https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

METEORv1.0.1: A novel framework for emulating multi-timescale regional climate responses

Marit Sandstad[1], Norman Julius Steinert[1], Susanne Baur[2,3], and Benjamin Mark Sanderson[1]
[1]CICERO Center for International Climate Research, Oslo 0349, Norway
[2]CECI, Université de Toulouse, CERFACS, CNRS, Toulouse, France
[3]CNRM, Université de Toulouse, Météo-France/CNRS, Toulouse, France
Correspondence: Marit Sandstad (marit.sandstad@cicero.oslo.no)

Abstract. Resolved spatial information for climate change projections is critical to any robust assessment of climate impacts and adaptation options. However, the range of spatially resolved future scenario assessments available is limited, due to the significant computational and human demands of Earth System Model (ESM) pipelines. In order to explore a wider variety of societal outcomes and to enable coupling of climate impacts into societal modeling frameworks, rapid spatial emulation of ESM response is therefore desirable. Existing linear pattern scaling methods assume spatial climate signals which scale linearly with global temperature change, where the pattern of response is independent of the nature and timing of emissions. However, this assumption may introduce biases in emulated climates, especially under net negative emissions and overshoot scenarios. To address these biases, we propose a novel emulation system, METEOR, which represents multi-timescale spatial climate responses to multiple climate forcers. The mapping of emissions to forcing is provided by the CICERO Simple Climate Model, combined with a calibration system which can be used to train model-specific pattern response engines using only core training simulations from CMIP. Here, we demonstrate that our fitted spatial emulation system is capable of rapidly and accurately predicting gridded responses to out-of-sample scenarios.

1 Introduction

Spatially resolved information is essential for informing robust assessments of mitigation and adaptation strategies in response to global climate change (IPCC, 2021, 2022). Accurate and detailed regional projections enable policymakers and stakeholders to understand potential impacts and to plan accordingly (IPCC, 2022). The Coupled Model Intercomparison Project (CMIP) aims to deliver this information, and is increasingly moving towards an operational procedure, wherein CMIP7 Earth System Models will be run on a semi-regular frequency, allowing updates of model complexity, historical forcing and future scenarios (Dunne et al., 2024). However, these pipelines are time-consuming and computationally intensive, resolving numerous physical, chemical, and biological processes at high spatial and temporal resolutions such that a modest number of scenario simulations with multiple models takes years to achieve (Eyring et al., 2016). Demand for regional climate information increasingly

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

requires more frequent updates for policy-relevant future scenarios. Limitations on time, computational and human resources needed for ESM simulations constrain the range of future scenarios and models that can be feasibly explored (Nicholls et al., 2020). In addition, there is increased need for fast spatial modeling frameworks where regional climate impacts are resolved in Integrated Assessment Models, to allow for the simulation of inequalities and impacts on social and natural systems (Ferrari et al., 2022).

To address these needs, spatial emulation techniques have been developed to rapidly approximate the output of ESMs under various scenarios (Zelazowski et al., 2018). Linear pattern scaling is one such approach, which assumes that spatial patterns of climate response scale linearly with global mean temperature change (Santer et al., 1990; Mitchell, 2003; Beusch et al., 2020). This method allows for quick estimations of regional climate change by scaling predefined spatial patterns according to projected global temperature changes (Tebaldi et al., 2021; Zhao et al., 2017).

Several pattern scaling techniques have been proposed and utilised in climate research. The SCENGEN tool, for instance, generates regional climate change scenarios by scaling standardised patterns derived from General Circulation Models (GCMs) (Hulme et al., 2000). Similarly, the MESMER framework employs pattern scaling to efficiently emulate temperature and precipitation fields from ESMs (Beusch et al., 2020, 2022). Likewise, the PRIME framework makes use of pattern scaling and subsequently provides the probabilistic spatial climate information as input for a Land surface model (Mathison et al., 2024), allowing more direct simulation of terrestrial ecosystems and downstream human and economic impacts. The STITCHES model uses a different approach, by splicing together portions of existing simulations with global mean temperatures corresponding to the desired prediction to emulate scenarios which have not yet been simulated (Tebaldi et al., 2022). This allows the model to represent, for example, differences between resolved patterns under warming and cooling climate states, but is limited by the finite number of climate analogs in the training dataset - which in CMIP is relatively sparse, and requires concatenation of segments of simulation which may produce unphysical step changes in the emulated climate. Further, the ESEm framework (Watson-Parris et al., 2021) provides options to build various types of more process agnostic machine learning based emulators for spatio-temporally resolved data, using approaches such as Random Forest, Gaussian Process and Neural Network regressions. Comparison of these methods to traditional linear pattern scaling shows good performance (Watson-Parris et al., 2022), but the setup used requires output which is not widely available for many models. Similarly, Mansfield et al. (2020) utilised the specific per species modelling intercomparison setup and output in PDRMIP (Myhre et al., 2017), to define short- and long-term ESM responses using Gaussian Process regression, which can then in turn be used to emulate long-term responses to new scenarios from short-term responses. Due to the data requirements for training to new models, this setup, though interesting, is not immediately applicable to CMIP6 or new emerging ESM datasets.

While each of these methods have increased capacity to produce regional climate projections with reduced computational demands, questions remain on how to emulate hysteresis, forcing dependency and nonlinear responses in future climate, which have been demonstrated to exist in Earth System Models but can be precluded by emulator assumptions. Models such as MESMER rely on the assumption that spatial patterns of climate response are a singular function of global mean temperature, insensitive to the forcing history and the emission trajectory (Collins et al., 2013; Zhao et al., 2017). Patterns of warming in PRIME are also subject to this linearity assumption, though the process-based land surface component could potentially

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

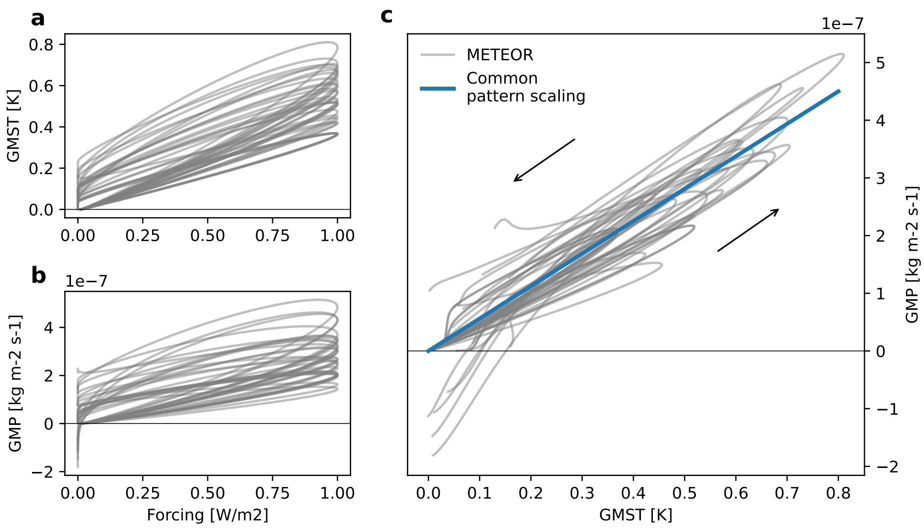

Figure 1. METEOR pattern scaling hysteresis. Illustration of a) global mean temperature (GMST) and b) precipitation (GMP) response to an idealised radiative forcing with a Gaussian ramp-up to 1 W/m^2 and subsequent ramp-down within 500 years. c) shows the relationship between GMST and GMP, which various patterns scaling approaches are based on. METEOR (grey), trained by CMIP6 models (see Methods), is able to emulate hysteresis behavior, compared to a common pattern scaling (blue) that uses linear regression of the relationship between GMST and GMP. Each individual grey line shows the response of METEOR trained to emulate a specific CMIP6 model.

represent memory in slow-timescale terrestrial processes. STITCHES can potentially resolve non-linear and time-emergent behavior to the degree that behavior is represented in the training scenarios, but is limited to the degree it can generalise hysteresis and climate reversibility dynamics. This linearity assumption allows reasonable performance under scenarios of gradual and monotonic climate change but can introduce significant biases under strong mitigation scenarios or scenarios involving overshoots in greenhouse gas concentrations (Herger et al., 2015; Good et al., 2015; Tebaldi and Knutti, 2007). Additionally, non-linearities in the climate system, such as feedback mechanisms and varying climate sensitivities over time, can lead to time-evolving and forcing-dependent spatial patterns that are not adequately captured by traditional pattern scaling approaches (Huntingford et al., 2000; Shiogama et al., 2010).

To address these limitations, we propose METEOR (Multivariate Emulation of Time-Evolving and Overlapping Responses), a novel emulation framework that accounts for spatial climate responses to a range of climate forcers emerging over different timescales by using impulse response assumptions applied to a spatially resolved basis set. Figure 1 shows how METEOR can produce non-linear and hysteresis responses to an idealised forcing trajectory. METEOR builds on the emissions-forcing engine from the CICERO Simple Climate Model (C-SCM; Sandstad et al., 2024), and can represent the temporal evolution and forcing dependency of spatial patterns, providing regional climate projections which preserve the hysteresis and time-evolving response to forcing present in the target Earth System Model, allowing more consistent representation of scenarios involving overshoots or pathways with diverse mixes of short and long lived climate forcers.

In this paper, we present the structure and validation of METEOR. We demonstrate how this approach can capture nonlinear and time-dependent aspects of regional climate change. As such, METEOR offers a practical tool for researchers to rapidly explore a wide variety of societal outcomes and to assess mitigation and adaptation options informing policymakers with greater confidence.

2 Methods

The methodology for METEOR is illustrated in Figure 2, and described below. The framework of METEOR is based on the assumption that a step change in a given climate forcer can induce a number of time-evolving patterns for a predicted output variable, each of which emerges on a specific timescale. The METEOR framework allows groups of climate forcers to be associated with a number of time-evolving pattern responses, which in the current version then combine linearly to give the total climate response.

The primary forcer component of climate change is the greenhouse gas (GHG) signal due to well-mixed greenhouse gases, which exert a forcing on the climate system. For METEOR, we assume that the time-evolving pattern of response to GHG forcing can be approximated by the response of the climate system to a step change in CO_2 concentrations. This is practical given the ready availability of *abrupt4x-*CO_2 simulations (in which atmospheric carbon dioxide levels are instantaneously quadrupled from pre-industrial levels and the system is allowed to evolve for at least 140 years) for all Earth System Models in the CMIP archive, and was found to be a reasonable approximation in prior process modeling studies which considered the pulse-response to a number of different greenhouse gases (Myhre et al., 2017).

METEOR assumes that the response to a step change in forcing from a given source can be represented by the sum of one or more impulse response patterns, each with its own timescale of emergence as represented by a decaying exponential timeseries for pattern saturation. Each timescale and corresponding pattern will capture elements of the physical response which emerge at different timescales, such that different spatial patterns can be associated, for example, with the warming of the shallow and deep ocean. Similarly, some forcing agents such as sulfate aerosols and black carbon are associated with markedly different warming patterns and timescales to those of well-mixed greenhouse gases (Myhre et al., 2017). The METEOR framework allows for individual forcers, or groups of forcers, to be associated with their own set of time-emergent patterns, allowing for the model to simulate and distinguish between the spatial pattern of climate change associated with different forcer types. As such, METEOR can be trained employing only steps 1-4 (Fig. 2), for a single forcer version, or steps 1-4 can be repeated for multiple forcers for which abrupt step-change forcing experiment data are available.

In practice, METEOR uses outputs of the *abrupt4x-*CO_2 experiment to fit the GHG response. Then, as separate step-change experiment data is not generally available in CMIP for other forcers, a residual signal from a *historical* and single *SSP* scenario run is used for inverse estimation of patterns and timescales for a sulfate aerosol response. We choose this technique to pick up an estimation of sulfate aerosol only, as this is the most different and impactful non-GHG forcer (Myhre et al., 2017). See also figure 7.7 of (Forster et al., 2021) where the aerosol forcing is the most uncertain and largest non-GHG contribution. In the setup we use, the totality of the aerosol-cloud-interaction is modelled using sulfate aerosol forcing, and a substantial part

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

of the aerosol radiation forcing comes from sulfate forcing. A further breakdown of forcer types requires additional dedicated experiments. For convenience, we will denote this part of the pattern as simply aerosol patterns, although it is both more specific (only fitted for sulfate aerosol) and less so (as it is a residual pattern, and will naturally also pick up other non-sulfate or even non-aerosol patterns). The obtained responses and patterns can then be used to emulate the spatial response to a previously unseen experiment for which there is emissions (or concentrations) data available.

This METEOR code framework is available at https://github.com/benmsanderson/METEOR and is importable as a Python library. The methodology does not come pre-trained, but includes tools to read local model output training data, and supporting functionality to download data on appropriate formats. In this article we present results for a selection of CMIP6 models applied to yearly temperature and precipitation, but the methodology is not limited to these variables, or to the set of ESM model output used here. Indeed any yearly variable output can in principle be emulated, though, the reliability of the fit must be assessed by the user. The computational time for training is relatively fast (order of less than minutes on a laptop), but performance depends on the resolution of the ESM model target for emulation, and local machine specifications including memory limitations. Overall the code allows for efficient emulation of models in the CMIP6 archive and computation and application to a variety of climate scenarios.

2.1 Training the model: Construction of transient spatial response patterns

In this section we describe how we use the *abrupt-4x*CO_2 experiment to find impulse response timescales and patterns for CO_2 that can be used to recreate and model GHG forcing response in various experiments. A combination of *historical* and scenario data is then employed to estimate separate timescales and patterns for sulfate aerosol forcing. These estimates are then added linearly to the GHG forcing and patterns to provide the composite response.

2.1.1 Greenhouse gas response estimation

We begin by obtaining the annual mean outputs for the target ESM we wish to emulate. The variables of interest in this study are surface air temperature (*tas*) and precipitation (*pr*). To obtain a greenhouse gas response signal, we use the *abrupt4x-*CO_2 experiment, in which atmospheric CO_2 concentrations are instantaneously quadrupled relative to pre-industrial levels, and the *piControl* (pre-industrial control) simulation (Eyring et al., 2016, Fig. 2, step 1). We obtain the radiative forcing time series for GHGs from the *abrupt4x-*CO_2 emissions by using the C-SCM emission-to-forcing module (Sandstad et al., 2024) that converts carbon emissions into climate forcing (Fig. 2, step 2).

For the given target variable, we estimate the gridded climate response to increased CO_2 by calculating the anomaly, subtracting the *piControl* climatology from the *abrupt4x*CO_2 simulation to yield $\mathbf{X}$, a matrix of dimensions $s \times t$, with s being the number of spatial grid points and t the number of years in the simulation with the values for the variable of interest in each point in space and time as its data. From $\mathbf{X}_{s,t}$ we calculate the global mean anomaly time series $\mathbf{x}_{\text{global}}(t)$ by performing an area-weighted average over all spatial grid points:

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

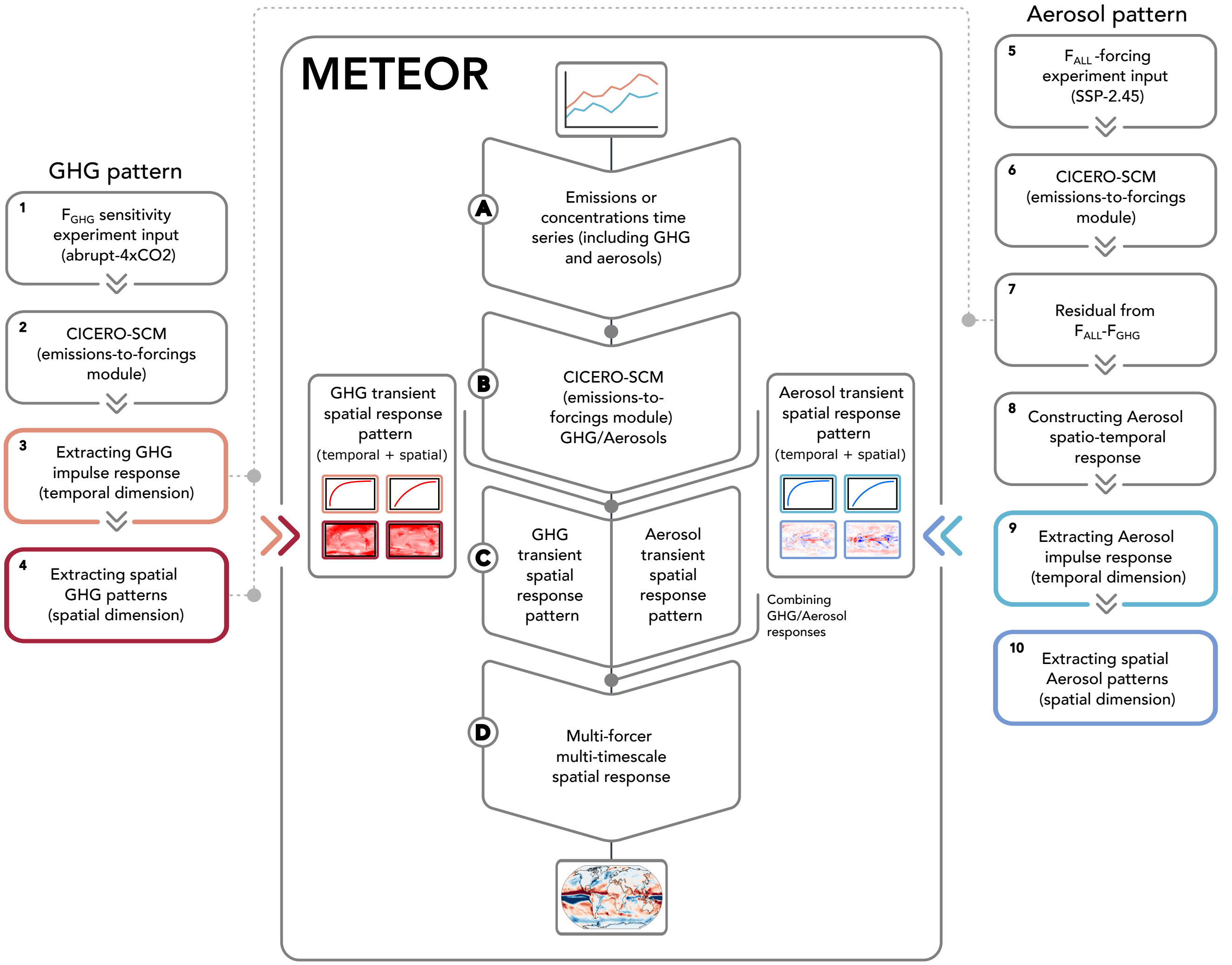

Figure 2. METEOR flowchart. Illustration of the concept and sequence of METEOR and its integration with the CICERO simple climate model. Note that the side panels represent the necessary preparatory steps for the creation of GHG (left) and aerosol (right) transient spatial climate response pattern. Here, the aerosol pattern calculations, specifically step 7, requires input from the GHG pattern calculation steps 3 and 4 (dotted gray line). Hence, steps 1–10 are required to train METEOR (center panels; also see section 2.1). Their outcome is used in step C in METEOR (see section 2.2). Note that METEOR partly integrates the C-SCM as it makes use of its emissions-to-forcings module.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

$$\mathbf{x}_{\text{global}}(t) = \frac{1}{s}\sum_{i=1}^{s} w_i \cdot \mathbf{X}_{i,t}, \tag{1}$$

where w_i is the area weight for pixel i, with a total number of pixels s.

We then find an approximate representation of this global mean response as a sum of n exponential decay functions, representing different climate response timescales:

$$\mathbf{x}_{\text{global}}(t) \approx \sum_{k=1}^{n} a_k \left(1 - e^{-t/\tau_k}\right), \tag{2}$$

where a_k are amplitudes and τ_k are the decay timescales to be determined for mode k.

METEORv1.0 can decompose the global mean response into a user-defined number of timescales, with each added timescale fitted in an exponentially longer and non-overlapping time range so that $\tau_k \in (10^k, 10^{k+1})$, minimised from an initial guess of $\tau_k^{\text{guess}} = 5 \cdot 10^k$. This allows clear separation of the timescales. Three timescales ($n = 3$), which will be referred to as inter-annual mode (1–10 years), inter-decadal mode (10–100 years) and inter-centennial mode (100–1000 years) response, is the configuration we will describe here. Note that choice of number of timescales (n) can be practically informed by assessing the point at which there is no further improvement in performance in the approximation detailed in Eq. 2 (see section 3 for an illustration).

For the timescale decomposition, we construct a matrix $\mathbf{T}_{\text{GHG}}$ of dimensions $n \times n_t$, where each row n corresponds to an exponential decay function with a specific timescale, and n_t refers to the number of simulation years (Fig. 2, step 3). $\mathbf{T}$ represents the temporal evolution of the global mean climate response across different timescales:

$$\mathbf{T}_{\text{GHG}\,k,t} = 1 - e^{-t/\tau_k}, \quad \text{for } k = 1, 2, \ldots, n. \tag{3}$$

Assuming that the spatiotemporal response can be represented as a sum of patterns which each emerge as a saturating exponential decay (Proistosescu and Huybers, 2017), we express the anomaly matrix $\mathbf{X}$ as a product of spatial patterns and temporal basis functions, where $\mathbf{B}$ is the spatial pattern matrix of dimensions $s \times n$:

$$\mathbf{X} = \mathbf{B}_{\text{GHG}} \cdot \mathbf{T}_{\text{GHG}}. \tag{4}$$

This assumption allows the calculation of the spatial response matrix $\mathbf{B}$ (Fig. 2, step 4), given that we have a prior estimate of the timescale matrix $\mathbf{T}$ and the full time-evolving output from the target model, $\mathbf{X}$. Hence, we solve for $\mathbf{B}$ with a least-squares estimate $\mathbf{T}^+$ (Barata and Hussein, 2012) of the matrix $\mathbf{T}$. Consequently, the derived values of $\mathbf{B}$ minimise the residuals between the observed and reconstructed anomalies:

$$\mathbf{B}_{\text{GHG}} = \mathbf{X} \cdot \mathbf{T}_{\text{GHG}}^{+}. \tag{5}$$

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

Finally, we normalise the spatial patterns by the effective radiative forcing $F_{4\times}$ associated with the quadrupling of CO_2 concentrations in the C-SCM, yielding a spatial pulse-response function $\widetilde{\mathbf{B}}_{\text{GHG}} = \frac{\mathbf{B}_{\text{GHG}}}{F_{4\times CO_2}}$ that represents the spatial climate response pattern per unit forcing.

2.1.2 Aerosol response estimation

Though the CO_2 pattern obtained using the methodology described above could in principle be used to emulate forcing from any component, there is reason to believe that other forcers affect various climate variables with a different spatial and temporal signature. In particular, the effect of aerosols, especially sulfate, on temperature and precipitation differs from that of greenhouse gases (Samset et al. (2018, 2019); Zhao et al. (2019); Monerie et al. (2022); Persad et al. (2023)). Estimating the climate response of any forcer species for METEOR in the same way as for CO_2 is possible, but direct estimation requires a species-specific forcing step-change experiment equivalent to the *abrupt4xCO$_2$*. Such experiments have only been performed for a limited number of CMIP5 generation of models (Myhre et al., 2017).

