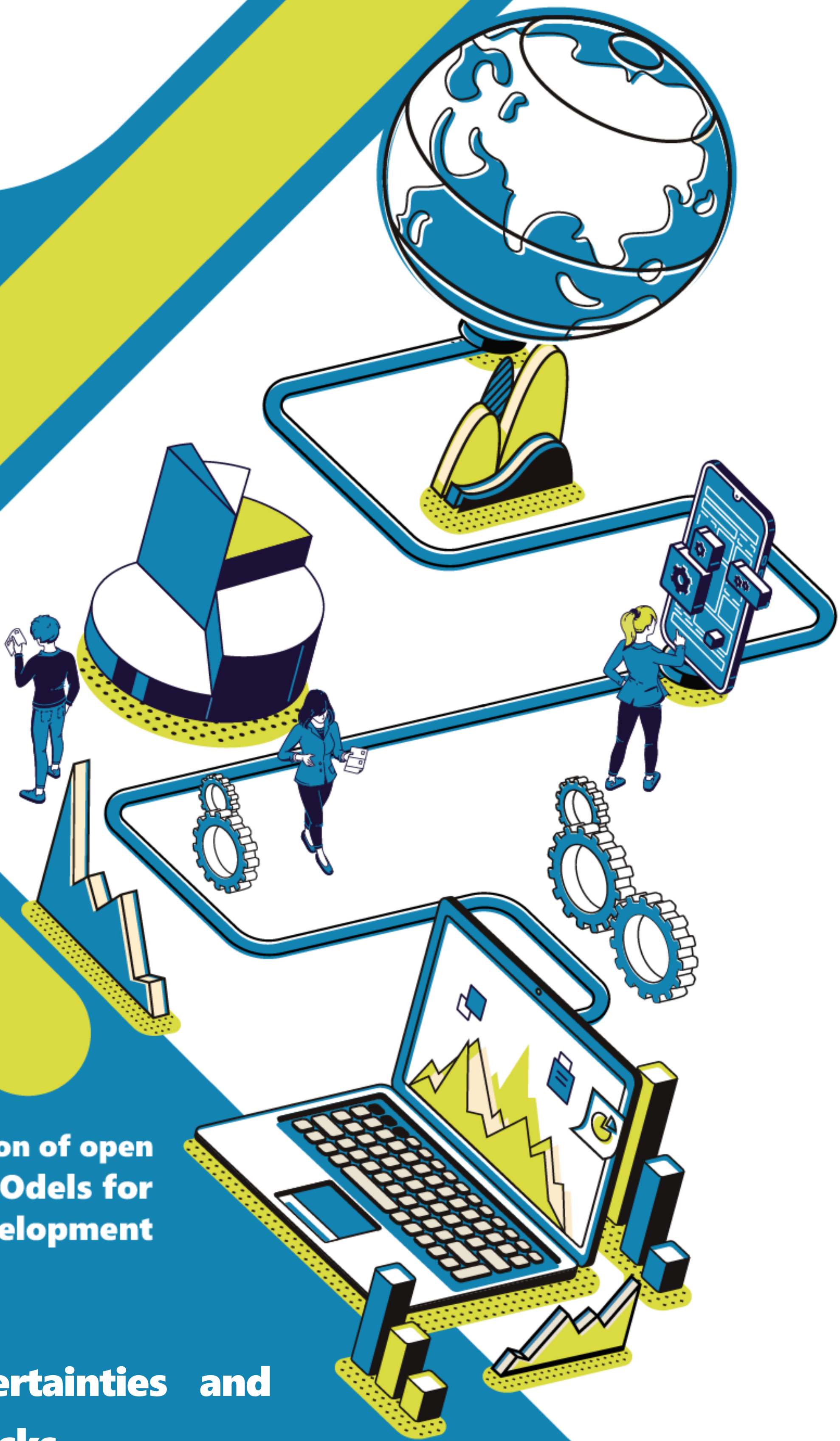
DIAMOND

www.climate-diamond.eu



Delivering the next generation of open Integrated Assessment MOdels for Net-zero, sustainable Development

# D4.4 – Climate uncertainties and carbon-climate feedbacks

WP4 – Integrate & Expand

Funded by the European Union

25/11/2025

## Disclaimer



## Copyright Message





| **Grant Agreement Number** | **101081179** | | **Acronym** | **DIAMOND** |
|---|---|---|---|---|
| **Full Title** | Delivering the next generation of open Integrated Assessment MOdels for Net-zero, sustainable Development | | | |
| **Topic** | HORIZON-CL5-2022-D1-02 | | | |
| **Funding scheme** | HORIZON EUROPE, RIA – Research and Innovation Action | | | |
| **Start Date** | December 2022 | **Duration** | 48 Months | |
| **Project URL** | http://www.climate-diamond.eu/ | | | |
| **EU Project Advisor** | Silvia Vaghi | | | |
| **Project Coordinator** | Institute of Communications and Computer Systems - ICCS | | | |
| **Deliverable** | D4.4 – Climate uncertainties and carbon-climate feedbacks | | | |
| **Work Package** | WP4 – Integrate & Expand | | | |
| **Date of Delivery** | **Contractual** 30/11/2025 | | **Actual** | 25/11/2025 |
| **Nature** | Report | **Dissemination** | Public | |
| **Lead Beneficiary** | CICERO Senter for Klimaforskning (CICERO) | | | |
| **Responsible Author** | Glen Peters | **Email** | glen.peters@cicero.oslo.no | |
| | CICERO | **Phone** | +4792891638 | |
| **Contributors** | Norman Steinert, Marit Sandstad, Benjamin Sanderson, Shivika Mittal (CICERO) | | | |
| **Reviewer(s)** | Natasha Frilingou, Konstantinos Koasidis, Alexandros Nikas (ICCS); Alaa Al Khourdajie (Imperial) | | | |
| **Keywords** | Climate uncertainties; climate impact; pattern scaling, METEOR | | | |

## EC Summary Requirements

# 1. Changes with respect to the DoA

Although integration with the climate and physical impact module was not foreseen in the OPEN-PROM model within the DoA, such a linkage was subsequently judged beneficial and therefore implemented.

No other changes with respect to the work described in the DoA.

# 2. Dissemination and uptake

This deliverable explains the structure, purpose, and use of various tools and pipelines to enable climate modelling feedback into IAMs. It can hence be used as a starting point to understand how coupling of a climate emulator with IAMs helps to incorporate climate impacts in the pathways modelled with IAMs. This approach can be used by IAM modellers inside and outside the DIAMOND consortium to couple an open-source climate emulator with IAMs. It also describes complexities faced and avenues for future developments and work (even beyond the scope of the project). All tools are open and links to open versions with persistent identifier versions are available, so uptake and use should be possible starting from this document.

# 3. Short summary of results (<250 words)

The assessment of climate impacts and adaptation options relies heavily on resolved spatial information from climate change projections. Current methods for generating these projections, particularly with the use of Earth System Models (ESMs), face challenges due to extensive computational and human resources requirements. To enhance the exploration of diverse societal outcomes and incorporate climate impacts into broader societal and economic models, rapid spatial emulation of ESM responses is essential. Existing linear scaling techniques often assume that spatial climate signals scale linearly with global temperature change, where the pattern of response is independent of the nature and timing of emissions. To address these issues, we developed the METEOR novel emulation system, representing multi-timescale climate responses to various climate forcers. On top of the description of the core METEOR model, we also expand on v1.5, which introduces a monthly climate variability system that accurately generates sub-annual climate sequences reflecting realistic seasonal cycles and interannual variability. This integrated framework enhances climate risk assessments by enabling extensive ensemble generation and supporting accurate sub-annual impact evaluations. Its open-source design allows for rapid development of sector-specific calculators, bridging the gap between global climate data and local decision-making. Future enhancements can expand the system's capabilities into sectors such as agriculture and health, along with improved analysis of compound events and seasonal effects under extreme warming conditions.

# 4. Evidence of accomplishment

This report along with:

- A scientific publication describing the METEOR model, which is published in the GMD journal.
- The code of the model, which is available on GitHub, and the latest version (v1.5) available here.
- Input data preparation to map IAMC formatted output to METEOR runnable inputs, which are available on GitHub as well.

# Preface

DIAMOND will update, upgrade, and fully open six Integrated Assessment Models (IAMs) that are emblematic in scientific and policy processes, improving their sectoral and technological detail, spatiotemporal resolution, and geographic granularity. It will further enhance modelling capacity to assess the feasibility and desirability of Paris-compliant mitigation pathways, their interplay with adaptation, circular economy, and other SDGs, their distributional and equity effects, and their resilience to extremes, as well as robust risk management and investment strategies. This will be done via integration of tools and insights from psychology, finance research, behavioural and labour economics, operational research, and physical science. The project will develop a transdisciplinary scientific approach to legitimise the implementation process and co-create research questions that stretch the frontiers of climate science, as well as establish vibrant communities of practice to transparently open model enhancements and to develop capacities, thereby lowering the entrance barriers to the established IAM community.

| | | |
|---|---|---|
| **ICCS** | INSTITUTE OF COMMUNICATIONS AND COMPUTER SYSTEMS | EL |
| **BC3** | ASOCIACION BC3 BASQUE CENTRE FOR CLIMATE CHANGE - KLIMA ALDAKETA IKERGAI | ES |
| **CESAR** | KRATENA KURT | AT |
| **CICERO** | CICERO SENTER FOR KLIMAFORSKNING | NO |
| **CYI** | THE CYPRUS INSTITUTE | CY |
| **E4SMA** | ENERGY ENGINEERING ECONOMIC ENVIRONMENT SYSTEMS MODELING AND ANALYSIS SRL | IT |
| **HOLISTIC** | HOLISTIC IKE | EL |
| **COMILLAS** | UNIVERSIDAD PONTIFICIA COMILLAS | ES |
| **ISINNOVA** | ISTITUTO DI STUDI PER L'INTEGRAZIONE DEI SISTEMI (I.S.I.S) - SOCIETA'COOPERATIVA | IT |
| **SEURECO** | SEURECO SOCIETE EUROPEENNE D'ECONOMIE SARL | FR |
| **UM** | UNIVERSITEIT MAASTRICHT | NL |
| **ESMIA** | ESMIA CONSULTANTS INC. | CA |
| **USMF** | THE UNIVERSITY OF MARYLAND FOUNDATIION INC | US |
| **UMD** | UNIVERSITY OF MARYLAND | US |
| **EPFL** | ECOLE POLYTECHNIQUE FEDERALE DE LAUSANNE | CH |
| **ETH** | EIDGENOESSISCHE TECHNISCHE HOCHSCHULE ZUERICH | CH |
| **UNIBAS** | UNIVERSITAT BASEL | CH |
| **Imperial** | IMPERIAL COLLEGE OF SCIENCE TECHNOLOGY AND MEDICINE | UK |
| **Oxford** | THE CHANCELLOR, MASTERS AND SCHOLARS OF THE UNIVERSITY OF OXFORD | UK |
| **UCL** | UNIVERSITY COLLEGE LONDON | UK |

# Executive Summary

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 resource demands of Earth System Model (ESM) pipelines. To explore a wider variety of societal outcomes and to enable coupling of climate impacts into societal and economic modelling frameworks, rapid spatial emulation of ESM response is therefore necessary. Existing linear pattern scaling methods assume that spatial climate signals 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 introduce and present 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 that can be used to train model-specific pattern response engines using only core training simulations from CMIP.

We then proceed to the presentation of METEOR v1.5, which introduces two major extensions.

First, a monthly climate variability system generates realistic sub-annual climate sequences with proper seasonal cycles and interannual variability. The system uses Principal Component Analysis (PCA) to decompose monthly climate anomalies into spatial modes, then models their temporal evolution using Vector Autoregressive with exogenous variables (VARX) processes that capture both internal climate dynamics and external forcing responses. Seasonal cycles are represented through harmonic analysis with temperature-dependent parameterisation, enabling realistic simulation of seasonal timing shifts and amplitude changes under warming.

Second, a modular impact assessment framework translates climate projections into societally relevant metrics through an extensible system of impact calculators. The framework standardises interfaces for ensemble processing and uncertainty quantification while allowing custom implementations. The initial system includes a comprehensive degree days calculator for energy applications, providing ensemble-based uncertainty quantification that propagates climate uncertainties through to impact projections.

The integrated framework offers key advantages for climate risk assessment: computational efficiency enables large ensemble generation; monthly resolution supports sub-annual impact assessments; and modular architecture facilitates rapid development of sector-specific calculators. The system bridges the gap between global climate projections and decision-making needs, enabling applications in climate risk analysis, planning, and sectoral impact evaluation. Hence, in the final part of the report, we elaborate on the pipeline set to establish coupling mechanisms between physical climate change and the economy, towards linking METEOR with integrated assessment models (IAMs) to enhance the representation of climate impacts in emissions mitigation scenarios.

Future developments include expansion to additional sectors (agriculture, water resources, health), compound event analysis capabilities, and enhanced seasonal representation for high warming scenarios. The open-source framework provides a foundation for operational climate services and supports growing demands for decision-relevant climate information.

# Contents



# Table of Figures

# Table of Tables

# 1 Introduction

One-way coupling of integrated assessment models (IAMs) to simple climate models provides a generally deterministic climate outcome of emissions and land use allocation that is used to iterate IAMs in finding least-cost solutions to climate targets. In reality, the climate mitigation nexus will likely be more complex: (i) the outcome of a given emissions pathway is associated with fundamental uncertainty, so representing the relative confidence of the climate outcome of mitigation strategies is critical; (ii) climate impacts can also affect the economy via modifications of land use productivity, human health, ecosystem damages, habitability, and socio-political conflict—i.e. feedbacks that are limited or non-existent in current IAMs; and iii) second order climate impacts driven by shifts in energy and land-use emissions resulting from mitigation or adaptation measures implemented in response to climate impacts. Climate change-related impacts, such as increasingly frequent climate and weather extremes, have led to widespread consequences for ecosystems, people, the economy, settlements, and infrastructure. Specific sectors such as agriculture, forestry, fisheries, energy and tourism - all with high direct climate exposure - are at risk of larger economic losses than other sectors, but many industrial and service sectors will suffer indirect effects through supply disruptions, especially from extreme events. These adverse consequences to the society and economy could slow down the speed of transition to reach net zero target. Detailed process Integrated assessment models (IAMs) that are widely assessed in the IPCC reports don't fully represent socioeconomic climate damages and consequently their impact on the speed and magnitude of the modelled pathways (P. Forster et al., 2018; Peters et al., 2023). Numerous studies have examined the impact of climate change on different sectors including the energy system, agriculture, and the financial sector but few studies using process detailed IAM have assessed the changes in emissions from energy and land system after considering climate impacts (Schoenberg et al., 2025; Schultes et al., 2021).

Spatially resolved information is essential for informing robust assessments of mitigation and adaptation strategies in response to global climate change (IPCC, 2021; IPCC, 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 requires more frequent updates for a wider range of 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., 2021). In addition, there is an increased need for fast spatial modelling frameworks where regional climate impacts are resolved to allow for the simulation of risks, inequalities, impacts on and interactions between social, economic and natural systems under climate change (Dietz et al., 2021; Kikstra et al., 2021).

To address these needs, spatial emulation techniques have been developed to rapidly approximate the output of ESMs under various scenarios (Zelazowski, Huntingford, Mercado, & Schaller, 2018). Linear pattern scaling is one such approach, which assumes that spatial patterns of climate response scale linearly with global mean temperature change (Lea Beusch, Lukas Gudmundsson, & Sonia I. Seneviratne, 2020; Mitchell, 2003) . This method allows for quick estimations of regional climate change by scaling predefined spatial patterns according to projected global temperature changes (B. Zhao, Reichler, Strong, & Penland, 2017).

