



RESEARCH ARTICLE

# A Dependent Multi-model Approach to Climate Prediction with Gaussian Processes

Marten Thompson<sup>1</sup> , Amy Braverman<sup>2</sup> and Snigdhasu Chatterjee<sup>1</sup> 

<sup>1</sup>School of Statistics, University of Minnesota, Minneapolis, 55455, MN, United States

<sup>2</sup>Jet Propulsion Laboratory, California Institute of Technology, Pasadena, 91109, CA, United States

\*Corresponding author. Email: [thom7058@umn.edu](mailto:thom7058@umn.edu)

**Received:** 10 January 2022

**Keywords:** Gaussian process; empirical Bayes; CMIP6

## Abstract

Simulations of future climate contain variability arising from a number of sources, including internal stochasticity and external forcings. However, to the best of our abilities climate models and the true observed climate depend on the same underlying physical processes. In this paper we simultaneously study the outputs of multiple climate simulation models and observed data, and we seek to leverage their mean structure as well as interdependencies that may reflect the climate's response to shared forcings. Bayesian modeling provides a fruitful ground for the nuanced combination of multiple climate simulations. We introduce one such approach whereby a Gaussian process is used to represent a mean function common to all simulated and observed climates. Dependent random effects encode possible information contained within and between the plurality of climate model outputs and observed climate data. We propose an empirical Bayes approach to analyze such models in a computationally efficient way. This methodology is amenable to the CMIP6 model ensemble, and we demonstrate its efficacy at forecasting global average near surface air temperature. Results suggest that this model and the extensions it engenders may provide value to climate prediction and uncertainty quantification.

## Impact Statement

Bayesian modeling provides a fruitful ground for the nuanced combination of multiple climate model outputs when predicting the Earth's climate characteristics for the coming years. We outline one such model and describe an empirical Bayes estimation approach that is computationally efficient. The proposed methodology, when applied to CMIP6 global temperature datasets, demonstrates that using empirical Bayesian techniques is better than using the simple “model democracy” approach of assigning equal weight to each climate model. We also obtain uncertainty bounds for the global temperature prediction problem.

## 1. Introduction

This paper is on predictive climate modeling using outputs of multiple models. The topic of how to assign weights to different climate models has been debated at length in the community. One approach is to assign equal weight to each model in multi-model climate ensembles, which has sometimes been labeled *model democracy* [15]. Other approaches for using information from multiple climate models may be found in [1, 7, 17, 26, 28, 10]. Bayesian approaches with uncertainty quantification may be

found in [40, 33] and elsewhere. In essence, most techniques directly or implicitly balance the performance of any candidate model output on how closely it emulates current and historical data, and on how well it imitates the outputs of other climate models. Technical details on different types of multi-model ensembles, including caveats about interpreting results from such ensembles, are available in [39, 18, 19, 16, 15, 35, 27, 17, 28, 21, 2, 13], among other places. A different approach is taken in [4, 6], where models are assigned scores based on their performance in capturing probabilistic properties of observed climate data. A review on different approaches of using climate model outputs may be found in [11].

A fundamental scientific reason for studying multi-model ensembles is that the same physical processes and forces are used for all reasonable climate models, and these processes also govern the observed climate data [22]. In this paper, we leverage this property of a shared physical basis for all climate model outputs and observed data concerning global temperature data. First, we adopt a statistical framework that assumes that the climate models' output and real data all share a common trend function over time. Second, we assume that the various model outputs and observed data are dependent on each other. This dependency is captured by framing the  $L$  climate model outputs and the observed data as an  $L+1$ -dimensional random vector at any point in time. This allows the entire dataset in  $L+1$ -dimensions over time to be viewed as a *vector Gaussian process*, which we discuss in some detail in Section 2.2. However, this leads to extremely challenging computations as well as a lack of transparency in the intermediate steps, hence we use an empirical Bayesian approach in this paper.

We analyze the climate model outputs associated with the Coupled Model Intercomparison Project, now in its sixth phase (CMIP6), which directs research in fundamental areas of climate science [9]. It organizes the federated efforts of many investigators whose model simulations examine natural climate variability and sensitivity. CMIP6 simulations occur under a wide range of scenarios characterized by a Shared Socioeconomic Pathway (SSP) and a level of increased radiative forcing. The Shared Socioeconomic Pathways describe a spectrum of possible futures in terms of human energy use, population, and sentiment towards climate stewardship [24]. They then inform the quantification of land use, energy consumption, and other factors used by climate modelers. Radiative forcing concerns the increase in energy transferred through Earth's atmosphere, relative pre-Industrial levels [5]. We analyze model output under SSP-5 with a radiative forcing value of 8.5 watts/m<sup>2</sup>, the most aggressive increase. The SSP-5 narrative describes an increase in social and economic capital driven by fossil fuel extraction and an overall technology-driven strategy to climate change mitigation.

