https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

Earth System Dynamics Discussions

Uncertainty-informed selection of CMIP6 Earth System Model subsets for use in multisectoral and impact models

Abigail Snyder[1], Noah Prime[1], Claudia Tebaldi[1], Kalyn Dorheim[1]

[1]Pacific Northwest National Laboratory, Joint Global Change Research Institute, College Park MD, 20740, USA

Correspondence to: Abigail Snyder (abigail.snyder@pnnl.gov)

Abstract. Earth System Models (ESMs) are heavily used to provide inputs to impact and multisectoral dynamic models. Therefore, representing the full range of model uncertainty, scenario uncertainty, and interannual variability that ensembles of ESMs capture, is critical to the exploration of the future co-evolution of the integrated human-Earth system. The pre-eminent source of these ensembles has been the Coupled Model Intercomparison Project (CMIP). With more modeling centers participating in each new CMIP phase, the size of the ESM archive is rapidly increasing, which can be intractable for impact modelers to effectively utilize due to computational constraints and the challenges of analyzing large datasets. In this work, we present a method to select a subset of the latest phase, CMIP6, models for use as inputs to a sectoral impact or multisectoral models, while still representing the range of model uncertainty, scenario uncertainty, and interannual variability of the full CMIP6 ESM results. This method is intended to help human-relevant impact and multisectoral modelers select climate information from the CMIP archive efficiently. This is particularly critical for large ensemble experiments of multisectoral dynamic models that may be varying additional features beyond climate inputs in a factorial design, thus putting constraints on the number of climate simulations that can be used. We focus on temperature and precipitation outputs of ESMs, as these are two of the most used variables among impact models and many other key input variables for impacts are at least correlated with one or both of temperature and precipitation (e.g. relative humidity). Besides preserving the multi-model ensemble variance characteristics, we prioritize selecting ESMs in the subset that preserve the very likely distribution of equilibrium climate sensitivity values as assessed by the latest IPCC report. This approach could be applied to other output variables of ESMs and, when combined with emulators, offers a flexible framework for designing more efficient experiments on human-relevant climate impacts. It can also provide greater insight into the properties of existing ESMs and the method may be informative for future experiment planning across ESMs.

1 Introduction

The future evolution of the integrated human-Earth system is highly uncertain, but one common approach to begin addressing this uncertainty is to use outputs from a variety of computationally expensive, highly detailed process-based Earth System Models (ESMs) run under different scenarios. This approach has been facilitated by the Couple Model Intercomparison Project (CMIP) (Eyring et al. 2016), which has organized experiments that are standardized across modeling centers. Scenario simulations from CMIP (most recently through ScenarioMIP, (O'Neill et al. 2016) are commonly used as inputs to downstream sectoral impact and multisector dynamic models, both by individual modeling efforts and by large, coordinated impact modeling projects, like AgMIP or ISIMIP (e.g. (Rosenzweig et al. 2013; Rosenzweig et al. 2014; Warszawski et al. 2014; Frieler et al. 2017)). Using such multi-model

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

Earth System Dynamics Discussions

ensembles captures the process and structural uncertainties represented by sampling across ESMs, scenario uncertainty, and, to the extent that an ESM runs multiple initial condition ensemble members for a scenario, internal variability of the individual ESM (Hawkins and Sutton 2009; Hawkins and Sutton 2011; Lehner et al. 2020). These Earth system uncertainties can then be propagated through an impact model (perhaps after bias-correction (Lange 2019)) to understand possible human-relevant outcomes.

From the Earth system modelers who produce climate data to the impact and multisectoral dynamic modelers who use it, each step in this process is computationally expensive. For Earth system modelers, variability across ESMs' projections of future climate variables can be significant (Hawkins and Sutton 2009; Hawkins and Sutton 2011; Lehner et al. 2020) and so the participation of multiple modeling centers running multiple scenarios is critical to understanding the future of the Earth system. Further, statistical evaluation (Tebaldi et al. 2021) suggests that 20-25 initial condition ensemble members for each scenario an ESM provides are needed to estimate the forced component of extreme metrics related to daily temperature and precipitation, which are key inputs to many impacts models covering hydrological, agricultural, energy and other sectors. Fortunately, emulation of ESM outputs to infill missing scenarios and enrich initial condition ensembles continues to improve (Beusch, Gudmundsson, and Seneviratne 2020; Nath et al. 2022; Quilcaille et al. 2022; Tebaldi, Snyder, and Dorheim 2022). This suggests that ESMs don't necessarily have to provide all of the runs desired for capturing possible futures, but instead a subset of scenarios including initial condition ensembles for emulator training. However the total burden across modeling centers to sample across ESMs and scenarios still remains high, even with this potential efficiency. Impact modelers often seek to understand future climate impacts in the context of ESM uncertainty by using the outputs of multiple ESMs under multiple scenarios as inputs to impact models (e.g. (Prudhomme et al. 2014; Müller et al. 2021)). In a world unburdened by time and computing constraints, an impact model would take as input every projected data set available to have a full understanding of possible outcomes. Our world includes those burdens, and when models require bias-corrected input climate data, the computational expense for impact modelers and the broader community only grows. This can quickly become an intractably-sized set of runs to perform and analyze for impact modelers. For the multisectoral dynamics community, whose modelers often attempt to integrate results from multiple impact models to understand interactions of different sectors of the integrated human-Earth system (Graham et al. 2020) this challenge multiplies. Finally, multisectoral dynamic models are beginning to run large ensemble experiments that vary additional features beyond climate inputs in a factorial design (e.g. (Dolan et al. 2021, 2022; Guivarch et al. 2022)) further adding to the computational costs to be faced.

For all communities involved, an efficient way to design and then use ESM runs is critical. While there is likely no perfect solution to balance the tension between these competing priorities, this work describes a method for selecting a subset of CMIP6 models that still faithfully represents the uncertainty characteristics of the entire data set, particularly in dimensions relevant to impact and multisectoral modelers. The subset of ESMs outlined here is merely one approach to make understanding the future of the human-Earth system more tractable. The calculations described in this paper may also serve as a useful characterization of ESM behavior for modelers in other contexts.

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

Earth System Dynamics Discussions

Finally, many of the choices made in implementing this method may be adaptable for other uses or priorities. We also briefly discuss ways that this work can be leveraged by Earth system modelers in future comparison exercises to more efficiently identify specific ESMs to focus on larger initial condition ensembles.

2 Methods

The overall steps of this method are summarized in Table 2. Sections 2.1 and 2.2 provide fuller details building up the method and justifying choices. Table 2 especially highlights the choices made for this particular effort, based on the authors' experience with multisectoral impact modeling. It is likely that the approach could be adapted with different regions of interest, indices of behavior, or ESM output variables, although validation of results in these cases would be necessary.

2.1 Data preparation and characterization

Impact models often require multiple output variables from an ESM on daily or monthly time scales, with temperature and precipitation being the most common variables needed. For tractability, we focus on the IPCC WG1 non-arctic land regions (Iturbide et al. 2022), as these regions are primarily where humans live, consume water, generate electricity, and grow food. I.e., the places most relevant in multisectoral models of the integrated human-Earth system. We also limit ourselves to ESMs that completed all four ScenarioMIP Tier 1 experiments (Table 1). This still results in more than 600 trajectories across models, scenarios, and ensemble members for each region. In this work, we are treating this collection of ESMs and scenario results in these regions as the full set of data of which we would like to faithfully represent the uncertainty characteristics, and then select a subset of ESMs for impact modelers to use, based on preserving those characteristics.

Table 1. Models and scenarios making up the full set of data, as well as their equilibrium climate sensitivity (ECS) values sourced from *(Meehl et al. 2020; Lovato et al. 2022; Scafetta 2022)*.

ESM	ECS	SSP126 Ensemble size	SSP245 Ensemble size	SSP370 Ensemble size	SSP585 Ensemble size
ACCESS-CM2	4.7	5	5	5	5
ACCESS-ESM1-5	3.9	40	10	30	40
BCC-CSM2-MR	3.0	1	1	1	1

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

Earth System Dynamics Discussions
Open Access
EGU

CAMS-CSM1-0	2.3	2	2	2	2
CESM2	5.2	3	3	3	3
CESM2-WACCM	4.8	1	3	1	3
CMCC-CM2-SR5	3.52	1	1	1	1
CMCC-ESM2	3.57	1	1	1	1
CanESM5	5.6	25	25	25	25
EC-Earth3-Veg-LR		3	3	3	3
FGOALS-f3-L	3.0	1	1	1	1
FGOALS-g3		4	4	4	4
GFDL-ESM4	2.6	1	3	1	1
INM-CM4-8	1.8	1	1	1	1
INM-CM5-0	1.9	1	1	5	1
IPSL-CM6A-LR	4.6	6	11	11	6
MIROC6	2.6	50	33	3	50
MPI-ESM1-2-HR	3.0	2	2	10	2
MPI-ESM1-2-LR	3.0	10	10	10	10
MRI-ESM2-0	3.2	5	5	5	5
NorESM2-MM	2.5	1	2	1	1
UKESM1-0-LL	5.3	13	14	13	5

105 For each scenario and region in each ESM, we extract the ensemble-average temperature and precipitation outputs: mid-century (2040-2059) average anomaly relative to that ESM's historical

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

Earth System Dynamics Discussions Open Access EGU

average (1995-2014), the end of century (2080-2099) anomaly relative to historical average, and the interannual standard deviation. These six indices (three for each of temperature and precipitation) per ESM-scenario-region combination are selected to result in data that represents the model uncertainty, scenario uncertainty, and interannual variability of our full set of data. The key question then, is how to efficiently characterize this collection of data in a way that suggests an efficient subsampling of models that still preserves the main dimensions of variations of the full ensemble.