Several models have exploited pattern scaling techniques 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 (L. Beusch, L. Gudmundsson, & S. I. Seneviratne, 2020; Beusch et al., 2022). Likewise, the PRIME framework makes use of pattern scaling and subsequently provides the 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 temperature and its derivative corresponding to the desired prediction to emulate scenarios which have not yet been simulated (Tebaldi, Snyder, & Dorheim, 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 analogues in the training dataset - which in CMIP is relatively sparse and requires concatenation of segments of simulations which may produce unphysical discontinuities in the emulated climate. Further, the ESEm framework (Watson-Parris, Williams, Deaconu, & Stier, 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 from experiments which have only been run by a smaller subset of the CMIP6 ESMs. 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 potentially powerful, 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 (Sanderson et al., 2024) but can be precluded by emulator assumptions. Pattern scaling 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; B. Zhao et al., 2017) . Patterns of warming in PRIME are also subject to this linearity assumption, though its process-based land surface model component could potentially represent memory in slow-timescale terrestrial processes if feedbacks were included. STITCHES can potentially resolve non-linear and time-emergent behaviour to the degree that behaviour 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 (Good et al., 2015; Herger, Sanderson, & Knutti, 2015; Tebaldi & Knutti, 2007). Additionally, non-linearities in the climate system, such as feedback mechanisms and varying climate sensitivities over time (Jonko, Shell, Sanderson, & Danabasoglu, 2013) , can lead to time-evolving and forcing-dependent spatial patterns that are not adequately captured by traditional pattern scaling approaches (Huntingford, 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. After describing the core structure and functionalities of the model, we delve into specific features included in the latest v1.5, while finally elaborating on a pipeline set to establish coupling mechanisms between physical climate change and the economy, towards linking METEOR with integrated assessment models to enhance the representation of climate impacts in emissions mitigation scenarios.

# 2 Climate feedbacks with spatial emulation: The METEOR Model

## 2.1 Uncertainty propagation from emissions to climate outcomes

Figure 1 illustrates 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.

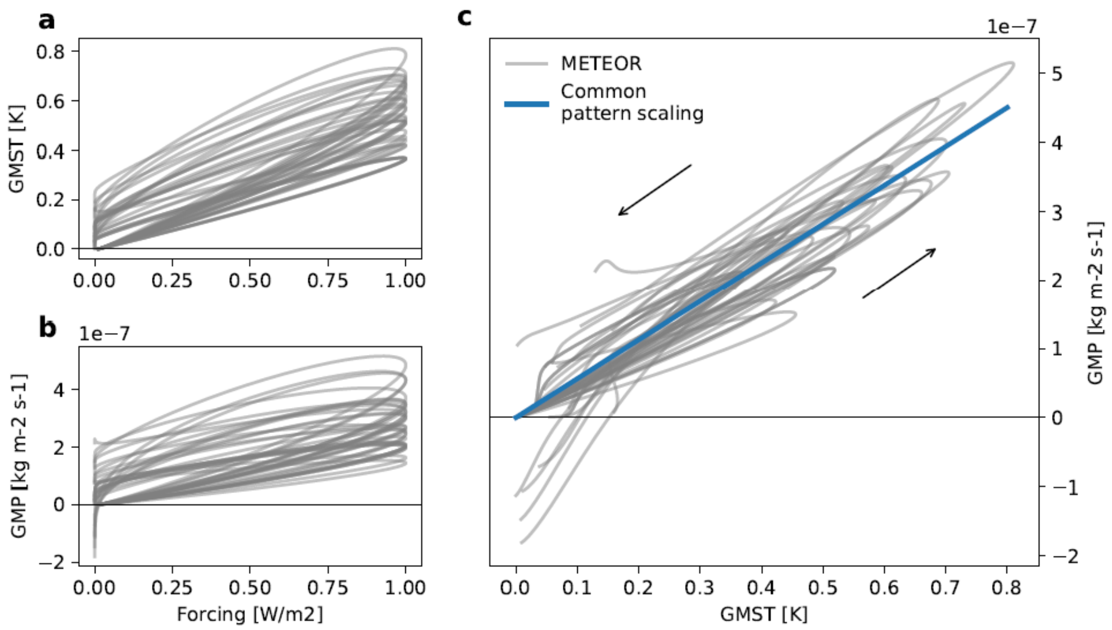



**Figure 1.** METEOR pattern scaling hysteresis.

*Note: Illustration of the emulated METEOR ensemble response showing a) global mean temperature (GMST) and b) precipitation (GMP) response to an idealised radiative forcing with a Gaussian ramp-up to 1 W/m² and subsequent ramp-down within 500 years. c) shows the emulated relationship between GMST and GMP, compared with common pattern scaling, which various patterns scaling approaches are based on. METEOR (grey), trained by CMIP6 models (see Methods), is able to emulate hysteresis behaviour, 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.*

In the following sections, 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.

## 2.2 Development of an open-source forcing- and timescale-dependent pattern scaling module (METEOR)

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 modelling 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 can capture elements of the physical response that 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 (Figure 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 future scenario run (O'Neill et al., 2016) is used for inverse estimation of patterns and timescales for a sulfate aerosol response. We choose this technique to distinguish sulfate aerosol only, as this is the most different and impactful non-GHG forcer (Forster et al., 2025; Forster et al., 2024; Myhre et al., 2022). See also Figure 7.7 of (Forster et al., 2021) where the aerosol forcing is currently the most uncertain and largest non-GHG contribution (though, this may change in the future (Adams et al., 2001; Bauer et al., 2007; Hauglustaine et al., 2014; Liao & Seinfeld, 2005; Liao et al., 2009), which is a source of structural uncertainty in the setup of METEOR demonstrated here). In the setup we use here, the totality of the aerosol-cloud-interactions are represented as a function of sulfate aerosol forcing, and a substantial part of the aerosol radiation forcing comes from sulfate forcing. All other forcer responses including tropospheric and stratospheric ozone, BC- and OC- aerosol direct forcing, and stratospheric water vapour forcing are mapped from forcing strength using the GHG-response patterns and timescales. A further breakdown of forcer types requires additional dedicated experiments. For convenience, we will denote this part of the pattern simply as 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 responses and patterns obtained can then be used to emulate the spatial response to a previously unseen experiment for which there is emissions (or concentrations) data available.

As emissions or concentration inputs, rather than forcing inputs, are available for both training and emulated scenarios, METEOR uses the emissions-forcing engine from C-SCM (Sandstad et al., 2024) to map from emissions or concentration inputs to forcing signals. For each modelled forcer type, the forcing input signal is scaled by C-SCM's own forcing strength in the training scenario.

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 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.

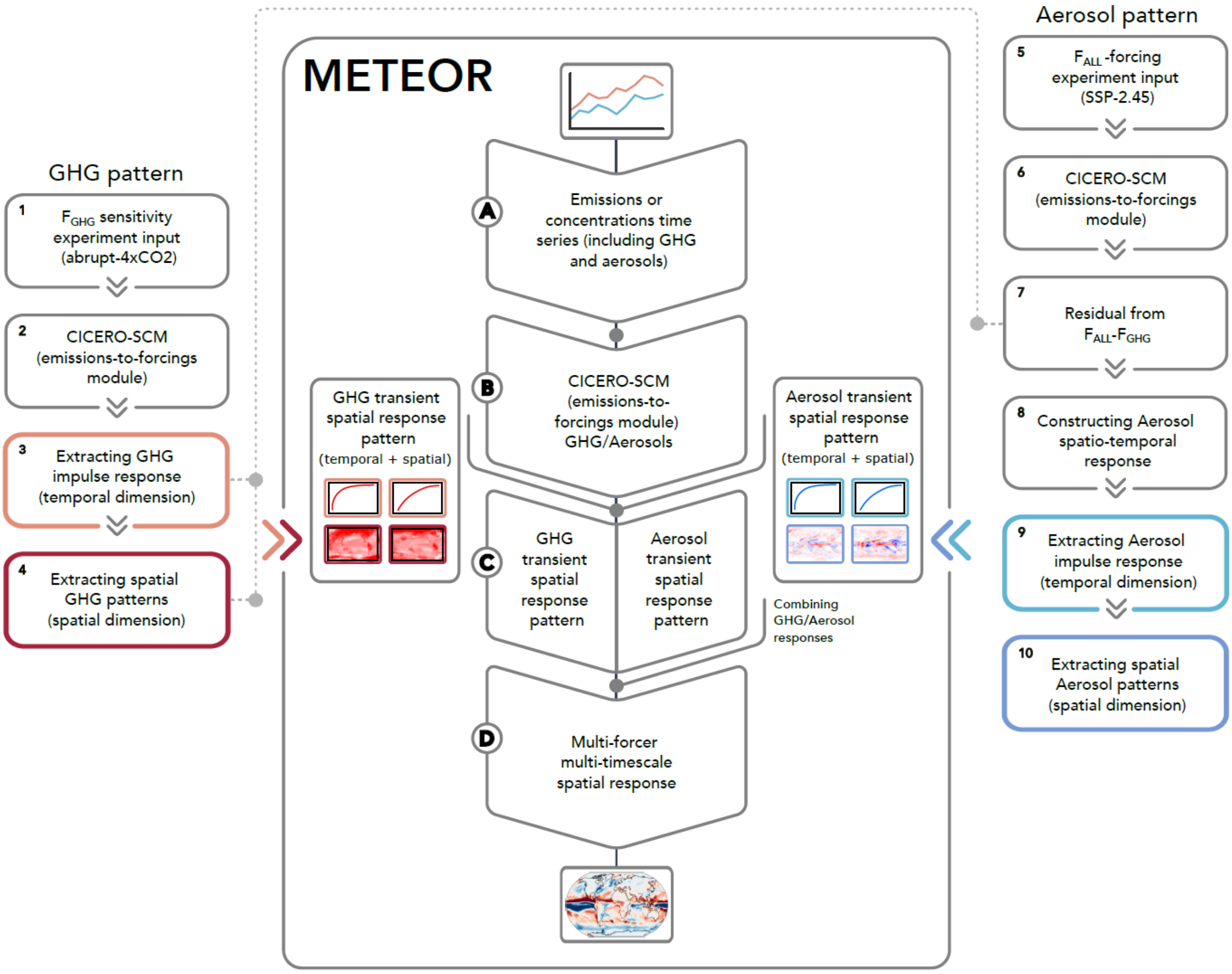



**Figure 2.** METEOR flowchart.

*Note: 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 grey line). Hence, steps 1–10 are required to train METEOR (centre panels). Their outcome is used in step C in METEOR. Note that METEOR partly integrates the C-SCM as it makes use of its emissions-to-forcings module.*

## 2.3 Training the model: Construction of transient spatial response patterns

In this section we describe how we use the *abrupt4x*$CO_2$ 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.

### 2.3.1 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-CO₂* 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) (Figure 2, step 1). We obtain the radiative forcing time series for GHGs from the *abrupt4x-CO₂* emissions by using the C-SCM emission-to-forcing module that converts carbon emissions into climate forcing (Figure 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 (full time-series mean) from the *abrupt4xCO₂* simulation to yield $\mathbf{X}$, a matrix of dimensions $s \times t$, with $s$ being the number of spatial grid points (of the input ESM resolution as METEOR works on the input data resolution with no regridding included) 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 (weighting by $\cos(\text{lat})$) over all spatial grid points:

**Eq. 1**

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

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:

**Eq. 2**

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

where $a_k$ are amplitudes and $\tau_k$ are the decay timescales to be determined for mode $k$. Note thatFreese, Giani, Fiore, and Selin (2024)  present a related approach to emulate from forcing to temperature using a Green-functions mapping which includes per grid-point impulse response functions, rather than METEOR's global impulse response functions.

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 for a structured 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 2.4 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 (Figure 2, step 3). $\mathbf{T}$ represents the temporal evolution of the global mean climate response across different timescales:

**Eq. 3**

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

Assuming that the spatiotemporal response can be represented as a sum of patterns which each emerge as a saturating exponential decay, 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 $s \times n$:

**Eq. 4**

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

This assumption allows the calculation of the spatial response matrix $\mathbf{B}$ (Figure 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 $\mathbf{T}^+$ (Moore-Penrose Pseudoinverse) of the matrix $\mathbf{T}$. Consequently, the derived values of $\mathbf{B}$ minimise the residuals between the observed and reconstructed anomalies:

**Eq. 5**

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

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.

Estimating the climate response of any forcer species for METEOR in the way described here for $CO_2$ is possible, but direct estimation requires a species-specific forcing step-change experiment equivalent to the *abrupt4xCO2*. Such experiments have only been performed for a limited number of CMIP5 generation of models (Myhre et al., 2017).

### 2.3.2 Aerosol response estimation

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 ($F_{\text{ALL}}$) experiment (where both greenhouse gases and aerosols are varying) and a synthetic GHG forcing ($F_{\text{GHG}}$) 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 2.4 we verify that this assumption yields reasonable results in the scenarios considered for CMIP6.

Similarly to how we found $\mathbf{X}$ for the *abrupt4xCO2*, 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* from CMIP6 ESM output (Figure 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 (Figure 2, step 6).

Further, let $F_{\text{GHG}}(t')$ and $F_{\text{aer}}(t')$ denote the GHG and aerosol forcings at time $t' = 1,2,\dots,n$, 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:

**Eq. 6**

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

or in matrix notation:

**Eq. 7**

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

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

**Eq. 8**

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

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 (Figure 2, step 7):

**Eq. 9**

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

As was done for the GHG response, we here assume the aerosol response can be represented with $n_{\mathrm{aer}} = 3$ timescales $\tau_{\mathrm{aer}}$ 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 (Figure 2, step 8):

**Eq. 10**

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

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:

**Eq. 11**

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

or in matrix notation:

**Eq. 12**

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

In order to find optimal values for $\tau_{\mathrm{aer}}$ (introduced in Section 2.3.1), we use an optimization algorithm to search for values of $\tau_{\mathrm{aer}}$ which minimise the error in the projection of the global mean target field residual timeseries (e.g. temperature, precipitation) onto the basis defined by $\mathbf{C}_{\mathrm{aer}}$ (Figure 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 (Barata & Hussein, 2012) $\mathbf{C}_{\mathrm{aer}}{}^{+}$ of the coefficients matrix $\mathbf{C}_{\mathrm{aer}}$:

**Eq. 13**

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

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 $\mathbf{B}_{\text{aer}}$ (Figure 2, step 10):

**Eq. 14**

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

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

The calculations of the transient spatial climate response patterns for GHG and aerosol forcing outlined in the previous sections 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 (Figure 2, step A) into forcing time series using the emissions-to-forcings module of the C-SCM (Figure 2, step B). Convolving these time series with the GHG and aerosol patterns (Figure 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 (Figure 2, step D):

**Eq. 15**

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

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

## 2.4 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. 4 for a full list) using a training dataset that consists of *abrupt4x-*$CO_2$, *piControl*, *historical* and *SSP2-4.5* modelling output in each case. Comparing to CMIP6 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: *SSP1-2.6*, *SSP3-7.0*, *SSP5-8.5* and *SSP5-3.4-over* to consider the out-of-sample performance (note that *SSP5-3.4-over* was only performed by a limited number of models).

Additionally, in the Annex 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. The Annex shows per model results including Figure 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.

