdock, T. Q. Downscaling Extremes: An Intercomparison of Multiple Methods for Future Climate. *Journal of Climate* **26**, 3429–3449 (2013).

[25] Chegwidden, O. S. *et al.* How Do Modeling Decisions Affect the Spread Among Hydrologic Climate Change Projections? Exploring a Large Ensemble of Simulations Across a Diversity of Hydroclimates. *Earth's Future* **7**, 623–637 (2019).

[26] Wang, H.-M., Chen, J., Xu, C.-Y., Zhang, J. & Chen, H. A Framework to Quantify the Uncertainty Contribution of GCMs Over Multiple Sources in Hydrological Impacts of Climate Change. *Earth's Future* **8**, e2020EF001602 (2020). URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020EF001602.

[27] Wootten, A., Terando, A., Reich, B., Boyles, R. & Semazzi, F. Characterizing Sources of Uncertainty from Global Climate Models and Downscaling Techniques. *Journal of Applied Meteorology and Climatology* **56**, 3245–3262 (2017).

[28] Wu, Y. *et al.* Quantifying the Uncertainty Sources of Future Climate Projections and Narrowing Uncertainties With Bias Correction Techniques. *Earth's Future* **10** (2022).

[29] 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). URL https://gmd.copernicus.org/articles/9/1937/2016/.

[30] IPCC. Summary for policymakers. In Pörtner, H. O. *et al.* (eds.) *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univer-

sity Press, Cambridge, UK and New York, NY, USA, 2022). URL https://www.ipcc.ch/report/ar6/wg2/downloads/report/IPCC_AR6_WGII_SummaryForPolicymakers.pdf.

[31] Lanzante, J. R., Dixon, K. W., Nath, M. J., Whitlock, C. E. & Adams-Smith, D. Some Pitfalls in Statistical Downscaling of Future Climate. *Bulletin of the American Meteorological Society* **99**, 791–803 (2017).

[32] Wootten, A. M., Dixon, K. W., Adams-Smith, D. J. & McPherson, R. A. Statistically downscaled precipitation sensitivity to gridded observation data and downscaling technique. *International Journal of Climatology* **41**, 980–1001 (2021).

[33] Rastogi, D., Kao, S.-C. & Ashfaq, M. How may the choice of downscaling techniques and meteorological reference observations affect future hydroclimate projections? *Earth's Future* **10**, e2022EF002734 (2022). URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022EF002734. E2022EF002734 2022EF002734.

[34] AghaKouchak, A. *et al.* Climate Extremes and Compound Hazards in a Warming World. *Annual Review of Earth and Planetary Sciences* **48**, 1–30 (2020).

[35] Howe, P. D., Marlon, J. R., Mildenberger, M. & Shield, B. S. How will climate change shape climate opinion? *Environmental Research Letters* **14**, 113001 (2019).

[36] McPhillips, L. E. *et al.* Defining Extreme Events: A Cross-Disciplinary Review. *Earth's Future* **6**, 441–455 (2018).

[37] Zscheischler, J. *et al.* A typology of compound weather and climate events. *Nature Reviews Earth & Environment* **1**, 333–347 (2020).

[38] Zelinka, M. D. *et al.* Causes of Higher Climate Sensitivity in CMIP6 Models. *Geophysical Research Letters* **47** (2020).

[39] Bartusek, S., Kornhuber, K. & Ting, M. 2021 North American heatwave amplified by climate change-driven nonlinear interactions. *Nature Climate Change* **12**, 1143–1150 (2022).

[40] Heeter, K. J. *et al.* Unprecedented 21st century heat across the Pacific Northwest of North America. *npj Climate and Atmospheric Science* **6**, 5 (2023).

[41] McKinnon, K. A. & Simpson, I. R. How Unexpected Was the 2021 Pacific Northwest Heatwave? *Geophysical Research Letters* **49** (2022).

[42] Thompson, V. *et al.* The 2021 western North America heat wave among the most extreme events ever recorded globally. *Science Advances* **8**, eabm6860 (2022).

[43] White, R. H. *et al.* The unprecedented Pacific Northwest heatwave of June 2021. *Nature Communications* **14**, 727 (2023).