The focus of our analysis is the output from a selection of 17 CMIP6 simulations [8, 44, 31, 42, 25, 36, 37, 20, 34, 38, 32, 29, 41, 43, 30, 3, 12]. All simulations were conducted under the SSP-5 8.5 scenario and produced monthly mean global near surface air temperature values spanning 1850 January to 2100 December. We remove seasonality by subtracting the mean monthly values from 1961-1990 [14]. The observed monthly mean global near surface air temperature was also available from 1850 January to 2020 December [23].

## 2. Methods

### 2.1. Hierarchical Bayesian Model

The CMIP6 simulations produce monthly global near surface air temperatures from 1850-2100. Let  $Y_{\ell,t}$  be the deseasonalized, mean-centered time series from  $\ell$ th CMIP6 climate simulation for  $\ell = 1, \dots, L$ . The observed monthly values through 2020 are denoted by  $Y_{0,t}$ . The vector  $\mathbf{Y}_t$  will represent all values at time  $t$ , and  $\mathbf{Y}_{\text{cmip},t}$  will represent the  $L$  climate simulations i.e.

$$\mathbf{Y}_t = \begin{bmatrix} Y_{0,t} \\ Y_{1,t} \\ \vdots \\ Y_{L,t} \end{bmatrix}; \quad \mathbf{Y}_{\text{cmip},t} = \begin{bmatrix} Y_{1,t} \\ \vdots \\ Y_{L,t} \end{bmatrix} \quad (1)$$

There are many ways to utilize these values when making predictions for  $Y_{0,t}$ . A simple method is to estimate a common mean function  $\mu(t)$  shared by all the climate simulations and observed data. Here, our goal is to also leverage the information contained within the plurality of these simulations beyond their mean. One such method is to assume

$$\mathbf{Y}_t \mid \mathbf{U}_t = \mu(t)\mathbf{1}_{L+1} + \mathbf{U}_t, \quad \text{where } \mathbf{U}_t \stackrel{\text{i.i.d.}}{\sim} N_{L+1}(\mathbf{0}, \Sigma). \quad (2)$$

where  $\mathbf{1}_{L+1}$  is an  $L+1$ -dimensional vector of 1's, and  $\mathbf{U}_t$  are independent random effects. These random elements are independent over time, but dependent between the  $L+1$  time series. Our assumption here is that these are Normally distributed, but this can be relaxed.

One may wonder if assuming the  $\mathbf{U}_t$  are independent over time, or identically distributed, is adequate for the available climate data. However, our extensive preliminary studies strongly indicate that there is no temporal dependency pattern in the  $\mathbf{U}_t$ : we have experimented using autoregressive integrated moving average (ARIMA) models, several kinds of Gaussian process-based models, conditionally heteroscedastic models and change-point detection procedures. All these preliminary studies indicate that (2) is the best option for modeling  $\mathbf{U}_t$ . However, note that our main modeling principle and ideas do not depend on the particular assumptions surrounding (2). We can easily replace the Gaussian distributional assumption or the independence assumption with other suitable conditions; however in such cases even the empirical Bayesian computations would be considerably lengthier and numerical integration-driven.

We represent the common monthly mean  $\mu(t)$  with a Gaussian process defined by the covariance kernel  $k_\alpha(\cdot, \cdot)$  parameterized by  $\alpha$ . While a variety of kernels exist, the squared exponential function is a common choice. It is parameterized by  $\alpha = (\sigma^2, \gamma)$ , the variance and lengthscale parameters. That is,

$$\mu(t) \sim \mathcal{GP}(0, k_\alpha) \quad \text{where } k_\alpha(t_1, t_2) = \sigma^2 e^{-\gamma(t_1 - t_2)^2}. \quad (3)$$

Together, Equations 2 and 3 define the hierarchical model given below.

$$\begin{aligned} \mathbf{Y}_t \mid \mu(t), \Sigma &\sim N_{L+1}(\mu(t)\mathbf{1}_{L+1}, \Sigma), \\ \mu(t) &\sim \mathcal{GP}(0, k_\alpha). \end{aligned} \quad (4)$$

