**METHODS PAPER**

# A dependent multimodel approach to climate prediction with Gaussian processes

Marten Thompson[1],[*] , Amy Braverman[2] and Snigdhansu Chatterjee[1]

[1]School of Statistics, University of Minnesota, Minneapolis, Minnesota, USA
[2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA
*Corresponding author. E-mail: thom7058@umn.edu

## 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 atmospheric science community. One approach is to assign equal weight to each model (Knutti, 2010). A more tailored approach is taken in Braverman et al. (2017) and Chatterjee (2019), where models are assigned scores based on their

performance in capturing probabilistic properties of observed climate data. Bayesian approaches with uncertainty quantification may be found in Tebaldi et al. (2005) and Smith et al. (2009). Other approaches for using information from multiple climate models may be found in Flato et al. (2014), Craigmile et al. (2017), Knutti et al. (2017), Sanderson et al. (2017), and Abramowitz et al. (2019). 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.

A fundamental scientific reason for studying multimodel 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 (Masson and Knutti, 2011). 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 (EB) 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 (Eyring et al., 2016). 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 SSPs describe a spectrum of possible futures in terms of human energy use, population, and sentiment toward climate stewardship (Riahi et al., 2017). 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 preindustrial levels (Intergovernmental Panel On Climate Change, 2007). We analyze model output under SSP-5 with a radiative forcing value of 8.5 watts/ $m^2$, the most aggressive increase. The SSP-5 narrative describes an increase in 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 (Bentsen et al., 2019; Dix et al., 2019; Good et al., 2019; Lijuan, 2019; Rong, 2019; Schupfner et al., 2019; Seland et al., 2019; Semmler et al., 2019; Shiogama et al., 2019; Swart et al., 2019a,b; Tachiiri et al., 2019; Wieners et al., 2019; Xin et al., 2019; Yukimoto et al., 2019; Ziehn et al., 2019; NASA Goddard Institute for Space Studies (NASA/GISS), 2020). 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 global monthly values from 1961 to 1990 (Jones et al., 1999). Constructing anomaly data and deseasonalizing it with a baseline of the averages over a 30-year time window is standard in climate data science. A 30-year period is often considered long enough for averaging out high-frequency weather-related variations, and the data between 1960 and 1990 is considered reliable, having adequate global coverage and modern enough for use as a baseline (Jones et al., 1999; Shaowu et al., 2004; Valev, 2006). With these, we forecast the observed monthly mean global near-surface air temperature, which is available from 1850 January to 2020 December (Morice et al., 2021). In all cases, we employed a simple (unweighted) mean and did not perform any regridding of the data beforehand.

## 2. Methods

### *2.1. Hierarchical Bayesian model*

The 17 CMIP6 simulations produce monthly global near-surface air temperatures from 1850 to 2100. Let $Y_{\ell,t}$ be the deseasonalized, mean-centered time series from $\ell$th CMIP6 climate simulation for $\ell = 1,\ldots,L = 17$ and month $t = 1,\ldots,T = 3{,}012$, where $t = 1$ corresponds to 1850 January. The observed

monthly values through 2020 are denoted by $Y_{0,t}$. The vector $\boldsymbol{Y}_t$ will represent all values at month $t$, and $\boldsymbol{Y}_{\text{CMIP},t}$ will represent the $L$ climate simulations, that is,

$$
\boldsymbol{Y}_t = \begin{bmatrix} Y_{0,t} \\ Y_{1,t} \\ \vdots \\ Y_{L,t} \end{bmatrix}; \quad \boldsymbol{Y}_{\text{CMIP},t} = \begin{bmatrix} Y_{1,t} \\ \vdots \\ Y_{L,t} \end{bmatrix}. \tag{1}
$$

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

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

where $\mathbf{1}_{L+1}$ is an $L+1$-dimensional vector of 1's, and $\boldsymbol{U}_t$ 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 $\boldsymbol{U}_t$ 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 $\boldsymbol{U}_t$: 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 $\boldsymbol{U}_t$. 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 EB 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 $\alpha$. The choice of kernel defines the class of functions over which we place our Gaussian process prior; here we have opted to use the squared exponential kernel. Since we have deseasoned our data, we may forgo explicitly encoding seasonality in the kernel (as in, e.g., Williams and Rasmussen, 2006, section 5.4). We also wanted to avoid placing any strict shape restrictions, for example, linear or polynomial. Finally, the Matérn $5/2$ kernel is a common alternative to the squared exponential, but our experimentation did not find a meaningful difference between the two. Please see Section 3 of the Supplementary Material for more details. The squared exponential kernel is parameterized by $\alpha = (\sigma^2, \gamma)$, the variance and lengthscale parameters. That is, $k_\alpha(t_1, t_2) = \sigma^2 e^{-\gamma(t_1 - t_2)^2}$. This and equation (2) define the hierarchical model given below:

$$
\begin{aligned}
\boldsymbol{Y}_t \,|\, \mu(t), \Sigma &\sim N_{L+1}(\mu(t)\mathbf{1}_{L+1}, \Sigma), \\
\mu(t) &\sim \mathcal{GP}(0, k_\alpha).
\end{aligned} \tag{3}
$$

## 2.2. Estimation: 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 (3) 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 $\underset{\sim}{\boldsymbol{Y}}$ 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,

$$
\begin{aligned}
\underset{\sim}{\boldsymbol{Y}} \,|\, \Sigma, K &\sim MN_{(L+1),T}\big(\mathbf{0}_{(L+1)\times T}, \Sigma, K\big), \text{ equivalently} \\
\text{vec}\Big(\underset{\sim}{\boldsymbol{Y}}\Big) \,|\, \Sigma, K &\sim N_{(L+1)\cdot T}\big(\mathbf{0}_{(L+1)\cdot T}, \Sigma \otimes K\big)
\end{aligned}, \tag{4}
$$

where $\otimes$ represents the Kronecker product.

With equation (4), standard properties of the multivariate normal distribution provide pointwise predictions and uncertainty quantification via its conditional mean and conditional variance for any subset of $\mathrm{vec}\left(\underset{\sim}{Y}\right)$. 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 $Y_{\ell,0:t_b}$ $\ell = 1,\ldots,L$ is also multivariate normal.

Unfortunately, the matrix $\Sigma \otimes K$ is quite large, with dimension $(L+1) \cdot T$ by $(L+1) \cdot T$. Inverting this is cubic in complexity and infeasible for large $L$ or $T$. Also, the manipulation of this size of matrix also incurs numerical overflow issues for sufficiently large dimension. For this final reason in particular we find it necessary to consider alternative estimation procedures.

### 2.3. Estimation: EB approach

While the vector Gaussian process model would lend itself to full Bayesian modeling, we propose an EB approach in order to reduce the computational overhead. Here, 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 $\overline{Y}_t$. That is, define $\overline{Y}_t = 1/(L+1)\sum_{\ell=0}^{L} Y_{\ell,t}$. Then, the parameters $\alpha = (\sigma^2, \gamma)$ are estimated by maximizing the marginal likelihood of $\overline{Y}_t$ 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 $\hat{U}_t = \left[\hat{U}_{0,t} \hat{U}_{1,t} \ldots \hat{U}_{L,t}\right]^T$. Use $\hat{U}_1,\ldots,\hat{U}_T$ 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

$$\hat{\Sigma} = \begin{pmatrix} \hat{\sigma}_0^2 & \hat{\Sigma}_0^T \\ \hat{\Sigma}_0 & \hat{\Sigma}_{\mathrm{CMIP}} \end{pmatrix}, \tag{5}$$

where $\hat{\sigma}_0^2$ is the variance of the observed data's random effects, and $\hat{\Sigma}_{\mathrm{CMIP}}$ 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 the months $t$ spanning 2021 to 2100 given the model outputs and historic observations. The distribution of these values follows from equation (3) and standard properties of the normal distribution. Given $\hat{\mu}(t)$ and simulated values $\boldsymbol{Y}_{\mathrm{CMIP},t}$

$$Y_{0,t} \mid \hat{\mu}(t), \boldsymbol{Y}_{\mathrm{CMIP},t} \sim N(m^*, \gamma^*), \tag{6}$$

$$m^* = \hat{\mu}(t) + \Sigma_0^T \Sigma_{\mathrm{CMIP}}^{-1}(\boldsymbol{Y}_{\mathrm{CMIP},t} - \hat{\mu}(t)\boldsymbol{1}_L), \tag{7}$$

$$\gamma^* = \sigma_0^2 - \Sigma_0^T \Sigma_{\mathrm{CMIP}}^{-1}\Sigma_0. \tag{8}$$