This full data can be written as a matrix $\boldsymbol{A}$ with *22 ESM * 4 Scenarios =88* rows and *6 indices * 43 IPCC regions = 258* columns. Principal components analysis (PCA) is then a natural technique to understand the variance of this full data set by forming the covariance matrix $\boldsymbol{S} = \boldsymbol{A}^T\boldsymbol{A}$. The eigenvectors of $\boldsymbol{S}$ are a set of orthogonal basis vectors (each vector is length 258) that are ordered by how much variance of the full data each eigenvector explains. Mathematically, this means that each row of $\boldsymbol{A}$, ($\{\overrightarrow{a_i} | i = 1 \dots 88\}$) representing the indices in all regions for a single ESM-scenario) can be projected into the space of eigenvectors $\{\overrightarrow{PC_i} | i = 1 \dots 88\}$ and written as $\overrightarrow{a_i} = \Sigma_j c_{ij} \overrightarrow{PC_j}$ for projection coefficients (coordinates in the basis of eigenvectors), c_{ij}. Thus $\overrightarrow{PC_1}$, for example represents some pattern of joint, spatiotemporal temperature and precipitation behaviors that explains the greatest variance across ESM-scenario observations. Each CMIP6 model-scenario combination has some contribution from this pattern described by its projection coefficient, c_{i1}. This projection can be done over all eigenvectors, or as is common with PCA, a small subset of the eigenvectors that explain the majority of variance. Figure 1 is a plot of the fraction variance explained by each of the first 15 eigenvectors. Based on this figure, we restrict ourselves to the first five eigenvectors for projections, explaining 71.8% of variance.

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

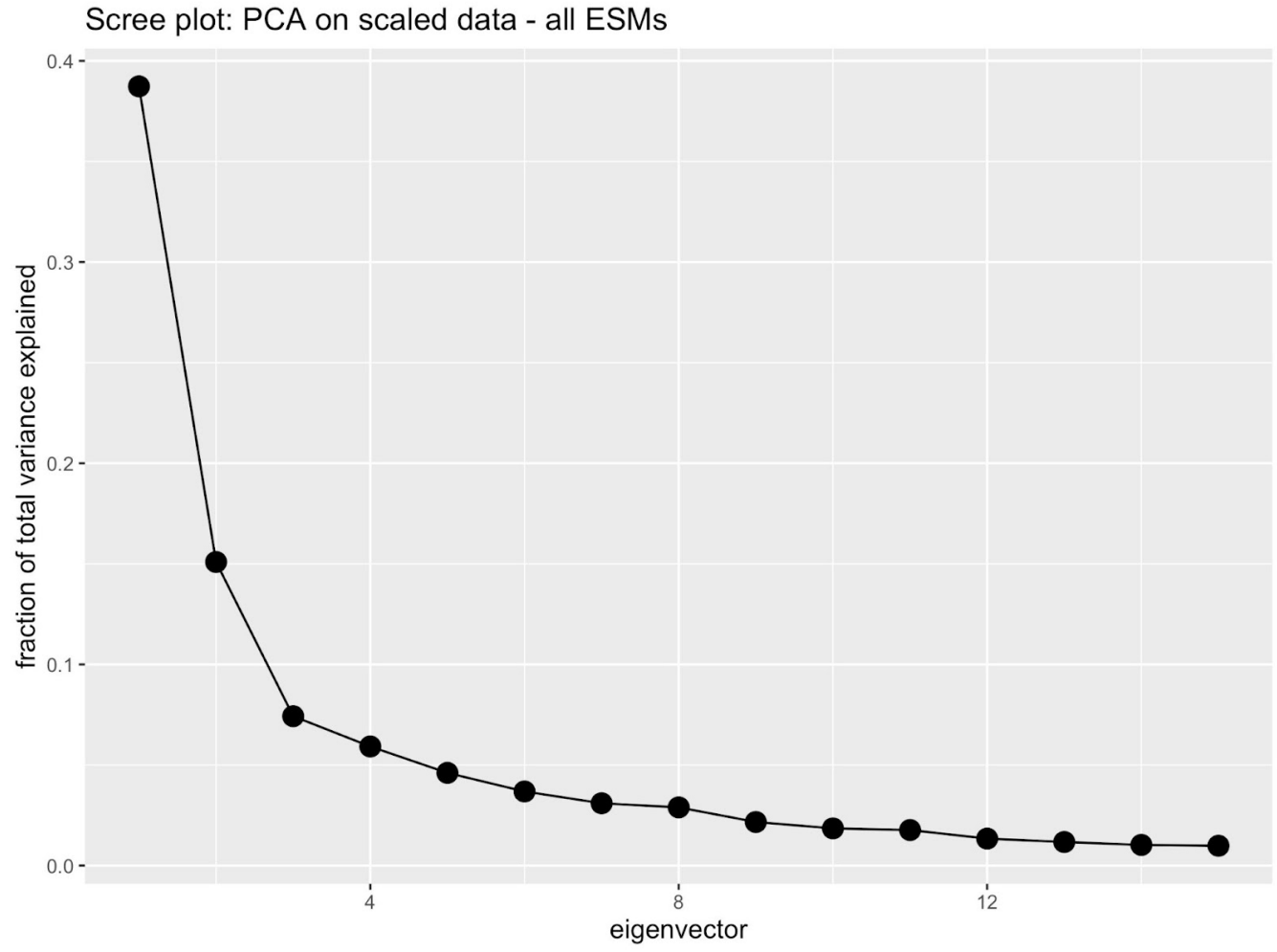


Figure 1: fraction of variance explained by each eigenvector for the full set of data, for the first 15 eigenvectors.

Figure 2 is a visual representation of these five eigenvectors. Each row is a map of all six indices for each vector. $\overrightarrow{PC_1}$ is dominated by temperature and, to a lesser extent, high latitude precipitation, highlighting that these features are responsible for 38.7% of the total variance of our full set of data (from Fig. 1). $\overrightarrow{PC_2}$ is dominated by temperature interannual variability and high latitude precipitation interannual variability. $\overrightarrow{PC_3}$ to $\overrightarrow{PC_5}$ feature a mix of the indices, with strong emphasis on precipitation related behaviors. Note that because we treated temperature and precipitation indices together in one matrix, the eigenvectors include joint temperature-precipitation behaviors that may be missed if the variables were treated separately.