## 2.2. Esimation: Full Bayesian Approach

Due to the nature of Gaussian processes,  $\mu(t)$  evaluated at points  $t = 1, \dots, T$  will also follow a multivariate normal distribution with mean zero and a covariance matrix  $K$  given by  $[K]_{ij} = k_\alpha(t_i, t_j)$ . Thus, Equation 4 may alternatively be expressed as a vector Gaussian process, written here as the matrix normal distribution this induces over time  $t = 1, \dots, T$ . Let  $\underline{\mathbf{Y}}$  be the  $(L+1) \times T$  matrix representing all observed and simulated time series. Let  $\mathbf{0}_{(L+1) \times T}$  be a  $(L+1) \times T$  matrix of zeros. Then,

$$\begin{aligned} \underline{\mathbf{Y}} \mid \Sigma, K &\sim MN_{(L+1), T}(\mathbf{0}_{(L+1) \times T}, \Sigma, K), \quad \text{equivalently} \\ \text{vec}(\underline{\mathbf{Y}}) \mid \Sigma, K &\sim N_{(L+1) \cdot T}(\mathbf{0}_{(L+1) \cdot T}, \Sigma \otimes K) \end{aligned} \quad (5)$$

where  $\otimes$  represents the Kronecker product.

With Equation 5, standard properties of the multivariate Normal distribution provide pointwise predictions and uncertainty quantification via its conditional mean and conditional variance for any subset of  $\text{vec}(\underline{\mathbf{Y}})$ . In particular, let  $t_a$  corresponds to 2020 December and  $t_b$  be 2100 December. The distribution of  $Y_{0,t_a:t_b}$  conditional on the observed time series  $Y_{0,0:t_a}$  and CMIP6 simulations  $Y_{\ell,0:t_b}$ ,  $\ell = 1, \dots, L$  is also multivariate normal.

Unfortunately, the matrix  $\Sigma \otimes K$  is large at  $(L+1) \cdot T$ -dimensional square. Inverting this is cubic in complexity and infeasible for large  $L$  or  $T$ . The manipulation of this large matrix also incurs numerical overflow issues for sufficiently large dimension.

### 2.3. Estimation: Empirical Bayesian Approach

While the vector Gaussian process model would lend itself to full Bayesian modeling, we propose an empirical Bayesian (EB) approach. This is done in order to reduce the computational overhead and for transparency. Below we describe the method by which  $\mu(t)$  and  $\Sigma$  are thus estimated, and how they afford prediction and uncertainty quantification.

We estimate the common mean  $\mu(t)$  as a Gaussian process on the time pointwise average of the model outputs. That is, define  $\bar{Y}_t = 1/(L+1) \sum_{\ell=0}^L Y_{\ell,t}$ . Then, the parameters  $\alpha = (\sigma^2, \gamma)$  are estimated by maximizing the marginal likelihood of  $\bar{Y}_t$  under the Gaussian process model. Denote the estimated values of the mean function at time  $t$  as  $\hat{\mu}(t)$ .

Next, take  $\hat{U}_{\ell,t} = Y_{\ell,t} - \hat{\mu}(t)$ . Then, let  $\hat{U}_t = [\hat{U}_{0,t} \ \hat{U}_{1,t} \dots \hat{U}_{L,t}]^\top$ . Use  $\hat{U}_1, \dots, \hat{U}_T$  to get  $\hat{\Sigma}$ , the  $(L+1) \times (L+1)$  variance-covariance matrix. This encodes the possible correlations between time series outside of the common mean, most importantly between the observed time series and the CMIP6 simulations' outputs. Let us write

$$\hat{\Sigma} = \begin{pmatrix} \hat{\sigma}_0^2 & \hat{\Sigma}_0^T \\ \hat{\Sigma}_0 & \hat{\Sigma}_{\text{cmip}} \end{pmatrix} \quad (6)$$

where  $\hat{\sigma}_0^2$  is the variance of the observed data's random effects, and  $\hat{\Sigma}_{\text{cmip}}$  is the  $L \times L$  covariance between random effects associated with the CMIP6 outputs.