### 2.4.1 GHG and aerosol multi-timescale pattern of climate change

Starting with the GHG response, Figure 3 shows the full emulation and contributions from the different timescale patterns obtained from the *abrupt4x-*$CO_2$ experiment for temperature and precipitation. Panels a and f show the CMIP6 multi-model mean outputs averaged over years 80–120 for the experiment. Panels b and g show the multi-model mean average of the METEOR emulations for the same models and time period. Below in panels c and h, d and i and e and j, of Figure 3 decomposes the model mean emulation into components associated with the inter-annual, inter-decadal and inter-centennial responses for temperature (c, d and e) and precipitation (h, i and j). The patterns show Arctic amplification (temperature) in both inter-annual and inter-decadal modes. In the inter-decadal temperature mode (d) METEOR picks up on a North-Atlantic relative cooling anomaly, which is induced by the nonlinear response of the Atlantic Meridional Overturning Circulation in some CMIP6 models. The fast mode response for precipitation is dominated by diverse responses over the tropical Pacific, with more widespread changes in the slower modes.

Figure 4 similarly shows multi-model mean CMIP6 data (panels a and g) and emulation (panels b and h), and the split into contributions from the various timescales and patterns. Here the average has been taken over years 1980–2020 in the combined historical and SSP2-4.5 experiment. Panels c and i show the combined GHG responses for temperature and precipitation, whereas panels d and j, e and k, and f and l show the inter-annual, inter-decadal and inter-centennial aerosol contributions. Broadly speaking, the inter-annual aerosol timescale provides an overall cooling or drying effect. The inter-decadal aerosol signal dampens the Arctic amplification and counteracts the spatial pattern of the precipitation signal. The inter-centennial patterns are largely weaker but negative versions of the inter-decadal patterns. For the net effect of all timescales, comparing the CMIP6 model output to the emulation output in Figures 3 and 4 show good agreement with the strength and spatial distribution of the signals.

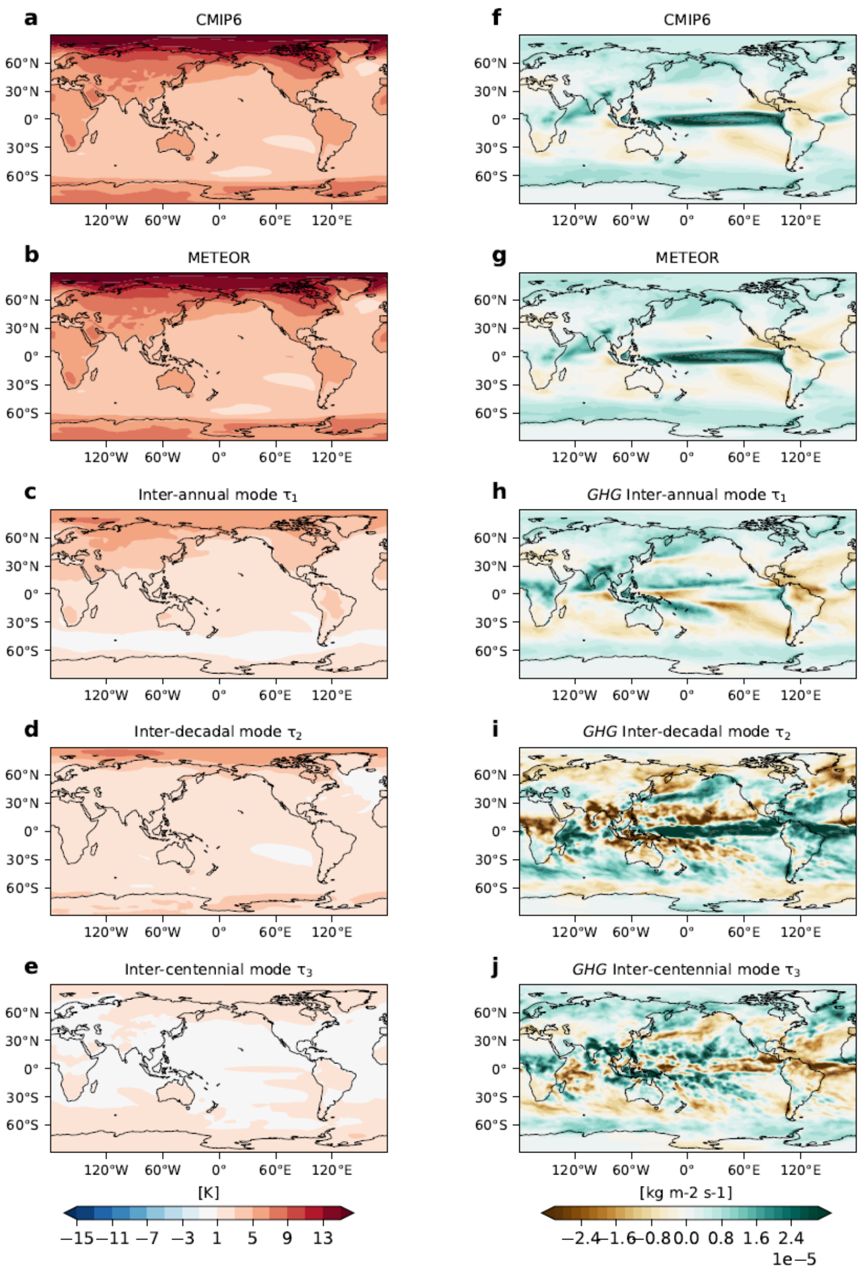



**Figure 3.** METEOR GHG pattern from CMIP6.

*Note: GHG patterns obtained from the abrupt4xCO2 scenario (CMIP6 model average) for temperature (left column) and precipitation (right column). Panels a and f show the CMIP6 model average in the 40-year average of the run between years 80 and 120. Panels b and g show the corresponding multi-model mean of the METEOR emulations for the scenario averaged over the same time period. Panels c and h, d and i and e and j show the contributions to the full emulations shown in b and g from the inter-annual, inter-decadal and inter-centennial modes, respectively.*

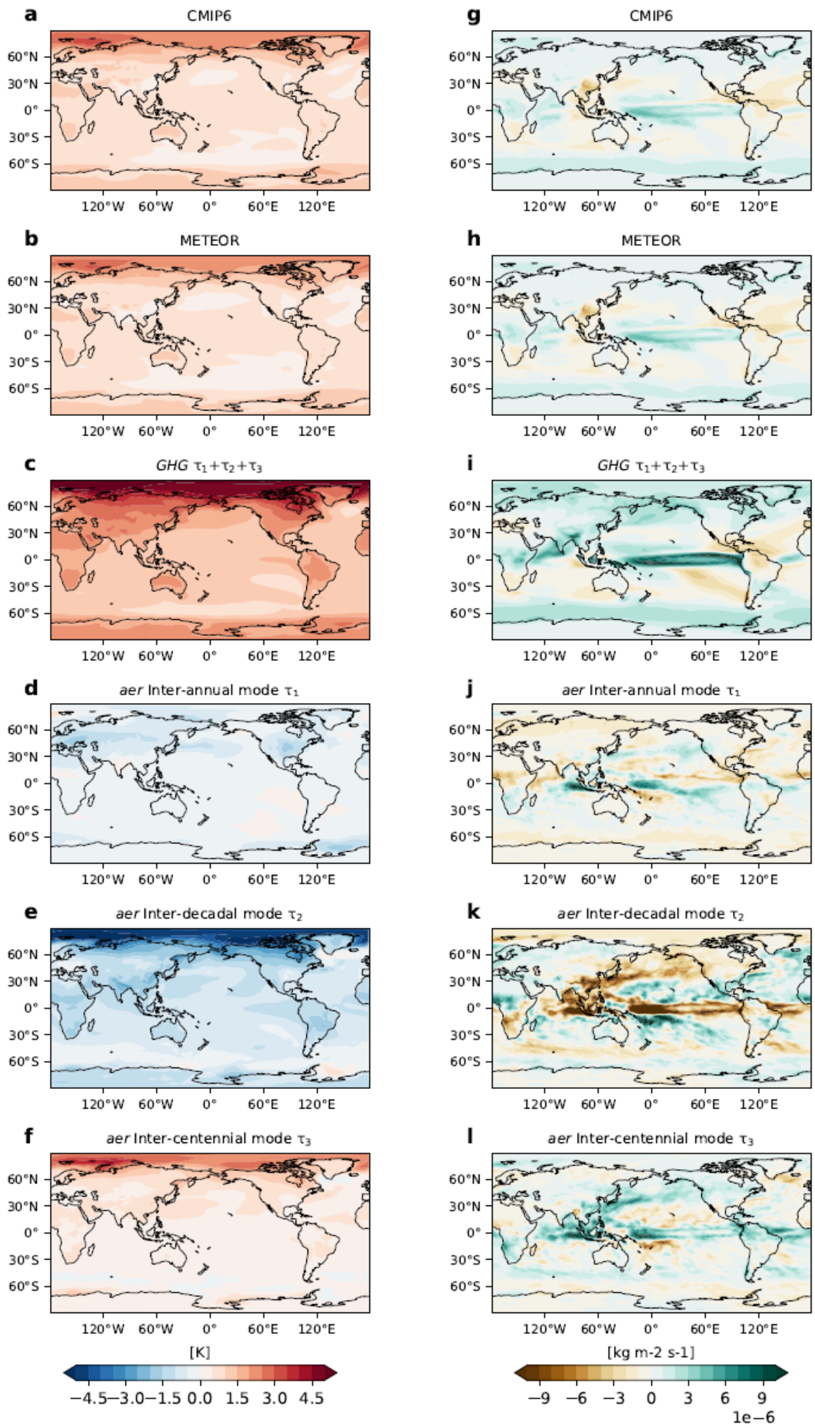



**Figure 4.** METEOR aerosol residuals pattern from CMIP6.

*Note: Residuals patterns obtained from the histroical+SSP2-4.5 scenario (CMIP6 model average) for temperature (left column) and precipitation (right column). Panels a and g show the CMIP6 model average in the 40-year average of the run between years 1980 and 2020. Panels b and h show the corresponding multi-model mean of the METEOR emulations for the scenario averaged over the same time period. Panels c and i show the combined emulated GHG forcing pattern from all non-sulfate aerosol forcing. Panels d and j, e and k and f and l show the contributions to the full emulations shown in b and h from the inter-annual, inter-decadal and inter-centennial sulfate aerosol driven modes, respectively.*

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 modelling 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.

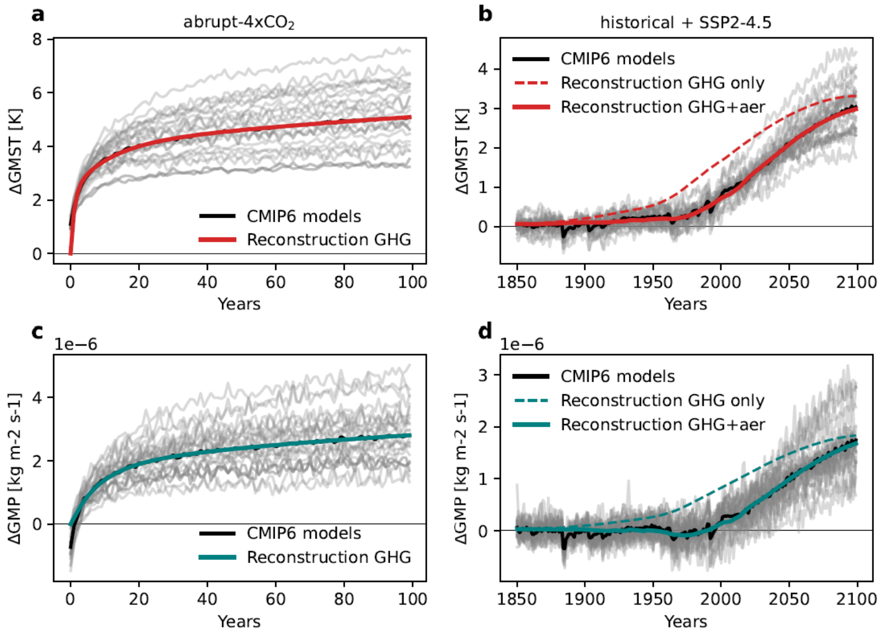



**Figure 5.** METEOR GHG and aerosol residual reconstructions from CMIP6 models.

*Note: METEOR reconstruction for abrupt4x-$CO_2$ (a, c) and historical+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.*

### 2.4.2 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 Figure 6,Figure 7) show both the response from greenhouse gas forcing alone (technically all forcers apart from sulfate aerosol), and the response from the model which includes the calibrated sulfate aerosol signal. The results demonstrate good performance of METEOR in out-of-sample scenarios, both in terms of the global mean and spatially resolved output for the multi-model mean of METEOR emulations compared to the CMIP6 multi-model mean. In the Annex we also show similar global mean fits for each model included, and overall, per model fidelity, though the time evolutions of some models are less well captured than others. In the global mean, the aerosol modes can 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 can 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.

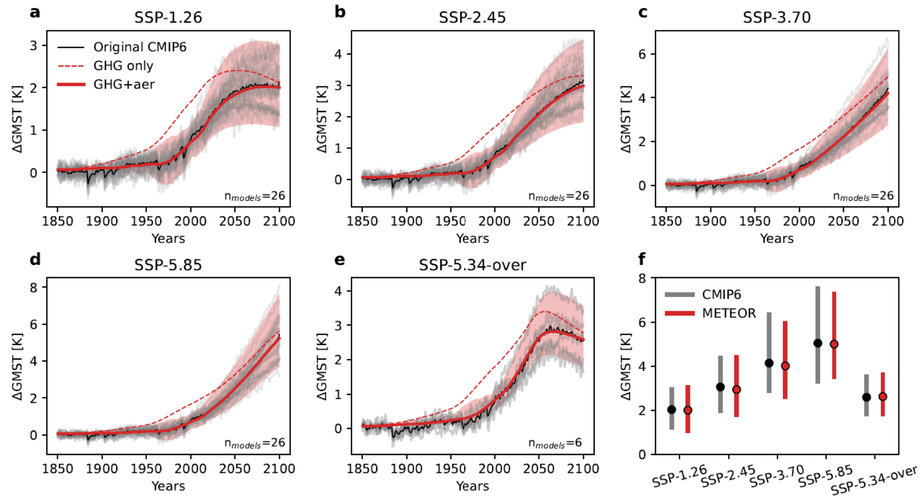



**Figure 6.** Global mean temperature reconstruction from METEOR vs CMIP6 data for various SSP scenarios.

*Note: 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 dashed), 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 modelling spread for the CMIP6 models (grey) and full METEOR reconstruction (red).*

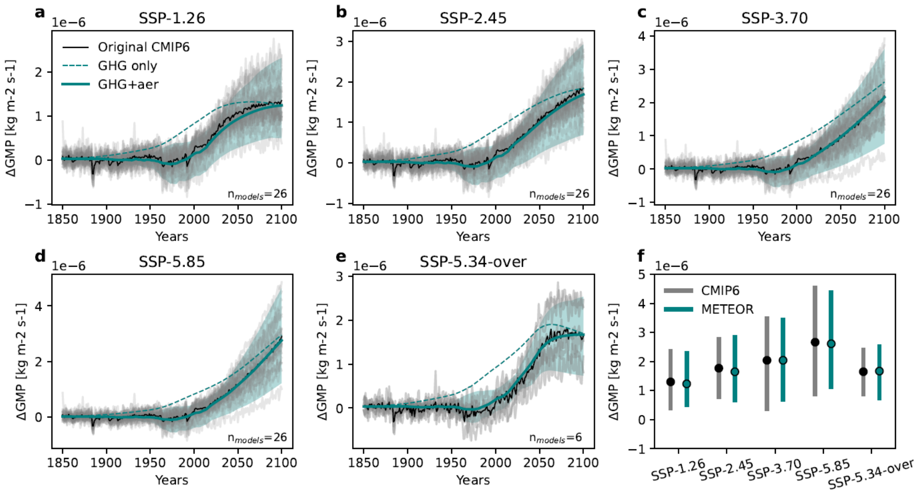



**Figure 7.** Global mean precipitation reconstruction from METEOR vs CMIP6 data for various SSP scenarios

*Note: 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 (teal dashed), 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 modelling spread for the CMIP6 models (grey) and full METEOR reconstruction (teal).*