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

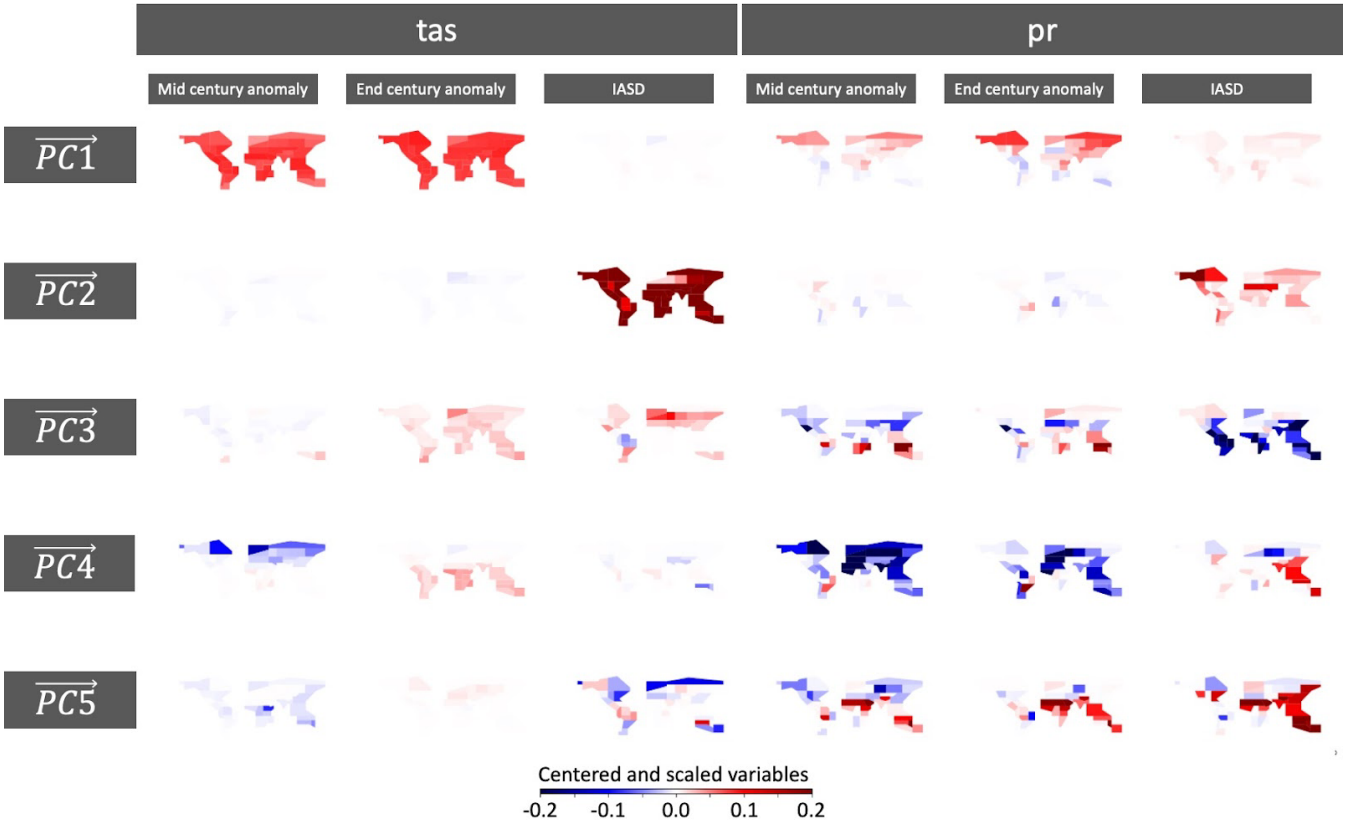


Figure 2: Maps of the first five eigenvectors of our full data. Each row is a single eigenvector, with maps presented for each of the six indices. Note that the colorbar scales are all standardized. A larger, landscape-oriented version of this figure is included in Appendix A (Fig. A1) for easier inspection.

145

By treating the span of these five eigenvectors as the representative space of full data, we can project all data into this space and visualize its behavior by two-dimensional plots of all five PCs combinations. Figure 3 shows these 2-d slices of the projection coefficients for each ESM and scenario into this space. If an impact modeler wished, they could run every model-scenario combination here for all available
150 ensemble members. In practice, however, this may not be computationally tractable to either run or analyze. This view also motivates our approach for selecting our subset of Earth System Models that preserve the uncertainty characteristics defined by this space.

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

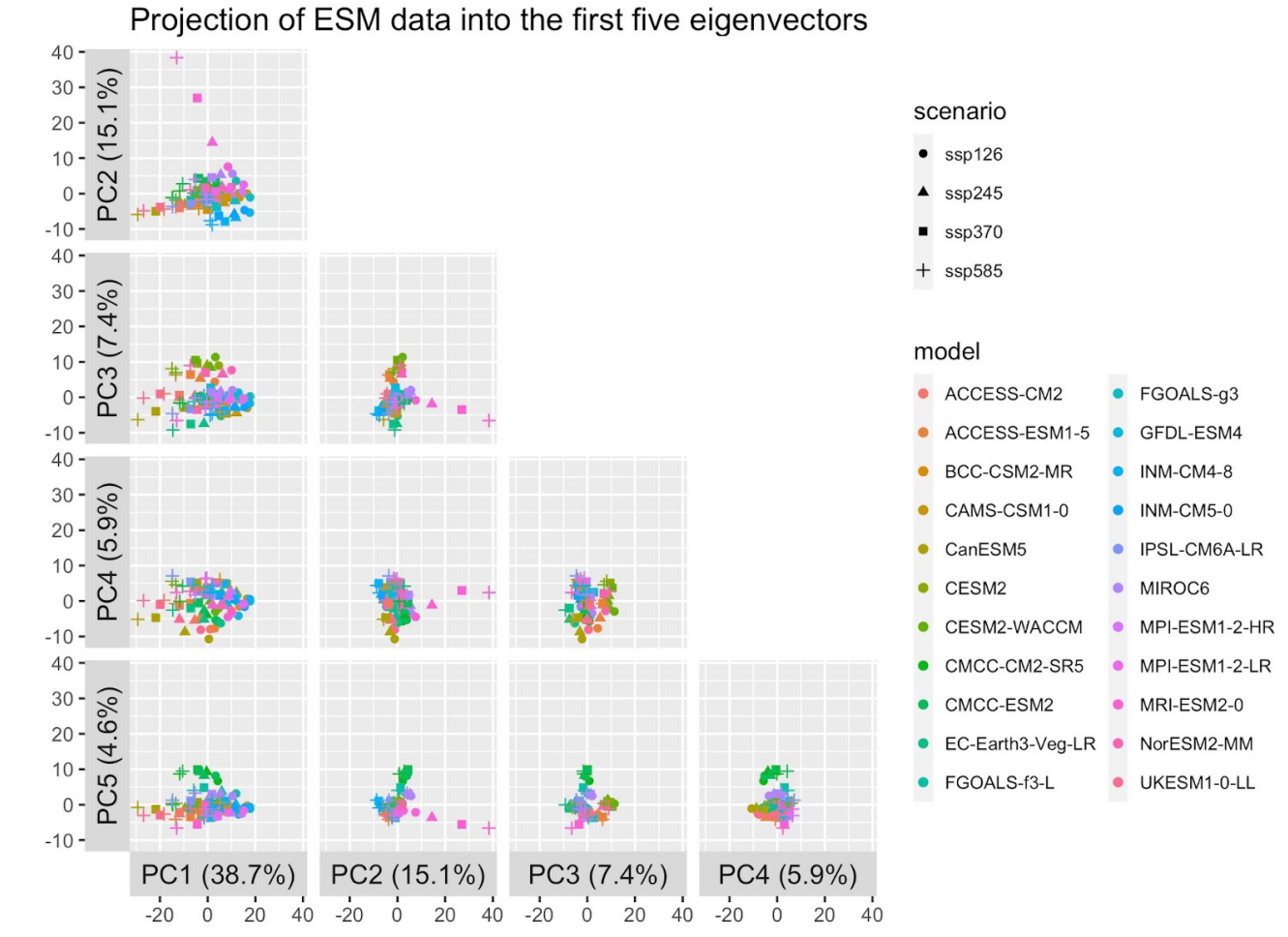


Figure 3: 2-d slices of the projection coefficients for each ESM-scenario combination into the space spanned by the first five eigenvectors.

2.2 Selection criteria of subset of ESMs

The subset of ESMs that minimizes distance to all other ESMs across this five-dimension space is the subset selected. In more detail, first, subsets of candidate ESMs are formed (in this work, five ESMs per subset, but the approach can be applied to any number of ESMs that we target as our subset size). While it would be possible to consider any combination of 5 ESMs from the full set of 22, in this work we add a pre-filtering step. From all 22 choose 5 potential subsets, we only consider as candidate subsets of ESMs those that roughly preserve the IPCC distribution of equilibrium climate sensitivity values (Core Writing Team & (eds.), 2023; Lovato et al., 2022; Meehl et al., 2020; Scafetta, 2022). Then for each subset, we step through each non-candidate ESM and calculate the minimum Euclidean distance to any

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

Earth System Dynamics Discussions

candidate ESM's coefficients. The summary metric for each subset of candidates is then the average over all non-candidate ESM minimum distances, and the subset of candidate ESMs with the smallest summary metric is the selected subset. Unlike many metrics (e.g. (Nash & Sutcliffe, 1970; Tebaldi et al., 2020)), there is unfortunately not a clear threshold for 'good enough' performance based on this metric and so in the results section, we provide further validation that the selected subset of ESMs for this setup is successful at preserving many of the major characteristics of the full data's uncertainty characterizations.

Table 2. Summary of method

Step	Description	Additional notes for this work
1	Identify relevant ESM output variables	Temperature, precipitation
2	Aggregate gridded time series to region-levels	IPCC WG1 non-arctic land
3	Identify and extract region indices for each variable, for each ESM-Scenario to capture characteristics of uncertainty of interest	Ensemble averaged: Mid-century anomaly, end of century anomaly, interannual standard deviation
4	Form a matrix of ESM*Scenario rows and Region*Indices columns for the full data and perform PCA; identify number of eigenvectors, *N*, responsible for majority of variance	*N=5* eigenvectors
5	Create candidate subsets of ESMs based on heuristic filters of interest	ESM subset size = 5; heuristic filter is that each subset must preserve the IPCC distribution of equilibrium climate sensitivity.
6	Calculate the summary metric for each subset and select the subset with the smallest value	Minimize distance from out-of-subset ESMs to a subset ESM across the *N=5* dimensions.