One of our goals is to predict  $Y_{0,t}$  for  $t \in 2021\text{--}2100$  given the model outputs and historic observations. The distribution of these values follows from Equation 4 and standard properties of the Normal distribution. Given  $\hat{\mu}(t)$  and simulated values  $\mathbf{Y}_{\text{cmip},t}$

$$Y_{0,t} \mid \hat{\mu}(t), \mathbf{Y}_{\text{cmip},t} \sim N(m^*, \gamma^*) \quad (7)$$

$$m^* = \hat{\mu}(t) + \Sigma_0^T \Sigma_{\text{cmip}}^{-1} (\mathbf{Y}_{\text{cmip},t} - \hat{\mu}(t) \mathbf{1}_L) \quad (8)$$

$$\gamma^* = \sigma_0^2 - \Sigma_0^T \Sigma_{\text{cmip}}^{-1} \Sigma_0 \quad (9)$$

These conditional mean and variance values provide natural point estimates and uncertainty quantification. Note that using  $\hat{\mu}(t)$  alone amounts to making predictions with the unconditional mean. In the next section, we will see that the conditional mean better reflects the month-to-month variation of these time series, as well as differences between them and the common mean.

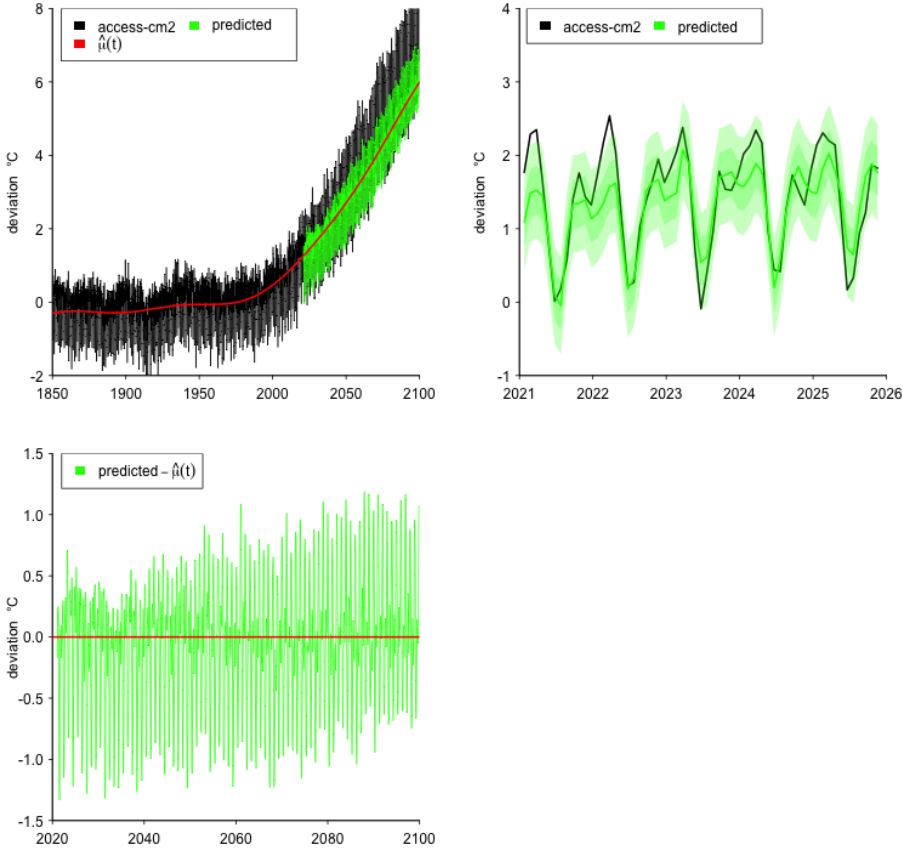
## 3. CMIP6 Results

### 3.1. Leave-One-Out Validation

A leave-one-out (LOO) style of evaluation allows us to investigate the validity of this approach. Setting aside the observed time series  $Y_{0,t}$ , the remaining CMIP6 simulations all span 1850 January to 2100 December. One by one, each of the available time series is singled out and treated like the “observed” values. That is, it is truncated to 2020 November, and we perform the model estimation approach from Section 2.3. This produces predictions through 2100, which we compare to the full time series's actual values.

As a tangible example, Figure 1 depicts the results when the access-cm2 CMIP6 simulation is being treated like the observed time series. Predictions begin in 2020 December and continue to 2100 December. They match the actual simulated values both in broad trend across the whole of their domain, and in local month-to-month fluctuations. Note also that while the common mean function  $\hat{\mu}(t)$  trends below the simulated values as time continues, the conditional mean (Equation 8) ameliorates this trend.

