PCIC SCIENCE BRIEF: CLIMATE MODEL GENEALOGY AND ITS RELATIONSHIP TO MODELLED CLIMATE PROPERTIES

Some of the global climate models used to study the Earth's climate system and make climate projections share components and computer code, much as if they were members of the same "families." This raises the question: to what extent do these family resemblances affect model results? Publishing in the *Journal of Advances in Modeling Earth Systems*, Kuma, Bender and Jönsson examined and quantified how this interdependence between models affects their simulated climate sensitivities, feedbacks and resulting projections of surface air temperature. Specifically, the authors found that models with shared code tend to have greater similarity in their climate sensitivities, strengths of feedbacks, and therefore in their projected surface temperatures. The authors also demonstrated that weighting ensembles of models according to their family resemblance resulted in a lower equilibrium climate sensitivity than when using a simple ensemble mean, and also reduced differences in climate sensitivity between the two most recent generations of climate models.

Introduction

Projections from ensembles of global climate models (GCMs) are used to understand how the Earth's climate may change in the future. These projections serve as the foundation for multiple sources of climate information, such as the downscaled climate scenarios and hydrologic projections developed at PCIC.

To inform the use and interpretation of these ensemble projections, it is useful to have a sense of how independent each of the models are from each other. How much do they have in common in terms of their description of climate system components and the resulting computer code, and how do these commonalities affect model results? In particular, how does this affect the strength of their simulated feedbacks[1] and their estimates of climate sensitivity[2]? It is these properties that will determine the severity of the climate impacts that will be experienced in various regions around the globe.

Kuma, Bender and Jönsson examined these questions through the lens of climate model ancestry in their article in the *Journal of Advances in Modeling Earth Systems*, the topic of this Science Brief. Answers to these questions can help characterize the uncertainty in model projections, helping us more wisely use the information they provide.

Shared Model Code

Certain sets of models share components, and thus portions of their underlying computer code. For instance, multiple models might share the same ocean or atmosphere components, or subcomponents, such as the code that describes chemical reactions within the atmosphere. These code components can either be shared between modelling groups directly, or duplicated by one group based on the published results of another. Tracing the use of shared code permits the grouping of models by fami-

1. Climate *feedbacks* are processes that can amplify or reduce the effect of an initial climate forcing. For example, increasing atmospheric greenhouse gas concentrations lead to warmer surface temperatures that hasten snow and sea ice melt, exposing more open water and ground to solar radiation. This leads to further warming that causes more ice and snow melt, and so on, constituting an amplification of the initial warming. This is known as the positive "ice-albedo feedback." Negative feedbacks also operate in the Earth system. For example, as the planet warms in response to greenhouse gas forcing, it radiates more longwave radiation back out into space (known as the "Planck feedback"). This cools the Earth, reducing somewhat the initial warming. It is the sum of feedbacks such as these that determines climate sensitivity. For more information on feedback processes, see Sherwood et al. (2020).
2. The Earth's *climate sensitivity* can be determined in several ways. Equilibrium climate sensitivity is the increase in global surface temperature that would eventually occur if the atmospheric concentration of carbon dioxide were doubled relative to the preindustrial period, and then held constant at that level indefinitely. However, determining equilibrium climate sensitivity requires running models for thousands of model years, which is computationally expensive. As a more practical alternative, effective climate sensitivity is often estimated instead. In this approach, the concentration of carbon dioxide in the modelled atmosphere is abruptly quadrupuled and the temperature change recorded after 150 simulated years. For the remainder of this Science Brief, we will follow Kuma and coauthors by referring to the metric they use, effective climate sensitivity, as simply "climate sensitivity."

lies and ancestry in a way that is analogous to how genealogy allows us to group people by family trees.

We can think of climate models as being, as the authors put it, "samples corresponding to representations of physical reality in a hypothesis space." That is, each model's structure is one representation of the climate system. In this way, we can think of models as being samples from a greater space of hypotheses. Ideally, models would be independent samples of this hypothesis space (Figure 1a). But in practice, due to their shared code and structure, they tend to cluster (Figure 1b), meaning that they do not sample uniformly across the hypothesis space.