3 Results and discussion

The selected subset of ESMs and their respective ECS values are provided in Table 3 and Fig. 4, which is identical to Fig. 3 but with the selected ESMs highlighted by black box outlines. We provide more quantitative validation in Section 3.1 based on the work of Hawkins and Sutton (Hawkins & Sutton, 2009, 2011).

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

Earth System Dynamics Discussions
Open Access
EGU

Table 3. Selected ESM subset and ECS values

ESM	ECS
ACCESS-CM2	*4.7*
ACCESS-ESM1-5	*3.9*
BCC-CSM2-MR	*3.0*
MIROC6	*2.6*
MRI-ESM2-0	*3.2*

185

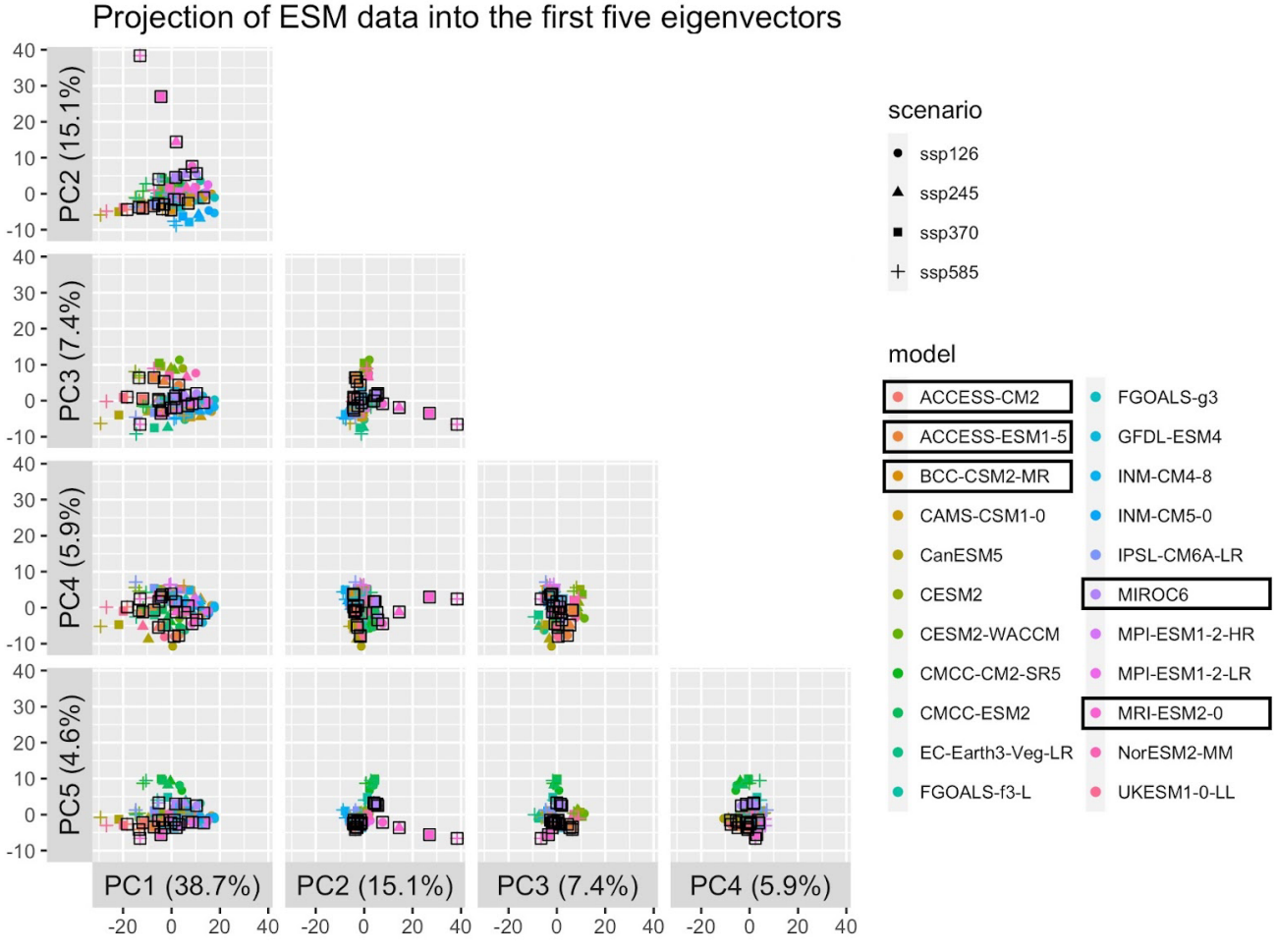


Figure 4: Same as Figure 3 but with the selected ESMs highlighted by black box outlines.

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

Earth System Dynamics Discussions

3.1 Validation

As noted in the introduction, the partitioning of uncertainty calculated by Hawkins and Sutton (Hawkins & Sutton, 2009, 2011) was critical in motivating this work and selecting regional indices of ESM behavior that address the different partitions for both temperature and precipitation. As we did not calculate the specific time series of Hawkins and Sutton (HS) fractions for internal variability (there, as here, quantified as interannual variability after detrending the annual mean time series), scenario uncertainty, and model uncertainty to form any part of our selection procedure, we can use these HS fractions as independent validation criteria. We calculate the time series of HS fractions for temperature and precipitation separately in each region, for the full set of data and over just our selected subset of data, i.e., over the full CMIP6 ensemble we started from, and only using the subset of 5 ESMs that our method identified. Details of these calculations are provided in Appendix B. To manage the inspection of three time series for each of 86 region-variable combinations, we use root mean square error (RMSE) to compare the full data time series and the subset data time series from 2040 onward (as that is the focus of our indices) for each uncertainty partition, for each variable in each region.

To identify specific region-variable combinations that are due for closer inspection, we set a threshold on the RMSE values for each uncertainty partition for each region-variable combination. If any of the three uncertainty partitions have RMSE>0.1, we flag that region-variable combination for closer inspection. While thresholds like this are often arbitrary to set, each uncertainty partition for the subset data explaining the fraction of total variance within 10% of the full data's partition seems a good place to start. We show in Appendix A the results of a less stringent choice. Lowering this threshold will of course flag more region-variables combinations, but as we point out below, a portion of the combinations flagged with a threshold of 0.1 still actually perform reasonably when plotted over time. Figure 5 provides a color-coded map of regions where temperature, precipitation, both, or neither have RMSE <= 0.1 for all three uncertainty partitions to give a sense of the spatial extent of performance. 54 of the 86 total region-variable combinations perform well based on this criterion alone. Note that many of these through Europe and East Asia are significant agricultural producers, regions where impacts often have critical implications for other regions and sectors in an integrated, multisectoral system.

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

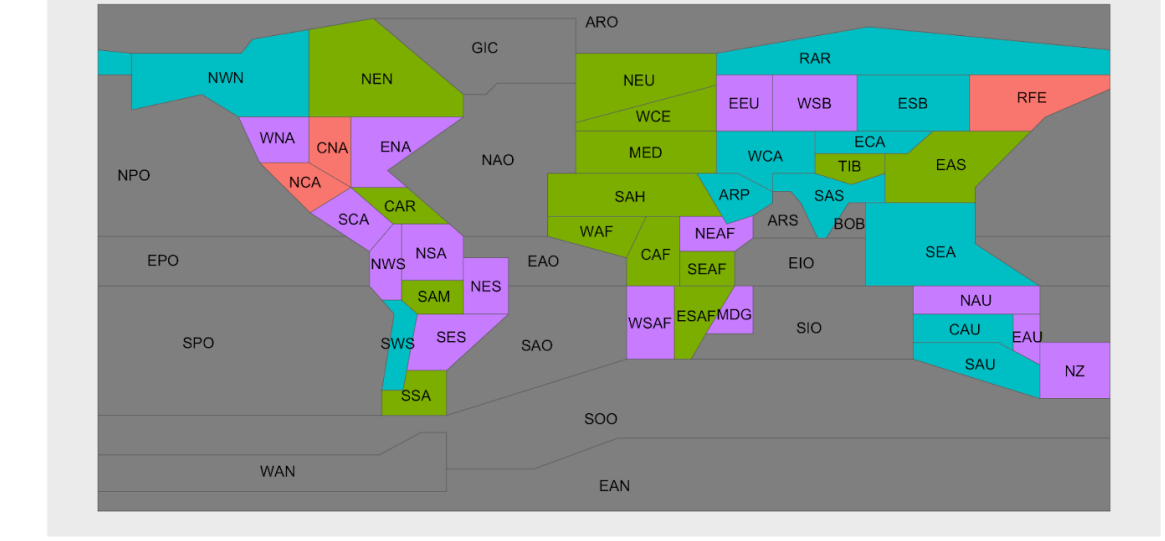


Figure 5. a color-coded map of regions where temperature, precipitation, both, or neither have RMSE <= 0.1 for all three uncertainty partitions.