Table 1 contains a summary of the LOO results. It includes the mean squared error between the actual output of the indicated simulation and the predictions of our empirical Bayes (EB) model. Compare



**Figure 1.** Amongst others, we used the access-cm2 CMIP6 time series to evaluate our model's predictive accuracy. The top left figure shows the whole of the training and test intervals, 1850 January - 2020 November and 2020 December - 2100 December, respectively. Just five years are shown in the top right figure, starting in 2021 January, to highlight the prediction's month-to-month accuracy. Light and dark shading indicate one and two standard deviations from the conditional mean, as found in Equations 8 and 9. On the top and bottom left we see the conditional mean also improves predictions by inducing a trend that better matches that of the test interval compared to  $\hat{\mu}(t)$ .

these to the MSE of the predictions found using just the common mean function  $\hat{\mu}(t)$  or pointwise average  $\bar{Y}_t$ . The empirical Bayes model is more performant than the common mean in every case.

### 3.2. Observed Time Series

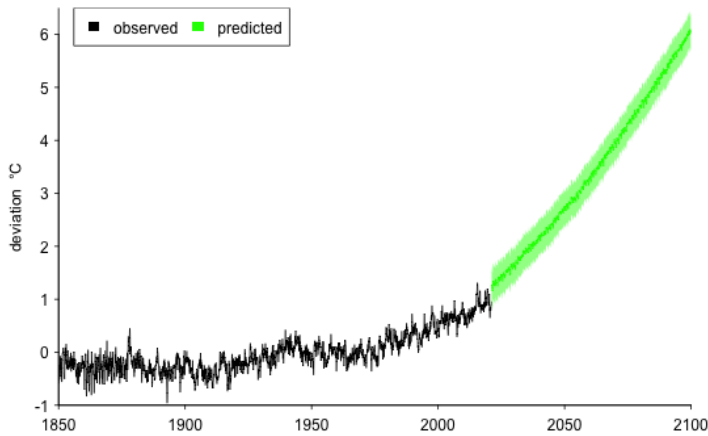
Making predictions for the observed climate follows the procedure outlined in Section 2.3. Figure 2 depicts the predicted values for the observed time series  $Y_{0,t}$  from 2020 December - 2100 December. The upward trend is driven by the common mean  $\hat{\mu}(t)$  estimated on the CMIP6 time series. The random effects induce the month-to-month variations, shown in Subsection 3.1 to better reflect CMIP6 time series.

Note that the observed time series has less variation than the CMIP6 simulations considered here. Over the shared period of 1850 January to 2020 November, the observed time series has variance 0.15 compared to 0.55 for the CMIP6 simulations on average. Comparing Figure 2 to the access-cm2 data in Figure 1 is representative of this. This is then reflected in the covariance matrix Equation 6; its

LOO Sim	$\text{MSE}_{\text{EB}}$	$\text{MSE}_{\hat{\mu}}$	$\text{MSE}_{\hat{Y}_t}$
access-cm2	0.44	1.01	0.38
access-esm1-5	0.16	0.41	0.09
bcc-csm2-mr	0.43	0.64	0.42
canesm5-canoe	3.57	4.09	3.72
canesm5	3.40	3.97	3.58
cesm2-waccm	0.59	1.03	0.58
cesm2	0.49	0.89	0.47
giss-e2-1-g	0.57	0.73	0.64
mcm-ua-1-0	0.39	0.72	0.64
miroc-es2l	0.64	1.06	0.66
miroc6	0.84	1.48	0.97
mpi-esm1-2-hr	0.83	1.08	0.98
mpi-esm1-2-lr	0.83	0.95	0.87
mri-esm2-0	0.34	0.69	0.26
noresm2-lm	1.30	1.68	1.29
noresm2-mm	1.12	1.45	1.20
ukesm1-0-ll	3.22	3.79	3.24

**Table 1.** Each CMIP6 simulation was treated like the observed time series, truncated to 2020 November, and predicted by various methods. The predictions produced by our empirical Bayes model are similar in accuracy to using  $\hat{Y}_t$  itself and also provide an estimate of variability. These predictions were superior to the common mean component  $\hat{\mu}(t)$  on every CMIP6 test simulation.

attenuating presence in Equation 8 explains how the predictions in Figure 2 show less variation than those in Figure 1.



**Figure 2.** Predicted values for mean global surface temperature using Equation 8, with shading to  $\pm 2$  standard deviations (Equation 9). This model leverages the common mean as well as individual correlations amongst CMIP6 simulations and historical observed data.