To construct the aerosol response patterns in the CMIP6 emulation, we instead make use of transient experiments to inversely calculate the aerosol spatial pulse response functions, constructing a residual between the ESM output for an all-forcing (F_{ALL}) experiment (where both greenhouse gases and aerosols are varying) and a synthetic GHG forcing (F_{GHG}) experiment, the response for which we estimate using the METEOR GHG pattern response above. This assumes that such a residual is primarily explained by the aerosol climate response. In Section 3 we verify that this assumption yields reasonable results in the scenarios considered for CMIP6.

Similarly to how we found $\mathbf{X}$ for the *abrupt-4xCO$_2$*, we obtain the full ESM scenario response $\mathbf{S}_{s,t'}$ (dimensions $s \times n'$, where n' is now the number of years in the scenario) from the CMIP model output corresponding to the emission scenario which will be used in the training process. For the all-forcing scenario, we utilise a combination of the *historical* experiment and the Shared Socioeconomic Pathway *SSP2-4.5* (Eyring et al., 2016) from CMIP6 ESM output (Fig. 2, step 5), though any experiment including all forcers could in principle be used. Here again, we obtain the aerosol forcing time series from the *historical* and *SSP2-4.5* emissions by using the C-SCM emission-to-forcing module (Fig. 2, step 6).

Further, let $F_{\text{GHG}}(t')$ and $F_{\text{aer}}(t')$ denote the GHG and aerosol forcings at time $t' = 1, 2, \ldots, n$, respectively. From the *SSP2-4.5* scenario, we compute the incremental changes in GHG forcing $\Delta F_{\text{GHG}}(t') = F_{\text{GHG}}(t') - F_{\text{GHG}}(t'-1)$, with $\Delta F_{\text{GHG}}(1) = F_{\text{GHG}}(1)$. We can then calculate the time-evolving coefficients by convolving the incremental GHG forcings with the exponential decay functions:

$$\mathbf{C}_{\text{GHG}i,t} = \sum_{t'=1}^{t} \Delta F_{\text{GHG}}(t') \cdot (1 - e^{-(t-t')/\tau_i}), \tag{6}$$

or in matrix notation:

$$\mathbf{C}_{\text{GHG}} = \{\mathbf{\Delta F_{GHG}} * \mathbf{T_{GHG}}\}(t), \tag{7}$$

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

where $\{\}$ indicates a convolution over the time dimension. From that, the estimated GHG-induced spatial response is then:

$$\mathbf{S}_{\mathrm{GHG}_{s,t'}} = \sum_{i=1}^{N} \widetilde{\mathbf{B}}_{\mathrm{GHG}_{s,i}} \cdot \mathbf{C}_{\mathrm{GHG}_{i,t'}}. \tag{8}$$

Using the total scenario response $\mathbf{S}_{s,t'}$ from the ESM simulations of *SSP2-4.5*, the aerosol-induced response can be estimated from the residual of subtracting the GHG response from the total response (Fig. 2, step 7):

$$\mathbf{S}_{\mathrm{resid}\,s,t'} = \mathbf{S}_{\mathbf{s},\mathbf{t}'} - \mathbf{S}_{\mathrm{GHG}\,s,t'}. \tag{9}$$

As was done for the GHG response, we here assume the aerosol response can be represented with $n_{\mathrm{aer}} = 3$ timescales τ_{aer} for the inter-annual, inter-decadal and inter-centennial responses, respectively. The time response matrix $\mathbf{T}_{\mathrm{aer}}$ for each timescale j for a step change in aerosol forcing is constructed as (Fig. 2, step 8):

$$T_{aer\,j,t'} = 1 - e^{-t'/\tau_{aer\,j}}. \tag{10}$$

The time-evolving coefficients for the aerosol pattern response in the scenario can then be calculated by convolving the aerosol forcing difference time series with the synthetic pulse response time series:

$$C_{aer\,j,t} = \sum_{t'=1}^{t} \Delta F_{\mathrm{aer}}(t') T_{aer\,j,t-t'}, \tag{11}$$

or in matrix notation:

$$\mathbf{C}_{\mathrm{aer}} = \{\mathbf{\Delta F}_{\mathrm{aer}} * \mathbf{T}_{\mathrm{aer}}\}(t). \tag{12}$$

In order to find optimal values for τ_{aer} (introduced in Section 2.1.2), we use an optimization algorithm to search for values of τ_{aer} which minimise the error in the projection of the global mean target field residual timeseries (e.g. temperature, precipitation) onto the basis defined by $\mathbf{C}_{\mathrm{aer}}$ (Fig. 2, step 9). Once $\mathbf{C}_{\mathrm{aer}}$ is known, we can create a least-squares estimate of the spatial patterns of aerosol response $\mathbf{B}_{\mathrm{aer}}$ as the product of the residual matrix with the Moore-Penrose pseudoinverse ${\mathbf{C}_{\mathrm{aer}}}^{+}$ of the coefficients matrix $\mathbf{C}_{\mathrm{aer}}$:

$$\mathbf{B}_{\mathrm{aer}} = \mathbf{S}_{\mathrm{resid}} \cdot {\mathbf{C}_{\mathrm{aer}}}^{+}. \tag{13}$$

The emulated spatial aerosol response $\mathbf{S}_{\mathrm{aer}}$ for a novel forcing timeseries $\mathbf{\Delta F}$ can then be computed by convolving with the timescale response matrix T_{aer} and taking the dot product with the aerosol spatial response patterns $\mathbf{B}_{\mathrm{aer}}$ (Fig. 2, step 10):

$$\mathbf{S}_{\mathrm{aer}} = \mathbf{B}_{\mathrm{aer}} \cdot \mathbf{C}_{\mathrm{aer}} = \mathbf{B}_{\mathrm{aer}} \cdot \{\mathbf{\Delta F} * \mathbf{T}_{\mathrm{aer}}\}(t) \tag{14}$$

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

2.2 Applying the model: Multi-forcer multi-timescale pattern scaling in out-of-sample scenarios

The previous sections yielded the transient spatial climate response patterns for GHG and aerosol forcing. These can now be used to emulate the spatio-temporal climate response of any climate scenario. For the application in a new (out-of-sample) emissions scenario, METEOR converts any given emission scenario (Fig. 2, step A) into forcing time series using the emissions-to-forcings module of the C-SCM (Fig. 2, step B). Convolving these time series with the GHG and aerosol patterns (Fig. 2, step C) can then be utilised to reconstruct the total multi-forcer multi-timescale emulated climate response by linearly combining the GHG and aerosol responses (Fig. 2, step D):

$$\mathbf{S}_{\text{emul}\,s,t'} = \mathbf{S}_{\text{GHG}\,s,t'} + \mathbf{S}_{\text{aer}\,s,t'}, \tag{15}$$

where $\mathbf{S}_{\text{GHG}\,s,t'}$ and $\mathbf{S}_{\text{aer}\,s,t'}$ are obtained according to equations (6, 8) and (11, 14) respectively, with forcing timeseries $\Delta F_{\text{GHG}}(t')$ and $\Delta F_{\text{aer}}(t')$ calculated for the emissions scenario of interest.

3 METEOR evaluation

To evaluate the performance of METEOR, we have trained the model on CMIP6 data for a large number of CMIP6 models (see tab. B1 for a full list) using a training set which consists of *abrupt4x-*CO_2, *piControl*, *historical* and *SSP-2.45* modeling output in each case. Comparing to modeling output from these experiments, we show the in-sample accuracy. Furthermore, we have applied the resulting emulation models to a number of additional scenarios from ScenarioMIP (Tebaldi et al., 2021): *SSP-1.26*, *SSP-3.70*, *SSP-5.85* and *SSP-5.34* to consider the out-of-sample performance (note that *SSP-5.34* was performed by a limited number of models).

Additionally, in appendix A we explore the effect and performance of the model for the *abrupt4x-*CO_2 experiment and the *historical* and *SSP2-4.5* experiment combination depending on the number of timescales used, showing that three timescales seem to yield good performance. Appendix B show per model results including table B1 which lists the values for the timescales obtained for GHG and aerosol response for Surface Air Temperature (*tas*) and Precipitation (*pr*) for each model and thus also serves as a reference of which models were included in the analysis.

3.1 GHG and aerosol multi-timescale pattern of climate change

Starting with the GHG response, Figure 3 show the timescales and patterns obtained from the *abrupt4x-*CO_2 experiment. Panels a and e, b and f, and c and g of Figure 3 show the model mean of the patterns for ($\mathbf{C}_{\text{GHG}}$) associated with the inter-annual, inter-decadal and inter-centennial responses for temperature (a, b and c) and precipitation (e, f and g). Panels d and h show the zonal mean values for each of the timescales for temperature and precipitation respectively. The patterns display Arctic amplification (temperature) in both inter-annual and inter-decadal modes, though Antarctic amplification is largely limited to the inter-centennial response. Fast mode response for precipitation is dominated by the tropical Pacific, with more widespread changes in the slower modes. Figure 4 similarly shows timescales and patterns obtained from the aerosol anomaly response

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

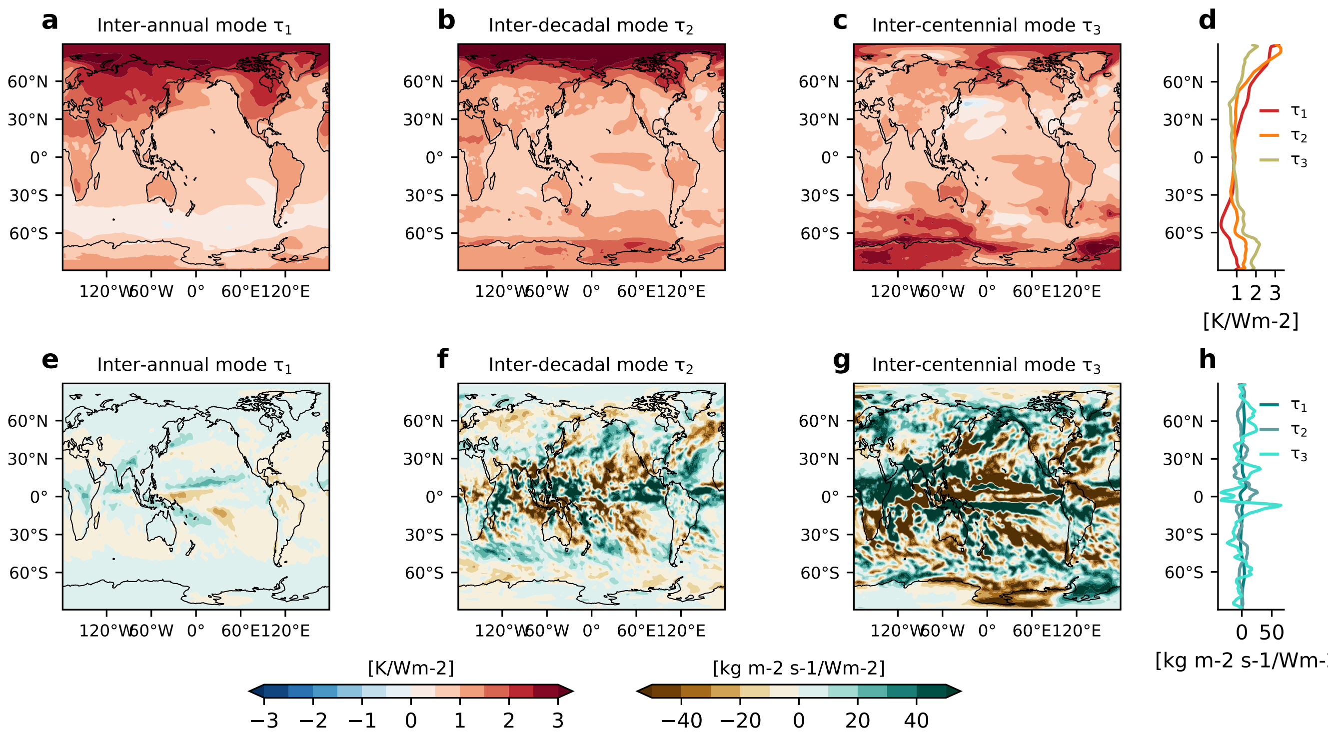

Figure 3. METEOR GHG pattern from CMIP6. GHG patterns obtained from the *abrupt-4xCO2* scenario (CMIP6 model average) for temperature (top row) and precipitation (bottom row). Panels a–c and e–g show the mean spatial patterns for the the inter-annual, inter-decadal and inter-centennial modes, respectively. Panels d and h show the zonal mean of the patterns associated with all timescales.

obtained using the *historical* and *SSP-2.45* and subtracting the GHG response obtained using the *abrupt4x-CO*$_2$ for temperature and precipitation. Again the patterns are shown in panels a and e (inter-annual modes), b and f (inter-decadal modes), and c and g (inter-centennial modes), with zonal average breakdown in panels d (temperature) and h (precipitation).

The global mean timeseries of the multi-model mean emulated fit is shown alongside the global mean timeseries of each ESM and the multi-model mean outputs in panels b (temperature) and d (precipitation) of Figure 5. The multi-model global mean response shown in panels a and b of Figure 5 matches the model mean calculated from the modeling output fairly well for both precipitation and temperature. As such, the multi-model mean of CMIP6 METEOR emulations is able to capture the global mean time evolution of the original ESM multi-model mean output well until 2100.

3.2 Climate response reconstruction in out-of-sample scenarios

In the main study detailed here, only the future *SSP2-4.5* scenario is used in the training of METEOR, so other scenarios can be used as out-of-sample test cases. Figures 6, 7, 8, and 9 show the global mean and spatial pattern performance of the emulation applied to the ScenarioMIP scenarios. The timeseries plots show response from greenhouse gas forcing alone, and response

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

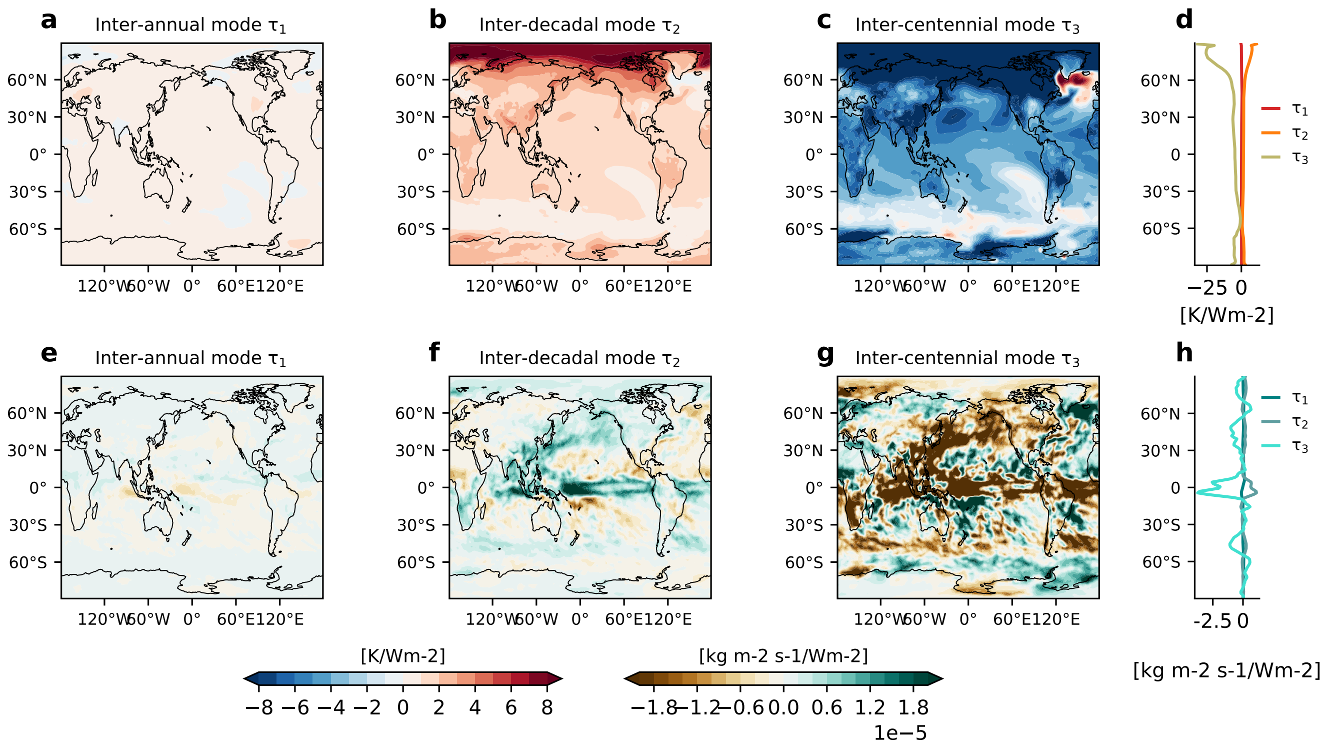

Figure 4. METEOR aerosol residuals pattern from CMIP6. Residuals patterns obtained from the *histroical+SSP2-4.5* scenario (CMIP6 model average) for temperature (top row) and precipitation (bottom row). Panels a–c and e–g show the mean spatial patterns for the the inter-annual, inter-decadal and inter-centennial modes, respectively. Panels d and h show the zonal mean of the patterns associated with all timescales.

from the model which includes the calibrated aerosol response. The results demonstrate good performance of METEOR in out-of-sample scenarios, both in terms of the global mean and spatially resolved output. In the global mean, the aerosol model is able to capture the broad temporal dynamics of the global mean aerosol effect: cooling in the late 20th Century, and a reduced effect in the future in all scenarios except *SSP3-7.0*.

Panel f of Figures 6 and 7 shows year 2100 temperature and precipitation change mean and model spread for all SSPs included. For each scenario, METEOR is able to capture the future spread, though multi-model mean end of 21st century warming is slightly underestimated in the high mitigation *SSP1-2.6* experiment. Panels a-e in both figures include dashed lines showing the METEOR reconstruction results obtained using only the GHG patterns and forcing.

Table 1 shows the Pearson's correlation coefficient and Root Mean Square Error (RMSE) for *tas* and*pr* between the METEOR pattern reconstruction and CMIP6 end of century (2080–2100) change for each of the SSPs. In general, out of sample performance is better for temperature than for precipitation. Note that limited ESM simulations are available for the *SSP5-3.4-over* scenario. Performance in-sample (for *SSP2-4.5*) is unsurprisingly greater than out-of-sample performance, with the largest

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

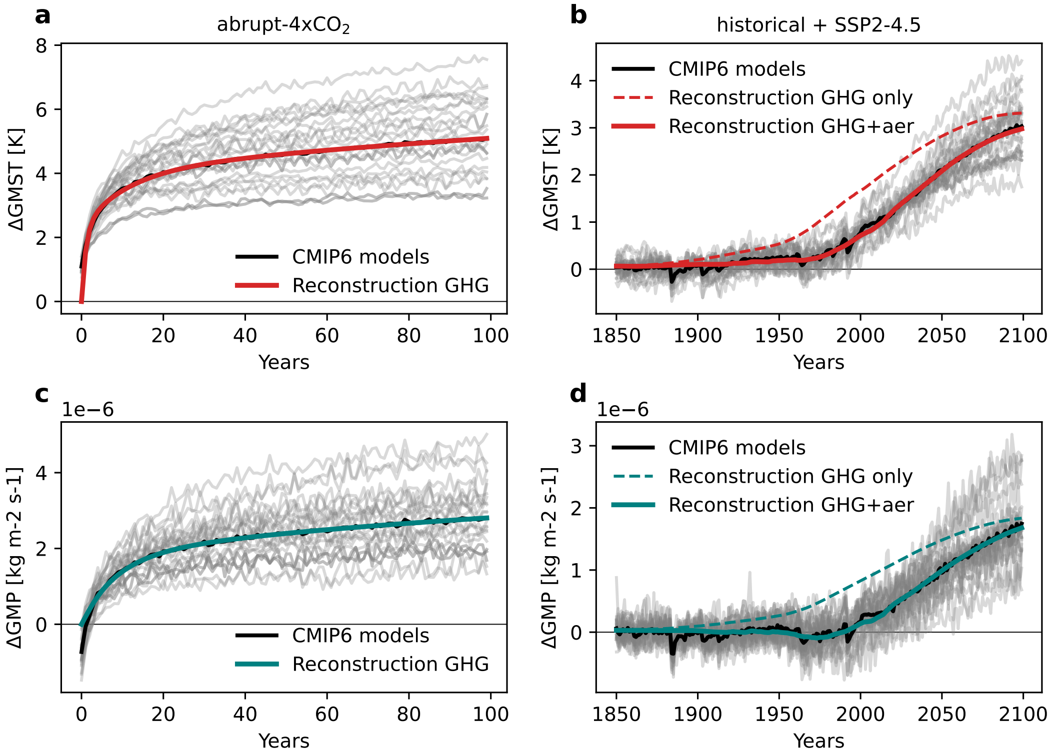

Figure 5. METEOR GHG and aerosol residual reconstructions from CMIP6 models. METEOR reconstruction for *abrupt4x-CO_2* (a, c) and *histroical+SSP2-4.5* (b,d) for global mean temperature (GMST, top row) and precipitation (GMP, bottom row) for each model (grey), the model mean (black) and the mean of the METEOR reconstructions (red for temperature and teal for precipitation). Panels b and d also include the mean reconstruction timeseries obtained using only the GHG pattern (stippled lines), the *abrupt4x-CO_2* reconstructions in panels a and c are identical with and without the aerosol pattern as only CO_2 is changing in this experiment.

Table 1. Root Mean Square Error (RMSE, in K for *tas* and in kg m-2 s-1 for *pr*) and Pearson correlation coefficient (Pearson) between the METEOR pattern reconstruction and CMIP6 end of century (2080–2100) change.

Metric	Variable	*SSP1-2.6*	*SSP2-4.5*	*SSP3-7.0*	*SSP5-8.5*	SSP-5.34-over
Pearson	*tas*	0.99	0.99	0.99	0.99	0.97
Pearson	*pr*	0.97	0.99	0.99	0.96	0.94
RMSE	*tas*	0.13	0.18	0.32	0.92	0.76
RMSE	*pr*	5.4e-7	3.9e-7	6.9e-7	26.4e-7	24.0e-7

errors indicated in the high emission *SSP5-8.5* scenario - but for all scenarios and variables considered, the correlation between spatial patterns of change exceeds 0.94. These fits are comparable to those reported for PRIME (see table 1 of Mathison et al. (2024)).