The time series of HS fractions for the remaining 32 region-variable combinations for which RMSE > 0.1 are plotted in Figure 6 (temperature in 14 regions) and Figure 7 (precipitation in 18 regions bottom panel). For temperature, we see that interannual variability is often performing well, with increasingly better performance over time? over time. The partitioning of model and scenario uncertainty is where the subset of ESMs begins to depart from the full data, although this too tends to have smaller discrepancies as time goes on. This is not surprising: in the full set of data, a good portion of model uncertainty is driven by different ECS values. As provided in Table 1, the values across ESMs that participated in tier 1 ScenarioMIP experiments do not match the IPCC very likely distribution. By contrast, we are only selecting subsets of ESMs that match this distribution, accounting for the departure. What we want to see, rather, is a qualitative agreement with the relevance of the three sources of uncertainty in the full ensemble. According to this criterion, most of the regions flagged by the application of the 0.1 threshold remain consistent with the full ensemble representation of the three uncertainty sources. Most frequently, regions feature a larger portion from interannual variability and a lesser portion from model uncertainty, again not surprisingly given our choice to reshape the distribution of ECS. For many of these regions (CNA, EAU, EEU, MDG, NAU, NCA, NSA, NWS, SCA, SES, WNA, WSB) the behavior of our subset approximates the full ensemble increasingly better as we move towards the later portion of the century. In Appendix A, we show the effect of relaxing the threshold to 0.2, resulting in nearly the full set of region/variable combinations passing our test (Fig. A2-A4).

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

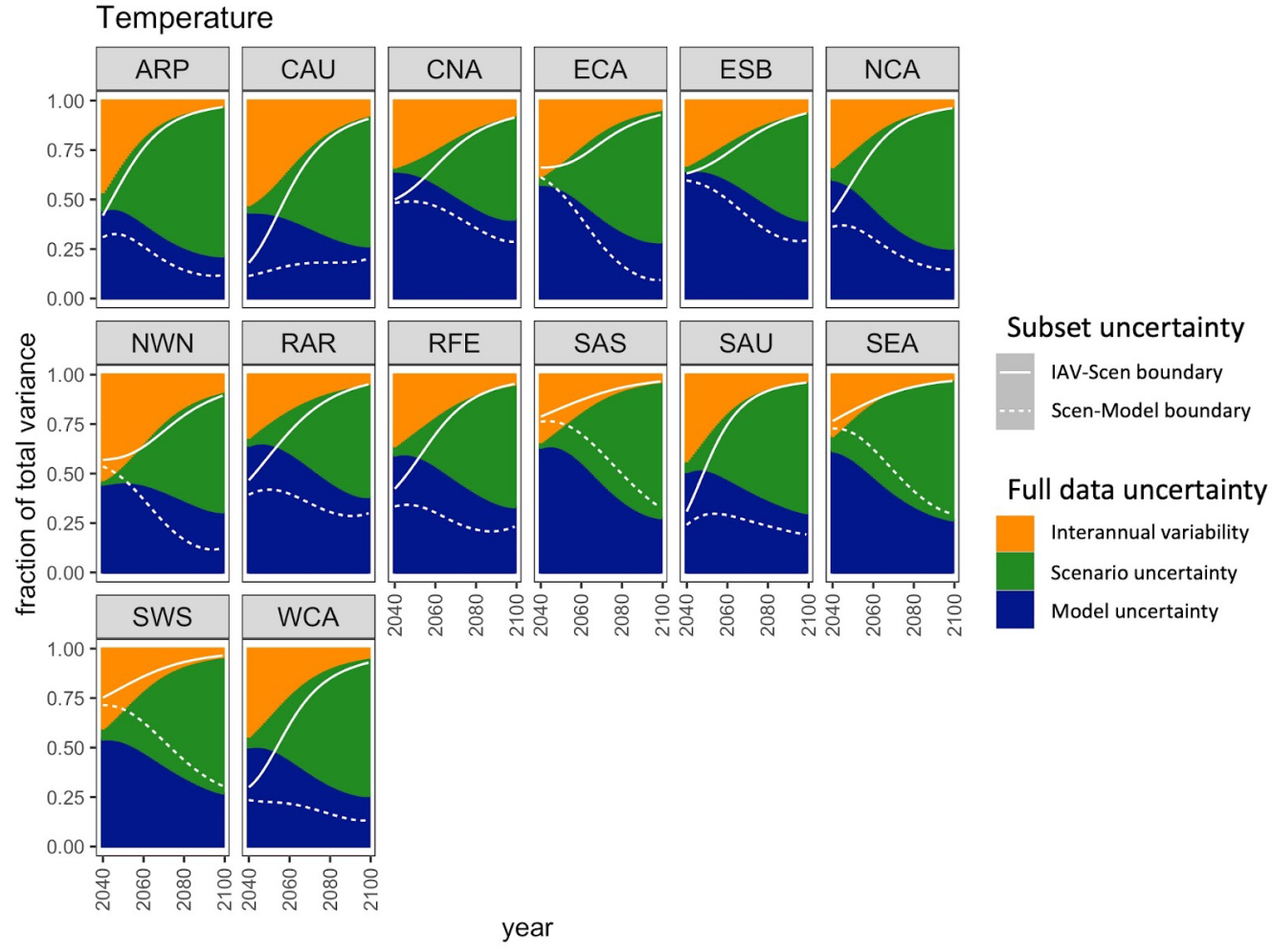


Figure 6: Regions flagged for closer inspection of their HS fraction time series for temperature. The color-blocked time series are the HS fractions from the full set of data, and the white curves overlaid are the respective boundaries for the subset data's uncertainty partitions.

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

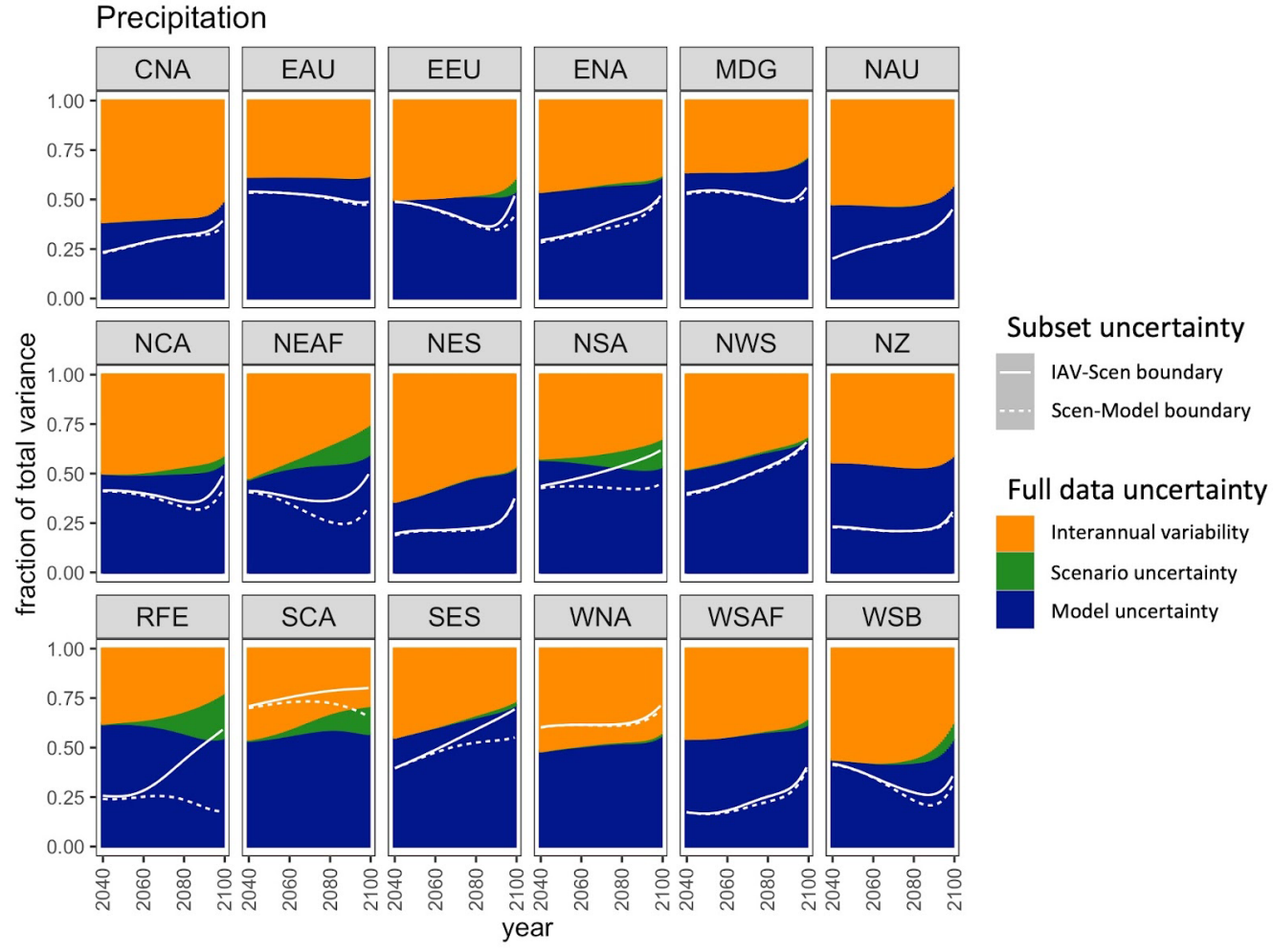


Figure 7: Same as Figure 6 but for precipitation.