[44] NOWData - NOAA Online Weather Data. URL `https://www.weather.gov/wrh/climate?wfo=sew`.

[45] Doss-Gollin, J. & Keller, K. A Subjective Bayesian Framework for Synthesizing Deep Uncertainties in Climate Risk Management. *Earth's Future* **11** (2023).

[46] Keller, K., Helgeson, C. & Srikrishnan, V. Climate Risk Management. *Annual Review of Earth and Planetary Sciences* **49**, 95–116 (2021).

[47] Maher, N., Milinski, S. & Ludwig, R. Large ensemble climate model simulations: introduction, overview, and future prospects for utilising multiple types of large ensemble. *Earth System Dynamics* **12**, 401–418 (2021).

[48] Gupta, R. S., Steinschneider, S. & Reed, P. M. Understanding contributions of paleo-informed natural variability and climate changes on hydroclimate extremes in the central valley region of california (2023). URL `https://doi.org/10.22541%2Fessoar.167870424.46495295%2Fv1`.

[49] Gutowski, W. J. *et al.* The ongoing need for high-resolution regional climate models: Process understanding and stakeholder information. *Bulletin of the American Meteorological Society* **101**, E664–E683 (2020).

[50] Lafferty, D. C. *et al.* Statistically bias-corrected and downscaled climate models underestimate the adverse effects of extreme heat on U.S. maize yields. *Communications Earth & Environment* **2**, 196 (2021).

[51] Pianosi, F. *et al.* Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software* **79**, 214–232 (2016).

[52] Hough, A. & Wong, T. E. Analysis of the evolution of parametric drivers of high-end sea-level hazards. *Advances in Statistical Climatology, Meteorology and Oceanography* **8**, 117–134 (2022).

[53] Razavi, S. *et al.* The Future of Sensitivity Analysis: An essential discipline for systems modeling and policy support. *Environmental Modelling & Software* **137**, 104954 (2021).

[54] Karimi, T., Reed, P., Malek, K. & Adam, J. Diagnostic Framework for Evaluating How Parametric Uncertainty Influences Agro-Hydrologic Model Projections of Crop Yields Under Climate Change. *Water Resources Research* **58** (2022).

[55] Mendoza, P. A. *et al.* Effects of Hydrologic Model Choice and Calibration on the Portrayal of Climate Change Impacts. *Journal of Hydrometeorology* **16**, 762–780 (2015).

[56] Müller, C. *et al.* Exploring uncertainties in global crop yield projections in a large ensemble of crop models and CMIP5 and CMIP6 climate scenarios. *Environmental Research Letters* **16**, 034040 (2021).

[57] Rising, J., Tedesco, M., Piontek, F. & Stainforth, D. A. The missing risks of climate change. *Nature* **610**, 643–651 (2022).

[58] Srikrishnan, V. *et al.* Uncertainty Analysis in Multi-Sector Systems: Considerations for Risk Analysis, Projection, and Planning for Complex Systems. *Earth's Future* **10** (2022).

[59] Lopez-Cantu, T., Webber, M. K. & Samaras, C. Incorporating uncertainty from downscaled rainfall projections into climate resilience planning in U.S. cities. *Environmental Research: Infrastructure and Sustainability* **2**, 045006 (2022).

[60] Kotamarthi, R. *et al.* *Downscaling Techniques for High-Resolution Climate Projections: From Global Change to Local Impacts* (Cambridge University Press, Cambridge, 2021).

[61] Maraun, D. *et al.* Towards process-informed bias correction of climate change simulations. *Nature Climate Change* **7**, 764–773 (2017).

[62] Schlenker, W. & Lobell, D. B. Robust negative impacts of climate change on african agriculture. *Environmental Research Letters* **5**, 014010 (2010). URL `https://dx.doi.org/10.1088/1748-9326/5/1/014010`.

[63] Sriver, R. L., Lempert, R. J., Wikman-Svahn, P. & Keller, K. Characterizing uncertain sea-level rise projections to support investment decisions. *PLOS ONE* **13**, 1–35 (2018). URL `https://doi.org/10.1371/journal.pone.0190641`.