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: 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 typical example, Figure 1 depicts the results when the ACCESS-CM2 CMIP6 simulation is being treated like the observed time series. Our choice of highlighting this particular example is arbitrary; please see the Supplementary Material for all such cases. 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
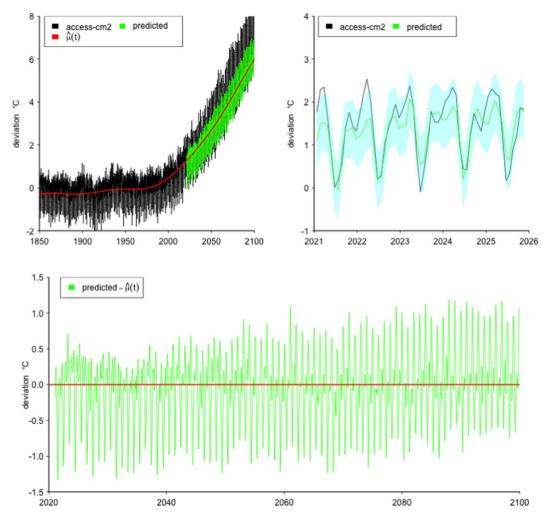


**Figure 1.** *Amongst others, we used the ACCESS-CM2 CMIP6 time series to evaluate our model's predictive accuracy. The top left figure shows the whole of the training and test intervals, 1850 January– 2020 November and 2020 December–2100 December, respectively. Just 5 years are shown in the top right figure, starting in 2021 January, to highlight the prediction's month-to-month accuracy. Shading (blue) indicates two standard deviations from the conditional mean, as found in equations (7) and (8). 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 $\hat{\mu}(t)$. The correlation terms "pull" the predicted values away from $\hat{\mu}(t)$ in a manner that reflects how the held-out time series' historic values differed from the others. In the case of ACCESS-CM2, this manifests in the predicted (green) trending above $\hat{\mu}(t)$ (red), which better matches the true future values.*

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 (7)) ameliorates this trend.

Table 1 contains a summary of the LOO results. It includes the mean squared error (MSE) between the actual output of the indicated simulation and the predictions of our EB model. We compare our method to the predictions found using just the common mean function $\hat{\mu}(t)$ and using the pointwise average $\overline{Y}_t$. The former amounts to only fitting the Gaussian process, omitting the random effects $U(t)$. The pointwise average merely takes the means of all training forecasts. The empirical Bayes model has a lower MSE than the common mean in every case.

### 3.2. Observed time series

Making predictions for the observed climate follows the procedure 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 Section 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 a variance of 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 (5); its

**Table 1.** Each CMIP6 simulation was treated like the observed time series, truncated to 2020 November, and predicted by various methods.

| LOO Sim | $MSE_{EB}$ | $MSE_{\hat{\mu}}$ | $MSE_{\overline{Y}_t}$ |
|---|---|---|---|
| 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 |

*Note.* The predictions produced by our empirical Bayes model are similar in accuracy to using $\overline{Y}_t$ 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.
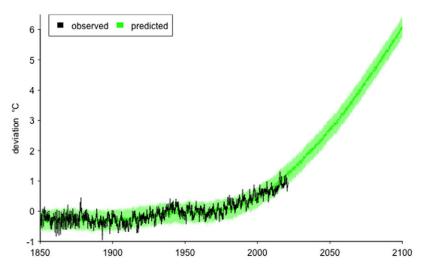
**Figure 2.** *Predicted values for mean global surface temperature using* equation (7)*, with shading to $\pm 2$ standard deviations (*equation (8)*). This model leverages the common mean as well as individual correlations amongst CMIP6 simulations and historical observed data.*

attenuating presence in equation (7) explains how the predictions in Figure 2 show less variation than those in Figure 1.

## 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 vis–a-vis a common mean. An EB 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.

**Provenance.** This article is part of the Climate Informatics 2022 proceedings and was accepted in *Environmental Data Science* on the basis of the Climate Informatics peer review process.

**Supplementary Materials.** To view supplementary material for this article, please visit http://doi.org/10.1017/eds.2022.24.

# References

**Abramowitz G**, **Herger N**, **Gutmann E**, **Hammerling D**, **Knutti,R**, **Leduc M**, **Lorenz R**, **Pincus R, and Schmidt GA** (2019) ESD reviews: Model dependence in multi-model climate ensembles: Weighting, sub-selection and out-of-sample testing. *Earth System Dynamics 10*(1), 91–105.