4 Conclusions

This work outlines and documents the success of a method for selecting a subset of ESMs from CMIP6 that overall preserve the uncertainty characteristics of the full set of CMIP data. Our methodology relies on pre-identifying regional indices of behavior for ESM output variables, as well as other filters (such as prioritizing preserving the IPCC distribution of ECS values) that are key to impact and multisectoral models. With these assumptions, far fewer climate inputs are needed to span the range of uncertainties seen in CMIP6, resulting in fewer impact model runs needing to be performed and analyzed. There are likely many situations in which a modeler could adapt the details of the method (outlined in Table 2) and code for their purposes, re-run to identify a subset of ESMs, and validate that new subset in much less time and with much fewer computing resources needed than simply running impact models with all

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

Earth System Dynamics Discussions
Open Access
EGU

scenarios and ensemble members available for the 22 ESMs documented in Table 1. For multisectoral modelers integrating multiple different impacts, or running large ensemble experiments, the time saved only grows, even when accounting for method adjustment and re-validation of results.
For Earth system modelers, there may be opportunities to identify fewer ESMs that would benefit from running more initial condition ensemble members. If all modeling centers performed a small number of ensemble members for a set of scenarios, this analysis could be repeated to identify specific models to run more ensemble members for. This can result in more efficient allocation of total computing resources in a model intercomparison exercise.

The methodology outlined in this paper is an adaptable approach to both retain the major uncertainty characteristics of a large collection of ESM data and to make changes (as we did to the ECS distribution). While there are resulting regions for both temperature and precipitation where the uncertainty partitions of the subset of ESMs differ from the full set of ESMs, these differences are expected based on the different ECS distribution represented by our subset ESMs compared to the full data. We hope that by providing detailed information about where the subset differs in Figures 5-7, impact modelers may be able to infer how results would change if the full set of data were used, with far lower computational burden than running all available data. Further, because the method is adaptable, an impact modeler particularly interested in a specific region could weight the outcomes in that region more heavily for selection of the subset.

Open Access
Earth System Dynamics Discussions

Appendix A Additional figures

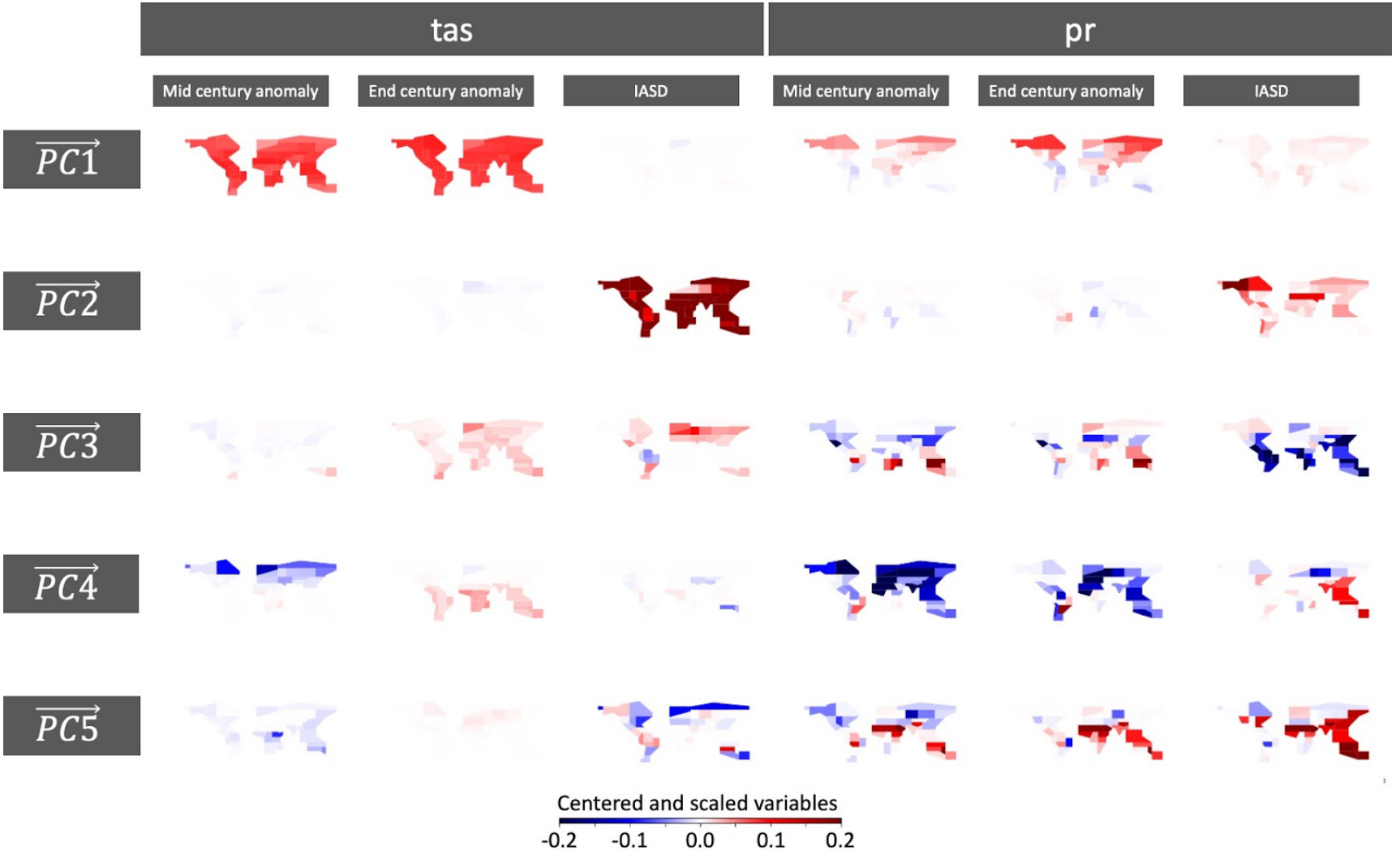
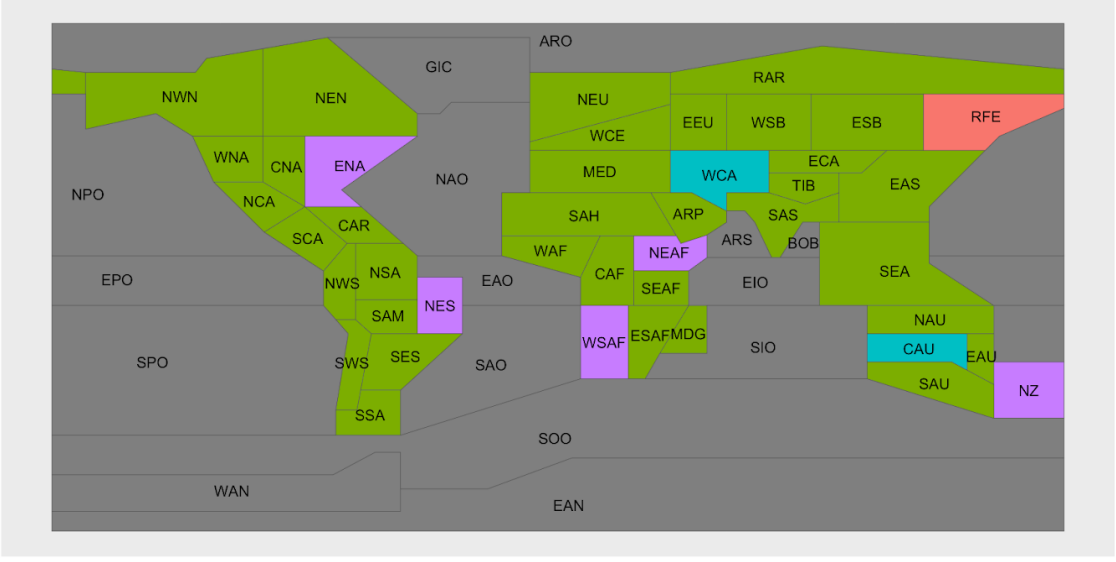


Figure A1: landscape-oriented version of Figure 2 for easier visual inspection.

285

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.