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 (Figures 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 low emissions and overshoot scenarios (*SSP1-2.6* and *SSP5-3.4-over*). This may be an imprint of, or slight over-fitting to, the *SSP2-4.5* 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 *SSP5-85* scenario for precipitation. These patterns can also be seen in the spatial RMSE maps of Figure 13. In addition, in the Annex we compare single model emulation for all the standard *SSP* scenarios for one of the models with the best METEOR emulations (NorESM2-MM) and one of the models with the worst METEOR emulations (CMCC-ESM2). Per model emulation to CMIP6 comparison maps for the *SSP5-3.4-over* scenarios for all models that we considered for that scenario are also shown in the Annex**.**

For a more regional breakdown, Figures 10 and 11 include CMIP6 regional averages for the last two decades of the 21st century for each of the *SSP* scenarios for 9 regions: Europe (0-45E, 37-75N), High Arctic (0-360E, 70-90N), Antarctic + Southern Ocean (0-360E, 55-90S), Tropics (0-360E, 20S-20N), South America (275-330E, 60S-15N), North America (200-310E, 10-70N), East Asia (65-150E, 5-60N), South East Asia and Australia (95-165E, 45S-5N), and Africa (0-60E, 40S-37N). METEOR's regional emulation of individual CMIP6 models is generally good for temperature across all regions (Figure 10). Polar amplification is captured well, and, despite larger absolute differences, the relative performance of METEOR is best in the High Arctic region, while the fit is worse for the Southern Hemisphere high latitudes. Scenario-differences tend to remain constant regardless of their forcing strength, which increases METEOR's emulation skill for scenarios with higher forcings. For precipitation, the results are a bit more diverse, albeit consistent across all regions (Figure 11). Again, the High Arctic region shows the best agreement between METEOR and CMIP6, while other regions show larger differences. For some regions and models, there is a contrasting direction of the precipitation response (i.e., opposite drying or wetting between METEOR and CMIP6), but this tends to be limited to low-magnitude precipitation changes.

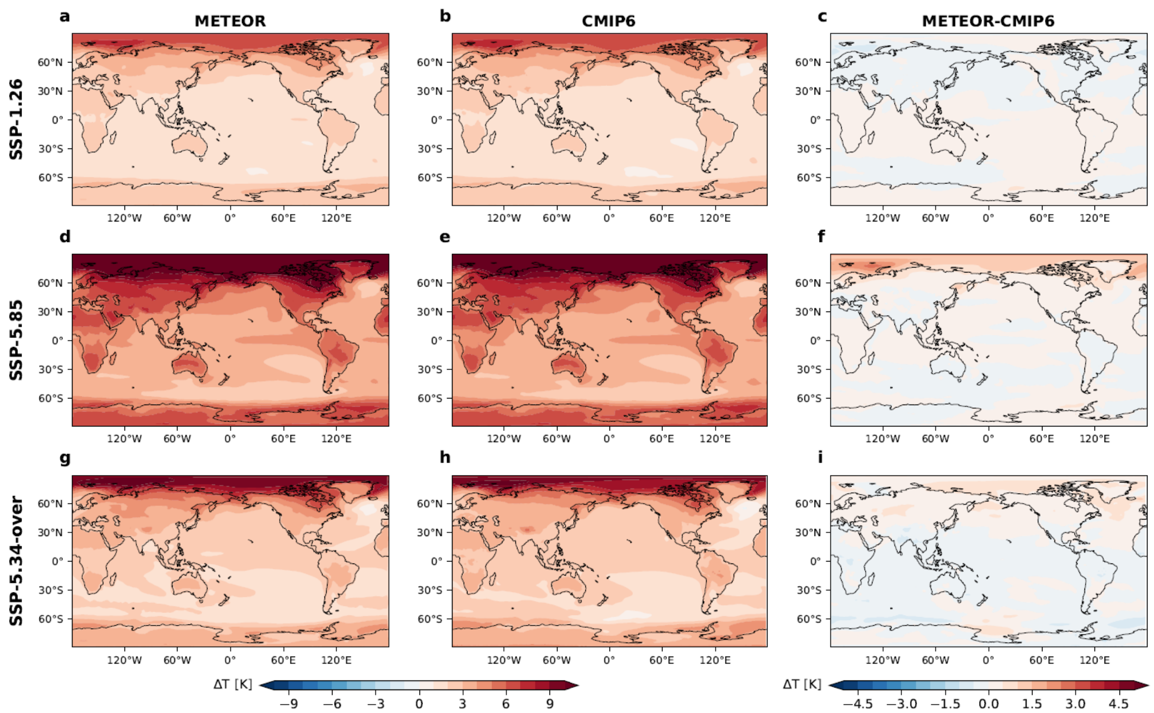



**Figure 8.** METEOR temperature reconstruction pattern vs CMIP6 data for various SSP scenarios.

*Note: 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 two decades 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.*

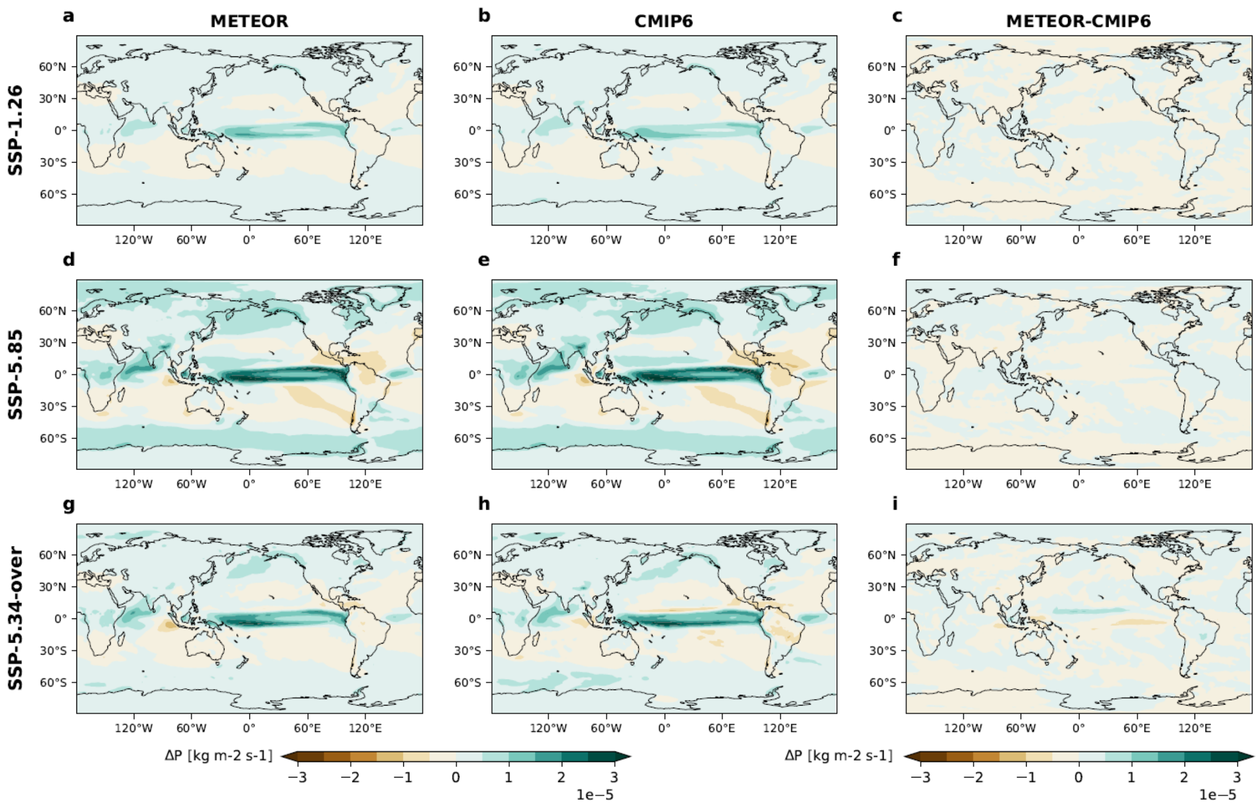



**Figure 9.** METEOR precipitation reconstruction pattern vs CMIP6 data for various SSP scenarios.

*Note: 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 two decades 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.*

Funded by
the European Union

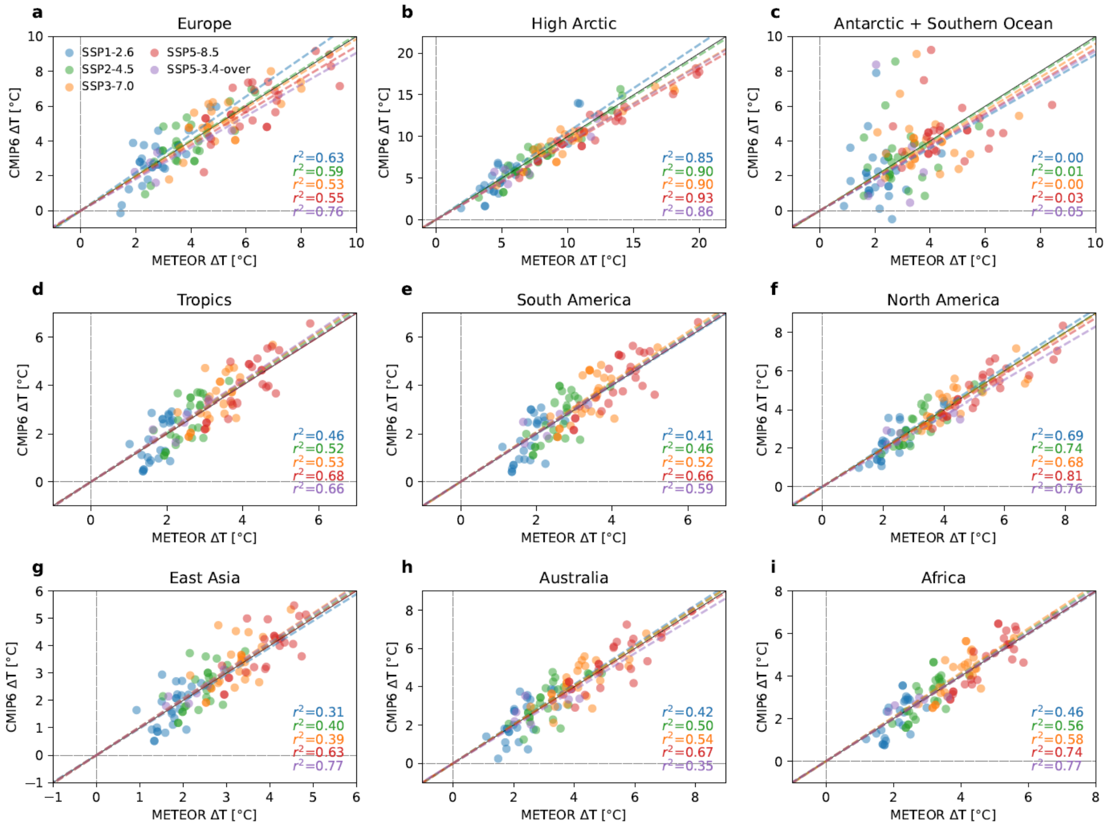



**Figure 10.** Regional CMIP6 vs. METEOR end of century temperature change.

*Note: Regionally averaged temperature change (tas) since pre-industrial for CMIP6 output and METEOR per emulation for each model. Panels show average end of century temperature change for Europe (a), High Arctic (b), Antarctic + Southern Ocean (c), Tropics (d), South America (e), North America (f), East Asia (g), Southeast Asia and Australia (h) and Africa (i). Each circle shows the CMIP6 model result (y-axis) as function of the METEOR emulation (x-axis) for one model and scenario combination averaged over the region. Blue points show results for SSP1-2.6, green for SSP2-4.5, orange for SSP3-7.0, red for SSP5-8.5 and purple for SSP5-3.4-over. The black diagonal shows the line where model result and emulations are equal, and dashed lines show* $r^2$ *distances from the linear regression per scenario.*

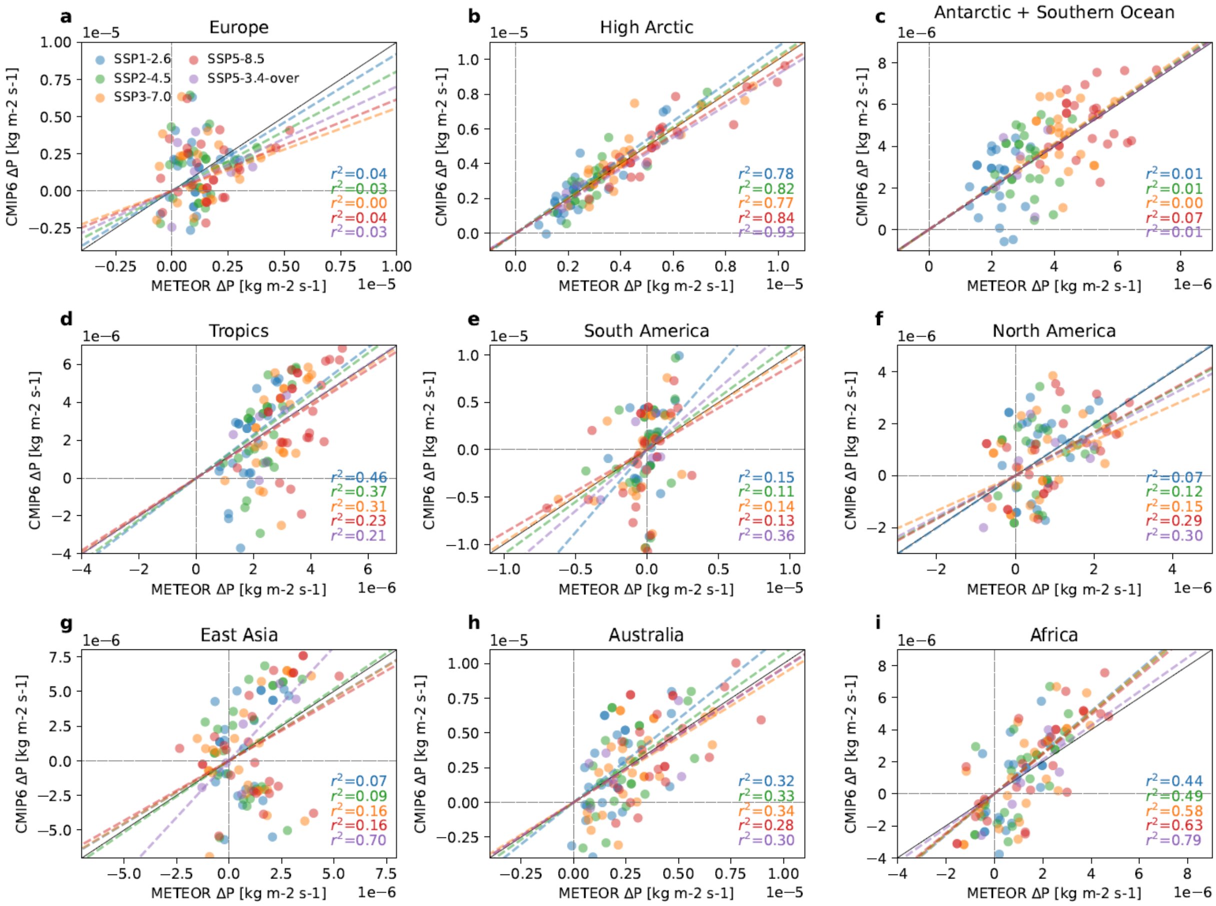



**Figure 11.** Regional CMIP6 vs. METEOR end of century precipitation change.

*Note: Regionally averaged precipitation change (pr) since pre-industrial for CMIP6 output and METEOR per emulation for each model. Panels show average end of century precipitation change for Europe (a), High Arctic (b), Antarctic + Southern Ocean (c), Tropics (d), South America (e), North America (f), East Asia (g), Southeast Asia and Australia (h) and Africa (i). Each circle shows the CMIP6 model result (y-axis) as function of the METEOR emulation (x-axis) for one model and scenario combination averaged over the region. Blue points show results for SSP1-2.6, green for SSP2-4.5, orange for SSP3-7.0, red for SSP5-8.5 and purple for SSP5-3.4-over. The black diagonal shows the line where model result and emulations are equal, and dashed lines show $r^2$ distances from the linear regression per scenario.*

## 2.5 Emulation of overshoot

The main motivation for the multiple timescale impulse response structure of METEOR is to allow for better emulation of overshoot including hysteresis behaviour in joint emulation of temperature and precipitation. We will therefore show some more detailed and per model results of the overshoot joint emulation through the example of the *SSP5-3.4-over* scenario.