The authors analysed the extent of this clustering, and considered how it affects the strength of simulated climate system feedbacks and estimates of climate sensitivity.

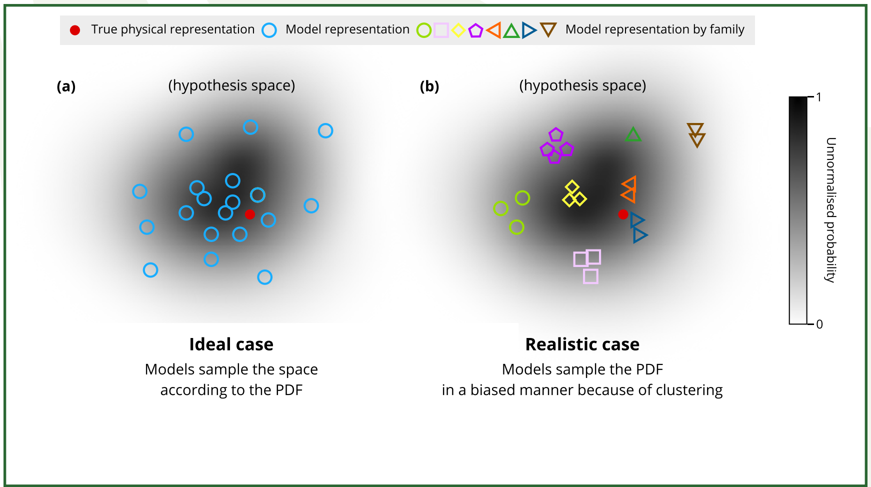

Figure 1: Conceptual examples illustrating how models sample from a hypothesis space (from Kuma et al., 2023).
In each figure panel, the shading indicates a probability density function (PDF) that quantifies our collective belief (based on our current best understanding) that a certain representation of the climate system is true, with darker and lighter shading representing greater and lesser belief in that representation, respectively. The red dot indicates the unknown true physical representation that need not coincide with the centre of the PDF. Model representations are given by the coloured shapes, as indicated. Panel (a) shows an ideal case, in which models sample from the hypothesis space in a relatively uniform or unbiased manner. Panel (b) shows the more realistic case, in which the models cluster because of shared components. Note also that this hypothesis space would actually have far more than two dimensions.

The Effect of Genealogy on Model Results

Kuma et al. examined 167 GCMs from 38 different modelling centres worldwide. The authors grouped the models into 14 families by shared model ancestry[3] (Figure 2), meaning that if two recent models B and C shared components with an older ancestor model A, then they would all be considered as belonging to the same family. The authors first found that the three largest model families account for about 70% of the multimodel ensemble from the sixth phase of the Coupled Model Intercomparison Project 6 (CMIP6), but only 52% of the earlier CMIP5 ensemble.[4] It is worth noting that the authors use two different but closely-related weighting schemes that they call "ancestry" and "family," but these tended to produce very similar results.[5]

Beginning with the topic of feedback processes in the models, the authors compared various weighting schemes, including weighting by family and ancestry, against a simple mean in which every model run is given equal weight (Figure 3). They found that the ancestry and family weightings gave very similar results and displayed the largest differences from the simple multi-model means for cloud feedbacks. The former weightings reduced the net (shortwave plus longwave) cloud feedbacks overall for CMIP6 models (Figure 3, Panel a), resulting in a lower climate sensitivity (by about 0.25°C) using these weightings. The corresponding difference in climate sensitivity using these weightings for CMIP5 models is negligible (Figure 3, Panel b). Weighting by model ancestry also reduces the differences in estimates of cloud feedbacks and climate sensitivity between CMIP6 and CMIP5 models (Figure 3, Panel c). The authors concluded that at least some of the differences in these metrics between the two CMIP phases is due to an overrepresentation of models from the largest families in CMIP6, primarily the Hadley Centre and NCAR Community Climate Model (CCM) groups. Furthermore, Kuma and co-