**Bentsen M**, **Olivie DJL**, **Seland Ø**, **Toniazzo T**, **Gjermundsen A**, **Graff LS**, **Debernard JB**, **Gupta AK**, **He Y**, **Kirkevag A**, **Schwinger J**, **Tjiputra J**, **Aas KS**, **Bethke I**, **Fan Y**, **Griesfeller J**, **Grini A**, **Guo C**, **Ilicak M**, **Karset IHH**, **Landgren OA**, **Liakka J**, **Moseid KO**, **Nummelin A**, **Spensberger C**, **Tang H**, **Zhang Z**, **Heinze C**, **Iversen T, and Schulz M** (2019) NCC NorESM2-MM model output prepared for CMIP6 ScenarioMIP ssp585. https://doi.org/10.22033/ESGF/CMIP6.8321

**Braverman A**, **Chatterjee S**, **Heyman M, and Cressie N** (2017) Probabilistic evaluation of competing climate models. *Advances in Statistical Climatology, Meteorology and Oceanography 3*(2), 93–105.

**Chatterjee S** (2019) The scale enhanced wild bootstrap method for evaluating climate models using wavelets. *Statistics & Probability Letters 144*, 69–73.

**Craigmile PF**, **Guttorp P and Berrocal V** (2017) Regional climate model assessment via spatio-temporal modeling. *Environmetrics 23*, 482–492.

**Dix M**, **Bi D**, **Dobrohotoff P**, **Fiedler R**, **Harman I**, **Law R**, **Mackallah C**, **Marsland S**, **O'Farrell S**, **Rashid H**, **Srbinovsky J**, **Sullivan A**, **Trenham C**, **Vohralik P**, **Watterson I**, **Williams G**, **Woodhouse M**, **Bodman R**, **Dias FB**, **Domingues C**, **Hannah N**, **Heerdegen A**, **Savita A**, **Wales S**, **Allen C**, **Druken K**, **Evans B**, **Richards C**, **Ridzwan SM**, **Roberts D**, **Smillie J**, **Snow K**, **Ward M, and Yang R** (2019) CSIRO-ARCCSS ACCESS-CM2 model output prepared for CMIP6 ScenarioMIP ssp585. https://doi.org/10.22033/ESGF/CMIP6.4332

**Eyring V**, **Bony S**, **Meehl GA**, **Senior CA**, **Stevens B**, **Stouffer RJ, and Taylor KE** (2016) Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development 9*(5), 1937–1958.

**Flato G**, **Marotzke J**, **Abiodun B**, **Braconnot P**, **Chou SC**, **Collins W**, **Cox P**, **Driouech F**, **Emori S, and Eyring V** (2014) 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: Cambridge University Press, pp. 741–866.

**Good P**, **Sellar A**, **Tang Y**, **Rumbold S**, **Ellis R**, **Kelley D, and Kuhlbrodt T** (2019) MOHC UKESM1.0-LL model output prepared for CMIP6 ScenarioMIP ssp585. https://doi.org/10.22033/ESGF/CMIP6.6405

**Intergovernmental Panel On Climate Change** (2007) Climate change 2007: The physical science basis. *Agenda 6*(07), 333.

**Jones PD**, **Horton E**, **Folland C**, **Hulme M**, **Parker D, and Basnett T** (1999) The use of indices to identify changes in climatic extremes. *Climatic Change 42*(1), 131–149.

**Jones PD**, **New M**, **Parker DE**, **Martin S, and Rigor IG** (1999) Surface air temperature and its changes over the past 150 years. Available at https://crudata.uea.ac.uk/cru/data/temperature/abs_glnhsh.txt (accessed 29 November 2021).

**Knutti R** (2010) The end of model democracy? *Climatic Change 102*(3), 395–404.

**Knutti R**, **Sedlacek J**, **Sanderson BM**, **Lorenz R**, **Fischer EM, and Eyring V** (2017) A climate model projection weighting scheme accounting for performance and interdependence. *Geophysical Research Letters 44*(4), 1909–1918.

**Lijuan L** (2019) *CAS FGOALS-g3 model output prepared for CMIP6 ScenarioMIP ssp585*. https://doi.org/10.22033/ESGF/CMIP6.3503

**Masson D and Knutti R** (2011) Climate model genealogy. *Geophysical Research Letters 38*(8), L08703.

**Morice CP**, **Kennedy JJ**, **Rayner NA**, **Winn J**, **Hogan E**, **Killick R**, **Dunn R**, **Osborn T**, **Jones P, and Simpson I** (2021) An updated assessment of near-surface temperature change from 1850: The HadCRUT5 data set. *Journal of Geophysical Research: Atmospheres 126*(3), e2019JD032361.