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

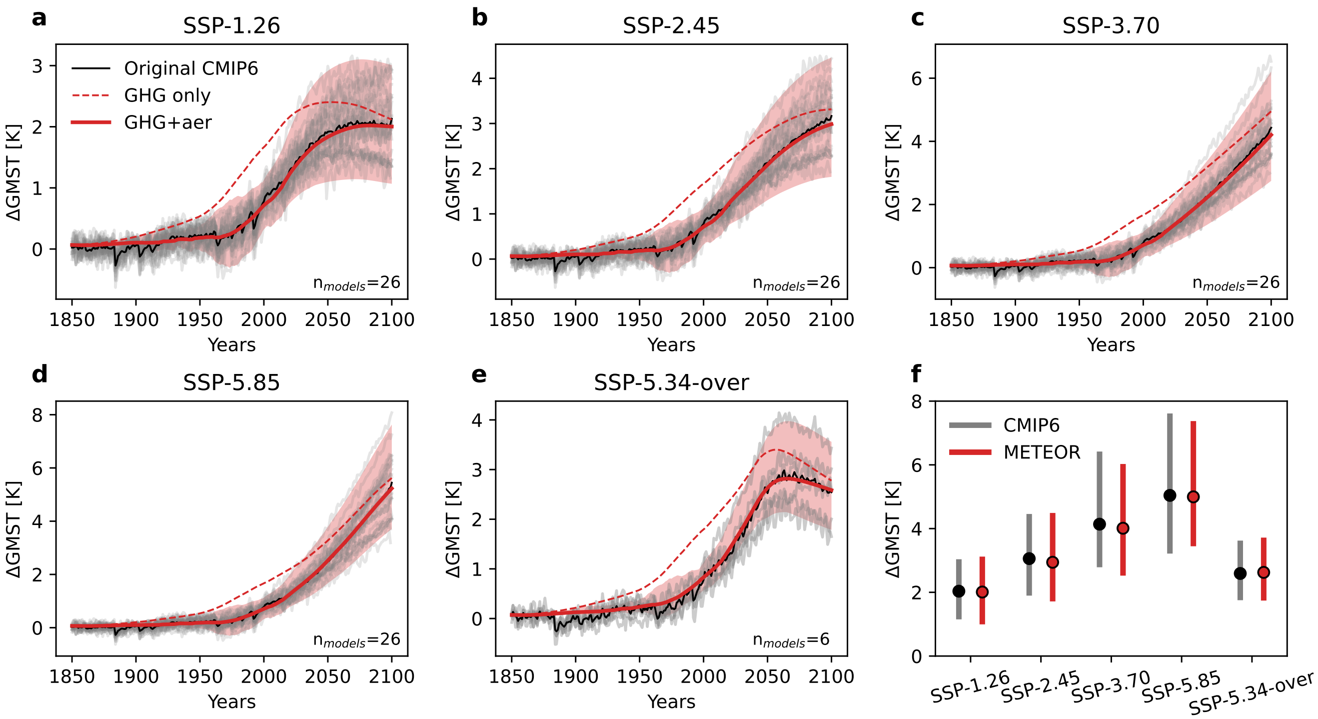

Figure 6. Global mean temperature reconstruction from METEOR vs CMIP6 data for various SSP scenarios. Panels a–f show global mean *tas* (GMST) change for each model (grey), the model mean (black) and the mean of the METEOR reconstructions using GHG only (red stippled), and the combined GHG and aerosol patterns (red solid). The red plume shows the distribution of the full METEOR reconstructions. Panel f shows the year 2100 *tas* change with mean and modeling spread for the CMIP6 models (grey) and full METEOR reconstruction (red).

We can illustrate this visually by considering the multi-model mean spatial patterns of change in CMIP6 and the out-of-sample METEOR reconstructions, where the emulated amplitude and patterns of temperature and precipitation change are highly consistent for each scenario considered (Figs. 8, 9). The spatial bias is largest for *SSP5-8.5* (the highest emissions scenario) for both temperature and precipitation. In the case of temperature there is cold bias for *SSP5-8.5* and more of a warm bias for the overshoot and low emissions scenarios (*SSP1-2.6* and *SSP5-3.4-over*). This may be an imprint or slight over-fitting to the *SSP-2.45* scenario, which is colder than the former, and hotter than the two latter scenarios at the end of the century. Note also that the colour scale in the bias plots (c, f and i) for temperature in Figure 8 has a smaller range, than the absolute value plots (a, b, d, e, g and h), whereas the scales are the same in Figure 9 mainly due to a stronger localised bias in the tropical pacific, especially for the *SSP5-8.5* scenario for precipitation.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

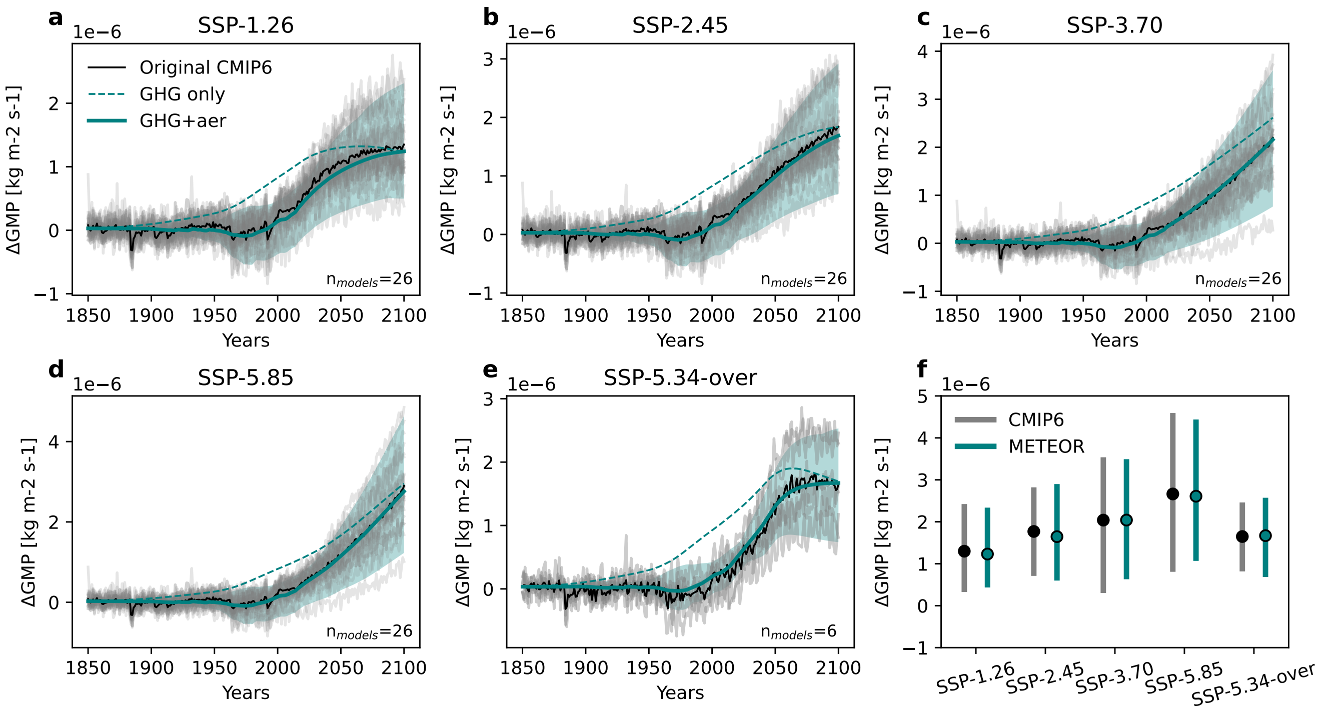

Figure 7. Global mean precipitation reconstruction from METEOR vs CMIP6 data for various SSP scenarios. Panels a–f show global mean *pr* (GMP) change for each model (grey), the model mean (black) and the mean of the METEOR reconstructions using GHG only (red stippled), and the combined GHG and aerosol patterns (teal solid). The teal plume shows the distribution of the full METEOR reconstructions. Panel f shows the year 2100 *pr* change with mean and modeling spread for the CMIP6 models (grey) and full METEOR reconstruction (teal).

4 Conclusions, discussion and outlook

Here, we have presented the METEOR (Multivariate Emulation of Time-Evolving and Overlapping Responses) v1.0 emulator framework for spatially resolved climate impacts. The framework allows for the reproduction of time-evolving response to a range of radiative forcers, allowing for the simulation of hysteresis dynamics and forcer-dependent responses.

We showed results of its training and application to CMIP6 models for experiments *abrupt4x-*CO_2, *piControl*, *historical*, *SSP2-4.5*, *SSP1-2.6*, *SSP3-7.0*, *SSP5-8.5* and *SSP5-3.4-over* for annual mean temperature (*tas*) and precipitation (*pr*). METEOR displays good overall performance for both in- and out-of-sample applications and though performance is slightly worse for out-of-sample applications (particularly for precipitation), the amplitude and spatial pattern of multi-model response is well captured for all scenarios considered. The model can be trained on new modeling output and to emulate new and unknown scenarios and reasonable accuracy can be expected.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

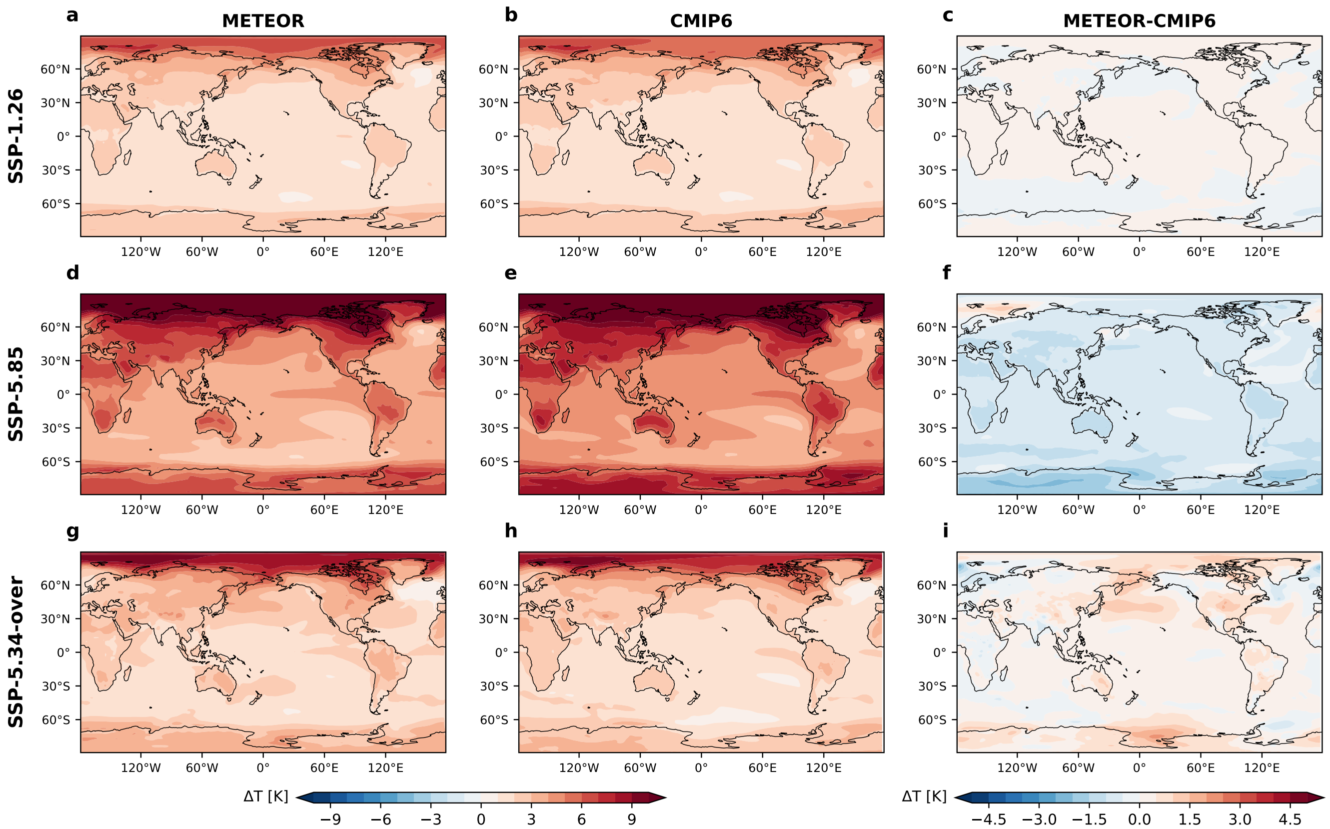

Figure 8. METEOR temperature reconstruction pattern vs CMIP6 data for various SSP scenarios. Panels in the first (a, d and g) and second (b, e and h) column show the the difference between the mean of the first 50 years of the *historical* experiment (1850–1900) and the last 20 years of the *SSP1-2.6*, *SSP5-8.5* and *SSP5-3.4-over* scenarios for the mean of the METEOR reconstructions and the CMIP6 model output, respectively. Panels in the third column (c, f and i) show the difference between the second and first column, illustrating the difference between the CMIP6 projections and the METEOR reconstruction.

The model is structured as an importable Python library with example Jupyter notebooks that demonstrate basic usage and reproduce the figures in this article. This should make the model accessible, useful and extendable for new uses for the wider climate research community.

4.1 Discussion

Though METEOR shows good fits, including the ability to fit hysteresis, it has limitations. For instance, the model assumes that increasing patterns and timescales are associated with increasingly longer and non-overlapping timescales. In practice, there might be several pattern effects going on on more closely related timescales, or on timescales which are shorter or longer than those considered here. However, allowing for additional modes in testing did not lead to notable increases in

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

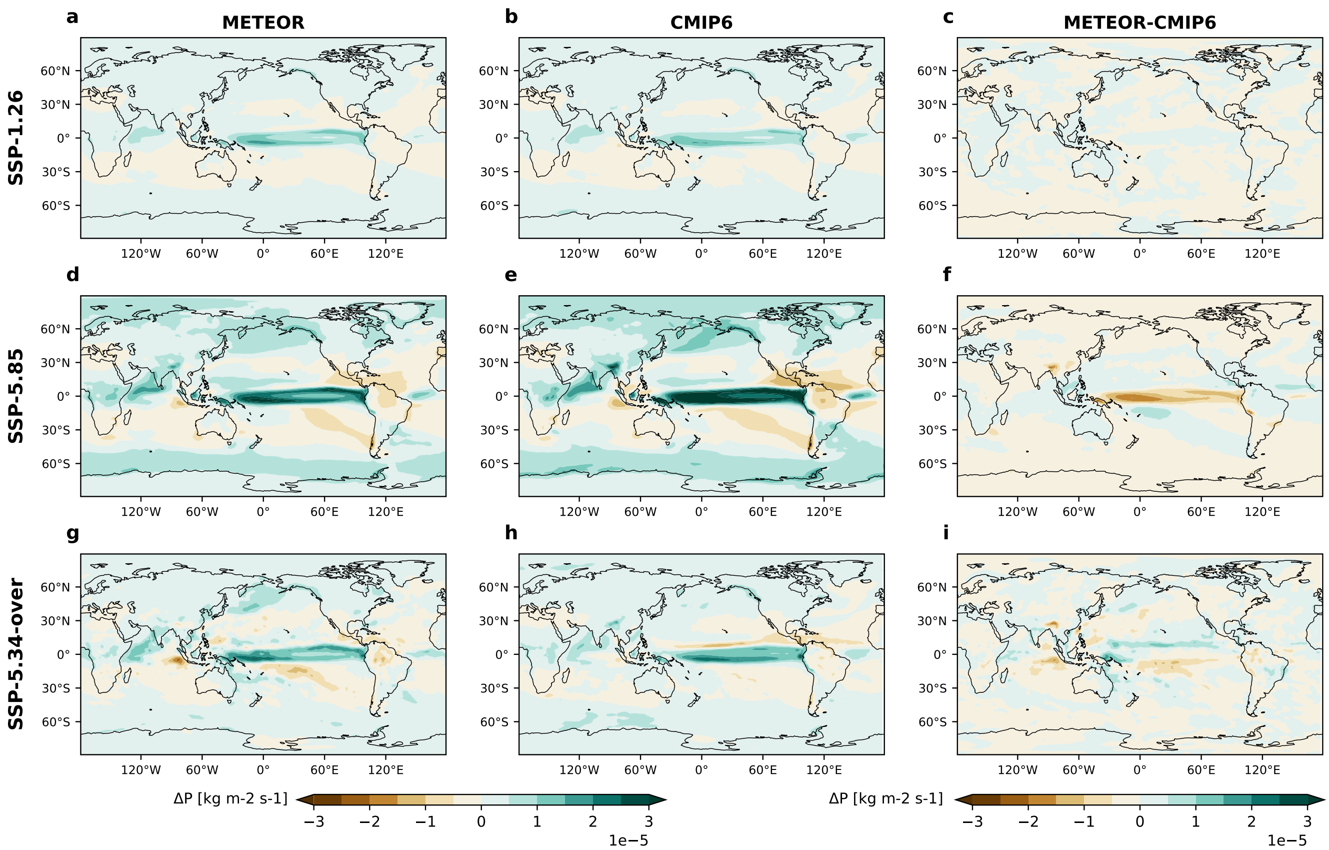

Figure 9. METEOR precipitation reconstruction pattern vs CMIP6 data for various SSP scenarios. Panels in the first (a, d and g) and second (b, e and h) column show the the difference between the mean of the first 50 years of the *historical* experiment (1850–1900) and the last 20 years of the *SSP1-2.6*, *SSP5-8.5* and *SSP5-3.4-over* scenarios for the mean of the METEOR reconstructions and the CMIP6 model output, respectively. Panels in the third column (c, f and i) show the difference between the second and first column, illustrating the difference between the CMIP6 projections and the METEOR reconstruction.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

performance. This is partly a function of data limits in training data. As with any emulator, performance is limited by the availability of training data; capturing multi-millennial timescale responses requires training data to unambiguously simulate those timescales. Similarly, gaining high confidence in shorter timescale responses requires large initial condition ensembles which are not generally available for the DECK simulations in CMIP. In general, the number of independent forcer responses and timescales could be improved with ESM datasets which isolate the effect of independent forcers.

Further, in METEORv1.0, we assume that the anomaly between the full target model response and the synthetic simulated greenhouse gas only response is due to sulfate aerosol forcing. However, in practice, anomalies will also be caused by any errors in the reconstruction of the synthetic GHG response, by other forcers which are not explicitly represented or by potential nonlinear interactions between forcers. The model framework is sufficiently flexible to test these possibilities, but data from current CMIP archives is limited. Hence, we isolate the effect of GHG and sulfate aerosol forcing leaning on results from the literature (Myhre et al., 2017; Samset et al., 2018, 2019; Persad et al., 2023; Wilcox et al., 2023) including the assessments of the last IPCC cycle (Forster et al., 2021), which highlight the strength, peculiarity and uncertainty of the aerosol and specifically sulfate aerosol forcing as a first order modifier from the pure GHG-driven response.

The strength and usefulness of an emulator such as METEOR lies particularly in its ability to produce impact results rapidly. In this paper we have focused on annual mean values of temperature, but for impact studies, seasonal cycle and extreme value data might be more useful, and some emulators already have extensions that allow for monthly or seasonal output (Nath et al., 2022; Schöngart et al., 2024; Tebaldi et al., 2022). Future versions of METEOR will seek to represent these additional dimensions, noting that METEOR's variable agnostic setup can facilitate, for example, training on seasonally varying data, i.e. building separate patterns for January temperatures, February temperatures, summer temperatures or maximum yearly temperatures.

As capturing hysteresis for modelling overshoot scenarios is an explicit motivation for the METEOR methodology, the apparent success of its fit to the *SSP5-3.4-over* scenario is reassuring. However, as only very few models have full available training data for this scenario, the robustness of METEOR for overshoot scenarios should be explored further. We note in particular that the experiment used to identify GHG timescales and patterns (*abrupt4x-*CO_2) only simulates the effects of a permanent increase in concentrations, so that in particular timescales associated with negative emissions from processes such as Carbon Dioxide Removal might differ. Experiments in CMIP7, such as flat10MIP (Sanderson et al., 2024) will provide assessments of potential asymmetries between responses to positive and negative emissions in Earth System Models, which could be used as better training data for METEOR.

For the modelling presented in this paper, we chose a combination of *historical* and *SSP2-4.5* to fit the residual aerosol patterns. In Figure 8 though quite small, we note a slight hot bias for *SSP1-2.6* and a slight cold bias for *SSP5-8.5*. This makes sense as the *SSP2-4.5* temperature trajectory lies between the two. This implies that there may be, on average, some nonlinear temperature response in CMIP models which is not captured in the pulse-response logic exploited here, and will likely also mean that fits are probably slightly more accurate when applied to scenarios that are not too far off from the scenario used for the residual fitting. In practice the biases here are small, and fits overall are good in the range considered, but if the intended application of METEOR is to scenarios in a very particular temperature trajectory range, choosing a residual to fit to that is

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

relatively centrally located for that range would make sense. For the purposes of this paper as well as most ordinary applications to CMIP6 data, we consider that *SSP2-4.5* is a reasonable choice.

To provide forcing input to METEOR, we have chosen to use the emissions to forcing module from the C-SCM model. Further development of the METEOR model could easily alleviate this, by allowing the model to be driven from forcing output which can derive from any simple climate model or other forcing output sources.

4.2 Outlook

The potential applications of METEOR are not limited to mean temperature and precipitation. It can be trained and used to model climate change response in any variable for which a strong connection between climate change forcing and variable evolution can be assumed. Examples can be impact related variables such as extreme value impact indicators (Quilcaille et al., 2022, 2023; Sillmann et al., 2013) or other climatological variables such as humidity, radiative balance or soil moisture. Further development would extend the model to more directly resolve impacts such as crop yields, human heat stress and sea level rise. However, we expect such indicators will require additional modules.

The scheme used to obtain the GHG response from *abrupt4x-*CO_2 can be employed in the same way for any forcer for which there is data for an abrupt step-change response. A collection of such experiments are contained in Myhre et al. (2017), however, that was performed with somewhat outdated modeling versions. METEOR can be applied to that dataset, or a similar updated one, to obtain separate per forcer timescales and patterns, allowing for further decomposition of spatial forcer response as a function of a wider range of species.

Similarly, the scheme for anomaly calculations that we employed to obtain the aerosol signal may also be performed for multiple forcers if data from experiments where only a single forcer is changed are available. The RAMIP (Regional Aerosol Model Intercomparison Project) dataset offers such data, which can even separate the forcing response from aerosols depending on their region of origin, which may in fact produce different results both in terms of spatial pattern and strength of climate results (Wilcox et al., 2023).

The timescales and patterns that METEOR obtains for any particular model, though mostly operational and for emulation purposes, may also indicate aspects of the underlying physics of the model, and comparison between parameters obtained for emulation of different models may itself provide indicators for understanding ESM differences (such as the role of fast and slow feedback processes in observed climate and implications for future climate commitments).

The emissions to forcing pipeline from the C-SCM is currently an integral part of METEOR. However, coupling METEOR to a similar pipeline from a different simple climate model, or allowing it to run directly from input forcing data is a logical future development goal, allowing uncertainty in the emissions-forcing pipeline to be decoupled from the forcing-pattern component. An alternative would be a higher level of integration within a simple climate model - such that the METEOR forcing-pattern mapping becomes an integral part of a wider model, including, for example, ecosystem components which could evolve as a function of regional climate as simulated within METEOR.