#### 4. Conclusion

This preliminary investigation affirms the use of our hierarchical Bayesian model in combining multi-model climate data. Our model reflects both the small and large scale deviational properties of a given

time series via a common mean. An empirical Bayesian treatment is an attractive approach to estimating this model; additional climate simulations can be considered with little added computational cost. This is especially important in applications where simulations produce high-dimensional output as is common in climate science.

Our model assumes that the properties of the common trend  $\mu(\cdot)$  can be adequately captured using a Gaussian process. This can be easily extended to a  $t$ -process or an elliptically contoured process with some additional computations. Also, our model extends to the case where there is a spatial component in the data or where we consider regional climate models: the principles outlined above extend simply to such cases.

**Funding Statement.** This research is partially supported by the US National Science Foundation (NSF) under grants 1939916, 1939956, and a grant from Cisco Systems Inc.

**Competing Interests.** The authors declare none.

**Data Availability Statement.** All materials and code necessary to reproduce these results may be found at [https://github.com/MartenThompson/climate\\_informatics\\_2022](https://github.com/MartenThompson/climate_informatics_2022). Please see relevant citations for CMIP6 [8, 44, 31, 42, 25, 36, 37, 20, 34, 38, 32, 29, 41, 43, 30, 3, 12] and observed data [14, 23].

**Ethical Standards.** The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

**Author Contributions.** **Marten Thompson:** Conceptualization, Methodology, Formal analysis, Software, Writing - Original Draft. **Amy Braverman:** Data Curation, Supervision. **Snigdhanu Chatterjee:** Conceptualization, Methodology, Supervision, Writing - Review & Editing.

**Supplementary Material.** None.

## References

- [1] Gab Abramowitz et al. “ESD Reviews: Model dependence in multi-model climate ensembles: weighting, sub-selection and out-of-sample testing”. In: *Earth System Dynamics* 10.1 (2019), pp. 91–105.
- [2] Gab Abramowitz et al. “Model dependence in multi-model climate ensembles: weighting, sub-selection and out-of-sample testing”. In: *Earth Syst. Dynam. Discuss* (2018).
- [3] Mats Bentsen et al. *NCC NorESM2-MM model output prepared for CMIP6 ScenarioMIP ssp585*. 2019. DOI: [10.22033/ESGF/CMIP6.8321](https://doi.org/10.22033/ESGF/CMIP6.8321). URL: <https://doi.org/10.22033/ESGF/CMIP6.8321>.
- [4] Amy Braverman et al. “Probabilistic evaluation of competing climate models”. In: *Advances in Statistical Climatology, Meteorology and Oceanography* 3.2 (2017), pp. 93–105.
- [5] Intergovernmental Panel On Climate Change. “Climate change 2007: The physical science basis”. In: *Agenda* 6.07 (2007), p. 333.
- [6] Snigdhanu Chatterjee. “The scale enhanced wild bootstrap method for evaluating climate models using wavelets”. In: *Statistics & Probability Letters* 144 (2019), pp. 69–73.
- [7] Peter F Craigmile, Peter Guttorp, and Veronica Berrocal. “Regional climate model assessment via spatio-temporal modeling”. In: *Environmetrics* 23 (2017), pp. 482–492.
- [8] Martin Dix et al. *CSIRO-ARCCSS ACCESS-CM2 model output prepared for CMIP6 ScenarioMIP ssp585*. 2019. DOI: [10.22033/ESGF/CMIP6.4332](https://doi.org/10.22033/ESGF/CMIP6.4332). URL: <https://doi.org/10.22033/ESGF/CMIP6.4332>.
- [9] Veronika Eyring et al. “Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization”. In: *Geoscientific Model Development* 9.5 (2016), pp. 1937–1958.
- [10] Gregory Flato et al. “Evaluation of climate models”. In: *Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 2014, pp. 741–866.
- [11] Kaibo Gong and Snigdhanu Chatterjee. “On Evaluation of Climate Models”. In: *Evaluating Climate Change Impacts*. Chapman and Hall/CRC, 2020, pp. 3–24.
- [12] Peter Good et al. *MOHC UKESM1.0-LL model output prepared for CMIP6 ScenarioMIP ssp585*. 2019. DOI: [10.22033/ESGF/CMIP6.6405](https://doi.org/10.22033/ESGF/CMIP6.6405). URL: <https://doi.org/10.22033/ESGF/CMIP6.6405>.
- [13] Nadja Herger et al. “Selecting a climate model subset to optimise key ensemble properties”. In: *Earth System Dynamics* 9.1 (2018), pp. 135–151.
- [14] Phil D Jones et al. *Surface air temperature and its changes over the past 150 years*. [https://crudata.uea.ac.uk/cru/data/temperature/abs\\_glnhsh.txt](https://crudata.uea.ac.uk/cru/data/temperature/abs_glnhsh.txt). Accessed 2021/11/29. 1999.
- [15] Reto Knutti. “The end of model democracy?” In: *Climatic change* 102.3 (2010), pp. 395–404.