Figure 13 shows the precipitation change as function of the temperature change globally and in each of 8 regions (Europe, High Arctic, Tropics, South America, North America, East Asia, Southeast Asia and Australia, and Africa). Globally, the emulation represents hysteresis effects evident in the CMIP6 ESMs well. At the regional scale, noise in the ESM simulations make this harder to assess, particularly in regions where the precipitation signal is relatively small. Nonetheless, in some regions such as the Tropics and South America, hysteresis features matching those of the ESMs are captured by METEOR.

Figure 13 show the end of century average maps of precipitation changes as function of temperature change in the *SSP5-3.4-over* scenario for each of the models. The results further illustrate METEOR's performance for the joint emulation on a per model basis. For some models like CESM2-WACCM and CanESM5, the regional co-evolution patterns seen in the emulation and the true model data are very similar, whereas for others there is a larger discrepancy. However, strong changes in the tropical Pacific are the dominant features in both models and emulation. The similarity of regional patterns of the temperature-dependent evolution of precipitation provides physics-based confidence in the joint evolution of temperature and precipitation in METEOR.

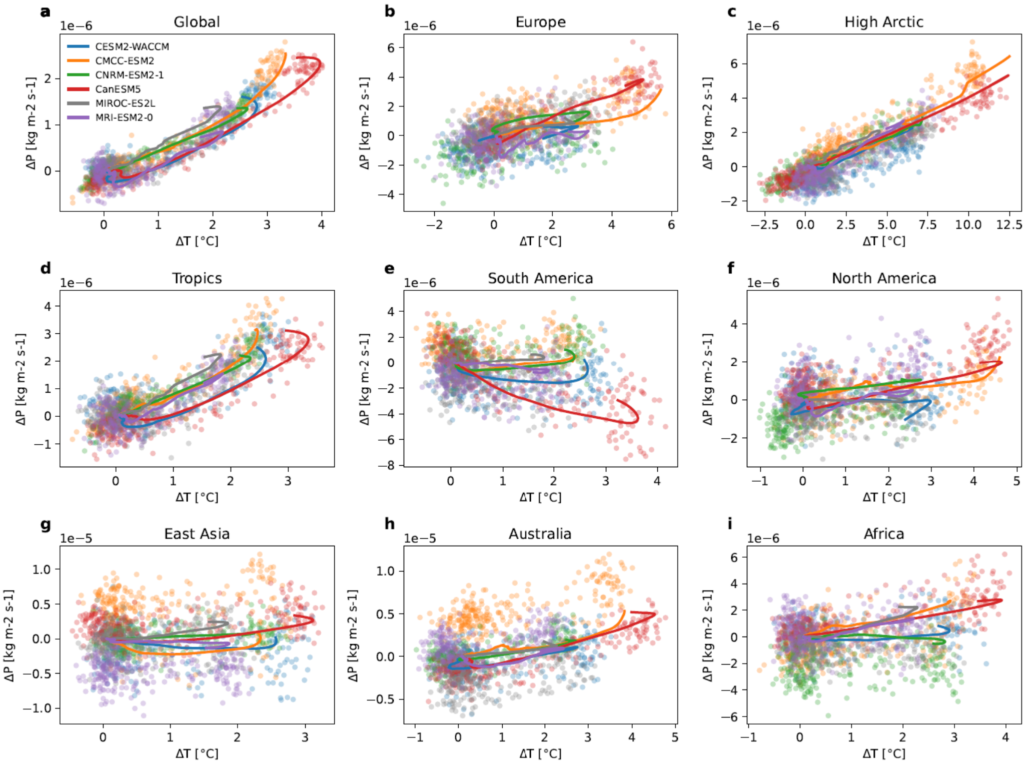



**Figure 12.** SSP5-3.4-over regional precipitation versus temperature.

*Note: Global and regional results for precipitation change as a function of temperature change in the SSP5-3.4-over scenario. Each dot represents results for a specific model in a specific year averaged over the regions Global (a), Europe (a), High Arctic (b), Tropics (d), South America (e), North America (f), East Asia (g), Australia (h) and Africa (i) for the SSP5-3.4-over scenario. Lines show the trajectory of emulation results over the same regions for each of the models CESM2-WACCM (blue), CMCC-ESM2 (orange), CNRM-ESM2-1 (green), CanESM5 (red), MIROC-ES2L (grey) and MRI-ESM2-0 (purple).*

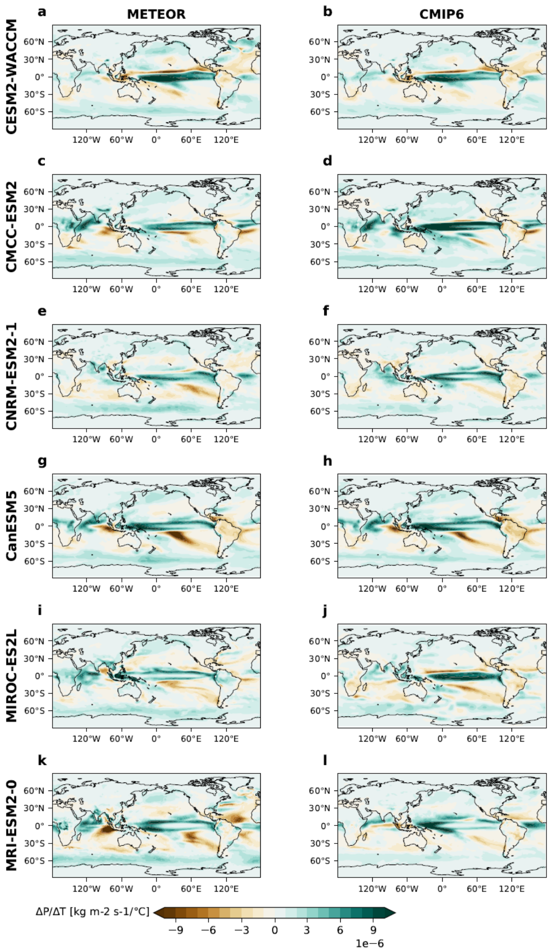



**Figure 13.** SSP5-3.4-over spatial precipitation versus temperature.

*Note: Spatial maps of end of century precipitation change as function of end of century temperature change for each model for the SSP5-3.4-over scenario. The left column (a, c, e, g, i and k) shows results of the METEOR emulation while the right column (b, d, f, h, j and l) shows the CMIP6 outputs directly.*

## 2.6 Performance and error metrics

Here we show some selected error and performance metrics for the METEOR evaluation to scenarios.

Table 1 shows the Pearson's correlation coefficient and Root Mean Square Error (RMSE) for temperature and precipitation 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 better than out-of-sample performance, with the largest 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).

**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* | SSP5-3.4-over |
|---|---|---|---|---|---|---|
| Pearson | *tas* | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| Pearson | *pr* | 0.97 | 0.99 | 0.99 | 0.96 | 0.94 |
| RMSE | *tas* | 0.13 | 0.15 | 0.34 | 0.50 | 0.32 |
| RMSE | *pr* | 5.7e-7 | 3.4e-7 | 7.1e-7 | 8.4e-7 | 13.1e-7 |

Table 2 shows comparisons to the ClimateBench evaluation suite, which combines temporal and spatial NRMSE values for emulations to the *SSP2-4.5* run of NorESM-LM specifically for several different emulation techniques. Results show that METEOR emulations are on par with the other emulation techniques. Note that NorESM-LM is not part of the CMIP6 multi modelled ensemble considered in the rest of this article, as it did not have available outputs for the *SSP5-8.5* scenario. Similar metrics for the models used in our CMIP6 model ensemble are listed in annex tables B2-B5. As overshoot is a special focus for METEOR, Table 3 shows the ClimateBench results for each of the models we have considered that ran the *SSP5-3.4-over* scenario.

**Table 2.** METEOR performance on NorESM-LM evaluated using and compared to the ClimateBench evaluation suite.(Watson-Parris, 2021; Watson-Parris et al., 2022).

| | $\mathrm{NRMSE}_s$ tas | $\mathrm{NRMSE}_g$ tas | $\mathrm{NRMSE}_t$ tas | $\mathrm{NRMSE}_s$ pr | $\mathrm{NRMSE}_g$ pr | $\mathrm{NRMSE}_t$ pr |
|---|---|---|---|---|---|---|
| METEOR ssp126 | 0.165 | 0.161 | 0.969 | 1.978 | 0.100 | 2.480 |
| METEOR ssp245 | 0.053 | 0.042 | 0.262 | 2.041* | 0.111* | 2.594* |
| METEOR ssp370 | 0.118 | 0.082 | 0.527 | 2.139 | 0.133 | 2.804 |
| Gaussian Process CB | 0.109 | 0.074 | 0.478 | 2.341 | 0.341 | 4.048 |
| Neural Network CB | 0.107 | 0.044 | 0.327 | 2.128 | 0.209 | 3.175 |
| Random Forest CB | 0.108 | 0.058 | 0.400 | 2.524 | 0.502 | 5.035 |
| Pattern Scaling CB | 0.080 | 0.048 | 0.320 | 2.006 | 0.331 | 3.662 |
| Variability CB | 0.052 | 0.072 | 0.414 | 1.350 | 0.268 | 2.691 |
| CMIP6 CB | 0.258 | 0.177 | 1.141 | 1.994 | 0.389 | 3.940 |

*Note: They compare normalised RMSE values obtained from comparing the average climate response over years 2080–2100 in model output and emulation ($\boldsymbol{NRMSE_s}$), normalising RMSE in the spatially averaged timeseries over the same years $\boldsymbol{NRMSE_g}$ and taking the linear combination between the two, dubbed the total NRMSE ($\boldsymbol{NRMSE_t} = \boldsymbol{NRMSE_s} + 5 \cdot \boldsymbol{NRMSE_g}$). *The exact methods used to derive the ClimateBench provided output for precipitation proved hard to replicate and hence to compare to our output, only the temperature outputs for SSP2-4.5 are compared directly to the ClimateBench output data itself. However, results were computed using the same metrics to the output for the ensemble member used for emulation taken from the CMIP6 database* *directly. Results for METEOR emulation errors for SSP1-2.6 and SSP3-7.0 were obtained in the same manner. The results for other emulations, internal variability and CMIP6 variability are taken directly from(Watson-Parris et al., 2022) . Note that the ClimateBench results are compared to the mean between three ensemble member outputs, hence some errors stemming from natural variability are dampened in the ClimateBench results and the result for SSP2-4.5 for METEOR compared to the other METEOR results.*

**Table 3.** METEOR performance for all models evaluated using the Climate bench evaluation metrics.

| | $\mathrm{NRMSE}_s$ tas | $\mathrm{NRMSE}_g$ tas | $\mathrm{NRMSE}_t$ tas | $\mathrm{NRMSE}_s$ pr | $\mathrm{NRMSE}_g$ pr | $\mathrm{NRMSE}_t$ pr |
|---|---|---|---|---|---|---|
| CESM2-WACCM ssp534-over | 0.229 | 0.207 | 1.264 | 0.074 | 0.009 | 0.117 |
| CMCC-ESM2 ssp534-over | 0.144 | 0.120 | 0.743 | 0.058 | 0.009 | 0.104 |
| CNRM-ESM2-1 ssp534-over | 0.289 | 0.251 | 1.542 | 0.070 | 0.013 | 0.137 |
| CanESM5 ssp534-over | 0.205 | 0.180 | 1.107 | 0.098 | 0.010 | 0.150 |
| MIROC-E2SL ssp534-over | 0.271 | 0.212 | 1.331 | * | * | * |
| MRI-ESM2-0 ssp534-over | 0.170 | 0.075 | 0.543 | 0.069 | 0.010 | 0.119 |

*Note: The comparison is to single ensemble members, so variability driven errors are included. *Precipitation metrics for MIROC-E2SL had issues.*

Finally, Figure 14 shows the maps of spatial RMSE for the end of the century emulation of CMIP6 models for four scenarios: *SSP1-2.6*, *SSP2-4.5*, *SSP3-7.0* and *SSP5-8.5*. The *SSP2-4.5* scenario overall shows the smallest RMSE, which is expected as it is METEOR's training scenario. For the higher emission scenarios *SSP3-7.0* and *SSP5-8.5*, the high Arctic and the Southern Ocean show larger RMSE for temperature, while for precipitation, largest RMSE are found over the tropical oceans. In the *SSP1-2.6* scenario, the RMSE for temperature shows overall the same pattern both for temperature and precipitation, but the former is much attenuated over the high Arctic. This indicates that, by training to *SSP2-4.5*, some non-linear Arctic ocean or sea-ice dynamics are not captured by METEOR in the higher emission out-of-sample scenarios for temperature, while the differences tend to scale more linearly for precipitation.