3. An exception was made for two very large model families that split in the 1980s, based on the open-source Community Climate Model [CCM] and based on the models from the European Centre for Medium-range Weather Forecasts, [ECMWF]. Despite sharing common ancestry, the authors treat them as two different families.
4. For more on the fifth phase of the Coupled Model Intercomparison Project, see Taylor et al. (2012) or here: http://cmip-pcmdi.llnl.gov/cmip5/. For more on the sixth phase of the Coupled Model Intercomparison Project, see Eyring et al. (2016) or here: https://pcmdi.llnl.gov/CMIP6/.
5. In the authors' weighting by ancestry, the oldest models were given the same weight, which was then subdivided between descendant models. So, if model A split into models B and C, but model C further split into models D and E, then the latter two models would each have a lower weight than model B. This differs from family weighting, where the weights are assigned equally to all members of the family. So, in the example above, models A, B, C, D and E would each have the same weight.

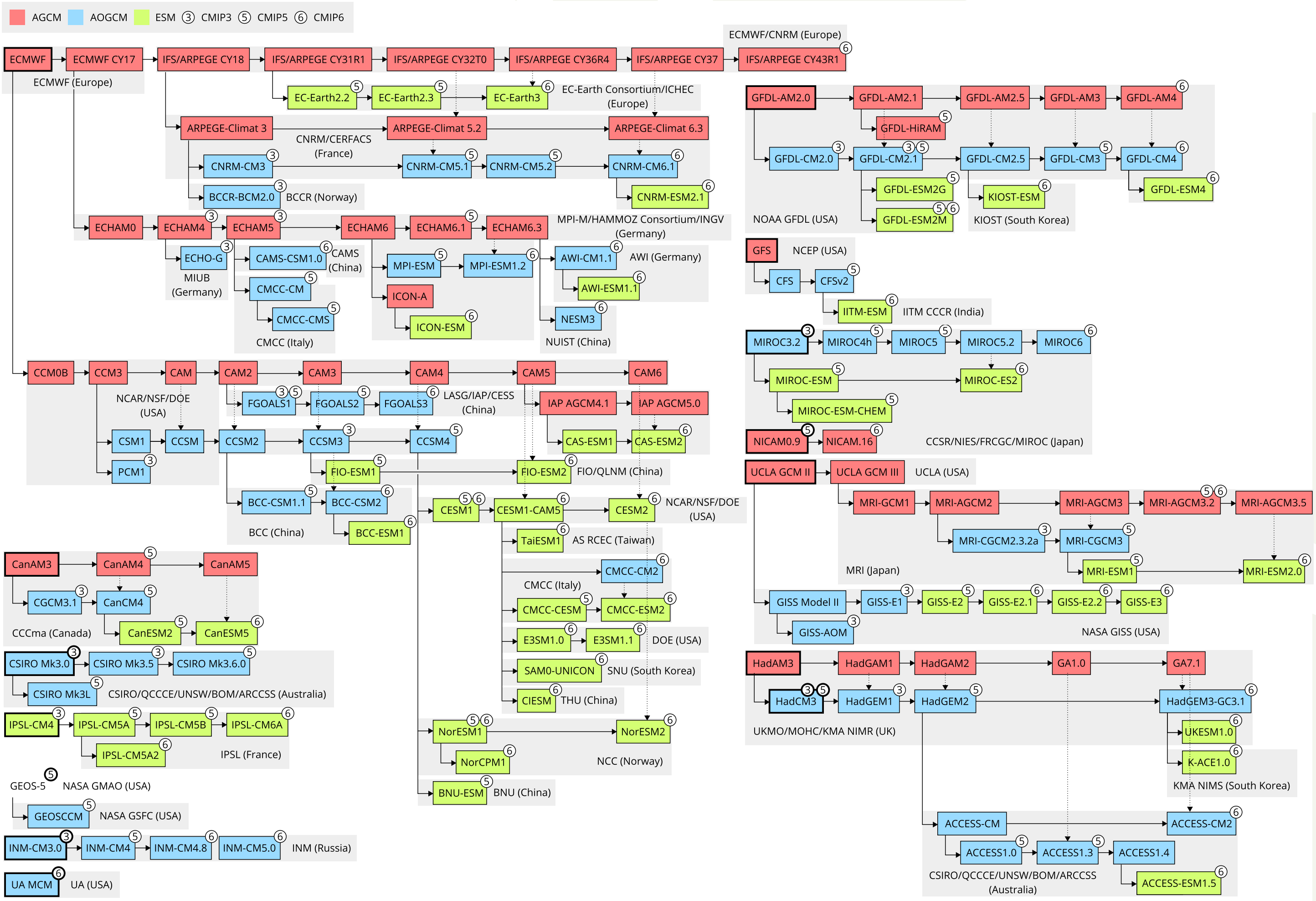

Figure 2: Model Code Genealogies (from Kuma et al., 2023).
This figure shows the model genealogies for CMIP models. Different model types—atmosphere general circulation models (AGCMs), atmosphere-ocean GCMs (AOGCMs) and Earth-system models (ESMs)—are distinguished by colours as in the legend at the top. Inheritance between different models is indicated by solid vertical arrows, inheritance by generation is indicated by solid horizontal arrows and inheritance between model types is indicated by dotted arrows. Grey shading indicates shared institutes or countries of development and numbers in circles indicate the phase of CMIP (i.e., 3, 5 or 6).

authors found that: 1) the higher mean climate sensitivity in the CMIP6 ensemble compared to CMIP5 can be attributed mainly to all 6 models from the Hadley family and a comparable number of models from the CCM family, and 2) this higher climate sensitivity is reduced under ancestry weighting by the smaller per-model weight assigned to models in large families.