Currently, METEOR does not have any representation of natural variability, nor is there native support for producing probabilistic output which spans uncertainty in either ESM training model, or in fitting uncertainty for a single model. Each of these

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

would be useful for impact applications, and would be logical extensions for future development. Goals for future versions include a probabilistic implementation in which a set of plausible METEOR configurations can be produced for a given ESM emulation, such that an ensemble of simulations can then provide risk guidance for climate impacts. Testing the robustness of timescales and patterns across ensemble runs of the same model would be beneficial in achieving this goal. Using emulations based on existing ensemble spread as in Schwaab et al. (2024) or including some variability into the emulator process pipeline itself similar to what has been done for the FAIR simple climate model (Bouabid et al., 2024) are both possible approaches to this.

METEORv1.0 is tailored towards annual data. However, for many applications, monthly or higher time resolution outputs are desirable. The methodology can be fairly easily generalised to monthly output (for example, by expanding the dimensionality of the spatial dimension to include different months), but further testing is required to ensure the realism of emulated timeseries. Additional modules are planned to model dominant modes of natural variability, and potentially extreme value indicators.

There may be indications that extremely fast (sub-yearly, near instantaneous), and very slow timescales (millennial) are not captured by the current setup. Aiming to fit the former could be a targeted extension, but the latter might require output from longer runs than are currently available for most models. The assumptions made of exponentially longer timescale lengths might also not be ideal, as local optima for several shorter or comparable timescales may not be found. A more thorough investigation of this may be needed, but Appendix A does include a discussion and investigation of the fits as a function of number of timescales included.

We provide METEOR as an open tool which we hope can be of wider benefit to the community, contributing to a wider body of emulation tools each of which provides unique advantages with ample scope for intercomparisons, coupling and ensemble studies. Finally, we look forward to community development of the METEOR platform to provide better integration into ecosystems of fast climate modeling tools which can be increasingly used in applications which require fast turnaround, such as simulating regional climate impacts in societal models.

Code availability. Code is openly available on github at https://github.com/benmsanderson/METEOR under the Apache-2.0 license at the v1.0.1 tag with doi 10.5281/zenodo.14967116 Sanderson et al. (2025)

Data availability. Emissions input data are from the Reduced Compexity Model Intercomparison Nicholls et al. (2020), data at Nicholls and Gieseke (2019). CMIP6 data are available through the Earth System Grid Federation (esgf) or via the zarr store google-api (https://storage.googleapis.com/cmip6/cmip6-zarr-consolidated-stores.csv).

Appendix A: Fitness depending on number of timescales

The METEOR model can be trained with an arbitrary number of timescales τ_k. In this paper we have shown results for a setup which assumes three timescales for each forcer. However, the user can specify to train and use and arbitrary number

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

of timescales and corresponding spatial patterns. Figure A1 shows the effect of increasing the number of timescales on the emulation fits. Panels a and d show with each model as connected points, the RMSE of the fit to the *abrupt4x-CO_2* as the number of timescales increase. Panels b an e show these values only for the multi model mean. The fitness increases substantially for both temperature and precipitation when going from one to two timescales, and a further slight increase is observed when increasing to three timescales.

Panels c and f of Fig. A1, show the RMSE model mean fit for temperature and precipitation with varrying numbers of both GHG and residual aerosol timescales to the *SSP2-4.5* experiment. Annual values per model for global mean reconstruction versus model output for each combination of GHG and aerosol timescales are shown in Figures A2 (temperature) and A3 (precipitation). From this, we observe that a combination of three timescales for both aerosols and GHG procides a reasonable balance. When a single aerosol timescale is used, we observe that the fit actually deteriorates as GHG timescales are increased from two to three. We believe that this is a consequence of the range constraints for timescales. With one or two very slow GHG timsecales, the residual pattern includes timescale signals which do not quite fit the *SSP2-4.5* scenario. With only one very short timescale with which to compensate for this, the residual can not compensate. This does not mean that sulfate aerosols have very long timescale mechanisms, but rather should serve to caution the user on the interpretability of the emulator outputs. Since the sulfate aerosol timescales and patterns are based on residual signals, they also include and have in them information which is just to do with the lack of accuracy of the *abrupt4x-CO_2* based GHG patterns and timescales to accurately map the effect of everything else that happens in the model in the *historical* and *SSP2-4.5* scenario runs. This compensating factor lead us to find it reasonable to use three timescales also for the aerosol patterns, and this shows the best overall fits by these measures.

Figure A4 shows the contributions of the different timescales and how they combine. Panels a (temperature) and c (precipitation) show the global mean time evolution of the reconstruction associated with each timescale, and we can see that all GHG patterns are associate with positive contributions, whereas the aerosol residuals patterns on inter-annual and inter-decadal scales have negative contributions, with the inter-centennial pattern giving a positive contribution. Panels b (temperature) and d (precipitation) show how the total reconstruction changes as subsequent patterns are added starting with the shortest timescale for GHG, adding longer GHG timescales first, before adding the aerosol patterns in order of timescale length.

Appendix B: Results for single model emulation

Here we include results per individual model emulated. Table B1 lists all models included and the timescales obtained for the main emulations of them. The model selection criteria was mainly one of convenience, including models that had data available from zarrr store CMIP6 google-api (https://storage.googleapis.com/cmip6/cmip6-zarr-consolidated-stores.csv) for all of the experiments *piControl*, *abrupt4x-CO_2*, *SSP1-2.6*, *SSP2-4.5*, *SSP3-7.0* and , *SSP5-8.5*. From this set, some models were excluded for various issues with the data available. Patterns associated with each of the timescales for each model are displayed in Figs. B1-B12, and global mean reconstructions for each of the four SSP experiments are shown in Figures B13–B15 (temperature) and Figures B17–B19 (precipitation). In addition, a smaller selection of models which also had data for

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

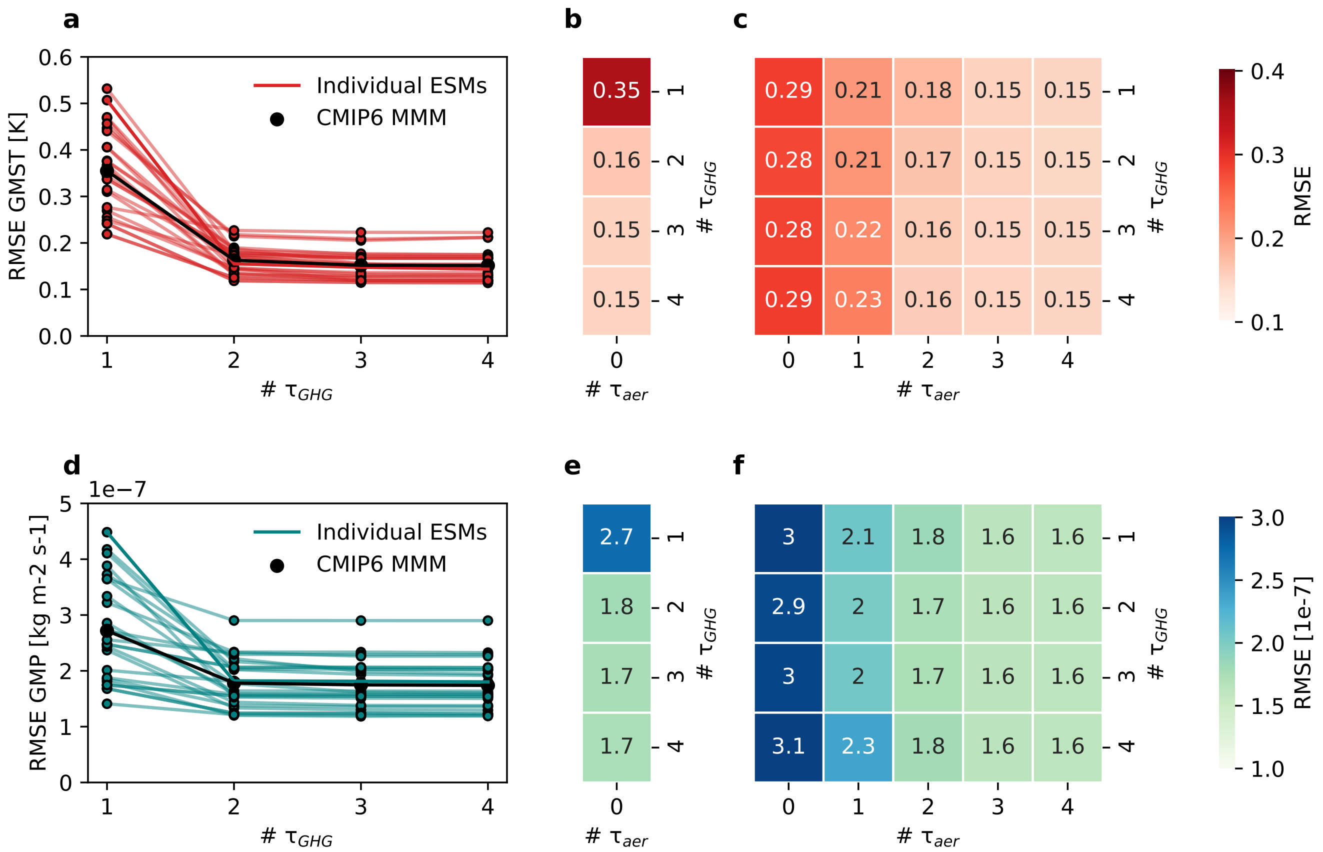

Figure A1. Evaluation of METEOR reconstruction as function of number of timescales. Panels a and d show the evolution of the RMSE parameter for temperature and precipitation respectively in the reconstruction of the *abrupt4x-CO_2* experiment as a function fo increased number of timescales for each individual CMIP6 model. Panels b and e show mean of the same numbers. Panels c and f show the same mean RMSE but applied to the reconstruction of the *SSP2-4.5* experiment and including a varying number of aerosol timescales modelled from residual.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

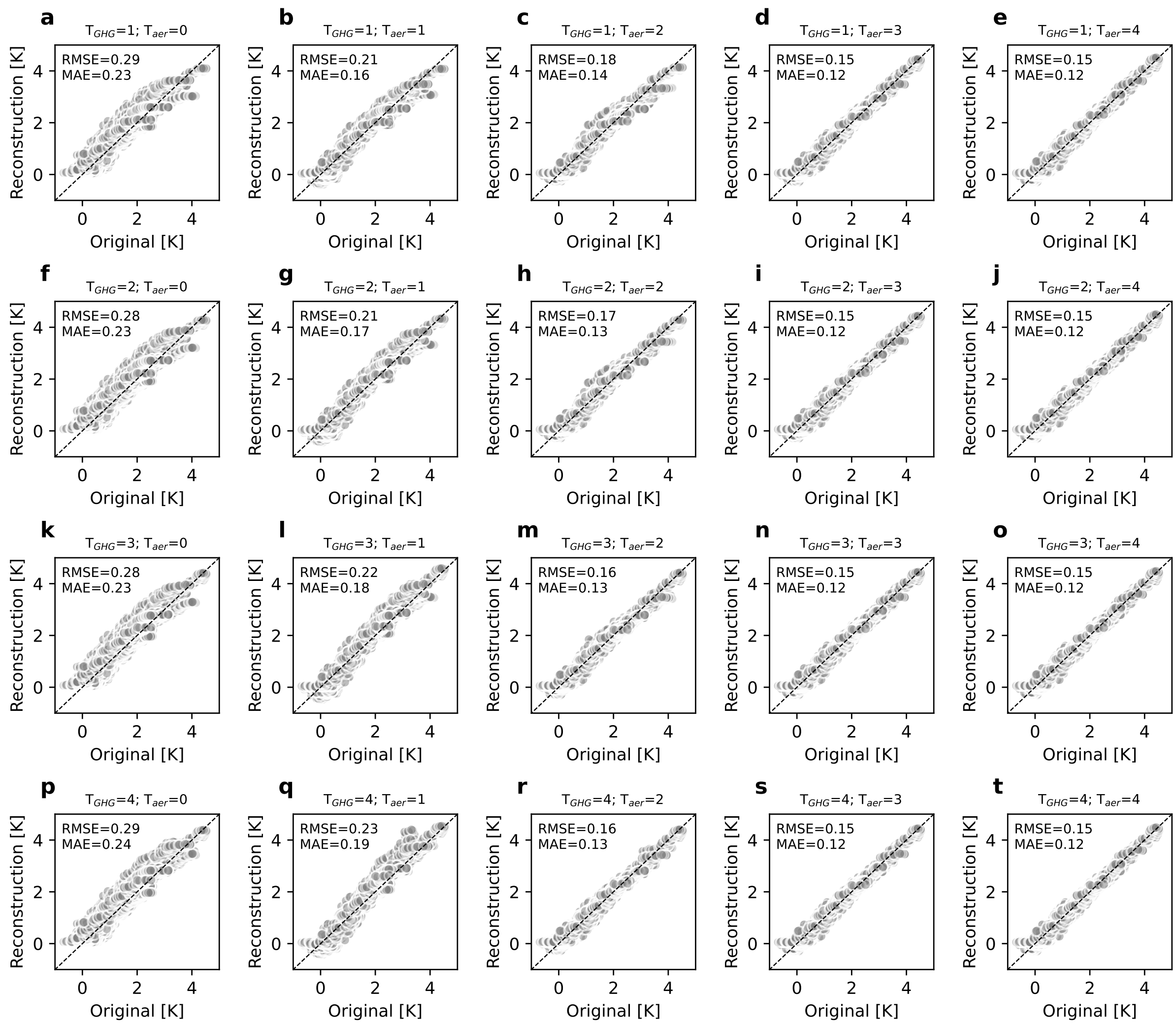

Figure A2. Evaluation of *tas* (original CMIP data vs.METEOR reconstruction) for varying number of timescales. Columns left to right show increasing numbers of aerosol timescales from 0 to 4, while rows from top to bottom show increasing numbers of GHG timescales from 1 to 4. Each point represents a global mean value for temperature for one model in the original data versus the reconstruction for *SSP2-4.5*. Model mean overall RMSE and mean absolute error (MAE) are displayed for each timescale combination.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

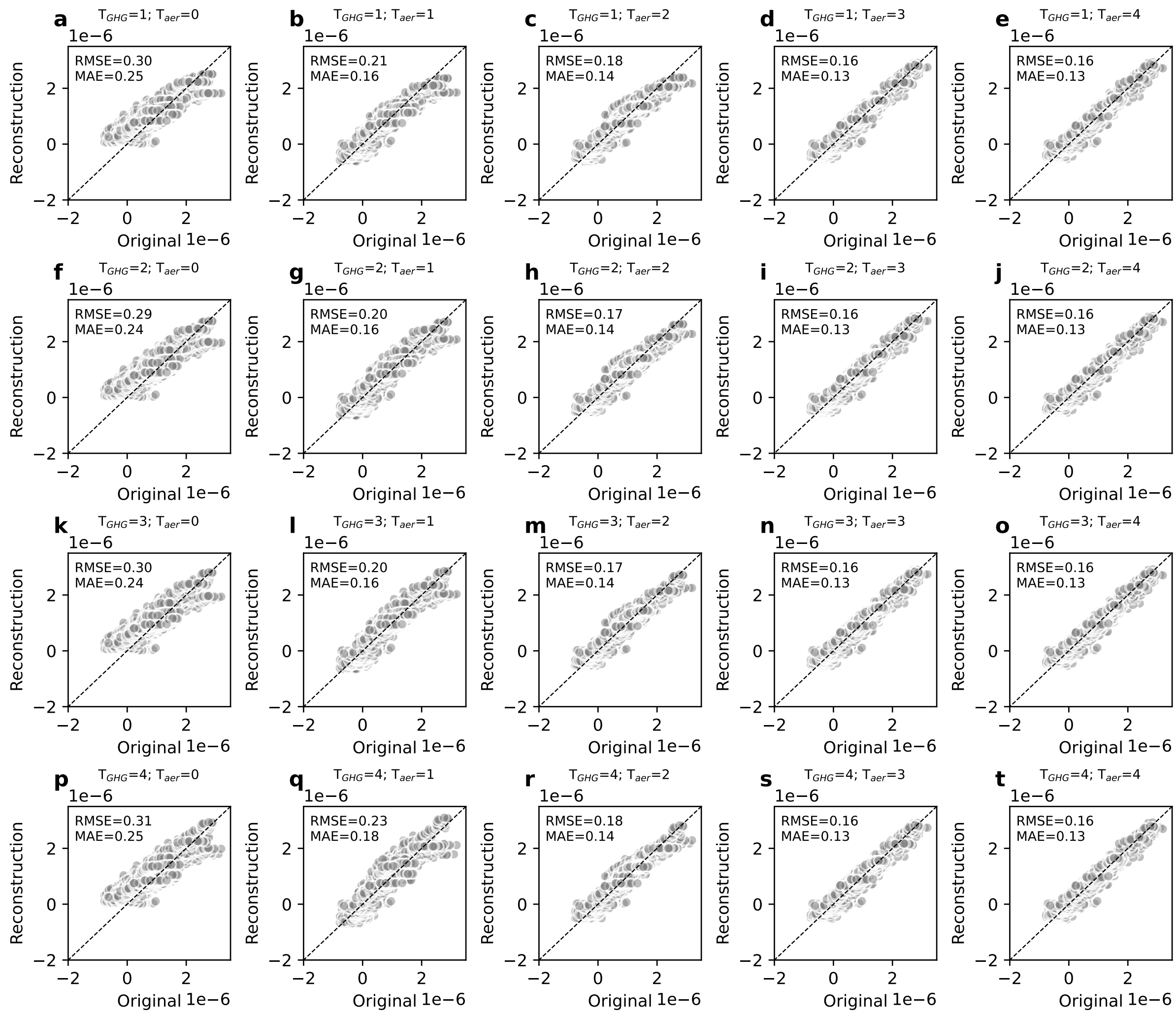

Figure A3. Evaluation of *pr* (original CMIP data vs. METEOR reconstruction) for varying number of timescales. Columns left to right show increasing numbers of aerosol timescales from 0 to 4, while rows from top to bottom show increasing numbers of GHG timescales from 1 to 4. Each point represents a global mean value for temperature for one model in the original data versus the reconstruction for *SSP2-4.5*. Model mean overall RMSE and mean absolute error (MAE) are displayed for each timescale combination.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

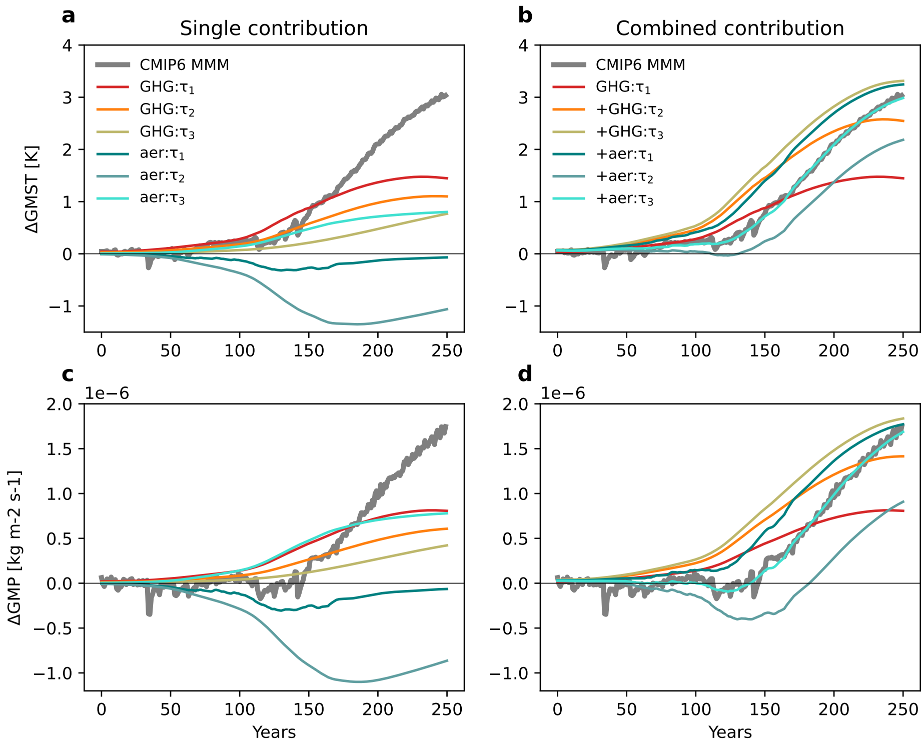

Figure A4. Multi model mean per timescale contributions. With a model trained with three GHG and three aerosol timescales, we show how the model mean global reconstruction contributions from each timescale pattern when emulating *SSP2-4.5*. Panels a (temperature) and c (precipitation) show the separate contributions from the inter-annual GHG (red), inter-decadal GHG (orange), inter-centennial GHG (green), inter-annual aerosol (dark blue), inter-decadal aerosol (medium blue) and inter-centennial aerosol (turquoise) overlayed on top of the original CMIP6 multi model data (grey). Panels b (temperature) and c (precipitation) show how the result of summing these patterns in this order.

SSP5-3.4-over were used to evaluate the performance of the emulation in overshoot scenarios. Global mean reconstruction results for these are shown individually in Figures B16 (temperature) and B20 (precipitation).

435 *Author contributions.* BMS designed the initial concept for the emulator and provided help throughout its development. SB contributed to the the initial concept and design. MS was the primary developer of the software infrastructure. NJS produced and made layouts for the figures and the code flow schematic. NJS, BMS, SB and MS all contributed to writing and reviewing all parts of the article.