**NASA Goddard Institute for Space Studies (NASA/GISS)** (2020) *NASA-GISS GISS-E2.1G model output prepared for CMIP6 ScenarioMIP ssp585*. https://doi.org/10.22033/ESGF/CMIP6.7460

**Riahi K**, **Van Vuuren DP**, **Kriegler E**, **Edmonds J**, **O'neill BC**, **Fujimori S**, **Bauer N**, **Calvin K**, **Dellink R, and Fricko O** (2017) The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change 42*, 153–168.

**Rong X** (2019) CAMS CAMS-CSM1.0 model output prepared for CMIP6 ScenarioMIP ssp585. https://doi.org/10.22033/ESGF/CMIP6.11052

**Sanderson BM**, **Wehner M, and Knutti R** (2017) Skill and independence weighting for multi-model assessments. *Geoscientific Model Development 10*(6), 2379–2395.

**Schupfner M**, **Wieners KH**, **Wachsmann F**, **Steger C**, **Bittner M**, **Jungclaus J**, **Fruh B**, **Pankatz K**, **Giorgetta M**, **Reick C**, **Legutke S**, **Esch M**, **Gayler V**, **Haak H**, **de Vrese P**, **Raddatz T**, **Mauritsen T**, **von Storch JS**, **Behrens J**, **Brovkin V**, **Claussen M**, **Crueger T**, **Fast I**, **Fiedler S**, **Hagemann S**, **Hohenegger, C**, **Jahns, T**, **Kloster, S**, **Kinne, S**, **Lasslop, G**, **Kornblueh, L**, **Marotzke, J**, **Matei D**, **Meraner K**, **Mikolajewicz U**, **Modali K**, **Muller W**, **Nabel J**, **Notz D**, **Peters-von Gehlen K**, **Pincus R**,

**Pohlmann H**, **Pongratz J**, **Rast S**, **Schmidt H**, **Schnur R**, **Schulzweida U**, **Six K**, **Stevens B**, **Voigt A, and Roeckner E** (2019) *DKRZ MPI-ESM1.2-HR model output prepared for CMIP6 ScenarioMIP ssp585*. https://doi.org/10.22033/ESGF/CMIP6.4403

**Seland Ø**, **Bentsen M**, **Olivie DJL**, **Toniazzo T**, **Gjermundsen A**, **Graff LS**, **Debernard JB**, **Gupta AK**, **He Y**, **Kirkevag A**, **Schwinger J**, **Tjiputra J**, **Aas KS**, **Bethke I**, **Fan Y**, **Griesfeller J**, **Grini A**, **Guo C**, **Ilicak M**, **Karset IHH**, **Landgren OA**, **Liakka J**, **Moseid KO**, **Nummelin A**, **Spensberger C**, **Tang H**, **Zhang Z**, **Heinze C**, **Iversen T, and Schulz M** (2019) *NCC NorESM2-LM model output prepared for CMIP6 ScenarioMIP ssp585*. https://doi.org/10.22033/ESGF/CMIP6.8319

**Semmler T**, **Danilov S**, **Rackow T**, **Sidorenko D**, **Barbi D**, **Hegewald J**, **Pradhan HK**, **Sein D**, **Wang Q, and Jung T** (2019) AWI AWI-CM1.1MR model output prepared for CMIP6 ScenarioMIP ssp585. https://doi.org/10.22033/ESGF/CMIP6.2817

**Shaouw W**, **Jinhong Z, and Jingning C** (2004) Interdecadal variability of temperature and precipitation in China since 1880. *Advances in Atmospheric Sciences 21*(3), 307–313.

**Shiogama H**, **Abe M, and Tatebe H** (2019) MIROC MIROC6 model output prepared for CMIP6 ScenarioMIP ssp585. https://doi.org/10.22033/ESGF/CMIP6.5771

**Smith RL**, **Tebaldi C**, **Nychka D, and Mearns LO** (2009) Bayesian modeling of uncertainty in ensembles of climate models. *Journal of the American Statistical Association 104*(485), 97–116.