With these results in hand, Kuma and colleagues analysed the model families to determine how interdependent they are. That is, do models from the same family tend to show similar values for feedbacks and climate sensitivity? And, if so, is this true for all model families or just some of them?

Kuma et al. found that while models from some large families tended to have similar values for certain feedbacks, others exhibited a wider spread of values.[6] In CMIP6, the six models descended from the Hadley Centre Atmosphere-only global Model (HadAM) tended to have very similar values for feedbacks, with total cloud and cloud shortwave feedbacks being larger than the mean from the ensemble, and longwave cloud feedbacks being smaller.

6. This feature arises from the fact that even when the GCMs in a family share the same component, e.g. an ocean model, the parameters determining the behaviour of the component are specific to each model. In other words, the parameters for these schemes have to be "tuned" in order to achieve optimal interactions with other components of the climate model, and this tuning can affect model behaviour substantially. Thus even models with the same components can have different characteristics.

The HadAM family, along with the IPSL family of models, also had an above-average climate sensitivity. By contrast, the UCLA family of five GCMs exhibited the opposite behaviour for all of the mentioned feedbacks, and had a below-average climate sensitivity.

The two largest model families in CMIP6, ECMWF (14 GCMs) and CCM (17 GCMs) showed a large and relatively even spread of climate sensitivity around the simple multi-model mean value. This is despite the fact that they exhibit roughly opposite behaviour in terms of their shortwave, longwave and total cloud feedbacks. Kuma et al. point out that in this case, ancestry or family weighting is unlikely to significantly affect either the mean or spread of climate sensitivity in the overall CMIP6 ensemble.