- [16] Reto Knutti, David Masson, and Andrew Gettelman. “Climate model genealogy: Generation CMIP5 and how we got there”. In: *Geophysical Research Letters* 40.6 (2013), pp. 1194–1199.
- [17] Reto Knutti et al. “A climate model projection weighting scheme accounting for performance and interdependence”. In: *Geophysical Research Letters* 44.4 (2017), pp. 1909–1918.
- [18] Reto Knutti et al. “Challenges in combining projections from multiple climate models”. In: *Journal of Climate* 23.10 (2010), pp. 2739–2758.
- [19] Reto Knutti et al. “Good Practice Guidance Paper on Assessing and Combining Multi Model Climate Projections”. In: *IPCC Expert Meeting on Assessing and Combining Multi Model Climate Projections: National Center for Atmospheric Research: Boulder, Colorado, USA: 25-27 January 2010: Meeting Report*. IPCC Working Group I Technical Support Unit, University of Bern. 2010.
- [20] Lijuan Li. *CAS FGOALS-g3 model output prepared for CMIP6 ScenarioMIP ssp585*. 2019. DOI: [10.22033/ESGF/CMIP6.3503](https://doi.org/10.22033/ESGF/CMIP6.3503). URL: <https://doi.org/10.22033/ESGF/CMIP6.3503>.
- [21] Ruth Lorenz et al. “Prospects and caveats of weighting climate models for summer maximum temperature projections over North America”. In: *Journal of Geophysical Research: Atmospheres* 123.9 (2018), pp. 4509–4526.
- [22] David Masson and Reto Knutti. “Climate model genealogy”. In: *Geophysical Research Letters* 38.8 (2011).
- [23] Colin P Morice et al. “An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set”. In: *Journal of Geophysical Research: Atmospheres* 126.3 (2021), e2019JD032361.
- [24] Keywan Riahi et al. “The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview”. In: *Global environmental change* 42 (2017), pp. 153–168.
- [25] Xinyao Rong. *CAMS CAMS-CSM1.0 model output prepared for CMIP6 ScenarioMIP ssp585*. 2019. DOI: [10.22033/ESGF/CMIP6.11052](https://doi.org/10.22033/ESGF/CMIP6.11052). URL: <https://doi.org/10.22033/ESGF/CMIP6.11052>.
- [26] Benjamin M Sanderson, Reto Knutti, and Peter Caldwell. “A representative democracy to reduce interdependency in a multimodel ensemble”. In: *Journal of Climate* 28.13 (2015), pp. 5171–5194.
- [27] Benjamin M Sanderson, Reto Knutti, and Peter Caldwell. “Addressing interdependency in a multimodel ensemble by interpolation of model properties”. In: *Journal of Climate* 28.13 (2015), pp. 5150–5170.
- [28] Benjamin M Sanderson, Michael Wehner, and Reto Knutti. “Skill and independence weighting for multi-model assessments”. In: *Geoscientific Model Development* 10.6 (2017), pp. 2379–2395.
- [29] Martin Schupfner et al. *DKRZ MPI-ESM1.2-HR model output prepared for CMIP6 ScenarioMIP ssp585*. 2019. DOI: [10.22033/ESGF/CMIP6.4403](https://doi.org/10.22033/ESGF/CMIP6.4403). URL: <https://doi.org/10.22033/ESGF/CMIP6.4403>.
- [30] Øyvind Seland et al. *NCC NorESM2-LM model output prepared for CMIP6 ScenarioMIP ssp585*. 2019. DOI: [10.22033/ESGF/CMIP6.8319](https://doi.org/10.22033/ESGF/CMIP6.8319). URL: <https://doi.org/10.22033/ESGF/CMIP6.8319>.