290 **Figure A2: A color-coded map of regions where temperature, precipitation, both, or neither have RMSE <= 0.2 for all three uncertainty partitions.**

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

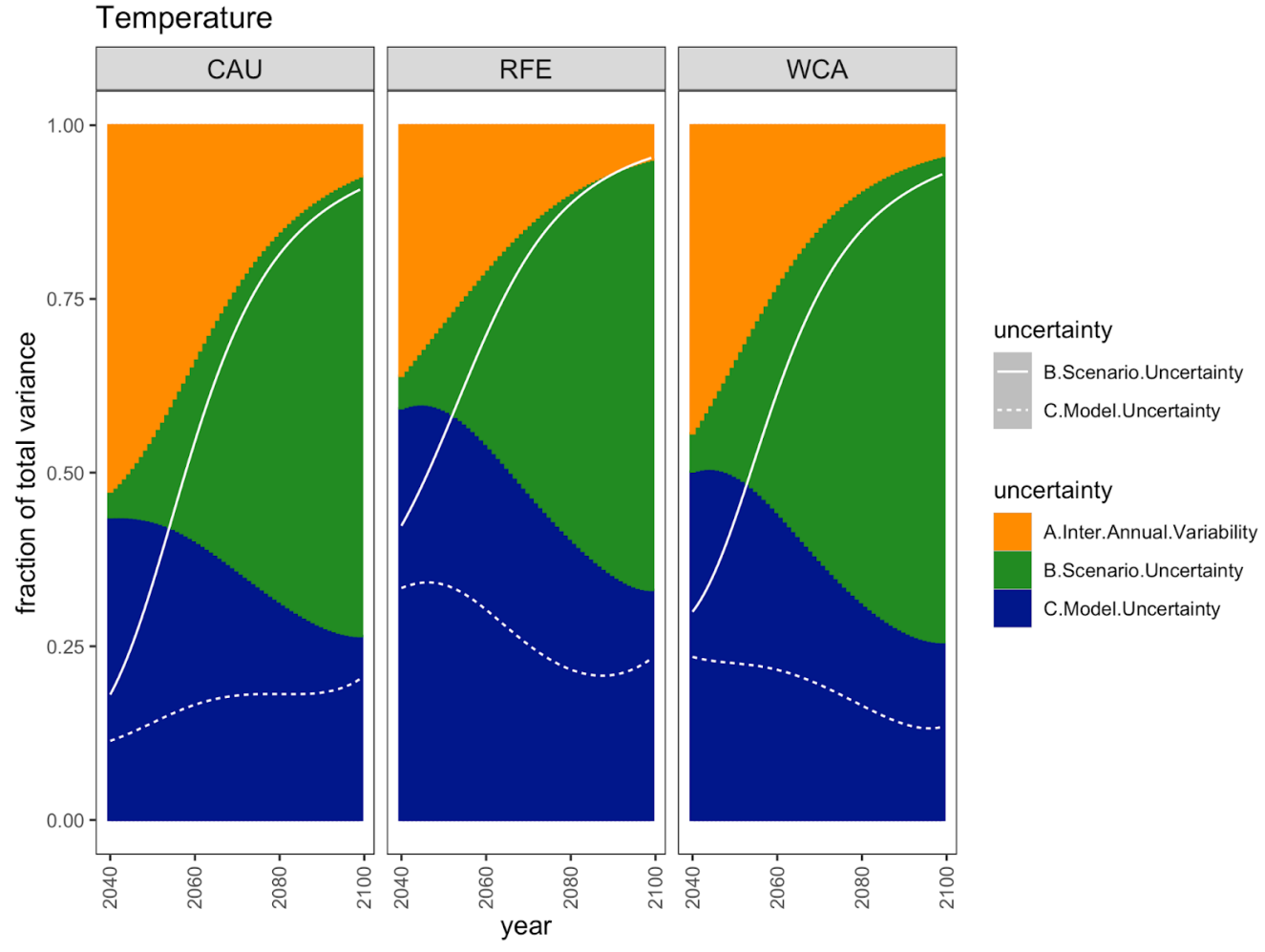


Figure A3: Same as Figure 6 but for RMS > 0.2 rather than 0.1

295

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

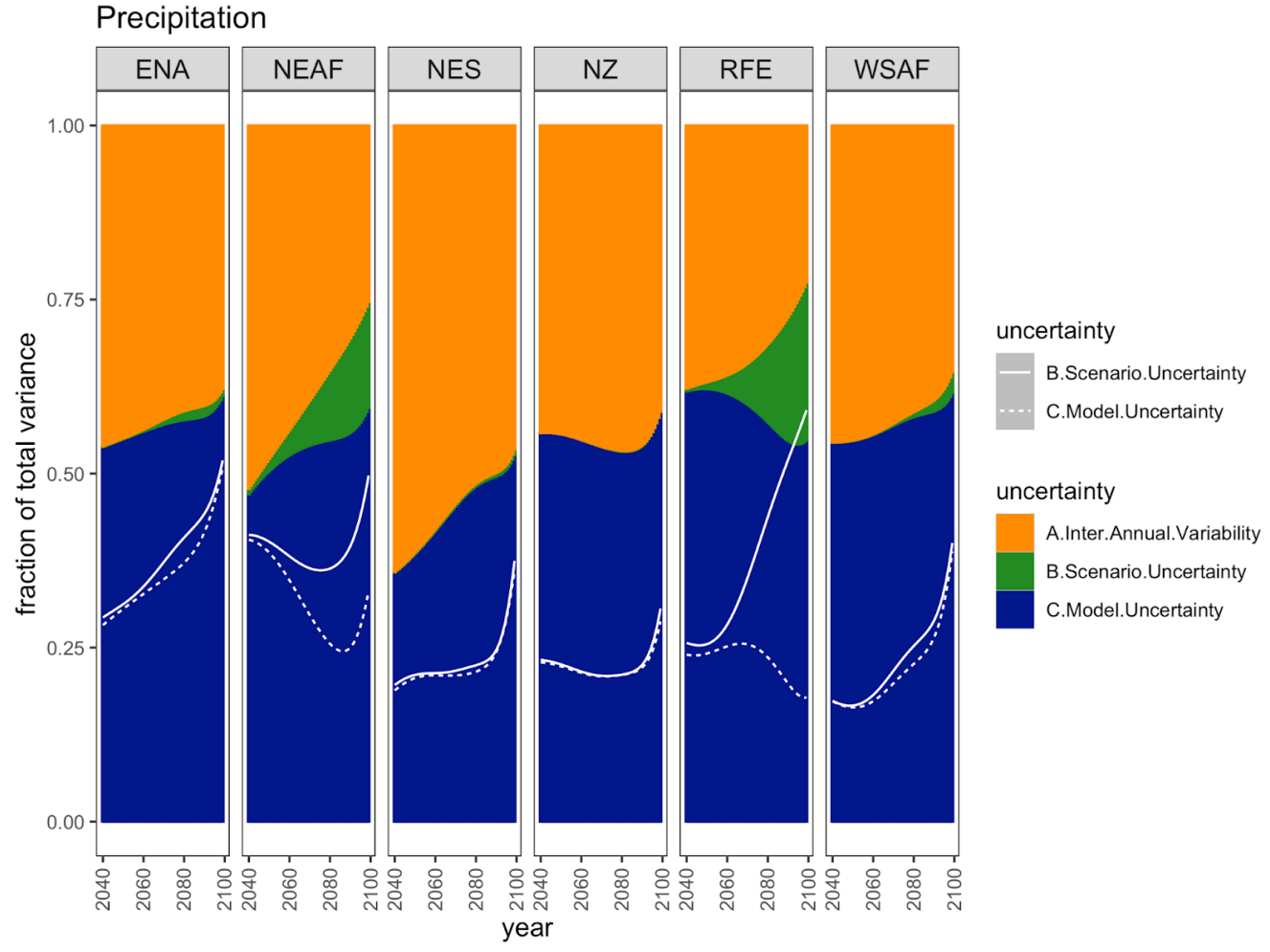


Figure A4: Same as Figure 7 but for RMS > 0.2 rather than 0.1.

Appendix B Hawkins and Sutton uncertainty calculations

Consider a set of trajectories for a given climate variable produced by various ESMs and scenarios. For example, this could be the annual average temperature or precipitation in a given world region. At each time step, *t,* there will be variation in the estimates from each observation in the set. The goal for a given set is to attribute a proportion of the variation or uncertainty at each time step to one of the three sources: interannual variation, model uncertainty, and scenario uncertainty. In our application, we want to do this for a "full" model set and compare the distribution of assigned variance to the same analysis on a selected subset of models.

The crux of this method for separating uncertainty is to write the raw predictions for each observation as $X_{m,s,t} = x_{m,s,t} + i_{m,s} + \varepsilon_{m,s,t}$, where $X_{m,s,t}$ is the raw prediction for model m scenario s at time t, $x_{m,s,t}$

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.