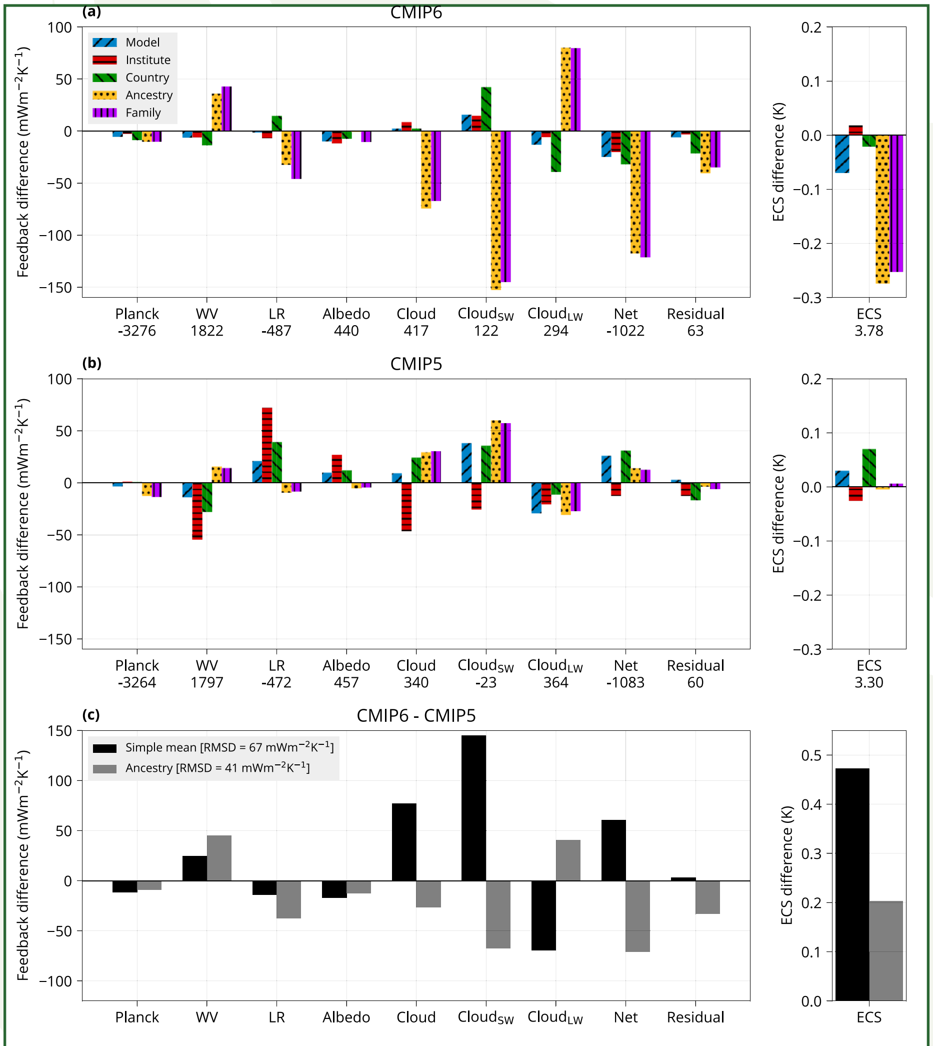

Figure 3: Feedbacks and effective climate sensitivities in CMIP6 and CMIP5 models under different weighting schemes (from Kuma et al., 2023). Panels a) and b) show differences in feedback strength (left panels) and effective climate sensitivity (ECS; right panels) for various types of weighting (as indicated by the colour and texture style in the legend) compared to a simple mean, for the CMIP6 (a) and CMIP5 (b) model ensembles. Note that shortwave cloud feedback, net feedback and ECS are all smaller in CMIP6 if model ancestry or family weighting is applied. Note also that these two forms of weighting give very similar results. Panel c) shows differences in feedback strength and ECS between CMIP6 and CMIP5 using a simple mean (in black) and weighted by ancestry (in grey). Note that the difference in ECS between CMIP6 and CMIP5 is reduced under model ancestry weighting.

The authors' results suggest that different families of models cluster together to differing degrees in terms of the strength of their simulated feedbacks and their climate sensitivities. Some families show strong interdependence while others have a relatively wide spread, especially for CMIP6. This raises the question of whether this translates into similar behaviour for their simulated climate variables. For example, if models cluster together in terms of climate sensitivity, then do they also display similar temperatures for a given concentration of greenhouse gases in the atmosphere?