**Swart NC**, **Cole JN**, **Kharin VV**, **Lazare M**, **Scinocca JF**, **Gillett NP**, **Anstey J**, **Arora V**, **Christian JR**, **Jiao Y**, **Lee WG**, **Majaess F**, **Saenko OA**, **Seiler C**, **Seinen C**, **Shao A**, **Solheim L**, **von Salzen K**, **Yang D**, **Winter B, and Sigmond M** (2019a) *CCCma CanESM5 model output prepared for CMIP6 ScenarioMIP ssp585*. https://doi.org/10.22033/ESGF/CMIP6.3696

**Swart NC**, **Cole JN**, **Kharin VV**, **Lazare M**, **Scinocca JF**, **Gillett NP**, **Anstey J**, **Arora V**, **Christian JR**, **Jiao Y**, **Lee WG**, **Majaess F**, **Saenko OA**, **Seiler C**, **Seinen C**, **Shao A**, **Solheim L**, **von Salzen K**, **Yang D**, **Winter B, and Sigmond M** (2019b) *CCCma CanESM5-CanOE model output prepared for CMIP6 ScenarioMIP ssp585*. https://doi.org/10.22033/ESGF/CMIP6.10276

**Tachiiri K**, **Abe M**, **Hajima T**, **Arakawa O**, **Suzuki T**, **Komuro Y**, **Ogochi K**, **Watanabe M**, **Yamamoto A**, **Tatebe H**, **Noguchi MA**, **Ohgaito R**, **Ito A**, **Yamazaki D**, **Ito A**, **Takata K**, **Watanabe S, and Kawamiya M** (2019) MIROC MIROC-ES2L model output prepared for CMIP6 ScenarioMIP ssp585. https://doi.org/10.22033/ESGF/CMIP6.5770

**Tebaldi C**, **Smith RL**, **Nychka D, and Mearns LO** (2005) Quantifying uncertainty in projections of regional climate change: A Bayesian approach to the analysis of multimodel ensembles. *Journal of Climate 18*(10), 1524–1540.

**Valev D** (2006) Statistical relationships between the surface air temperature anomalies and the solar and geomagnetic activity indices. *Physics and Chemistry of the Earth 31*, 109–112.

**Wieners K-H**, **Giorgetta M**, **Jungclaus J**, **Reick C**, **Esch M**, **Bittner M**, **Gayler V**, **Haak H**, **de Vrese P**, **Raddatz T**, **Mauritsen T**, **von Storch JS**, **Behrens J**, **Brovkin V**, **Claussen M**, **Crueger T**, **Fast I**, **Fiedler S**, **Hagemann S**, **Hohenegger C**, **Jahns T**, **Kloster S**, **Kinne S**, **Lasslop G**, **Kornblueh L**, **Marotzke J**, **Matei D**, **Meraner K**, **Mikolajewicz U**, **Modali K**, **Muller W**, **Nabel J**, **Notz D**, **Peters-von Gehlen K**, **Pincus R**, **Pohlmann H**, **Pongratz J**, **Rast S**, **Schmidt H**, **Schnur R**, **Schulzweida U**, **Six K**, **Stevens B**, **Voigt A, and Roeckner E** (2019) MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 ScenarioMIP ssp585. https://doi.org/10.22033/ESGF/CMIP6.6705

**Williams CK and Rasmussen CE** (2006) *Gaussian Processes for Machine Learning*, vol. *2*. Cambridge, MA: MIT press, p. 3.

**Xin X**, **Wu T**, **Shi X**, **Zhang F**, **Li J**, **Chu M**, **Liu Q**, **Yan J**, **Ma Q, and Wei M** (2019) BCC BCC-CSM2MR model output prepared for CMIP6 ScenarioMIP ssp585. https://doi.org/10.22033/ESGF/CMIP6.3050

**Yukimoto S**, **Koshiro T**, **Kawai H**, **Oshima N**, **Yoshida K**, **Urakawa S**, **Tsujino H**, **Deushi M**, **Tanaka T**, **Hosaka M**, **Yoshimura H**, **Shindo E**, **Mizuta R**, **Ishii M**, **Obata A, and Adachi Y** (2019) *MRI MRI-ESM2.0 model output prepared for CMIP6 ScenarioMIP ssp585*. https://doi.org/10.22033/ESGF/CMIP6.6929

**Ziehn T**, **Chamberlain M**, **Lenton A**, **Law R**, **Bodman R**, **Dix M**, **Wang Y**, **Dobrohotoff P**, **Srbinovsky J**, **Stevens L**, **Vohralik P**, **Mackallah C**, **Sullivan A**, **O'Farrell S, and Druken K** (2019) *CSIRO ACCESS-ESM1.5 model output prepared for CMIP6 ScenarioMIP ssp585*. https://doi.org/10.22033/ESGF/CMIP6.4333