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

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

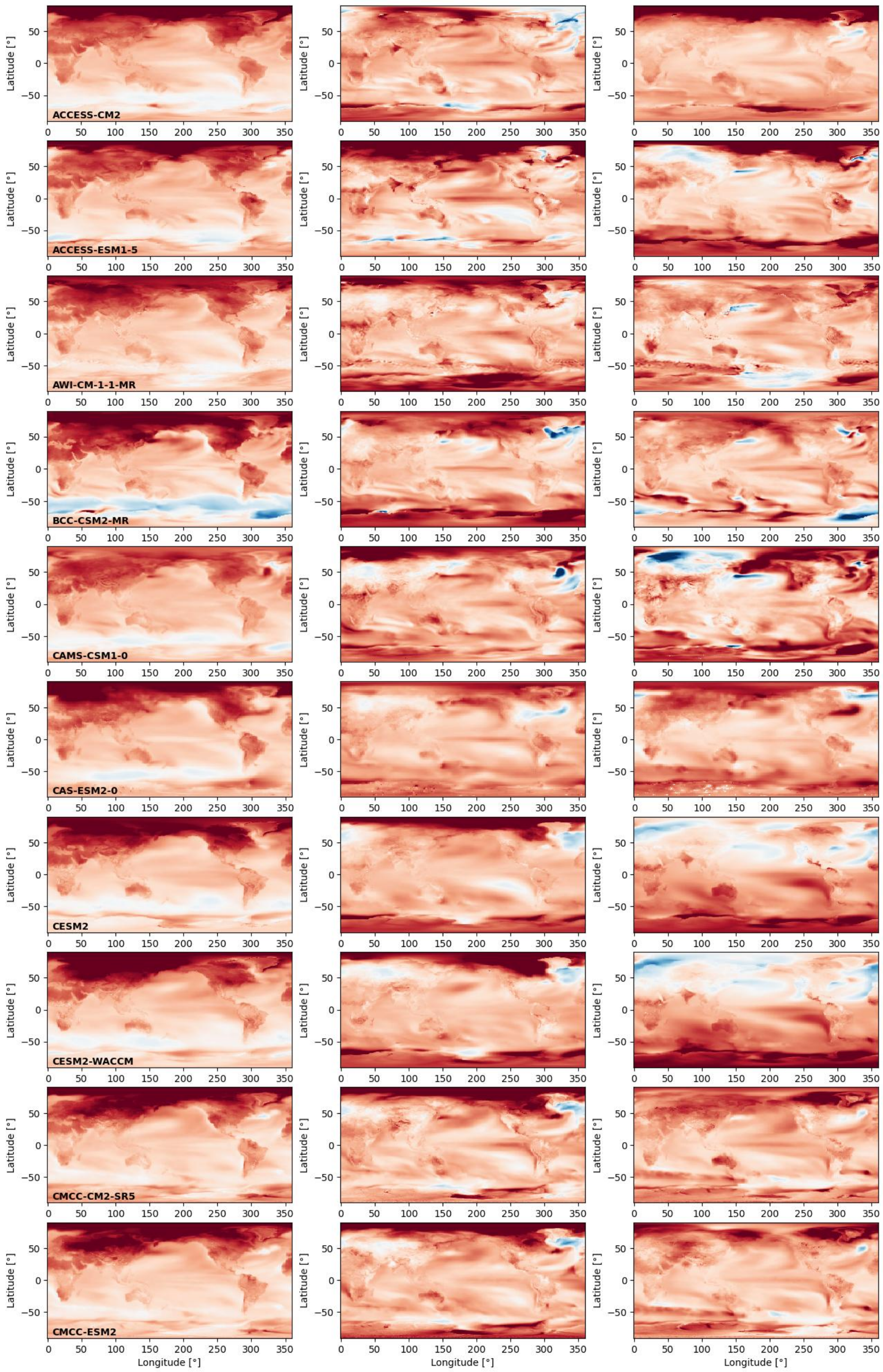

Figure B1. GHG patterns of temperature for single models, one model per row. The columns show from left to right the patterns associated with the inter-annual mode, the inter-decadal mode and the inter-centennial mode.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

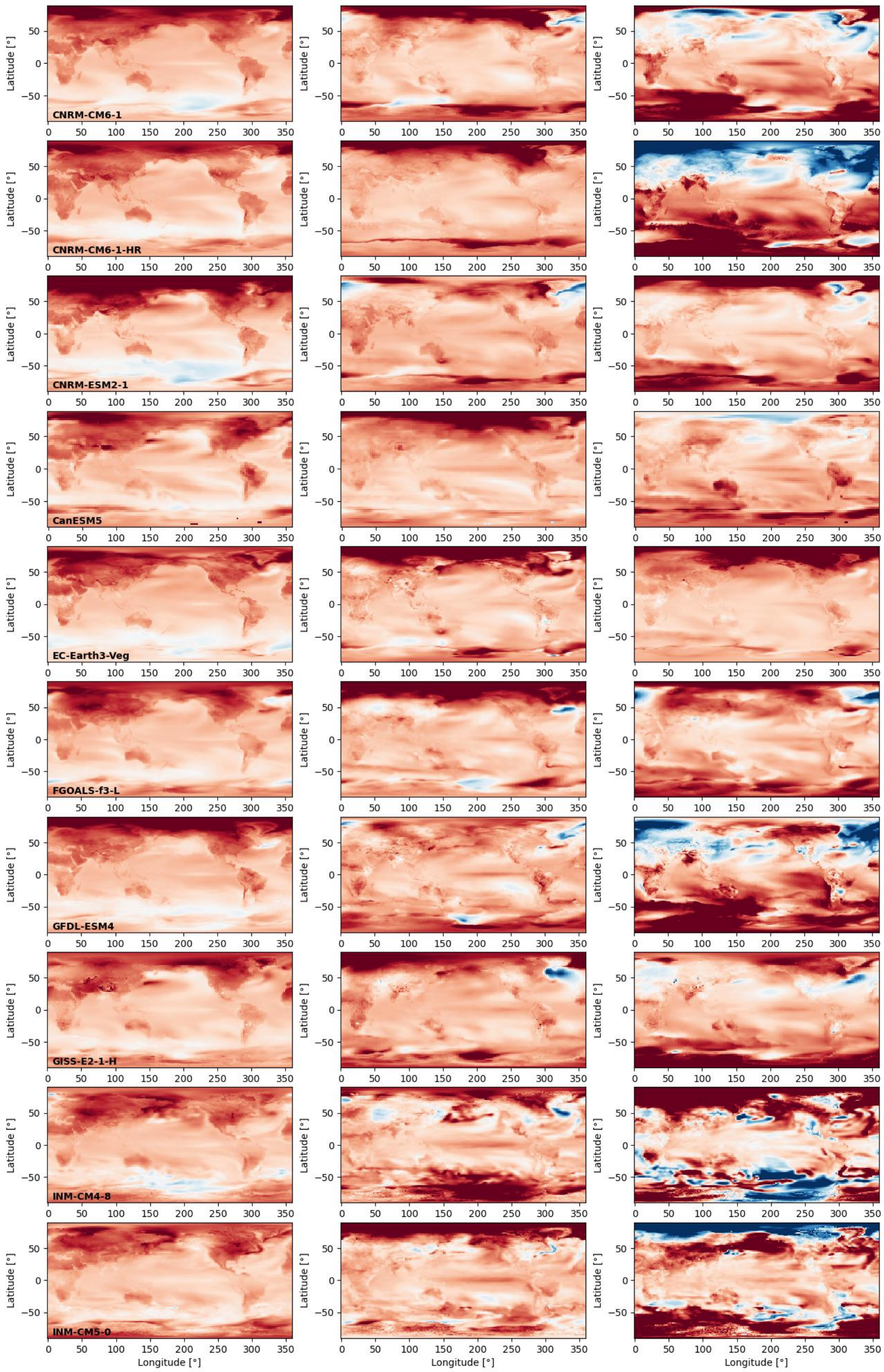

Figure B2. GHG patterns of temperature for single models, one model per row. The columns show from left to right the patterns associated with the inter-annual mode, the inter-decadal mode and the inter-centennial mode.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

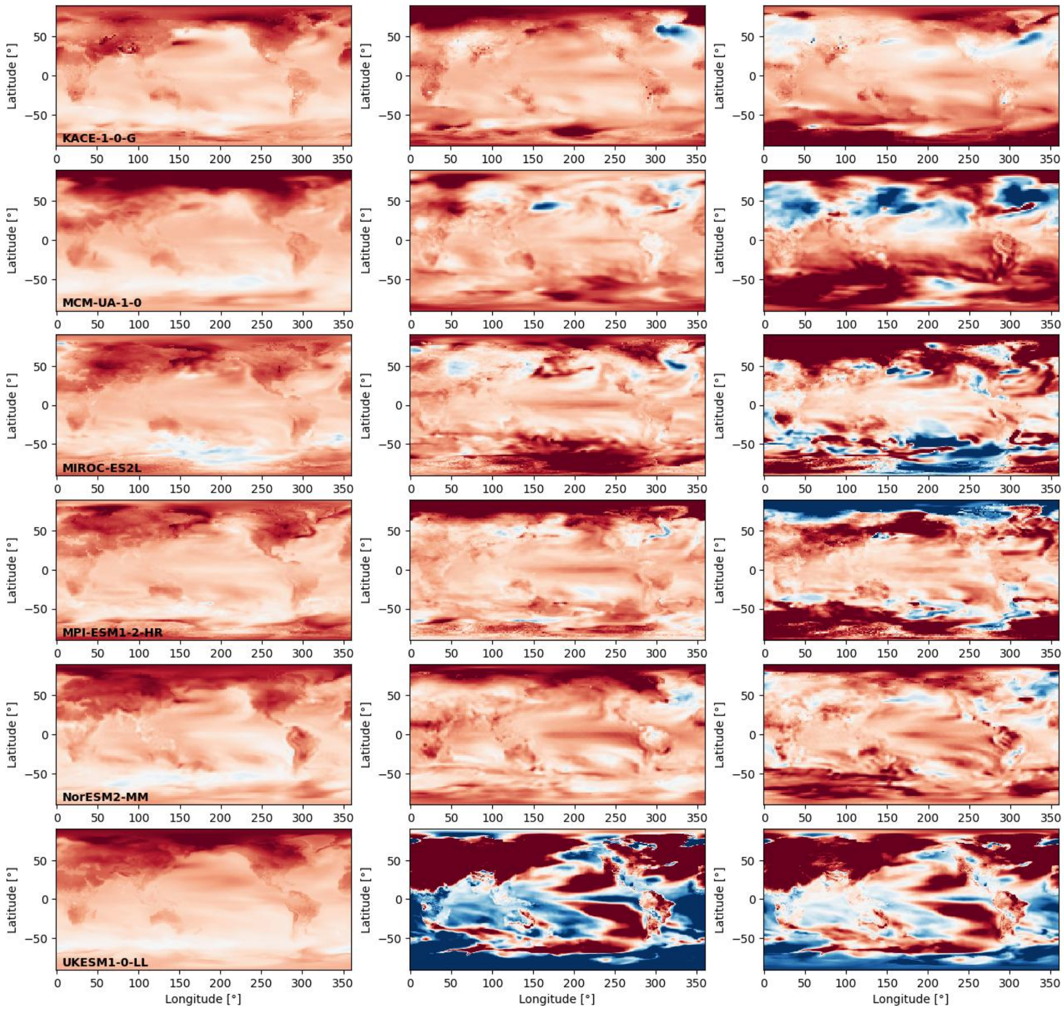

Figure B3. GHG patterns of temperature for single models, one model per row. The columns show from left to right the patterns associated with the inter-annual mode, the inter-decadal mode and the inter-centennial mode.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

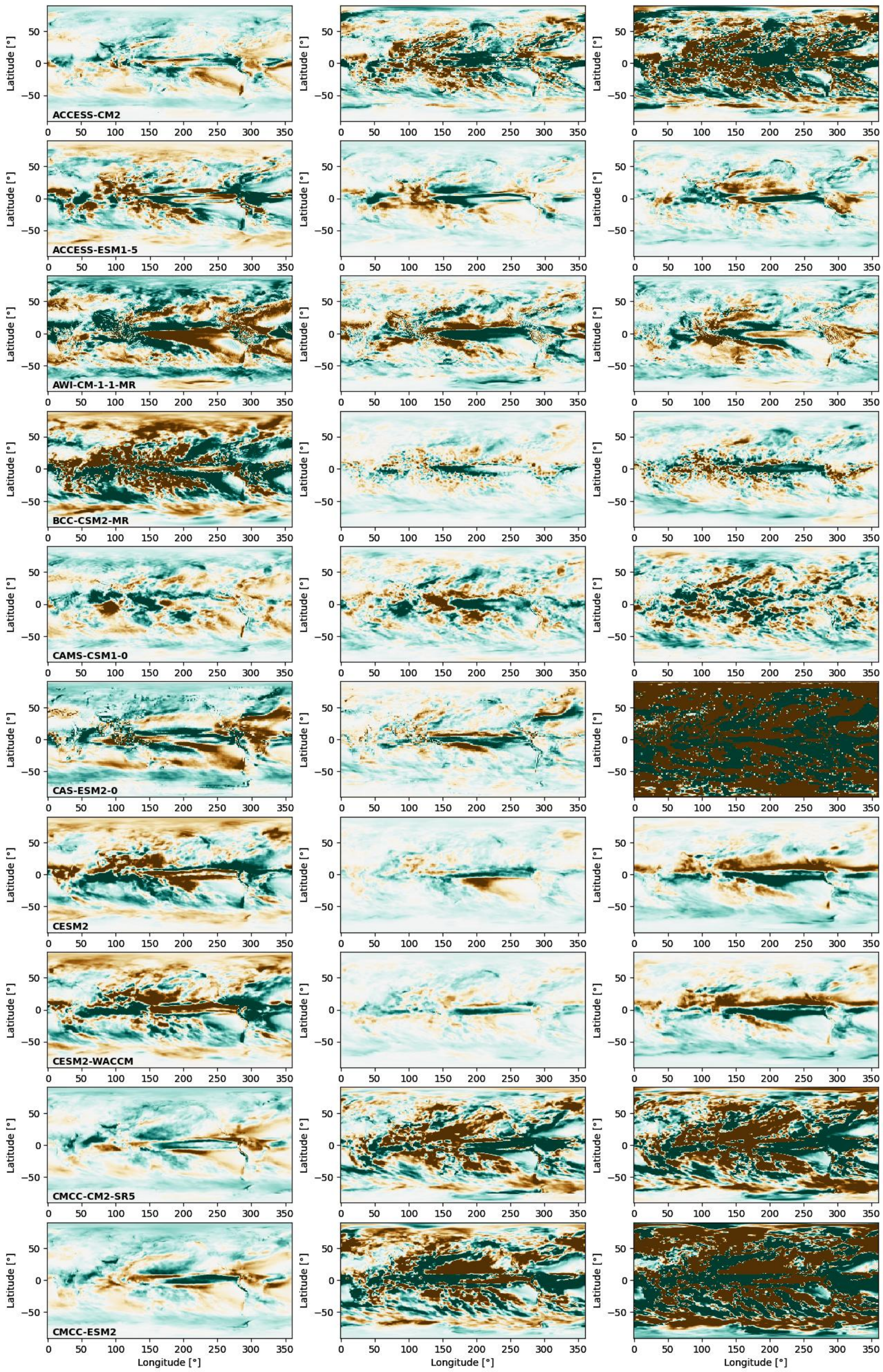

Figure B4. GHG patterns of precipitation for single models, one model per row. The columns show from left to right the patterns associated with the inter-annual mode, the inter-decadal mode and the inter-centennial mode.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

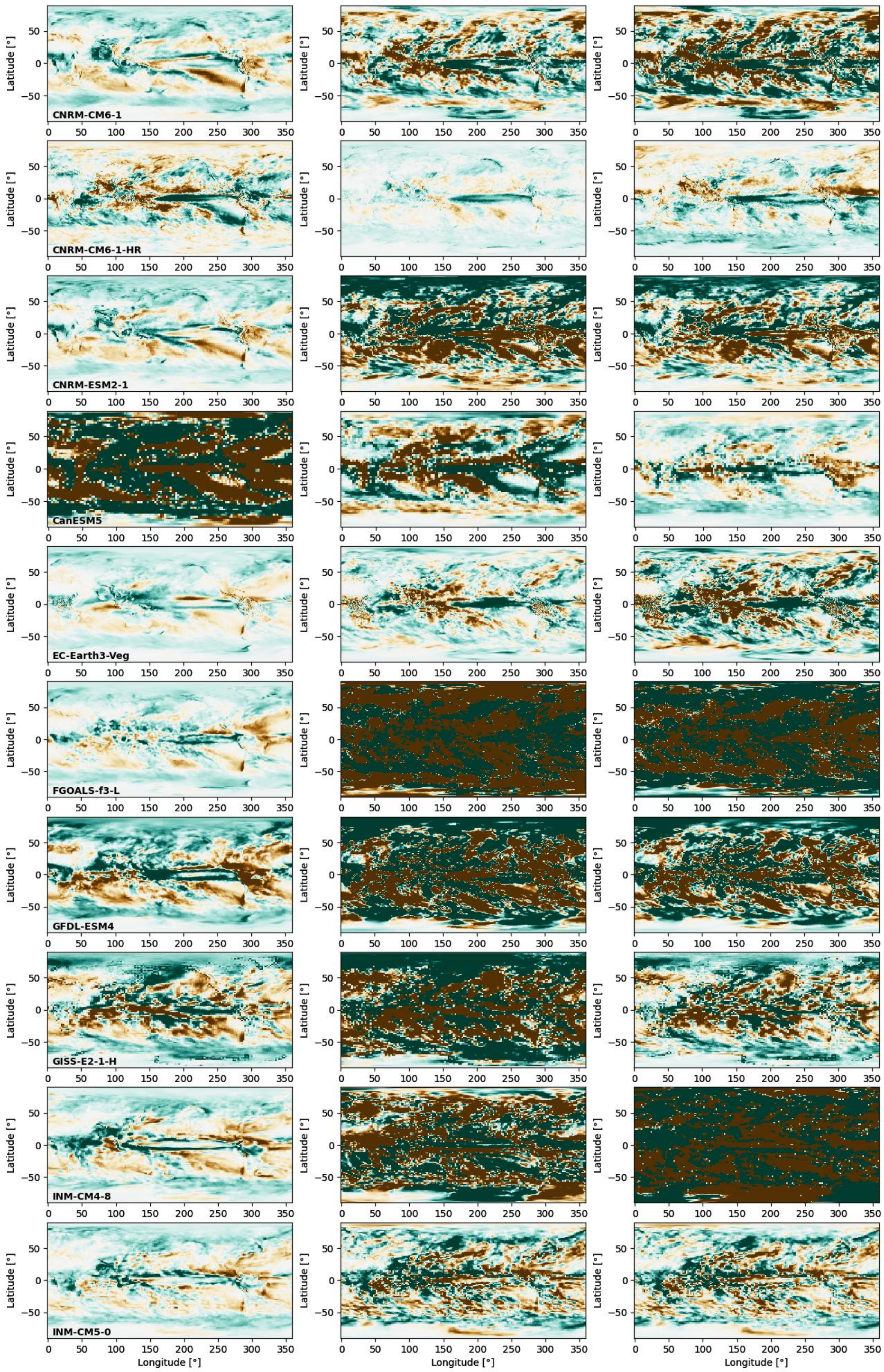

Figure B5. GHG patterns of precipitation for single models, one model per row. The columns show from left to right the patterns associated with the inter-annual mode, the inter-decadal mode and the inter-centennial mode.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

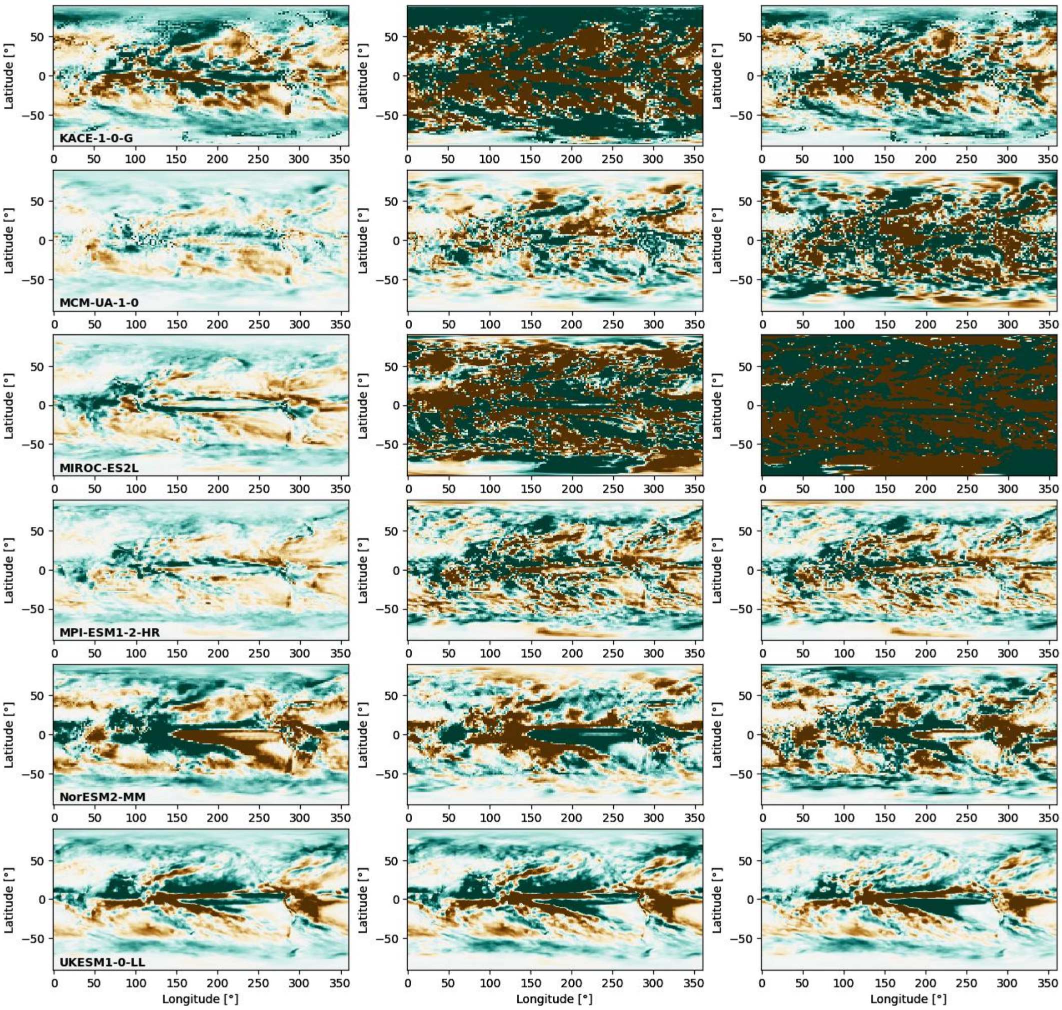

Figure B6. GHG patterns of precipitation for single models, one model per row. The columns show from left to right the patterns associated with the inter-annual mode, the inter-decadal mode and the inter-centennial mode.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