[64] Thrasher, B. *et al.* NASA Global Daily Downscaled Projections, CMIP6. *Scientific Data* **9**, 262 (2022).

[65] Gergel, D. R. *et al.* Global downscaled projections for climate impacts research (GDPCIR): preserving extremes for modeling future climate impacts. *EGUsphere* **2023**, 1–35 (2023). URL `https://egusphere.copernicus.org/preprints/2023/egusphere-2022-1513/`.

[66] Lange, S. Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0). *Geoscientific Model Development* **12**, 3055–3070 (2019). URL `https://gmd.copernicus.org/articles/12/3055/2019/`.

[67] Lange, S. ISIMIP3BASD (2022). URL `https://doi.org/10.5281/zenodo.7151476`.

[68] Chegwidden, O. *et al.* Open data and tools for multiple methods of global climate downscaling (2022). URL `https://carbonplan.org/research/cmip6-downscaling-explainer`.

[69] Gutmann, E. D. *et al.* En-GARD: A Statistical Downscaling Framework to Produce and Test Large Ensembles of Climate Projections. *Journal of Hydrometeorology* **23**, 1545–1561 (2022).

[70] Matwin, S. *et al.* DeepSD. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 1663–1672 (2017).

[71] 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). URL `https://www.sciencedirect.com/science/article/pii/S0959378016300681`.

[72] 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). URL `https://doi.org/10.1175/JCLI3790.1`.

[73] Hersbach, H. *et al.* The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* **146**, 1999–2049 (2020). URL `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803`.

[74] Cucchi, M. *et al.* WFDE5: bias-adjusted ERA5 reanalysis data for impact studies. *Earth System Science Data* **12**, 2097–2120 (2020). URL `https://essd.copernicus.org/articles/12/2097/2020/`.

[75] Lange, S. *et al.* WFDE5 over land merged with ERA5 over the ocean (W5E5 v2.0) (2021). URL https://doi.org/10.48364/ISIMIP.342217.

[76] Brekke, L. & Barsugli, J. Uncertainties in Projections of Future Changes in Extremes. In AghaKouchak, A., Easterling, D., Hsu, K., Schubert, S. & Sorooshian, S. (eds.) *Extremes in a Changing Climate*, vol. 65, 309–346 (Springer, Dordrecht, 2013). URL https://doi.org/10.1007/978-94-007-4479-0_11.

[77] Sippel, S. *et al.* Uncovering the forced climate response from a single ensemble member using statistical learning. *Journal of Climate* **32**, 5677 – 5699 (2019). URL https://journals.ametsoc.org/view/journals/clim/32/17/jcli-d-18-0882.1.xml.

[78] Yip, S., Ferro, C. A. T., Stephenson, D. B. & Hawkins, E. A Simple, Coherent Framework for Partitioning Uncertainty in Climate Predictions. *Journal of Climate* **24**, 4634–4643 (2011).

[79] Brunner, L. *et al.* Comparing Methods to Constrain Future European Climate Projections Using a Consistent Framework. *Journal of Climate* **33**, 8671–8692 (2020).

[80] Qasmi, S. & Ribes, A. Reducing uncertainty in local temperature projections. *Science Advances* **8** (2022).

[81] Brunner, L. *et al.* Reduced global warming from CMIP6 projections when weighting models by performance and independence. *Earth System Dynamics* **11**, 995–1012 (2020).

[82] Wootten, A. M., Massoud, E. C., Sengupta, A., Waliser, D. E. & Lee, H. The Effect of Statistical Downscaling on the Weighting of Multi-Model Ensembles of Precipitation. *Climate* **8**, 138 (2020).

[83] Lorenz, R. *et al.* Prospects and Caveats of Weighting Climate Models for Summer Maximum Temperature Projections Over North America. *Journal of Geophysical Research: Atmospheres* **123**, 4509–4526 (2018).

[84] Microsoft Open Source, McFarland, M., Emanuele, R., Morris, D. & Augspurger, T. microsoft/PlanetaryComputer: October 2022 (2022). URL https://doi.org/10.5281/zenodo.7261897.