- [31] Tido Semmler et al. *AWI AWI-CM1.1MR model output prepared for CMIP6 ScenarioMIP ssp585*. 2019. DOI: [10.22033/ESGF/CMIP6.2817](https://doi.org/10.22033/ESGF/CMIP6.2817). URL: <https://doi.org/10.22033/ESGF/CMIP6.2817>.
- [32] Hideo Shiogama, Manabu Abe, and Hiroaki Tatebe. *MIROC MIROC6 model output prepared for CMIP6 ScenarioMIP ssp585*. 2019. DOI: [10.22033/ESGF/CMIP6.5771](https://doi.org/10.22033/ESGF/CMIP6.5771). URL: <https://doi.org/10.22033/ESGF/CMIP6.5771>.
- [33] Richard L Smith et al. “Bayesian modeling of uncertainty in ensembles of climate models”. In: *Journal of the American Statistical Association* 104.485 (2009), pp. 97–116.
- [34] NASA Goddard Institute for Space Studies (NASA/GISS). *NASA-GISS GISS-E2.1G model output prepared for CMIP6 ScenarioMIP ssp585*. 2020. DOI: [10.22033/ESGF/CMIP6.7460](https://doi.org/10.22033/ESGF/CMIP6.7460). URL: <https://doi.org/10.22033/ESGF/CMIP6.7460>.
- [35] Maria Antonia Sunyer et al. “A Bayesian approach for uncertainty quantification of extreme precipitation projections including climate model interdependency and nonstationary bias”. In: *Journal of Climate* 27.18 (2014), pp. 7113–7132.
- [36] Neil Cameron Swart et al. *CCCma CanESM5 model output prepared for CMIP6 ScenarioMIP ssp585*. 2019. DOI: [10.22033/ESGF/CMIP6.3696](https://doi.org/10.22033/ESGF/CMIP6.3696). URL: <https://doi.org/10.22033/ESGF/CMIP6.3696>.
- [37] Neil Cameron Swart et al. *CCCma CanESM5-CanOE model output prepared for CMIP6 ScenarioMIP ssp585*. 2019. DOI: [10.22033/ESGF/CMIP6.10276](https://doi.org/10.22033/ESGF/CMIP6.10276). URL: <https://doi.org/10.22033/ESGF/CMIP6.10276>.
- [38] Kaoru Tachiiri et al. *MIROC MIROC-ES2L model output prepared for CMIP6 ScenarioMIP ssp585*. 2019. DOI: [10.22033/ESGF/CMIP6.5770](https://doi.org/10.22033/ESGF/CMIP6.5770). URL: <https://doi.org/10.22033/ESGF/CMIP6.5770>.
- [39] Claudia Tebaldi and Reto Knutti. “The use of the multi-model ensemble in probabilistic climate projections”. In: *Philosophical transactions of the royal society A: mathematical, physical and engineering sciences* 365.1857 (2007), pp. 2053–2075.
- [40] Claudia Tebaldi et al. “Quantifying uncertainty in projections of regional climate change: A Bayesian approach to the analysis of multimodel ensembles”. In: *Journal of Climate* 18.10 (2005), pp. 1524–1540.
- [41] Karl-Hermann Wieners et al. *MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 ScenarioMIP ssp585*. 2019. DOI: [10.22033/ESGF/CMIP6.6705](https://doi.org/10.22033/ESGF/CMIP6.6705). URL: <https://doi.org/10.22033/ESGF/CMIP6.6705>.
- [42] Xiaoge Xin et al. *BCC BCC-CSM2MR model output prepared for CMIP6 ScenarioMIP ssp585*. 2019. DOI: [10.22033/ESGF/CMIP6.3050](https://doi.org/10.22033/ESGF/CMIP6.3050). URL: <https://doi.org/10.22033/ESGF/CMIP6.3050>.
- [43] Seiji Yukimoto et al. *MRI MRI-ESM2.0 model output prepared for CMIP6 ScenarioMIP ssp585*. 2019. DOI: [10.22033/ESGF/CMIP6.6929](https://doi.org/10.22033/ESGF/CMIP6.6929). URL: <https://doi.org/10.22033/ESGF/CMIP6.6929>.
- [44] Tilo Ziehn et al. *CSIRO ACCESS-ESM1.5 model output prepared for CMIP6 ScenarioMIP ssp585*. 2019. DOI: [10.22033/ESGF/CMIP6.4333](https://doi.org/10.22033/ESGF/CMIP6.4333). URL: <https://doi.org/10.22033/ESGF/CMIP6.4333>.