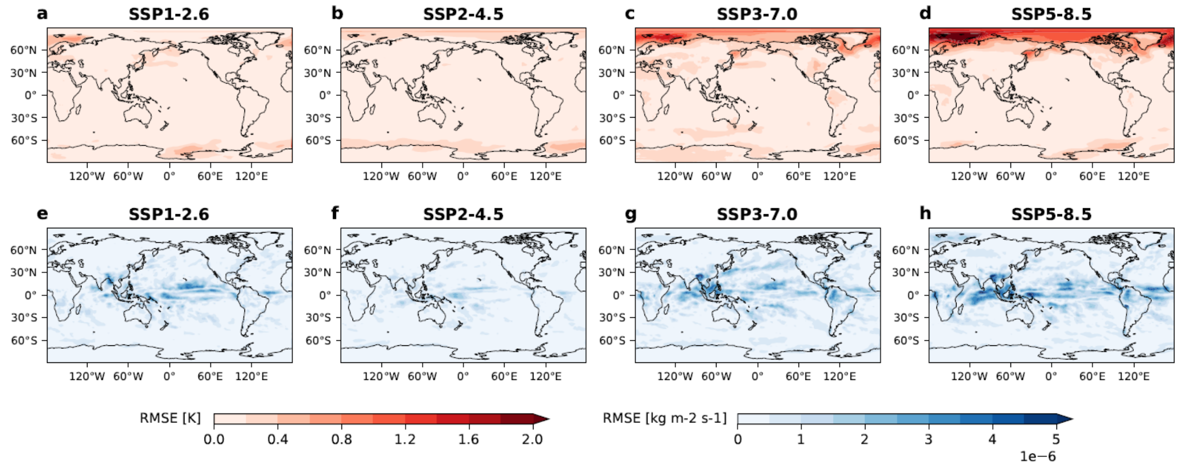



**Figure 14.** *Spatial RMSE.*

*Note: End of century spatial RMSE for temperature (top row) and precipitation (bottom row) in each of the standard scenarios SSP1-2.6 (a, e), SSP2-4.5 (b, f) SSP3-7.0 (c, g) and SSP5-8.5 (d, h).*

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 (unsurprisingly) 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. Beyond the demonstrations here, the model can easily be trained on new modelling output and to emulate new and unknown scenarios and reasonable accuracy can be expected.

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.

## 2.7 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 occurring on more closely related timescales, or on timescales which are shorter or longer than those considered here. Allowing for additional modes in testing did not lead to notable increases in performance. This is partly a function of data limits in training data. In particular, the number of models which have the shortest possible timescale of 1 year indicate that a shorter mode might yield better results. However, timescales of less than 1 year cannot be represented given the annual mean data used for training in this study, and training simulations are generally run for 150–250 years, hence shorter or longer timescales cannot be meaningfully fitted with these 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. Fitting shorter (sub-annual) timescales requires an extension of the methodology considered here and any meaningful approach would require a treatment of the seasonal cycle and internal variability, which is planned for future releases of METEOR. Further, gaining high confidence in shorter timescale responses would require large initial condition ensembles, which would allow the assessment of forced response given presence of internal variability with large impacts on the estimation of faster response timescales. In general, the number of independent forcer responses and timescales could be improved with ESM datasets which isolate the effect of independent forcers.

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 too limited to conduct these refinements for the full range of models considered here. There exist experiments which could be used to provide additional information for some models, such as the DAMIP experiments: ssp245-aer ssp245-GHG, ssp245-CO2, ssp245-stratO3, which we wish to exploit in the future. However, for this demonstration, we prioritise showing METEOR's performance using only experiments which have been performed by most CMIP6 models, and therefore demonstrating METEOR's versatility. Hence, we isolate the effect of GHG and sulfate aerosol forcing leaning on results from the literature (Myhre et al., 2017; Persad et al., 2023; Samset et al., 2019; B. H. Samset et al., 2018; Wilcox et al., 2023; Zhao et al., 2019) including the assessments of the last IPCC cycle , 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., 2024; Nathet al., 2022; 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 specific 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 additional 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 sulfate 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 relatively centrally located for that range is advisable. 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. Similarly, the aerosol profile in the future may have low emissions of sulfate aerosol, while nitrate aerosol emissions might increase (Adams et al., 2001; Bauer et al., 2007; Hauglustaine et al., 2014; Liao & Seinfeld, 2005) . The forcing responses to nitrate aerosols are currently not well constrained, and can be highly dependent on the location and height of the emissions source (Aamaas et al., 2016). In addition to possible biases from all non-CO2 forcers that are treated as GHG, but which may in practice feed into the residual pattern, nitrate aerosol effects might be a particular bias to look out for in future projections.

# 3 METEORv1.5 - expanding the METEOR framework to produce climate impacts

The version of METEOR documented in Sandstad et al. (2025) emulated gridded, annual mean variables – and is unique in pattern scaling models for differentiating between the spatial climate effects of aerosols and greenhouse gases, for resolving the emergence of signals on multiple timescales, and for its capacity to represent irreversible climate impacts during an overshoot scenario.

While useful for representing climate evolution and indicators, many climate impacts require resolution of shorter time intervals, and simulation of stochastic variability (or weather). With this in mind, METEORv1.5 builds on the climate patterns simulated by METEOR to represent the seasonal cycle and weather variability, together with a modular climate impacts representation (Sanderson et al., 2025). These additions enable METEOR to generate realistic monthly climate projections with proper seasonal cycles and short-term variability, while providing tools to translate these climate outputs into societally relevant impact metrics. The framework maintains computational efficiency while adding the temporal resolution and impact assessment capabilities necessary for many climate risk applications.

## 3.1 Seasonal model

METEORv1.5 represents seasonal cycles through a combination of harmonic analysis and temperature-dependent parameterization. Annual and semi-annual harmonic components are extracted from the monthly climate data and modelled as functions of the evolving global temperature trajectory, allowing seasonal amplitudes and phases to shift under warming scenarios.

The seasonal model captures both the mean seasonal cycle and its variability, including changes in seasonal timing (such as earlier spring onset) and amplitude modifications (such as reduced winter-summer temperature contrasts in polar regions). Temperature-dependent seasonality parameters are derived from the CMIP6 training data by fitting relationships between seasonal characteristics and background climate state. This enables the model to generate monthly climate sequences where seasonal patterns evolve consistently with the underlying warming trajectory, rather than simply imposing fixed seasonal cycles onto annual projections.

The seasonal component at each spatial location can be represented as:

**Eq. 16**

$$S(m,T) = A_1(T)\cos\left(\frac{2\pi m}{12} + \phi_1(T)\right) + A_2(T)\cos\left(\frac{4\pi m}{12} + \phi_2(T)\right)$$

where S(m,T) is the seasonal anomaly for month \$m\$ given global temperature T, $A_1(T)$ and \$$A_2(T)$ are the temperature-dependent amplitudes of the annual and semi-annual harmonics respectively, and $\phi_1(T)$ and $\phi_2(T)$ are the corresponding temperature-dependent phase shifts. The temperature dependencies are parameterized in a linear regression framework as:

**Eq. 17**

$$A_i(T) = A_{i,0} + \alpha_i \cdot T \quad \text{and} \quad \phi_i(T) = \phi_{i,0} + \beta_i \cdot T$$

where the coefficients $\alpha_i$ and $\beta_i$ are derived from the CMIP6 training data by fitting relationships between seasonal characteristics and background climate state. This enables the model to generate monthly climate sequences where seasonal patterns evolve consistently with the underlying warming trajectory, rather than simply imposing fixed seasonal cycles onto annual projections.

## 3.2 Noise Model

The monthly climate variability model employs a two-stage approach combining Principal Component Analysis (PCA) with Vector Autoregressive with eXogenous variables (VARX) modelling (Penm et al., 1993). The system first decomposes monthly climate anomalies from CMIP6 training data into spatial modes using PCA, typically retaining 8-12 dominant modes that capture the majority of spatial variance. These spatial modes represent the dominant patterns of climate variability and reduce the dimensionality of the problem while preserving the essential spatial structure.

The temporal evolution of these modes is then modelled using a VARX process, which extends traditional Vector Autoregressive (VAR) models to include external forcing variables. A VARX model captures both the internal dynamics of the climate system (through autoregressive terms) and the response to external forcing (through exogenous variables). For a system with *p* spatial modes, the VARX(k,s) model can be written as:

**Eq. 18**

$$\mathbf{Z}_t = \mathbf{c} + \sum_{i=1}^{k} \mathbf{A}_i \mathbf{Z}_{t-i} + \sum_{j=0}^{s} \mathbf{B}_j \mathbf{X}_{t-j} + \boldsymbol{\epsilon}_t$$

where Zt is the p-dimensional vector of principal component scores at time t, (Principal components are the coefficients of orthogonal spatial patterns which can be used to approximate the gridded variability of a climate model using a reduced basis). Xt represents the exogenous forcing variables (typically global temperature trajectory), Ai are p by p autoregressive coefficient matrices capturing lagged climate dependencies between principal components, Bj are p by q coefficient matrices relating external forcing to the modes, and (ϵt) is a vector of white noise. The lag orders k and s are by default chosen to be 1 and 3 months, capturing the memory of the climate system

This approach preserves the spatial-temporal covariance structure of natural climate variability while allowing the seasonal cycles to evolve with changing background climate states. The VARX framework naturally handles the interaction between internal climate dynamics and external forcing, enabling realistic simulation of how natural variability patterns change under different warming scenarios. METEORv1.5 represents seasonal cycles and noise for both temperature and precipitation by default, though any CMIP6 variable can be added through configuration changes (however, performance will need to be assessed on a variable-by-variable basis).

## 3.3 Ensemble Generation and Uncertainty Quantification

The fitted climate, noise and seasonal models allow for the rapid generation of large ensembles of climate futures for any given scenario. The noise model generates multiple realizations of monthly climate by sampling from the fitted VARX process, providing ensemble-based uncertainty quantification that complements METEOR's existing multi-model and multi-scenario capabilities. Users can generate ensembles with potentially hundreds of members using full climate realizations that include both the forced response and natural variability. The synthetic climate preserves key statistical properties of the original Earth System Model climates including spatial correlations, temporal persistence, and seasonal amplitude variations. This ensemble capability enables probabilistic impact assessments and risk analysis applications that require characterisation of climate uncertainty ranges.

## 3.4 Modular Impact Assessment Framework

METEOR 1.5 introduces an extensible `ImpactCalculator` base class that standardizes the interface for converting climate data into impact metrics, allowing for easy construction of additional climate impact parameterisations for either direct assessment or for incorporation into integrated assessment modelling frameworks. This modular design allows users to implement custom impact calculations while benefiting from common infrastructure for ensemble processing, metadata handling, and statistical analysis. The framework includes an `ImpactEnsemble` class that provides ensemble statistics and tools for comparing impact distributions across scenarios or time periods. The system handles both deterministic calculations from single climate realizations and probabilistic assessments from climate ensembles, with automatic propagation of uncertainty from climate inputs through to impact outputs.

## 3.5 Degree Days Application and Validation

The initial implementation of the impacts framework includes a comprehensive degree days calculator that demonstrates the framework's capabilities for energy-sector climate impacts. The calculator implements both heating degree days (HDD) and cooling degree days (CDD) with configurable base temperatures and seasonally varying standard deviations that account for sub-monthly temperature variability. The implementation follows established methodologies for translating monthly temperature data into degree day accumulations while preserving the statistical properties necessary for energy demand modelling. This application serves as both a practical tool for energy impact assessment and a template for developing additional impact calculators within the framework. The degree days calculator has been validated against observational data and demonstrates the system's ability to produce realistic impact projections that capture both long-term trends and year-to-year variability.

Together these additions allow for modelling of monthly data which can then be used to model impact related variables such as HDD (Heating Degree Days) and CDD (Cooling Degree Days) using one or more stochastic realisations for the noise (Figure 15).

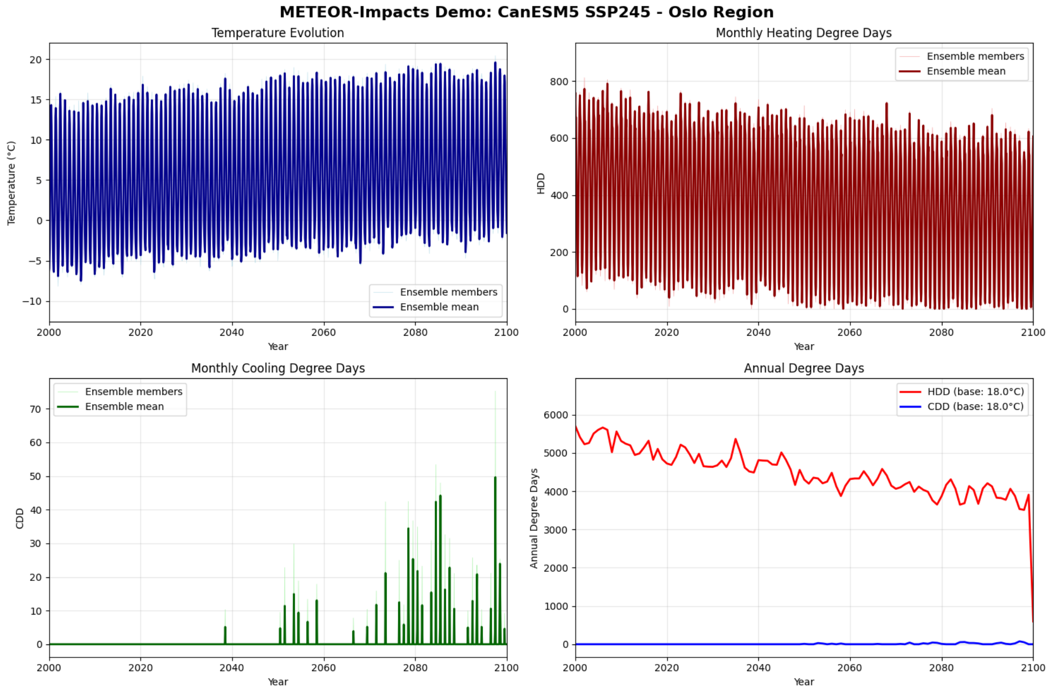



**Figure 15.** *METEOR-Impacts demonstration for Oslo region under SSP2-4.5 scenario using METEOR's ensemble emulation of CanESM5, a CMIP6 Earth System Model used in training.*

*Note: This four-panel figure illustrates the complete METEOR-Impacts workflow combining pattern scaling, monthly climate variability, and impact assessment. (a) Temperature Evolution: Monthly temperature projections (°C) showing ensemble spread (light blue lines) and ensemble mean (dark blue) from 2000-2100. The noise model generates realistic inter-annual variability around the forced warming trend. (b) Monthly Heating Degree Days: Monthly HDD accumulation (base temperature 18°C) with individual ensemble members (light red) and ensemble mean (dark red). HDD decreases over time as winters warm, with substantial month-to-month and year-to-year variability preserved. (c) Monthly Cooling Degree Days: Monthly CDD accumulation showing increasing summer cooling demand as temperatures rise above the 18°C threshold. Individual realizations (light green) capture natural variability while the ensemble mean (dark green) shows the forced trend. (d) Annual Degree Days Summary: Annual totals for both HDD (red) and CDD (blue) illustrating the long-term energy sector implications - declining heating demand and increasing cooling demand under warming. The methodology demonstrates how METEOR's pattern scaling combined with the new noise generator produces climate projections with realistic seasonal cycles and natural variability, enabling probabilistic impact assessment through ensemble analysis.*

# 4 Linking METEOR with IAMs

## 4.1 Coupling mechanisms between physical climate change and the economy

METEOR can aid in the establishment of a coupling framework between physical climate change and the economy in IAMs by providing regionally resolved inputs of physical climate variables. METEOR is explicitly motivated by the need to improve climate-economy feedback integration. While current IAMs often neglect impacts such as land-use productivity shifts, human health damages, or socio-political conflicts, METEOR offers a scalable bridge: by rapidly generating probabilistic regional climate information, it allows IAMs to systematically explore feedback pathways. The open-source implementation of METEOR ensures accessibility and interoperability. As such METEOR helps bridging the gap between ESM-scale outputs and IAM input needs. In practice, this means IAMs could integrate regionally-resolved climate projections to better capture climate–economy feedbacks. This enables rapid, computationally efficient production of spatially explicit climate information and directly tackles the limitations of one-way deterministic coupling by embedding uncertainty, non-linear dynamics, and forcing-dependent responses into the climate-economy interface.