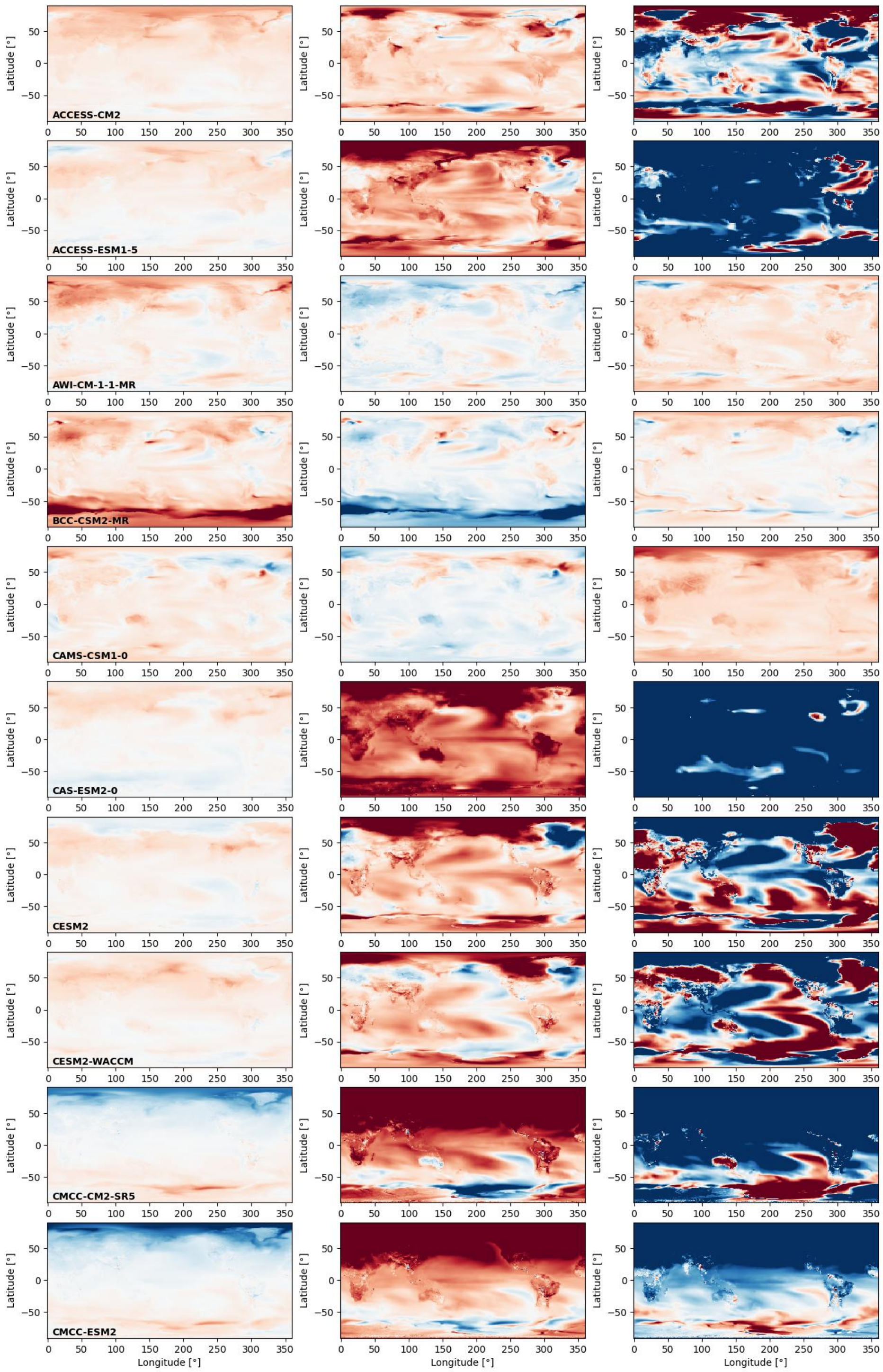

Figure B7. Residual sulfate aerosol patterns of temperature for single models, one model per row. The columns show from left to right the patterns associated with the inter-annual mode, the inter-decadal mode and the inter-centennial mode.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

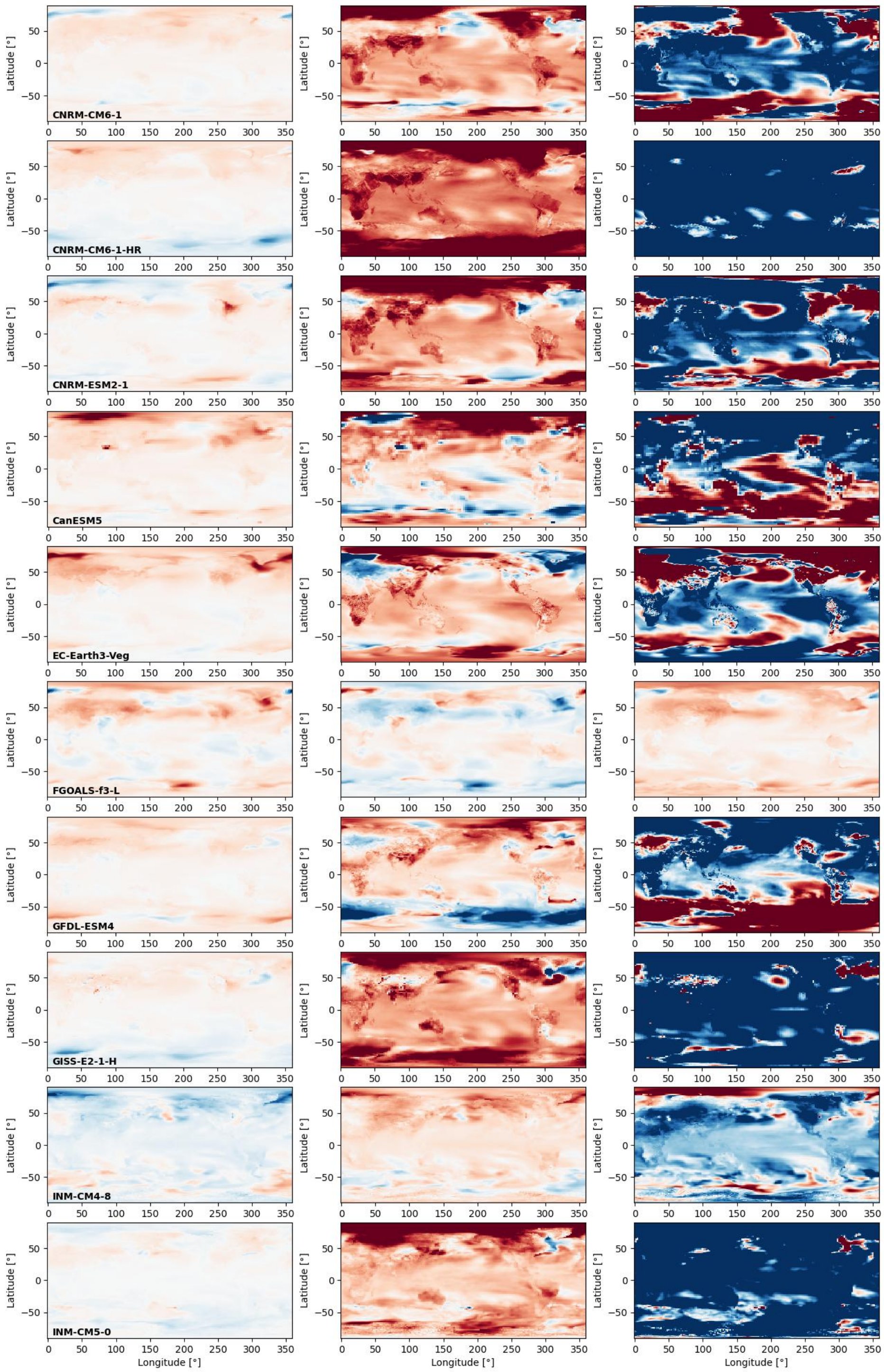

Figure B8. Residual sulfate aerosol patterns of temperature for single models, one model per row. The columns show from left to right the patterns associated with the inter-annual mode, the inter-decadal mode and the inter-centennial mode.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

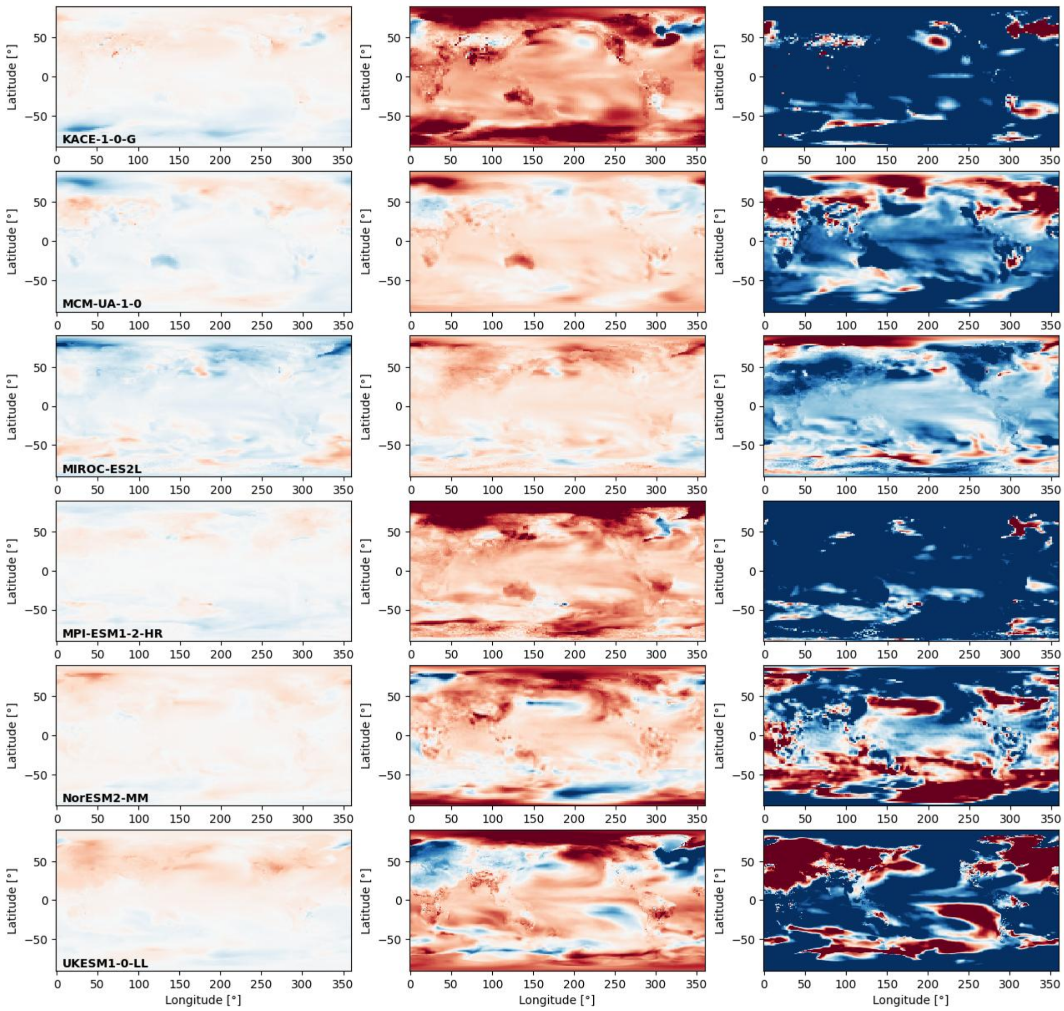

Figure B9. Residual sulfate aerosol patterns of temperature for single models, one model per row. The columns show from left to right the patterns associated with the inter-annual mode, the inter-decadal mode and the inter-centennial mode.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

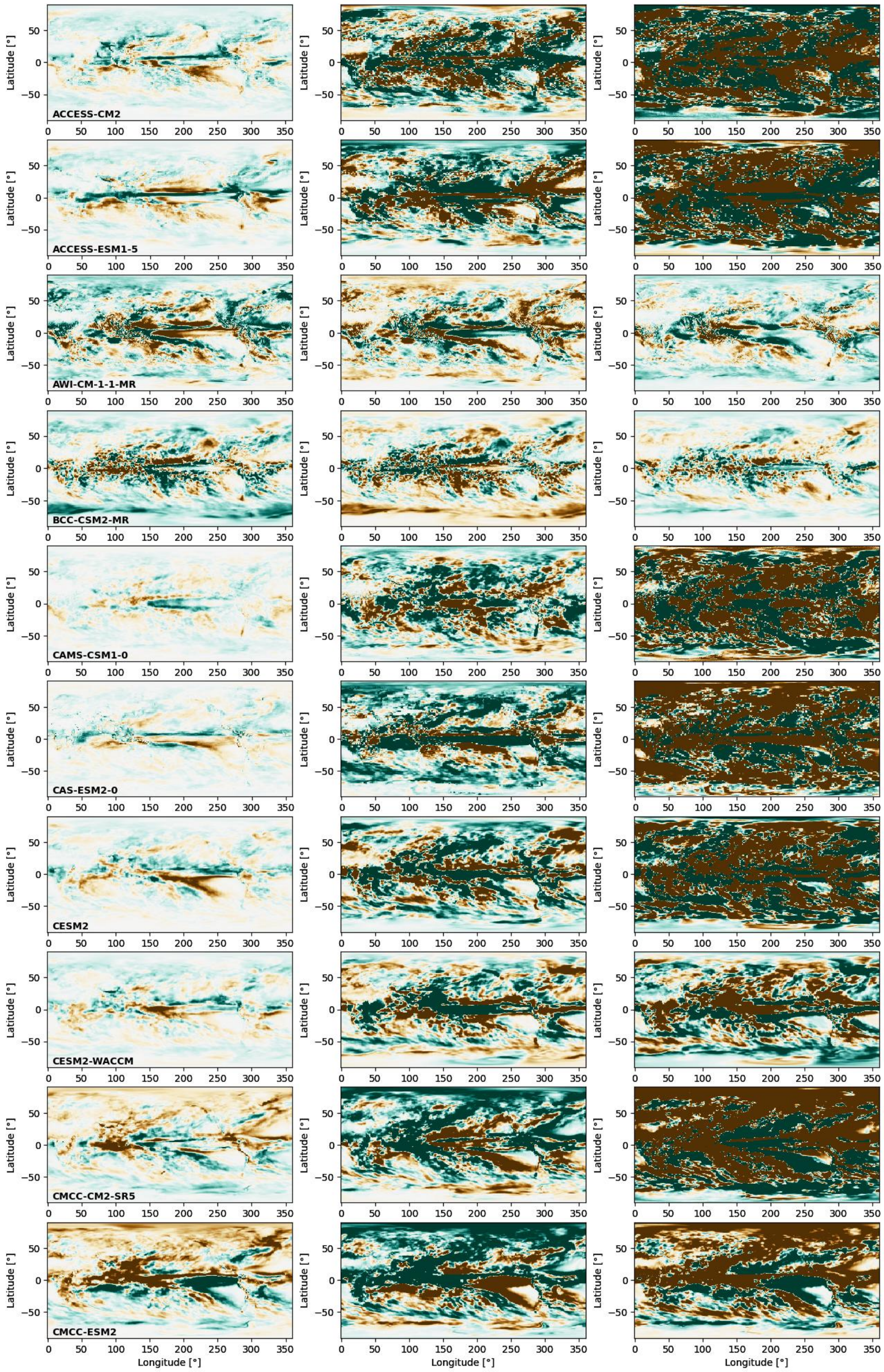

Figure B10. Residual sulfate aerosol patterns of precipitation for single models, one model per row. The columns show from left to right the patterns associated with the inter-annual mode, the inter-decadal mode and the inter-centennial mode.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

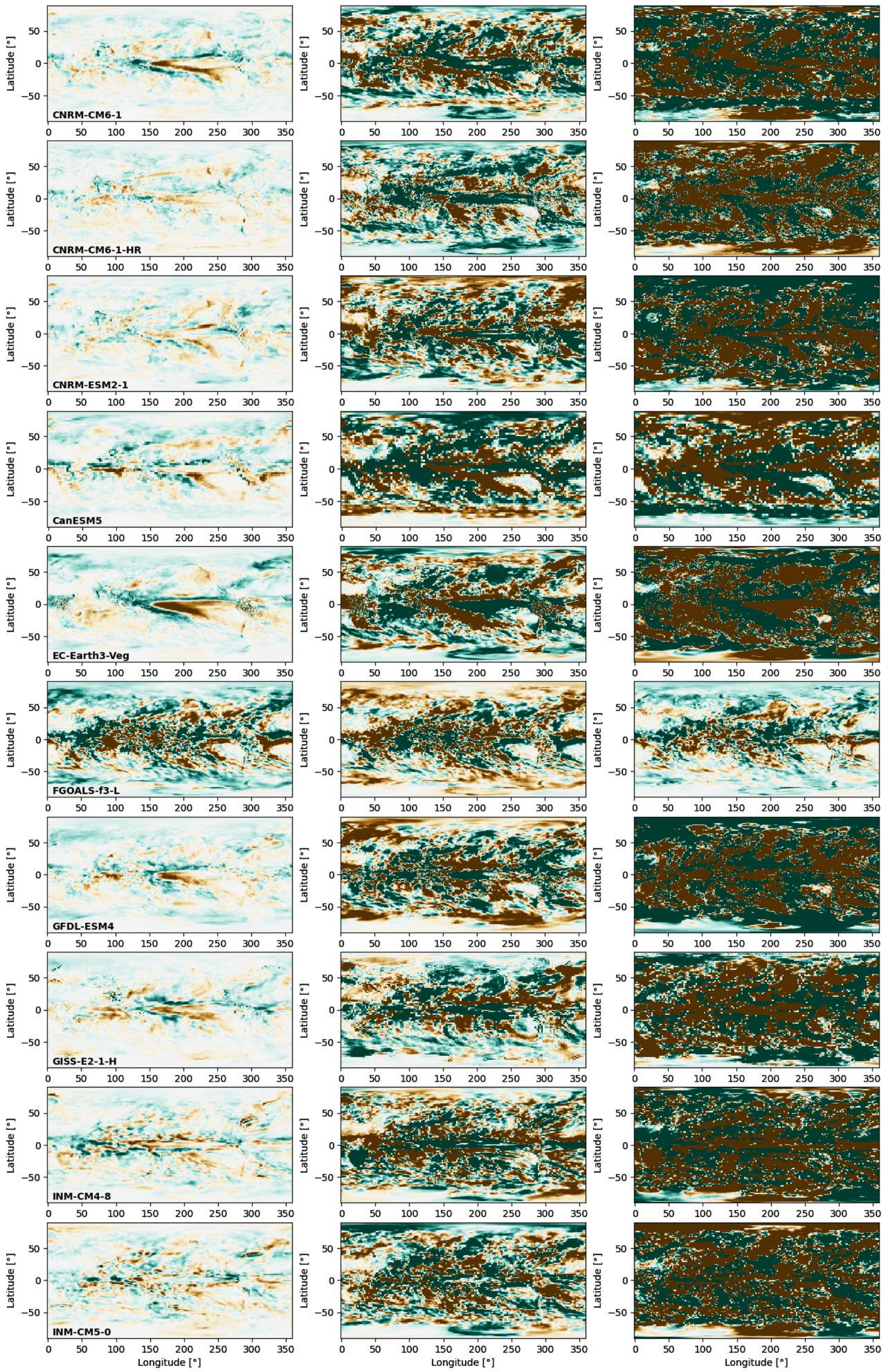

Figure B11. Residual sulfate aerosol patterns of precipitation for single models, one model per row. The columns show from left to right the patterns associated with the inter-annual mode, the inter-decadal mode and the inter-centennial mode.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

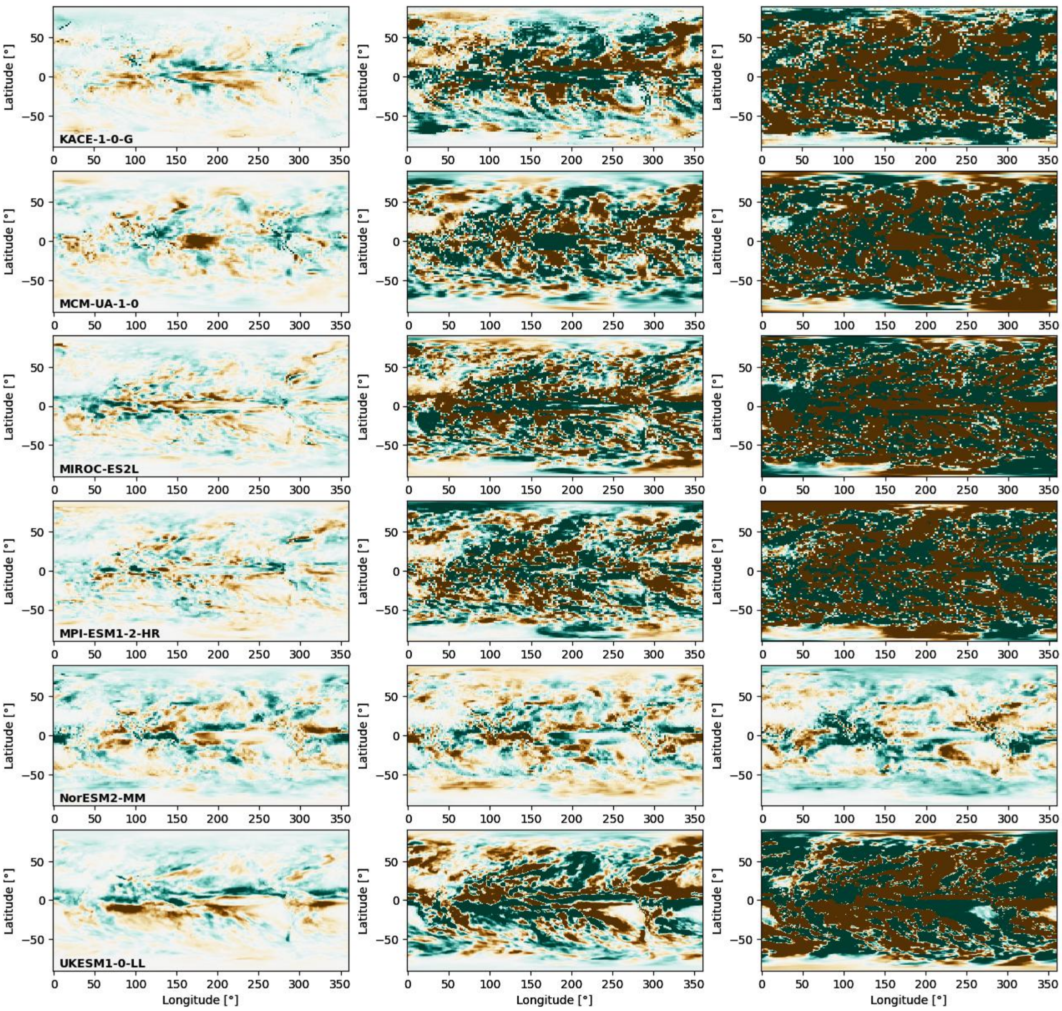

Figure B12. Residual sulfate aerosol patterns of precipitation for single models, one model per row. The columns show from left to right the patterns associated with the inter-annual mode, the inter-decadal mode and the inter-centennial mode.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