Earth System Dynamics Discussions
Open Access
EGU

s a smoothed fit of the variable anomaly with reference period 1995-2014, $i_{m,s}$ is the average variable value over the reference period, and $\varepsilon_{m,s,t}$ is the residual, representing interannual variation.

We can then essentially calculate the interannual variation component as the variance of all ε, the model uncertainty component at each time step as the variation in x over the different models, and the scenario uncertainty at each time step as the variation in x over the different scenarios. The variance calculations each have a weighting component. Models who more closely match the trend of observational data (W5E5v2.0 (Lange et al., 2021)) over the historic period will have their observations hold more weight. The weighting is as follows: $w_m = \frac{1}{x_{obs}+|x_m - x_{obs}|}$, where x_{obs} is the warming observed from 1995-2014 in the observation dataset (calculated as the difference in the smooth fit polynomial at the ends of that period), and x_m is the same thing but for the given model m. Weights are normalized ($W_m = \frac{w_m}{\sum_m w_m}$) to give the interannual variability component $V = \sum_m W_m \, var_{s,t}(\varepsilon_{m,s,t})$. The model uncertainty component is $M(t) = \frac{1}{N_s}\sum_s var_m^W(x_{m,s,t})$ for the number of scenarios used N_s (four in this study) and using the weighted variance function (var^W). The scenario uncertainty component is $S(t) = var_s(\sum_m W_m \, x_{m,s,t})$.

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.


Code and data availability: All code and data are available via a Github metarepository (https://github.com/JGCRI/SnyderEtAl2023_uncertainty_informed_curation_metarepo) and minted with a permanent DOI (https://doi.org/10.57931/2223040)

Author contributions: CT conceived of the project, AS led design of the methodology and performed analysis, NP performed analysis, KD provided data; all authors contributed to methodology, analysis, and the writing of the paper.

Competing interests: The authors declare that they have no conflict of interest.

Acknowledgements: This research was supported by the U.S. Department of Energy, Office of Science, as part of research in MultiSector Dynamics, Earth and Environmental System Modeling Program. The Pacific Northwest National Laboratory is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830. The views and opinions expressed in this paper are those of the authors alone.

References

Beusch, L., Gudmundsson, L., & Seneviratne, S. I. (2020). Emulating Earth system model temperatures with MESMER: from global mean temperature trajectories to grid-point-level realizations on land. *Earth System Dynamics*, *11*(1), 139–159.

Core Writing Team, H. L., & (eds.), J. R. (2023). *IPCC, 2023: Summary for Policymakers. In: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. https://doi.org/10.59327/IPCC/AR6-9789291691647.001

Dolan, F., Lamontagne, J., Calvin, K., Snyder, A., Narayan, K. B., Di Vittorio, A. V., & Vernon, C. R. (2022). Modeling the economic and environmental impacts of land scarcity under deep uncertainty. *Earth's Future*, *10*(2), e2021EF002466.

Dolan, F., Lamontagne, J., Link, R., Hejazi, M., Reed, P., & Edmonds, J. (2021). Evaluating the economic impact of water scarcity in a changing world. *Nature Communications*, *12*(1), 1915.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Frieler, K., Lange, S., Piontek, F., Reyer, C. P. O., Schewe, J., Warszawski, L., Zhao, F., Chini, L., Denvil, S., Emanuel, K., & Others. (2017). Assessing the impacts of 1.5 C global warming--simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b). *Geoscientific Model Development*, *10*(12), 4321–4345.

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.


Graham, N. T., Hejazi, M. I., Chen, M., Davies, E. G. R., Edmonds, J. A., Kim, S. H., Turner, S. W. D., Li, X., Vernon, C. R., Calvin, K., & Others. (2020). Humans drive future water scarcity changes across all Shared Socioeconomic Pathways. *Environmental Research Letters: ERL [Web Site]*, *15*(1), 014007.

Guivarch, C., Le Gallic, T., Bauer, N., Fragkos, P., Huppmann, D., Jaxa-Rozen, M., Keppo, I., Kriegler, E., Krisztin, T., Marangoni, G., & Others. (2022). Using large ensembles of climate change mitigation scenarios for robust insights. *Nature Climate Change*, *12*(5), 428–435.

Hawkins, E., & Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, *90*(8), 1095–1108.

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

Iturbide, M., Fernández, J., Gutiérrez, J. M., Pirani, A., Huard, D., Al Khourdajie, A., Baño-Medina, J., Bedia, J., Casanueva, A., Cimadevilla, E., & Others. (2022). Implementation of FAIR principles in the IPCC: the WGI AR6 Atlas repository. *Scientific Data*, *9*(1), 629.

Lange, S. (2019). Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1. 0). *Geoscientific Model Development*, *12*(7), 3055–3070.

Lange, S., Menz, C., Gleixner, S., Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Schmied, H. M., Hersbach, H., Buontempo, C., & Cagnazzo, C. (2021). *WFDE5 over land merged with ERA5 over the ocean (W5E5 v2.0)* [dataset]. ISIMIP Repository. https://doi.org/10.48364/ISIMIP.342217

Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E. M., Brunner, L., Knutti, R., & Hawkins, E. (2020). Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. *Earth System Dynamics*, *11*(2), 491–508.

Lovato, T., Peano, D., Butenschön, M., Materia, S., Iovino, D., Scoccimarro, E., Fogli, P. G., Cherchi, A., Bellucci, A., Gualdi, S., & Others. (2022). CMIP6 simulations with the CMCC Earth system model (CMCC-ESM2). *Journal of Advances in Modeling Earth Systems*, *14*(3), e2021MS002814.

Meehl, G. A., Senior, C. A., Eyring, V., Flato, G., Lamarque, J.-F., Stouffer, R. J., Taylor, K. E., & Schlund, M. (2020). Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models. *Science Advances*, *6*(26), eaba1981.

Müller, C., Franke, J., Jägermeyr, J., Ruane, A. C., Elliott, J., Moyer, E., Heinke, J., Falloon, P. D., Folberth, C., Francois, L., & Others. (2021). Exploring uncertainties in global crop yield projections in a large ensemble of crop models and CMIP5 and CMIP6 climate scenarios. *Environmental Research Letters: ERL [Web Site]*, *16*(3), 034040.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290.

Nath, S., Lejeune, Q., Beusch, L., Seneviratne, S. I., & Schleussner, C.-F. (2022). MESMER-M: an Earth system model emulator for spatially resolved monthly temperature. *Earth System Dynamics*, *13*(2), 851–877.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., & Sanderson, B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*(9), 3461–3482. https://doi.org/10.5194/gmd-9-3461-2016

Prudhomme, C., Giuntoli, I., Robinson, E. L., Clark, D. B., Arnell, N. W., Dankers, R., Fekete, B. M., Franssen, W., Gerten, D., Gosling, S. N., & Others. (2014). Hydrological droughts in the 21st century,

https://doi.org/10.5194/esd-2023-41
Preprint. Discussion started: 5 January 2024
© Author(s) 2024. CC BY 4.0 License.


hotspots and uncertainties from a global multimodel ensemble experiment. *Proceedings of the National Academy of Sciences*, *111*(9), 3262–3267.

Quilcaille, Y., Gudmundsson, L., Beusch, L., Hauser, M., & Seneviratne, S. I. (2022). Showcasing MESMER-X: Spatially Resolved Emulation of Annual Maximum Temperatures of Earth System Models. *Geophysical Research Letters*, *49*(17), e2022GL099012.

Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A. C., Müller, C., Arneth, A., Boote, K. J., Folberth, C., Glotter, M., Khabarov, N., & Others. (2014). Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proceedings of the National Academy of Sciences*, *111*(9), 3268–3273.

Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., Boote, K. J., Thorburn, P., Antle, J. M., Nelson, G. C., Porter, C., Janssen, S., & Others. (2013). The agricultural model intercomparison and improvement project (AgMIP): protocols and pilot studies. *Agricultural and Forest Meteorology*, *170*, 166–182.

Scafetta, N. (2022). Advanced Testing of Low, Medium, and High ECS CMIP6 GCM Simulations Versus ERA5-T2m. *Geophysical Research Letters*, *49*(6), e2022GL097716.

Tebaldi, C., Armbruster, A., Engler, H. P., & Link, R. (2020). Emulating climate extreme indices. *Environmental Research Letters: ERL [Web Site]*, *15*(7), 074006.

Tebaldi, C., Dorheim, K., Wehner, M., & Leung, R. (2021). Extreme metrics from large ensembles: investigating the effects of ensemble size on their estimates. *Earth System Dynamics*, *12*(4), 1427–1501.

Tebaldi, C., Snyder, A., & Dorheim, K. (2022). STITCHES: creating new scenarios of climate model output by stitching together pieces of existing simulations. *Earth System Dynamics*, 1–58. https://doi.org/10.5194/esd-2022-14

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., & Schewe, J. (2014). The inter-sectoral impact model intercomparison project (ISI--MIP): project framework. *Proceedings of the National Academy of Sciences*, *111*(9), 3228–3232.