The proposed pipeline establishes a two-way, templated process for coupling IAM scenarios with regionalised climate impacts via the open-source METEOR model. The idealised process pipeline is described as followed:

1. <u>Scenario generation in IAMs</u>: An IAM produces emissions trajectories using the IAMC standard format. These scenarios define (and quantify) the underlying socio-economic and policy assumptions.

2. <u>Preprocessing for SCMs</u>: A dedicated repository (Sandstad 2025) converts IAMC-formatted emissions into a format compatible with the CICERO Simple Climate Model (C-SCM). This ensures smooth translation of IAM outputs into the climate emulator framework. The repository is being documented and updated, ensuring visibility in project workflows.

3. <u>Running METEOR</u>: METEOR, either newly trained or using an existing trained version—depending on the use case and set of available scenario training data—receives the emissions input (via C-SCM's emissions-to-forcing engine) and produces gridded climate outputs. These include temperature and precipitation fields with timescale- and forcing-dependent patterns and can be extended to additional impact-relevant variables (see Section 2.2).

4. <u>Processing to IAM-readable format:</u> The gridded outputs are translated back into a format suitable for IAMs. Depending on the specific model, this may involve applying region masks, aggregating variables, or reducing dimensionality to quantities directly usable by the IAM (e.g., regional averages, impact indices). While the details are model-dependent, this step generally requires applying common post-processing routines.

5. <u>Feedback to IAMs:</u> The processed impact data is fed back into the IAM, which can then rerun the scenario under the new information.

This loop can be repeated iteratively to capture feedbacks between climate impacts and socio-economic pathways. Each step of the pipeline is open-source and accessible, enabling users to reproduce and adapt the workflow. In practice, however, the complexity of handling formats, dependencies, and documentation means that guidance and active support will be required to ensure smooth uptake across the IAM and SCM communities.

The advances made with the development of METEOR mark a substantial step toward risk-informed, feedback-inclusive IAM–SCM coupling. METEOR can enable IAMs to move from targeting single climate outcomes (i.e., temperature for use in damage functions) to assessing the risk distribution of climate outcomes. In addition to probabilistic outcomes from the parameter space of the training data, METEOR can be used to generate ensembles of plausible configurations, providing IAMs with risk-based climate information rather than deterministic projections.

## 4.2 Overview of impacts that can be represented in IAMs

Numerous analyses have examined the impacts of climate change on the energy system including impacts on residential and commercial heating and cooling demand, capacity factors of power plants, and potential of different energy sources (Byers et al., 2024; Isaac & van Vuuren, 2009; Yalew et al., 2020; Zapata et al., 2022). Additionally, a substantial body of literature has explored how climate change effects may influence the economy at a macroeconomic level. This includes both sector-by-sector analyses (e.g., Takakura et al., 2019) and broader empirical studies that assess the entire economy (e.g., Burke et al., 2015; Burke et al., 2023), highlighting that different types of climate change impacts influence the economy through various channels.

In this section, we conducted a survey to understand the types of impacts that can be included in the DIAMOND models and modules, through a pipeline such as the one presented in Section 2.9. Our sample included 9 models—including the six core IAMs (i.e., OMNIA, GCAM-Europe, OPEN-PROM, GEMINI-E3-EU, CLEWs-EU, NEMESIS-World), openTEPES, MIO-ES (a macroeconomic input-output model, which, although not directly part of DIAMOND, has influenced other project activities (e.g., the development of the household heterogeneity database), and NEMESIS-EU (since it constitutes a separate version from the global one)—with one response being collected per model. Figures 16, 17, 18 and 19 show the response to questions on whether various impacts can be included in the models.

Reponses show that impacts on heating and cooling degree days followed by those on crop yields are the ones that can be represented by most models, hence we have focused our efforts on these variables. There are several impact equations in the literature that utilise mean temperature and precipitation to assess the impact of climate change on overall GDP, which can be included in all the models (Burke et al., 2015; Kalkuhl & Wenz, 2020). This resultant target GDP can be achieved by including supply side or demand side shocks in the economic models similar to the methodology adopted by the Network for Greening the Financial System (NGFS) studies (Richters et al., 2024).

Population is an exogenous variable in the model and the dynamics of population—including migration, fertility, and mortality—are not included in the models involved in the consortium. This makes it challenging to assess the impact of climate change on human migration and its impact on the regional energy needs in the models, despite a couple of theoretical "Yes" responses, indicating a possibility to do so under strict assumptions.

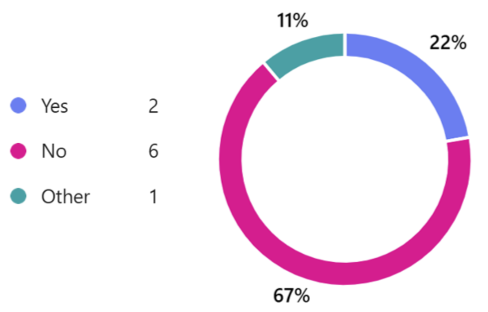



**Figure 16.** Responses from the survey indicate whether models involved in DIAMOND can include the impact of climate change on human migration in their models.

Most models that have detailed representation of energy technologies can incorporate the effects of climate change on heating and cooling energy demands. In these models, climate parameters such as heating degree days and cooling degree days are taken into account when estimating heating and cooling demand. This approach helps to establish a connection between the physical emulator and integrated assessment models (IAM). In CGE models cooling and heating degree days are not disaggregated, potentially requiring short- and long-term temperature elasticities (e.g., De Cian et al. 2012) to be used to assess the impact of temperature on the fuel mix.

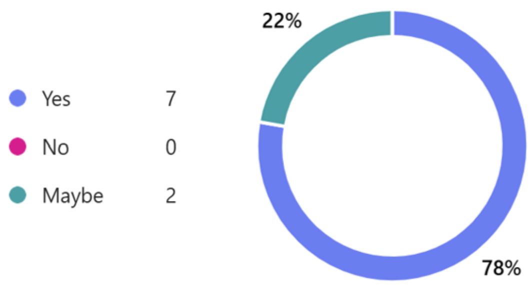



**Figure 17.** Responses from the survey indicate whether models involved in DIAMOND can include the impact of climate change on heating and cooling in their models.

Temperature and Precipitation at the spatial scale can be utilised in the CLEWs-EU land and water module. A simple water balance is included in the land and water module of the model, which allows the assessment of water availability impact on crops. For the baseline scenario, inputs for the model include precipitation, evapotranspiration, crop yield, and water deficit (crop and country specific), taken from the GAEZ database. Any change in the precipitation due to climate change from METEOR can be included in the model which would lead to lower water availability for crops. Lower availability of water in some regions can also result in increasing energy demand for desalination. OMNIA and OPEN-PROM model are linked with the land use model MAgPIE. The MAgPIE model of agriculture and land use simulates investments in agricultural research and development (R&D) endogenously, with R&D levels determined by both the degree of warming and the crop model used within the MAgPIE framework. Currently, only the impact of increased $CO_2$ concentrations on crop yields is included in the crop model, while the impact of temperature is not. This might result in the underestimation of adaptation spending on the agricultural system to protect crop yields as the world warms, or the impact on food prices due to food storages. To overcome this limitation, the generated climate data from METEOR model can be used, along with empirical assumptions to scale agricultural and forestry yields within MAgPIE, analysing resulting shifts in land use, production costs, bioenergy potential, and land-related emissions ($CO_2$, $CH_4$, $N_2O$). Several statistical models for assessing the impacts of climate on crop yields have been highlighted in the work of Roberts et al., (2017). However, the crop model linked to MAgPIE does not incorporate other climate stressors, such as beetles, drought mortality, and interactive fire, which makes it challenging to have a comprehensive representation, despite significant improvements offered through the METEOR link.

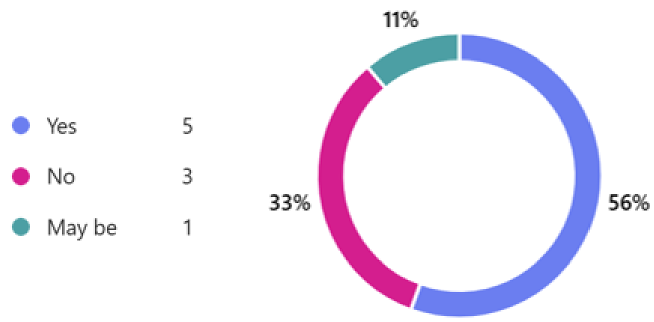



**Figure 18.** Responses from the survey indicate whether models involved in DIAMOND can include the impact of climate change on crop yield in their models.

Due to higher temperatures and rising humidity caused by climate change, workers are increasingly susceptible to illnesses and more likely to make errors and have accidents. This is largely a result of decreased concentration (OECD, 2024), which ultimately leads to reduced productivity. The thermal comfort of workers, along with their physical and physiological limits, determines work ability and directly influences labour productivity. In a CGE model, specific amounts of labour, capital, and resources are required to produce one unit of output. If climate-change-induced stress reduces productivity, then, for example, more labour per unit of output would be required. This can be represented in the CGE model by multiplying labour input by a factor greater than 1. We take the same approach as in Knittel et al 2020 to assess the impact of climate change on labour productivity and therefore on overall economy of the region. The relationship between wet bulb temperature and work ability is estimated by Bröde et al. (2017):

$$WA = 0.1 + 0.9/(1 + (WBGT/\alpha_1)^{\alpha_2})$$

Depending on the work intensity, $(\alpha_1, \alpha_2)$ can take the values of (34.64, 22.72) for light work, (32.93, 17.81) for medium work and (30.94, 16.64) for heavy work. Wet bulb temperature can be calculated using the temperature and relative humidity from the METEOR model that can be used to derive the work ability of labour in different scenarios.

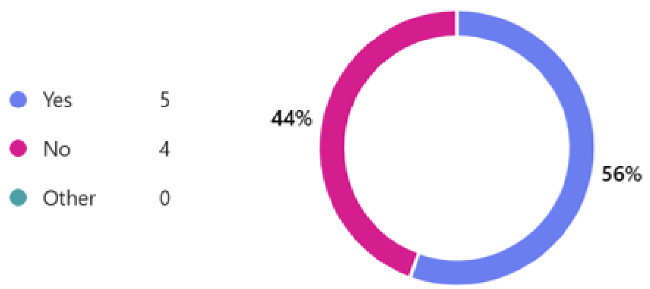



**Figure 19.** Responses from the survey indicate whether models involved in DIAMOND can include the impact of climate change on labour productivity in their models.

Davidson and Janssens (2006) shows that carbon stored in the soil is released into the atmosphere due to accelerated decomposition caused by warming, which can create a positive feedback loop that exacerbates climate change. On the other hand, if the increase in plant-derived carbon inputs to the soil surpasses the rise in decomposition rates, the feedback would be negative. Despite extensive research, there is still no consensus on how temperature affects the decomposition of soil carbon and not many models can represent these dynamics, going beyond the focus on this task.

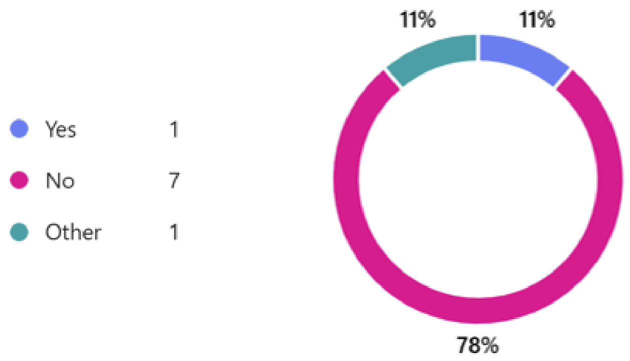



**Figure 20.** Responses from the survey indicate whether models involved in DIAMOND can include the impact of climate change on soil organic carbon in their models.

## 4.3 Illustration of coupling between METEOR and OPEN-PROM

In this section, we describe a workflow using global greenhouse gas (GHG) emissions from the OPEN-PROM model under a no-policy scenario to estimate annual precipitation and temperature (Figure 21). This workflow can be used for other IAM models too. These estimates can then be used to assess various impacts that can be incorporated into the model.

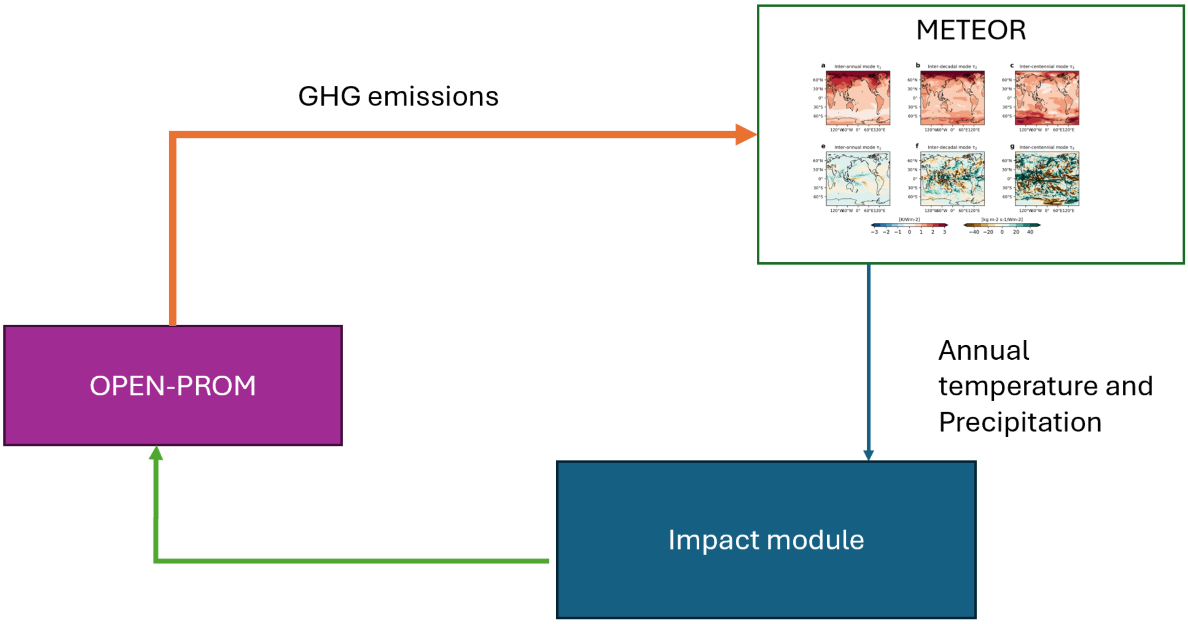



**Figure 21.** Generalised workflow diagram to showcase the coupling between the spatially resolving emulator METEOR and IAM model (example using OPEN-PROM)

Figure 22 shows the emissions for the baseline scenario from the OPEN-PROM model, while Figure 23 illustrates the projected changes in temperature and precipitation for the years 2050 and 2100 as estimated by the METEOR model at grid level. Impacts can then be incorporated back into the IAM based on the processes described in Section 3.1.