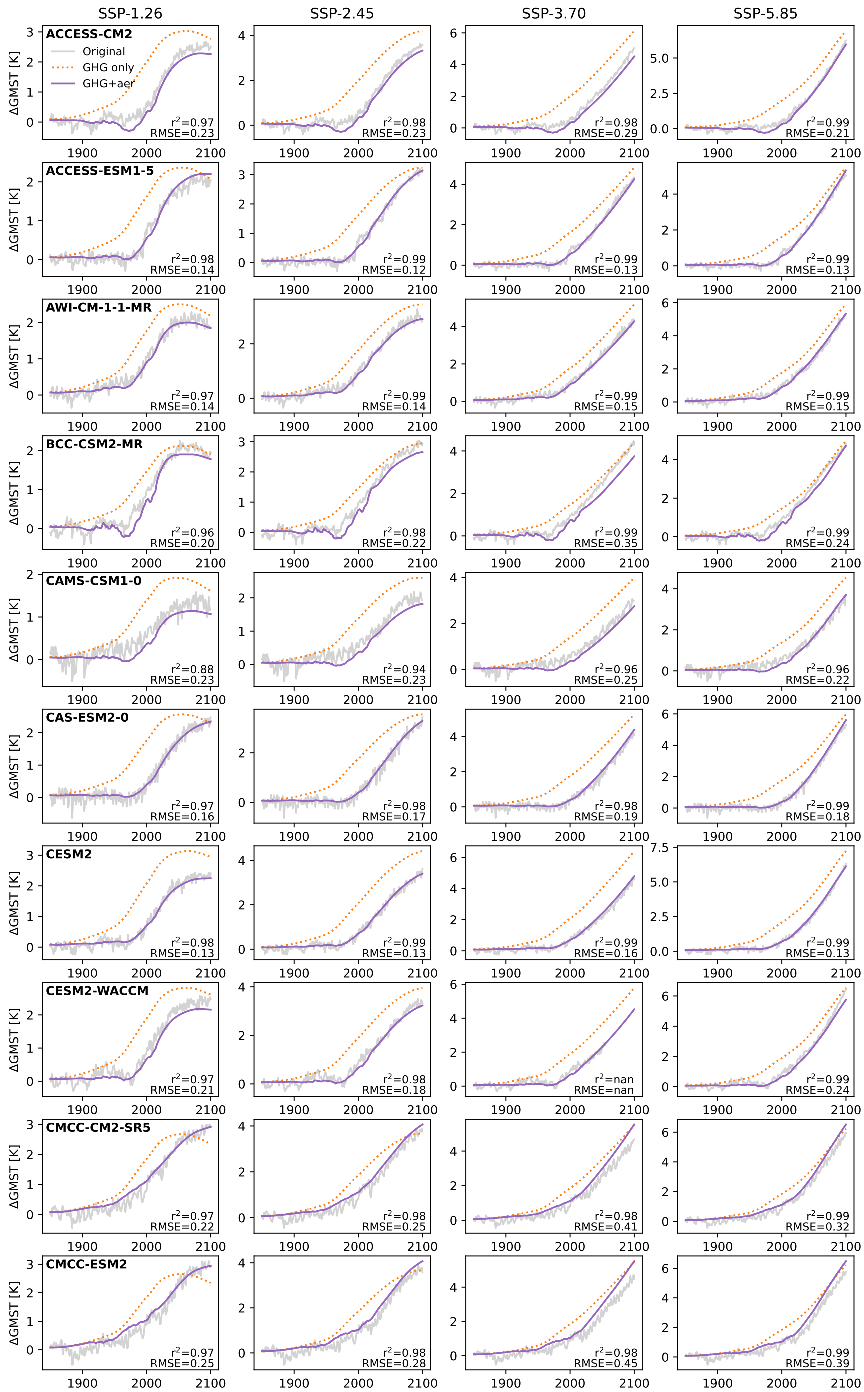

Figure B13. GMST original data (grey), fit reconstruction for only GHG (orange) and full reconstruction (purple) for single models, one model per row. The results were obtained using *SSP2-4.5* to fit the aerosol sulfate residual. From left to right show the columns therefore show *SSP1-2.6* (out-of-sample), *SSP2-4.5* (in-sample), *SSP3-7.0* (out-of-sample) and *SSP5-8.5* (out-of-sample) fits.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

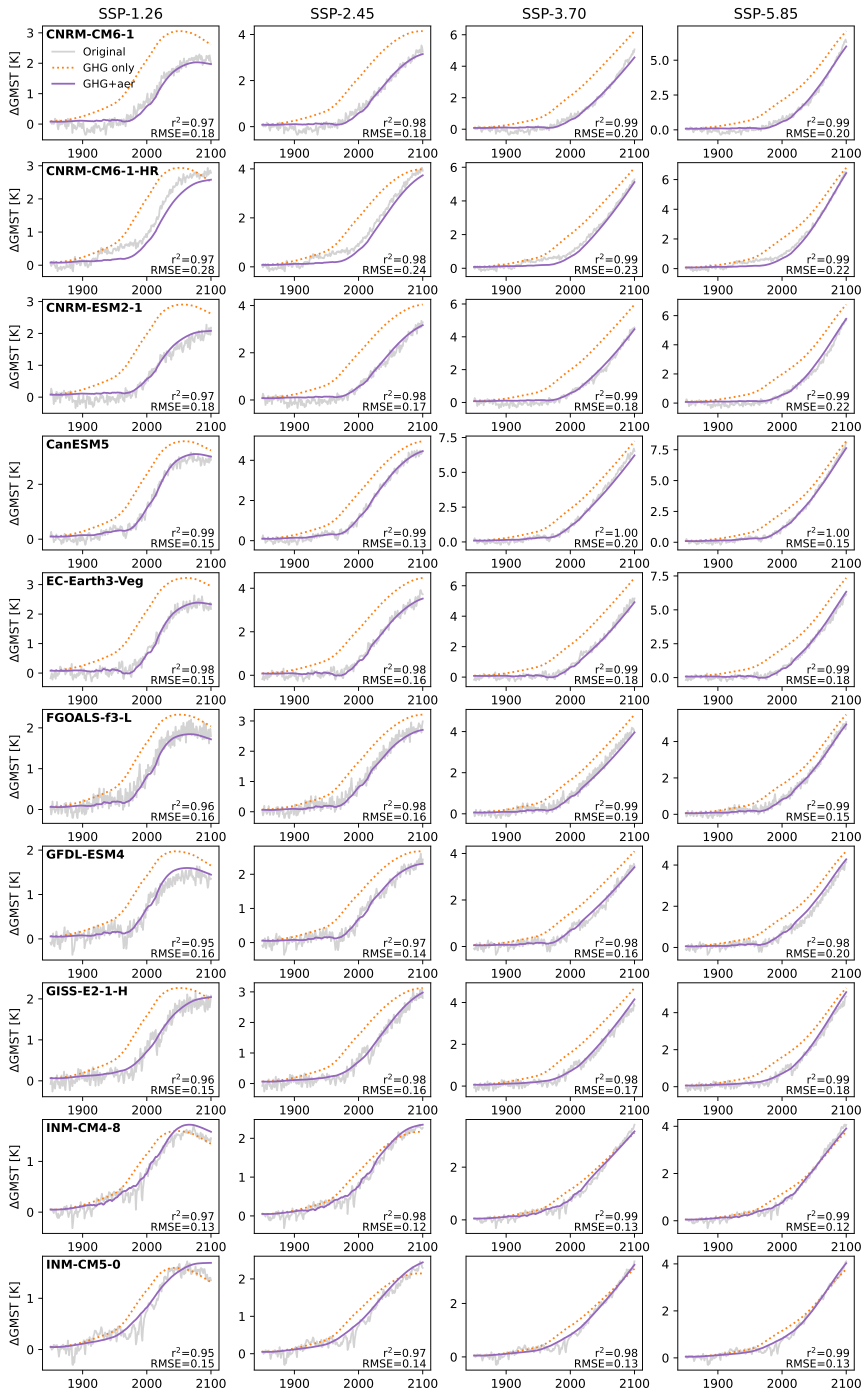

Figure B14. GMST original data (grey), fit reconstruction for only GHG (orange) and full reconstruction (purple) for single models, one model per row. The results were obtained using *SSP2-4.5* to fit the aerosol sulfate residual. From left to right show the columns therefore show *SSP1-2.6* (out-of-sample), *SSP2-4.5* (in-sample), *SSP3-7.0* (out-of-sample) and *SSP5-8.5* (out-of-sample) fits.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

Table B1. METEOR pattern scaling timescales τ split into inter-annual (0–10 years), inter-decadal (10–100 years) and inter-centennial (100-1000 years) timescales provided for each CMIP6 model, for temperature and precipitation and greenhouse gas (GHG) and residuals (aer) responses, respectively. Note that in METEORv1.0.1, the respective timescales are forced to fit within the specified ranges.

Earth System Model	**Temperature**						**Precipitation**					
	GHG			aer			GHG			aer		
	τ_1	τ_2	τ_3	τ_1	τ_2	τ_3	τ_1	τ_2	τ_3	τ_1	τ_2	τ_3
ACCESS-CM2	1.0	12.7	100.0	1.1	73.4	998.1	6.4	100.0	1000.0	1.2	95.4	719.1
ACCESS-ESM1-5	1.2	12.9	100.0	1.8	100.0	1000.0	1.0	10.0	996.0	2.4	100.0	1000.0
AWI-CM-1-1-MR	1.0	10.0	490.0	2.3	18.9	102.8	3.0	10.0	1000.0	2.0	10.0	163.4
BCC-CSM2-MR	1.0	10.0	974.1	1.0	10.0	100.0	1.0	10.0	1000.0	1.0	11.3	100.0
CAMS-CSM1-0	1.0	10.9	980.2	1.0	10.0	100.0	2.0	12.0	1000.0	1.0	100.0	1000.0
CAS-ESM2-0	1.0	12.2	298.0	1.4	100.0	1000.0	4.7	31.6	837.3	1.0	100.0	1000.0
CESM2	1.1	10.0	993.4	2.3	100.0	1000.0	1.0	10.0	995.4	3.9	100.0	1000.0
CESM2-WACCM	1.4	10.0	999.9	1.1	100.0	1000.0	1.0	10.0	999.8	2.3	58.0	197.2
CMCC-CM2-SR5	1.0	10.0	123.3	1.0	100.0	538.3	6.9	100.0	1000.0	8.6	59.5	1000.0
CMCC-ESM2	1.0	10.0	212.9	1.0	59.7	221.6	7.2	100.0	1000.0	10.0	41.0	179.6
CNRM-CM6-1	1.0	15.0	100.0	1.0	100.0	1000.0	7.7	100.0	1000.0	1.0	100.0	1000.0
CNRM-CM6-1-HR	1.1	15.1	1000.0	1.0	100.0	1000.0	1.0	10.5	687.8	1.0	100.0	1000.0
CNRM-ESM2-1	1.2	10.0	999.9	1.0	100.0	1000.0	10.0	99.0	711.3	1.0	100.0	1000.0
CanESM5	1.2	13.0	429.9	5.7	100.0	1000.0	10.0	12.8	1000.0	4.5	99.1	281.6
EC-Earth3-Veg	1.0	10.9	100.0	1.0	100.0	1000.0	4.9	100.0	1000.0	1.0	100.0	1000.0
FGOALS-f3-L	1.0	10.0	975.6	1.6	10.0	100.0	6.7	100.0	116.8	7.2	16.9	100.0
GFDL-ESM4	1.0	10.0	999.9	1.6	100.0	1000.0	10.0	100.0	791.5	1.0	100.0	1000.0
GISS-E2-1-H	1.0	10.0	939.4	2.5	100.0	1000.0	10.0	24.0	100.0	1.3	100.0	1000.0
INM-CM4-8	1.0	14.2	964.7	1.0	10.0	1000.0	8.1	100.0	1000.0	9.3	100.0	1000.0
INM-CM5-0	1.0	10.7	1000.0	1.0	100.0	1000.0	10.0	49.5	100.0	10.0	100.0	1000.0
KACE-1-0-G	1.0	10.0	939.4	2.5	100.0	1000.0	10.0	24.9	100.0	1.3	100.0	1000.0
MCM-UA-1-0	1.0	10.0	759.0	1.0	45.0	1000.0	2.9	18.9	999.9	4.3	100.0	1000.0
MIROC-ES2L	1.0	14.2	964.7	1.0	10.0	1000.0	8.1	100.0	1000.0	9.3	100.0	1000.0
MPI-ESM1-2-HR	1.0	10.7	1000.0	1.0	100.0	1000.0	10.0	49.5	100.0	10.0	100.0	1000.0
NorESM2-MM	1.0	10.0	1000.0	1.2	100.0	1000.0	4.2	13.3	1000.0	1.0	15.1	100.0
UKESM1-0-LL	1.6	100.0	1000.0	2.5	99.9	999.1	10.0	16.1	100.0	5.2	100.0	1000.0

Acknowledgements. The European Union through the Horizon 2020 project PROVIDE (grant agreement No. 101003687) and the Horizon Europe Research and Innovation programme project DIAMOND (grant no. 101081179) funded this work. We thank Bjørn Hallvard Samset and Camilla Weum Stjern for clarifying discussions on the topic of aerosols. ChatGPT was used in the initial formulation of the introduction all of which has since been checked and rewritten.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

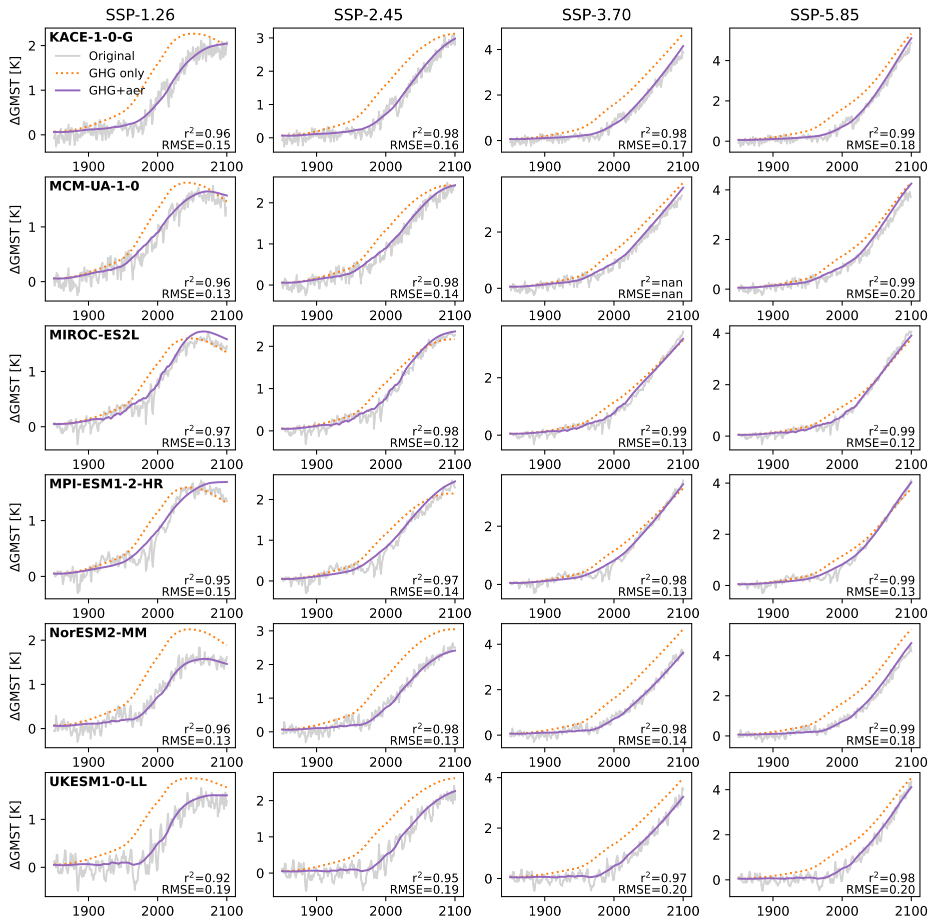

Figure B15. GMST original data (grey), fit reconstruction for only GHG (orange) and full reconstruction (purple) for single models, one model per row. The results were obtained using *SSP2-4.5* to fit the aerosol sulfate residual. From left to right show the columns therefore show *SSP1-2.6* (out-of-sample), *SSP2-4.5* (in-sample), *SSP3-7.0* (out-of-sample) and *SSP5-8.5* (out-of-sample) fits.

References

Barata, J. C. A. and Hussein, M. S.: The Moore–Penrose pseudoinverse: A tutorial review of the theory, Brazilian Journal of Physics, 42, 146–165, 2012.

Beusch, L., Gudmundsson, L., and Seneviratne, S. I.: Emulating Earth system model temperatures with MESMER: from global mean temperature trajectories to grid-point-level realizations on land, Earth System Dynamics, 11, 139–159, https://doi.org/10.5194/esd-11-139-2020, 2020.

Beusch, L., Nicholls, Z., Gudmundsson, L., Hauser, M., Meinshausen, M., and Seneviratne, S. I.: From emission scenarios to spatially resolved projections with a chain of computationally efficient emulators: coupling of MAGICC (v7.5.1) and MESMER (v0.8.3), Geoscientific Model Development, 15, 2085–2103, https://doi.org/10.5194/gmd-15-2085-2022, 2022.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

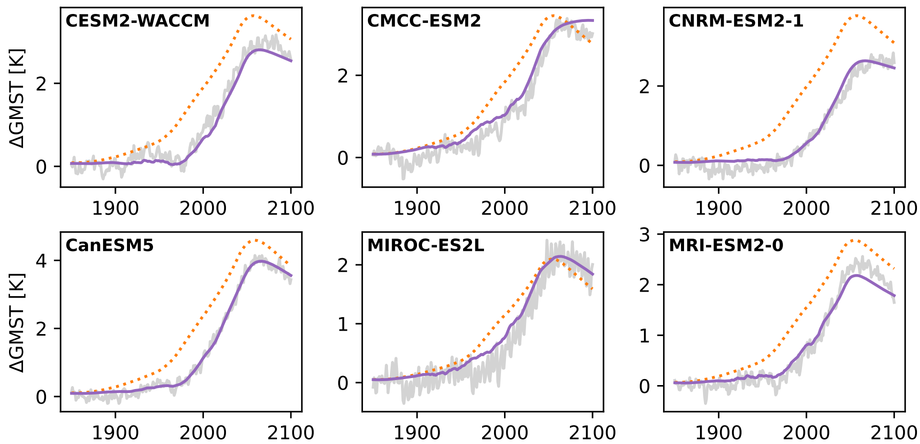

Figure B16. GMST original data (grey), fit reconstruction for only GHG (orange) and full reconstruction (purple) for single models that had reasonable data quality for the overshoot scenario *SSP5-3.4-over*. The results were obtained using *SSP2-4.5* to fit the aerosol sulfate residual. Each sub-plot show results for a single model.

Bouabid, S., Sejdinovic, D., and Watson-Parris, D.: FaIRGP: A Bayesian Energy Balance Model for Surface Temperatures Emulation, Journal of Advances in Modeling Earth Systems, 16, e2023MS003 926, https://doi.org/https://doi.org/10.1029/2023MS003926, e2023MS003926 2023MS003926, 2024.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A., and Wehner, M.: Long-term Climate Change: Projections, Commitments and Irreversibility, 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, edited by Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., book section 12, p. 1029–1136, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, ISBN ISBN 978-1-107-66182-0, https://doi.org/10.1017/CBO9781107415324.024, 2013.

Dunne, J. P., Hewitt, H. T., Arblaster, J., Bonou, F., Boucher, O., Cavazos, T., Durack, P. J., Hassler, B., Juckes, M., Miyakawa, T., Mizielinski, M., Naik, V., Nicholls, Z., O'Rourke, E., Pincus, R., Sanderson, B. M., Simpson, I. R., and Taylor, K. E.: An evolving Coupled Model Intercomparison Project phase 7 (CMIP7) and Fast Track in support of future climate assessment, EGUsphere, 2024, 1–51, https://doi.org/10.5194/egusphere-2024-3874, 2024.

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

Ferrari, L., Carlino, A., Gazzotti, P., Tavoni, M., and Castelletti, A.: From optimal to robust climate strategies: expanding integrated assessment model ensembles to manage economic, social, and environmental objectives, Environmental Research Letters, 17, 084 029, 2022.

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., Lunt, D., Mauritsen, T., Palmer, M., Watanabe, M., Wild, M., and Zhang, H.: The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

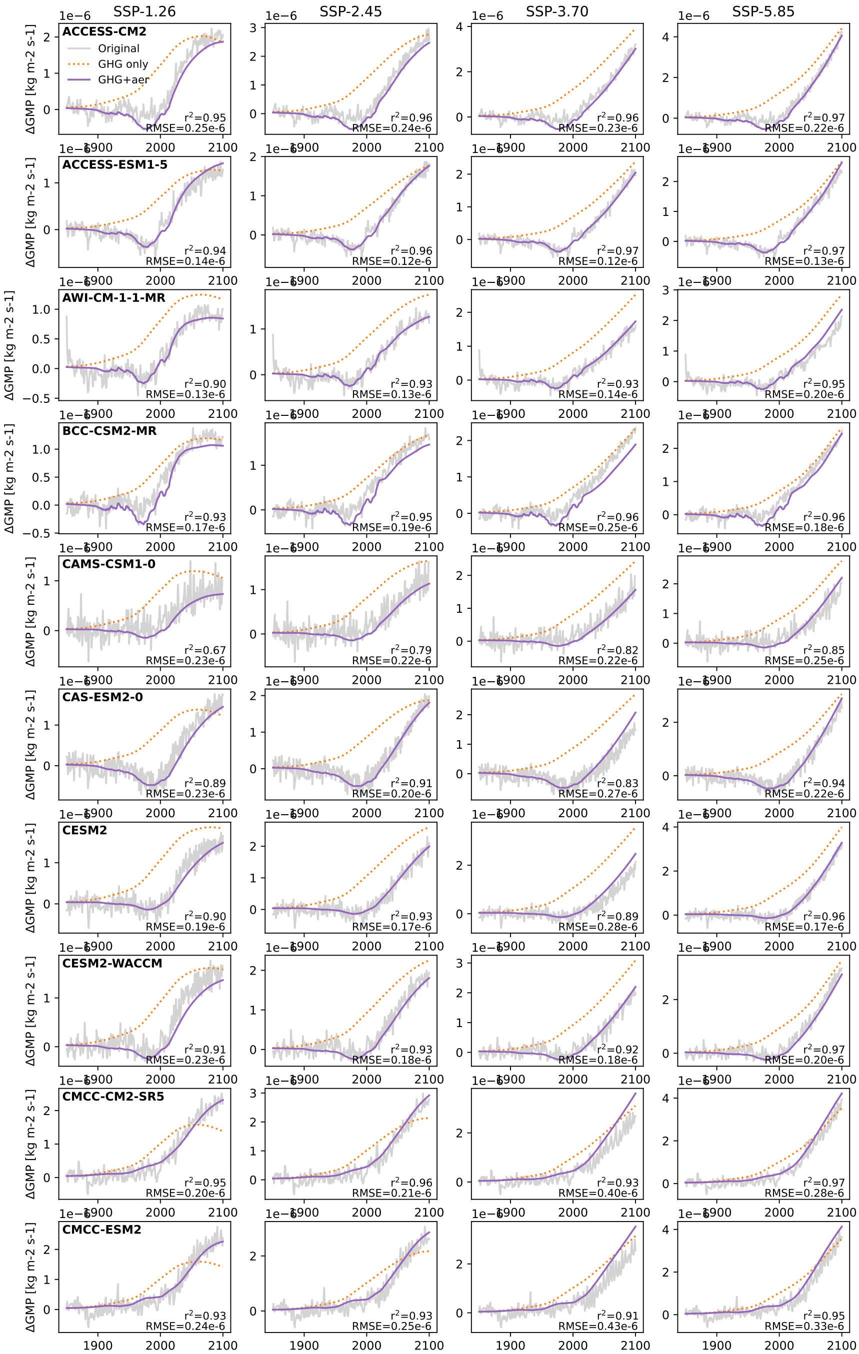

Figure B17. GMP original data (grey), fit reconstruction for only GHG (orange) and full reconstruction (purple) for single models, one model per row. The results were obtained using *SSP2-4.5* to fit the aerosol sulfate residual. From left to right show the columns therefore show *SSP1-2.6* (out-of-sample), *SSP2-4.5* (in-sample), *SSP3-7.0* (out-of-sample) and *SSP5-8.5* (out-of-sample) fits.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