With this question in mind, Kuma and coauthors concluded their study with an examination of model-simulated, global mean surface air temperature by CMIP6 GCMs under four different emissions scenarios. In the first experiment, the models were driven using historical greenhouse gas emissions from the years 1860 to 2000. In the second, historical emissions were specified up to the year 2015 and those from the moderate SSP2-4.5 emissions scenario up to the year 2100. In the third, the atmospheric concentration of carbon dioxide was abruptly quadrupled from preindustrial levels (taken to be those from the year 1850) and held steady for 150 years. In the final experiment, carbon dioxide levels were raised more gradually, again from the preindustrial level, by 1% per year for 150 years.

The two largest model families, CCM and ECMWF, which had large spreads in climate sensitivity, also exhibited a large spread in surface air temperature change (from the preindustrial value). Their family-mean temperature change also fell near the middle of the range for all CMIP6 models, consistent with the climate sensitivity results. Also consistent with these previous results were the temperature changes in the UCLA (lower than the CMIP6 average) and HadAM (above-average for CMIP6) families. Finally, model families with only a few members tended to exhibit very similar temperature change over the simulated period in all experiments.

Summary and Outlook

Anticipating the issue of model independence that forms the focus of Kuma and coauthors' work, and drawing on prior work by Brunner et al. (2020), PCIC researchers recently employed an initial screening process to generate model subsets for downscaled CMIP6 scenarios over Canada and its subregions.[7] PCIC's initial subset of 16 models based on model ancestry covers 10 of the 14 main families identified by Kuma and colleagues, with the largest fraction of models coming from the CCM (about one-third of the Earth system models) and ECMWF (about one-fifth of the Earth system models), together accounting for half of the models in PCIC's initial subset. These were the two families described by Kuma et al. as having the largest spread in climate sensitivity and surface air temperature, as well as "middle-of-the-range" mean temperatures.

The work of Kuma, Bender and Jönsson, which examines how commonalities between climate models affect their simulated climate feedbacks, climate sensitivity and projections of surface air temperature, is an important contribution to applied climate research. There is an urgent need to provide the users of climate model information with a means of reducing the large range of possible outcomes described in projections from the CMIP5 and CMIP6 ensembles, as this can slow or hinder decision-making. Kuma et al. found that weighting CMIP6 climate models according to model ancestry resulted in a reduction in overall cloud feedbacks and a lower climate sensitivity than a simple ensemble mean. It also reduced the difference in estimates of climate sensitivity between the most recent generation of models participating in CMIP6 and the previous generation of models from CMIP5. This is important because many existing regional climate assessments are based on the previous intercomparison: hence, closer agreement between results from the two ensembles maintains the relevance of those assessments. Since climate sensitivity, in particular, is an important determinant of the magnitude of climate impacts, better constraining its range via the consideration of model interdependence is an important step forward.

References:

Brunner, L., et al., 2020: Reduced global warming from CMIP6 projections when weighting models by performance and independence. *Earth System Dynamics*, **11**, 4, 995-1012, doi:10.5194/esd-11-995-2020.

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

Kuma, P., F. A.-M. Bender and A.R. Jönsson,, 2023: Climate model code genealogy and its relation to climate feedbacks and sensitivity. *Journal of Advances in Modeling Earth Systems*, **15**, e2022MS003588, doi:10.1029/2022MS003588.

Sherwood, S. C., et al., 2020: An assessment of Earth's climate sensitivity using multiple lines of evidence. *Reviews of Geophysics*, **58**, e2019RG000678, doi:10.1029/2019RG000678.

Taylor, K.E. et al., 2012: An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, **93**, 4, 485-498, doi: 10.1175/BAMS-D-11-00094.1.

7. For more information on the models used, see the Statistically Downscaled Climate Scenarios page on our Data Portal: http://pacificclimate.org/data/statistically-downscaled-climate-scenarios.