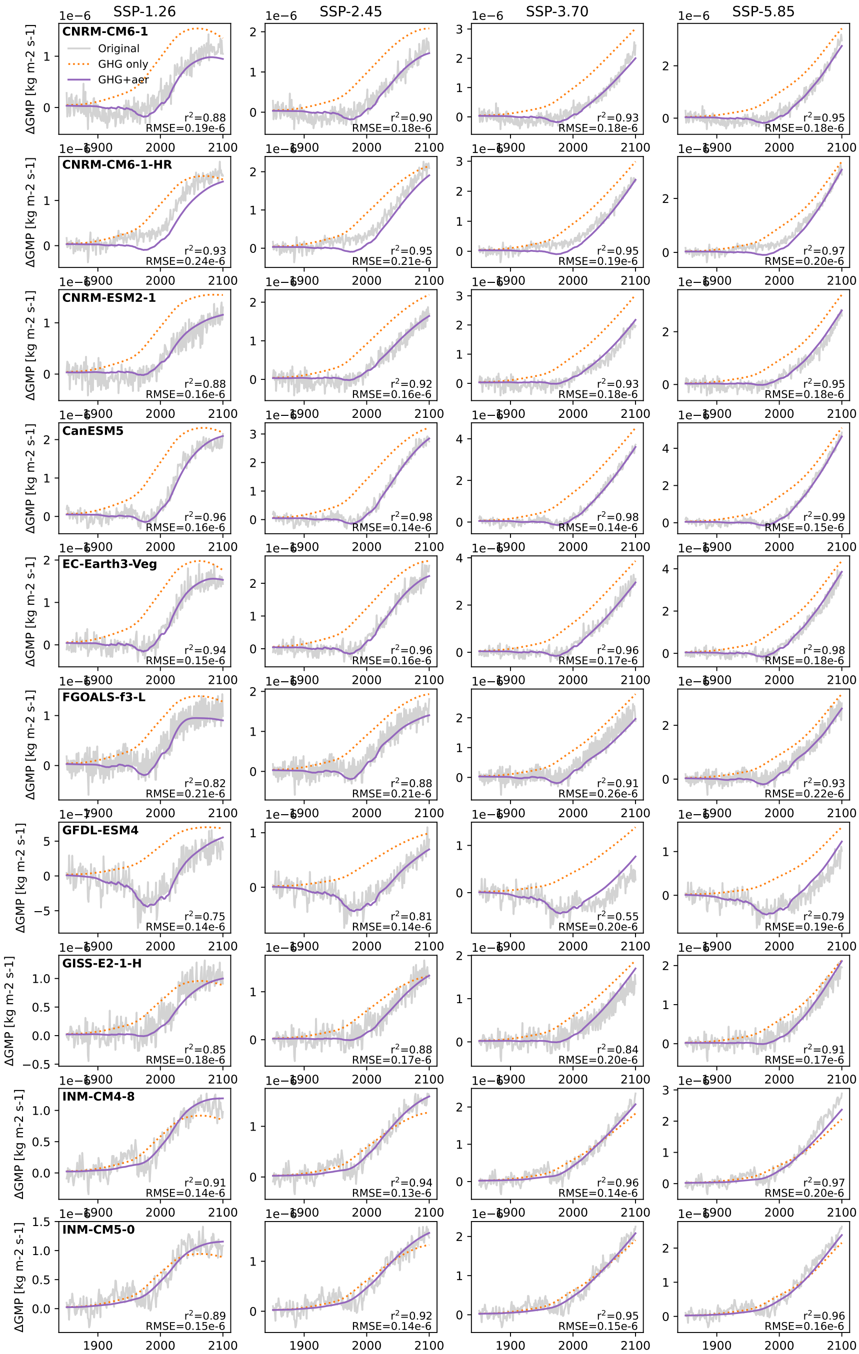

Figure B18. GMP original data (grey), fit reconstruction for only GHG (orange) and full reconstruction (purple) for single models, one model per row. The results were obtained using *SSP2-4.5* to fit the aerosol sulfate residual. From left to right show the columns therefore show *SSP1-2.6* (out-of-sample), *SSP2-4.5* (in-sample), *SSP3-7.0* (out-of-sample) and *SSP5-8.5* (out-of-sample) fits.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

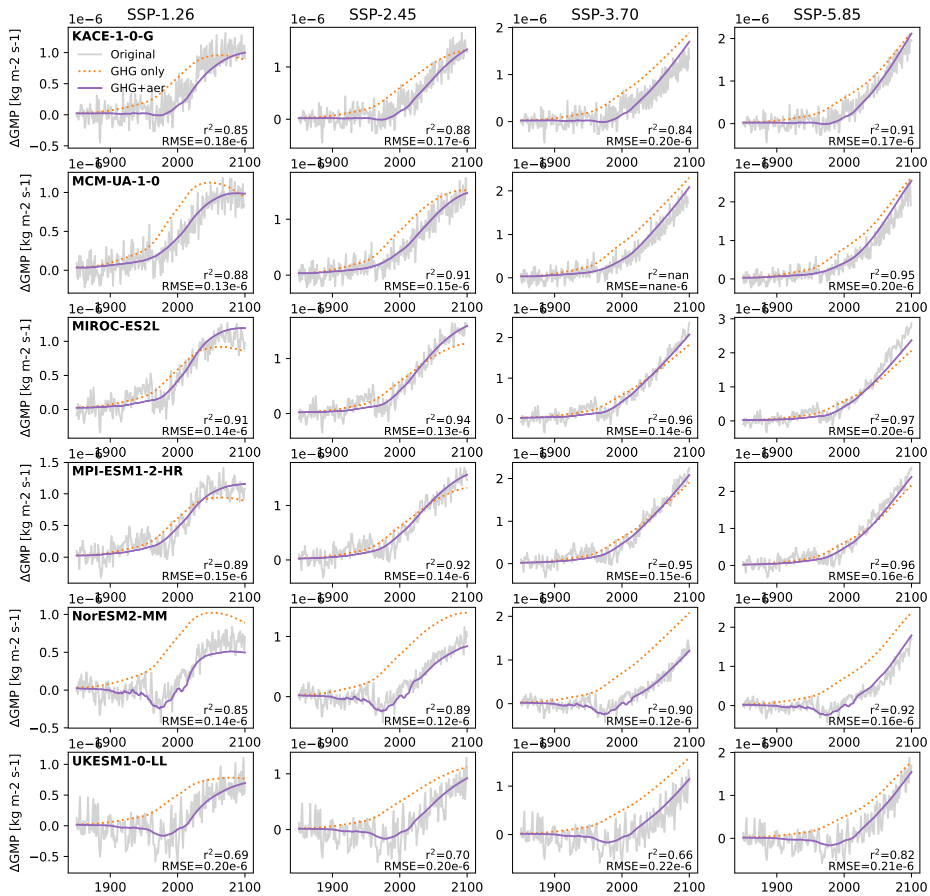

Figure B19. GMP original data (grey), fit reconstruction for only GHG (orange) and full reconstruction (purple) for single models, one model per row. The results were obtained using *SSP2-4.5* to fit the aerosol sulfate residual. From left to right show the columns therefore show *SSP1-2.6* (out-of-sample), *SSP2-4.5* (in-sample), *SSP3-7.0* (out-of-sample) and *SSP5-8.5* (out-of-sample) fits.

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., book section 7, pp. 923–1054, Cambridge University Press, Cambridge, UK and New York, NY, USA, https://doi.org/10.1017/9781009157896.009, 2021.

Good, P., Lowe, J. A., Andrews, T., Wiltshire, A., Chadwick, R., Ridley, J. K., Menary, M. B., Bouttes, N., Dufresne, J. L., Gregory, J. M., Schaller, N., and Shiogama, H.: Nonlinear regional warming with increasing CO2 concentrations, Nature Climate Change, 5, https://doi.org/10.1038/nclimate2498, 2015.

Herger, N., Sanderson, B. M., and Knutti, R.: Improved pattern scaling approaches for the use in climate impact studies, Geophysical Research Letters, 42, 3486–3494, https://doi.org/https://doi.org/10.1002/2015GL063569, 2015.

Hulme, M., Wigley, T. M. L., Barrow, E. andRaper, S., Centella, A., Smith, S., and Chipanshi, A.: Using a Climate Scenario Generator for Vulnerability and Adaptation Assessments, in: MAGICC and SCENGEN Version 2.4 Workbook and CD-ROM, edited by Lim, B.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

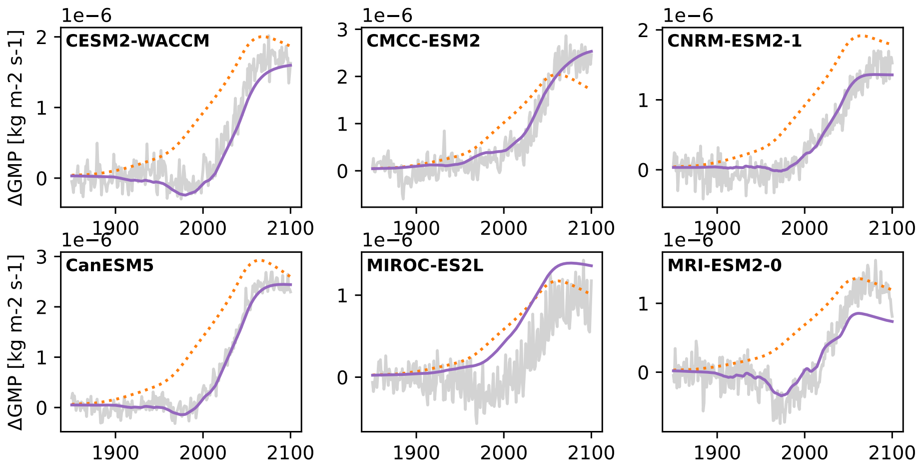

Figure B20. GMP original data (grey), fit reconstruction for only GHG (orange) and full reconstruction (purple) for single models that had reasonable data quality for the overshoot scenario *SSP5-3.4-over*. The results were obtained using *SSP2-4.5* to fit the aerosol sulfate residual. Each sub-plot show results for a single model.

and Smith, J., Reprint, East Anglia: University of East Anglia. Climatic Research Unit, https://opensky.ucar.edu/islandora/object/books%3A487, 2000.

Huntingford, C. et al.: Climate sensitivity analysis of forest growth uncertainties, Global Ecology and Biogeography, 9, 359–374, 2000.

IPCC: Climate Change 2021: The Physical Science Basis, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021.

IPCC: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, USA, https://doi.org/10.1017/9781009325844, 2022.

Mansfield, L. A., Nowack, P. J., Kasoar, M., Everitt, R. G., Collins, W. J., and Voulgarakis, A.: Predicting global patterns of long-term climate change from short-term simulations using machine learning, npj Climate and Atmospheric Science, 3, https://doi.org/10.1038/s41612-020-00148-5, 2020.

Mathison, C. T., Burke, E., Kovacs, E., Munday, G., Huntingford, C., Jones, C., Smith, C., Steinert, N., Wiltshire, A., Gohar, L., and Varney, R.: A rapid application emissions-to-impacts tool for scenario assessment: Probabilistic Regional Impacts from Model patterns and Emissions (PRIME), EGUsphere [preprint], pp. 1–28, https://doi.org/https://doi.org/10.5194/egusphere-2023-2932, 2024.

Mitchell, T. D.: Pattern Scaling: An Examination of the Accuracy of the Technique for Describing Future Climates, Climatic Change, 60, 217–242, https://doi.org/10.1023/A:1026035305597, 2003.

Monerie, P.-A., Wilcox, L. J., and Turner, A. G.: Effects of Anthropogenic Aerosol and Greenhouse Gas Emissions on Northern Hemisphere Monsoon Precipitation: Mechanisms and Uncertainty, Journal of Climate, 35, 2305 – 2326, https://doi.org/10.1175/JCLI-D-21-0412.1, 2022.

Myhre, G., Forster, P. M., Samset, B. H., Hodnebrog, Ø., Sillmann, J., Aalbergsjø, S. G., Andrews, T., Boucher, O., Faluvegi, G., Fläschner, D., Iversen, T., Kasoar, M., Kharin, V., Kirkevåg, A., Lamarque, J.-F., Olivié, D., Richardson, T. B., Shindell, D., Shine, K. P., Stjern, C. W.,

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

Takemura, T., Voulgarakis, A., and Zwiers, F.: PDRMIP: A Precipitation Driver and Response Model Intercomparison Project—Protocol and Preliminary Results, Bulletin of the American Meteorological Society, 98, 1185 – 1198, https://doi.org/10.1175/BAMS-D-16-0019.1, 2017.

Nath, S., Lejeune, Q., Beusch, L., Seneviratne, S. I., and Schleussner, C.-F.: MESMER-M: an Earth system model emulator for spatially resolved monthly temperature, Earth System Dynamics, 13, 851–877, https://doi.org/10.5194/esd-13-851-2022, 2022.

Nicholls, Z. and Gieseke, R.: RCMIP Phase 1 Data (v2.0.0), https://doi.org/10.5281/zenodo.4016613, https://doi.org/https://doi.org/10.5281/zenodo.4016613, dataset, 2019.

Nicholls, Z. R. J., Meinshausen, M., Lewis, J., Gieseke, R., Dommenget, D., Dorheim, K., Fan, C.-S., Fuglestvedt, J. S., Gasser, T., Golüke, U., Goodwin, P., Hartin, C., Hope, A. P., Kriegler, E., Leach, N. J., Marchegiani, D., McBride, L. A., Quilcaille, Y., Rogelj, J., Salawitch, R. J., Samset, B. H., Sandstad, M., Shiklomanov, A. N., Skeie, R. B., Smith, C. J., Smith, S., Tanaka, K., Tsutsui, J., and Xie, Z.: Reduced Complexity Model Intercomparison Project Phase 1: introduction and evaluation of global-mean temperature response, Geoscientific Model Development, 13, 5175–5190, https://doi.org/10.5194/gmd-13-5175-2020, 2020.

Persad, G., Samset, B. H., Wilcox, L. J., Allen, R. J., Bollasina, M. A., Booth, B. B. B., Bonfils, C., Crocker, T., Joshi, M., Lund, M. T., Marvel, K., Merikanto, J., Nordling, K., Undorf, S., van Vuuren, D. P., Westervelt, D. M., and Zhao, A.: Rapidly evolving aerosol emissions are a dangerous omission from near-term climate risk assessments, Environmental Research: Climate, 2, 032 001, https://doi.org/10.1088/2752-5295/acd6af, 2023.

Proistosescu, C. and Huybers, P. J.: Slow climate mode reconciles historical and model-based estimates of climate sensitivity, Science Advances, 3, e1602 821, 2017.

Quilcaille, Y., Gudmundsson, L., Beusch, L., Hauser, M., and Seneviratne, S. I.: Showcasing MESMER-X: Spatially Resolved Emulation of Annual Maximum Temperatures of Earth System Models, Geophysical Research Letters, 49, e2022GL099 012, https://doi.org/https://doi.org/10.1029/2022GL099012, e2022GL099012 2022GL099012, 2022.

Quilcaille, Y., Gudmundsson, L., and Seneviratne, S. I.: Extending MESMER-X: a spatially resolved Earth system model emulator for fire weather and soil moisture, Earth System Dynamics, 14, 1333–1362, 2023.

Samset, B. H., Sand, M., Smith, C. J., Bauer, S. E., Forster, P. M., Fuglestvedt, J. S., Osprey, S., and Schleussner, C.-F.: Climate Impacts From a Removal of Anthropogenic Aerosol Emissions, Geophysical Research Letters, 45, 1020–1029, https://doi.org/https://doi.org/10.1002/2017GL076079, 2018.

Samset, B. H., Lund, M. T., Bollasina, M., Myhre, G., and Wilcox, L.: Emerging Asian aerosol patterns, Nature Geoscience, 12, 582–584, https://doi.org/10.1038/s41561-019-0424-5, 2019.

Sanderson, B., Sandstad, M., and normansteinert: benmsanderson/METEOR: v1.0.1 (v1.0.1), Zenodo, https://doi.org/10.5281/zenodo.14967116, 2025.

Sanderson, B. M., Brovkin, V., Fisher, R., Hohn, D., Ilyina, T., Jones, C., Koenigk, T., Koven, C., Li, H., Lawrence, D., et al.: flat10MIP: An emissions-driven experiment to diagnose the climate response to positive, zero, and negative CO2 emissions, EGUsphere, 2024, 1–39, 2024.

Sandstad, M., Aamaas, B., Johansen, A. N., Lund, M. T., Peters, G. P., Samset, B. H., Sanderson, B. M., and Skeie, R. B.: CICERO Simple Climate Model (CICERO-SCM v1.1.1) – an improved simple climate model with a parameter calibration tool, Geosci. Model Dev., 17, 6589–6625, 2024.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

Santer, B. D., Wigley, T. M., Schlesinger, M. E., and Mitchell, J. F.: Developing Climate Scenarios from Equilibrium GCM Results, Max-Planck-Institut für Meteorologie report, Max-Planck-Institut für Meteorologie, https://api.semanticscholar.org/CorpusID:127017512, 1990.

Schöngart, S., Gudmundsson, L., Hauser, M., Pfleiderer, P., Lejeune, Q., Nath, S., Seneviratne, S. I., and Schleussner, C.-F.: Introducing the MESMER-M-TPv0.1.0 module: spatially explicit Earth system model emulation for monthly precipitation and temperature, Geoscientific Model Development, 17, 8283–8320, https://doi.org/10.5194/gmd-17-8283-2024, 2024.

Schwaab, J., Hauser, M., Lamboll, R. D., Beusch, L., Gudmundsson, L., Quilcaille, Y., Lejeune, Q., Schöngart, S., Schleussner, C.-F., Nath, S., Rogelj, J., Nicholls, Z., and Seneviratne, S. I.: Spatially resolved emulated annual temperature projections for overshoot pathways, Scientific Data, 11, https://doi.org/10.1038/s41597-024-04122-1, 2024.

Shiogama, H., Emori, S., Takahashi, K., Nagashima, T., Ogura, T., Nozawa, T., and Takemura, T.: Emission Scenario Dependency of Precipitation on Global Warming in the MIROC3.2 Model, Journal of Climate, 23, 2404 – 2417, https://doi.org/10.1175/2009JCLI3428.1, 2010.

Sillmann, J., Kharin, V. V., Zwiers, F. W., Zhang, X., and Bronaugh, D.: Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections, Journal of geophysical research: atmospheres, 118, 2473–2493, 2013.

Tebaldi, C. and Knutti, R.: The use of the multi-model ensemble in probabilistic climate projections, Philosophical Transactions of the Royal Society A, 365, 2053–2075, https://doi.org/10.1098/rsta.2007.2076, 2007.

Tebaldi, C., Debeire, K., Eyring, V., Fischer, E., Fyfe, J., Friedlingstein, P., Knutti, R., Lowe, J., O'Neill, B., Sanderson, B., van Vuuren, D., Riahi, K., Meinshausen, M., Nicholls, Z., Tokarska, K. B., Hurtt, G., Kriegler, E., Lamarque, J.-F., Meehl, G., Moss, R., Bauer, S. E., Boucher, O., Brovkin, V., Byun, Y.-H., Dix, M., Gualdi, S., Guo, H., John, J. G., Kharin, S., Kim, Y., Koshiro, T., Ma, L., Olivié, D., Panickal, S., Qiao, F., Rong, X., Rosenbloom, N., Schupfner, M., Séférian, R., Sellar, A., Semmler, T., Shi, X., Song, Z., Steger, C., Stouffer, R., Swart, N., Tachiiri, K., Tang, Q., Tatebe, H., Voldoire, A., Volodin, E., Wyser, K., Xin, X., Yang, S., Yu, Y., and Ziehn, T.: Climate model projections from the Scenario Model Intercomparison Project (ScenarioMIP) of CMIP6, Earth System Dynamics, 12, 253–293, https://doi.org/10.5194/esd-12-253-2021, 2021.

Tebaldi, C., Snyder, A., and Dorheim, K.: STITCHES: creating new scenarios of climate model output by stitching together pieces of existing simulations, Earth System Dynamics, 13, 1557–1609, https://doi.org/10.5194/esd-13-1557-2022, 2022.

Watson-Parris, D., Williams, A., Deaconu, L., and Stier, P.: Model calibration using ESEm v1.1.0 – an open, scalable Earth system emulator, Geoscientific Model Development, 14, 7659–7672, https://doi.org/10.5194/gmd-14-7659-2021, 2021.

Watson-Parris, D., Rao, Y., Olivié, D., Seland, Ø., Nowack, P., Camps-Valls, G., Stier, P., Bouabid, S., Dewey, M., Fons, E., Gonzalez, J., Harder, P., Jeggle, K., Lenhardt, J., Manshausen, P., Novitasari, M., Ricard, L., and Roesch, C.: ClimateBench v1.0: A Benchmark for Data-Driven Climate Projections, Journal of Advances in Modeling Earth Systems, 14, e2021MS002 954, https://doi.org/https://doi.org/10.1029/2021MS002954, e2021MS002954 2021MS002954, 2022.

Wilcox, L. J., Allen, R. J., Samset, B. H., Bollasina, M. A., Griffiths, P. T., Keeble, J., Lund, M. T., Makkonen, R., Merikanto, J., O'Donnell, D., Paynter, D. J., Persad, G. G., Rumbold, S. T., Takemura, T., Tsigaridis, K., Undorf, S., and Westervelt, D. M.: The Regional Aerosol Model Intercomparison Project (RAMIP), Geoscientific Model Development, 16, 4451–4479, https://doi.org/10.5194/gmd-16-4451-2023, 2023.

Zelazowski, P., Huntingford, C., Mercado, L. M., and Schaller, N.: Climate pattern-scaling set for an ensemble of 22 GCMs – adding uncertainty to the IMOGEN version 2.0 impact system, Geoscientific Model Development, 11, 541–560, https://doi.org/10.5194/gmd-11-541-2018, 2018.

https://doi.org/10.5194/egusphere-2025-1038
Preprint. Discussion started: 13 March 2025

Zhao, A. D., Stevenson, D. S., and Bollasina, M. A.: The role of anthropogenic aerosols in future precipitation extremes over the Asian Monsoon Region, Climate Dynamics, 52, 6257–6278, https://doi.org/10.1007/s00382-018-4514-7, 2019.

Zhao, B., Reichler, T., Strong, C., and Penland, C.: Simultaneous Evolution of Gyre and Atlantic Meridional Overturning Circulation Anomalies as an Eigenmode of the North Atlantic System, Journal of Climate, 30, 6737 – 6755, https://doi.org/10.1175/JCLI-D-16-0751.1, 2017.