Funded by the European Union

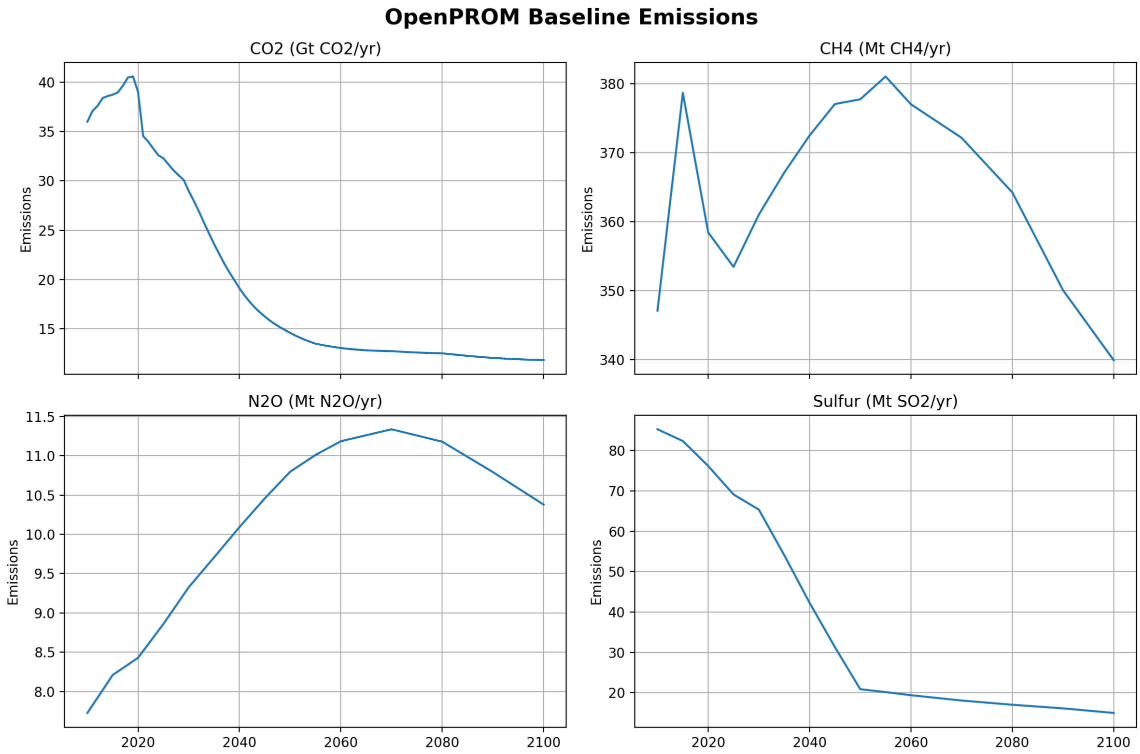



**Figure 22.** Emissions trajectories for $CO_2$, $CH_4$, $N_2O$ and Sulfur from the OpenPROM baseline DIAMOND run.

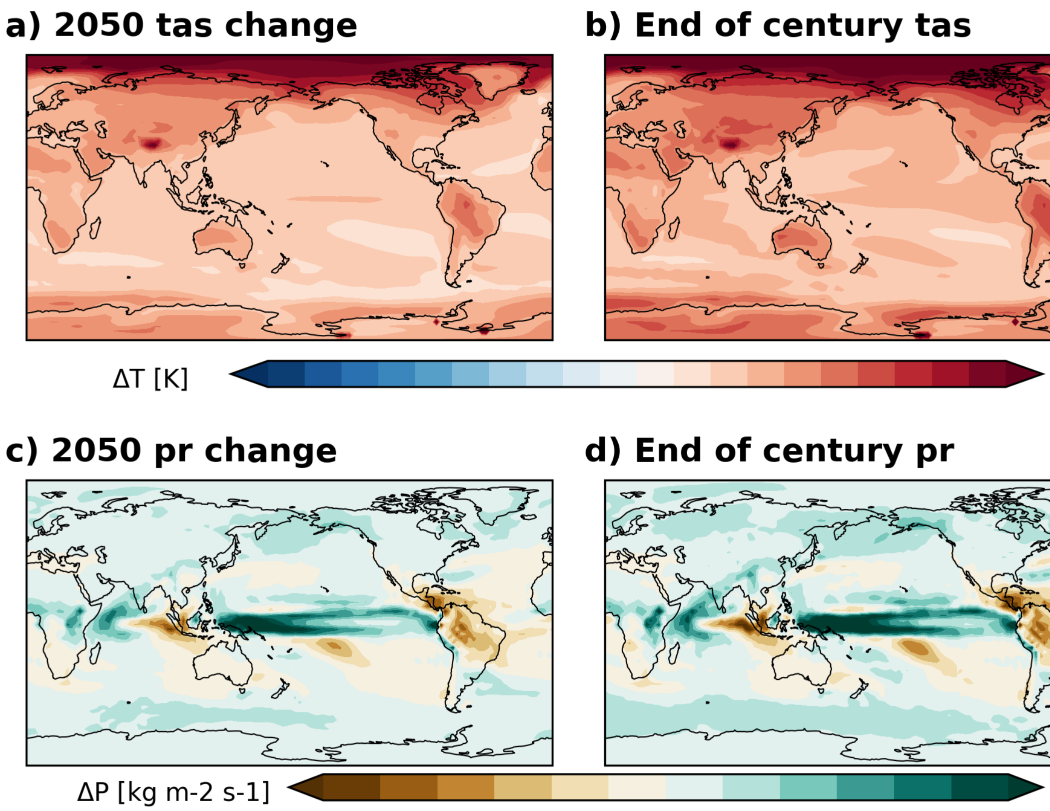



**Figure 23.** METEOR prediction maps of annual mean change.

*Note: This map showsannual mean change since 1850-1900 average in the 20-year period from 2040-2060 (i.e. 2050) in a and c, and in 2080-2100 (end of century) in b and d, using forcing from OPEN-PROM's baseline DIAMOND run and emulating to the CanESM5 CMIP6 climate model. Colourscales are symmetrical ranging from -10 to 10 degrees K in temperature change, and from* $-3 \times 10^{-5}$ *to* $-3 \times 10^{-5}$ *kg m-2 s-1 for precipitation changes.*

## 4.4 Heating and Cooling demand

Several papers have looked at the impacts of variations in heating and cooling of buildings due to future temperature changes (Byers et al., 2024; Colelli & Cian, 2020; Yalew et al., 2020). The seasonal and variability enhanced version of METEOR-Noise can be used to generate such estimates to feed into IAM models to provide climate feedbacks. Figure 24 shows the METEOR generated cooling degree days, estimated for a single stochastic realisation of a grid cell over Oslo for the SSP2-4.5 scenario using the HDD and CDD parametrisation described in Isaac and van Vuuren (2009) to go from local monthly temperature outputs from METEOR-Noise to monthly HDD or CDD predictions. At the global level, the climate change impact on heating and cooling service demand and additional investments requirements to adapt to these changes might not be significant, but at the regional level including such dynamics helps adapt electricity planning to the high peak demand period. In the NEMESIS-World model, there is no representation of total energy consumption in households by its uses (heating, cooling and lighting etc), so this issue needs to be addressed by fixing the share of different end uses from a technologically detailed model for the baseline scenario. The variation of HDD and CDD in the GEMINI-E3-EU model can be incorporated by adjusting technical progress associated with fossil energy and electricity consumption for residential purpose. This adjustment is based on the assumption that changes in HDD and CDD are equivalent to a decrease or increase in the energy required to maintain the same level of comfort for heating and cooling.

Figure 25 shows the linkages between different models (a link to NEMESIS-World may still be possible but would require a three-way connection with a technologically detailed model).

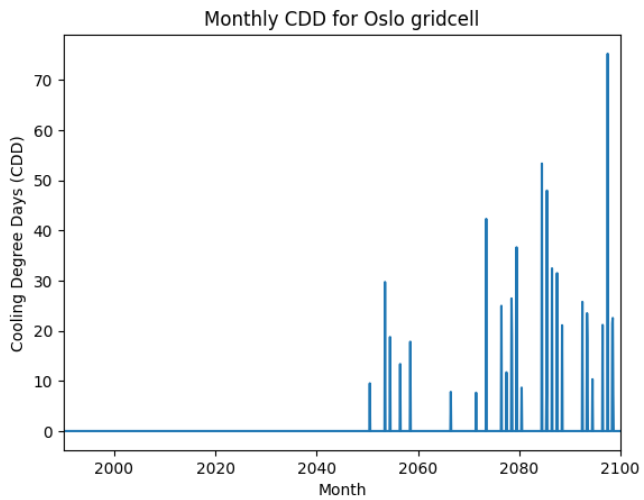



**Figure 24.** Simulated stochastic cooling degree days for Oslo, generated from a single stochastic realisation of METEOR-Noise.

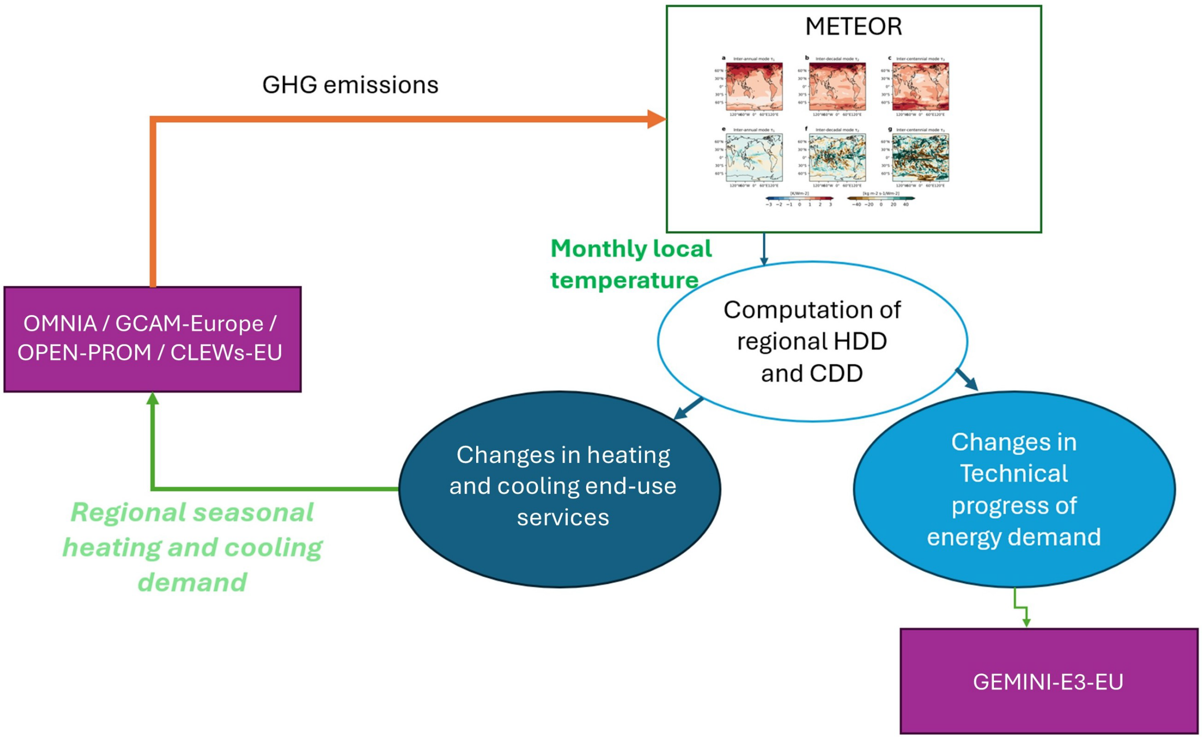



**Figure 25.** Illustration demonstrates the flow of emissions data from IAMs (Integrated Assessment Models) to METEOR and how this data can be used to compute regional Heating Degree Days (HDD) and Cooling Degree Days (CDD) values for different scenarios that these values are used to estimate the changes in heating and cooling end-use demand feeding back to IAMs.

## 4.5 Outlook

The potential applications of METEOR are not limited to mean temperature and precipitation. The model 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 of such expansions can include estimating the impact from variables such as extreme value impact indicators 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. Although, we expect such indicators to require additional modules, such advancements may further assist towards capturing climate change implications on mitigation pathways in IAMs.

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 is contained in PDRMIP. However, PDRMIP was performed with older versions of modelling frameworks. METEOR can be applied to that dataset, or similar updated ones, 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.

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. For instance, the DAMIP (The Detection and Attribution Model Intercomparison Project) (Gillett et al., 2016) and RAMIP (Regional Aerosol Model Intercomparison Project) (Wilcox et al., 2023) datasets offer such data, the latter of 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 would be useful for impact applications and would be logical extensions for future development. Probabilistic spatial information is implemented using various techniques from the MESMER framework ( Beusch et al., 2020; Quilcaille et al., 2022; Quilcaille et al., 2023). Goals for future versions of METEOR include probabilistic implementations 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 on a global scale for the FAIR simple climate model (Bouabid, Sejdinovic, & Watson-Parris, 2024) are both possible approaches to this.

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.

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. This on top of the envisaged use of the model in project activities as part of the link with IAMs (see WP5 activities). Finally, we look forward to community development of the METEOR platform to provide better integration into ecosystems of fast climate modelling tools which can be increasingly used in applications which require rapid turnaround, such as simulating regional climate impacts in societal models.

# References


Aamaas, B., Berntsen, T. K., Fuglestvedt, J. S., Shine, K. P., & Bellouin, N. (2016). Regional emission metrics for short-lived climate forcers from multiple models. *Atmospheric Chemistry and Physics, 16*, 7451-7468. doi:10.5194/acp-16-7451-2016

Adams, P. J., Seinfeld, J. H., Koch, D., Mickley, L., & Jacob, D. (2001). General circulation model assessment of direct radiative forcing by the sulfate-nitrate-ammonium-water inorganic aerosol system. *Journal of Geophysical Research: Atmospheres, 106*, 1097-1111. doi:10.1029/2000JD900512

Barata, J. C. A., & Hussein, M. S. (2012). The Moore–Penrose pseudoinverse: A tutorial review of the theory. *Brazilian Journal of Physics, 42*, 146-165. doi:10.1007/s13538-011-0052-z

Bauer, S. E., Koch, D., Unger, N., Metzger, S. M., Shindell, D. T., & Streets, D. G. (2007). Nitrate aerosols today and in 2030: a global simulation including aerosols and tropospheric ozone. *Atmospheric Chemistry and Physics, 7*, 5043-5059. doi:10.5194/acp-7-5043-2007

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

Beusch, L., Gudmundsson, L., & Seneviratne, S. I. (2020). Emulating Earth system model temperatures with MESMER: from global mean temperature trajectories to grid-point-level realizations on land. *Earth System Dynamics, 11*(1), 139-159. doi:10.5194/esd-11-139-2020

Beusch, L., Nicholls, Z., Gudmundsson, L., Hauser, M., Meinshausen, M., & Seneviratne, S. I. (2022). 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. doi:10.5194/gmd-15-2085-2022

Bouabid, S., Sejdinovic, D., & Watson-Parris, D. (2024). FaIRGP: A Bayesian Energy Balance Model for Surface Temperatures Emulation. *Journal of Advances in Modeling Earth Systems, 16*. doi:10.1029/2023MS003926

Burke, M., Hsiang, S. M., & Miguel, E. (2015). Global non-linear effect of temperature on economic production. *Nature, 527*(7577), 235-239. doi:10.1038/nature15725

Burke, M., Zahid, M., Diffenbaugh, N., & Hsiang, S. M. (2023). Quantifying Climate Change Loss and Damage Consistent with a Social Cost of Greenhouse Gases. *NBER Working Papers, 31658*. Retrieved from

Byers, E., Meng, M., Mastrucci, A., van Ruijven, B., & Krey, V. (2024). Flexible emulation of the climate warming cooling feedback to globally assess the maladaptation implications of future air conditioning use. *Environmental Research: Energy, 1*(3). doi:10.1088/2753-3751/ad6f11

Colelli, F. P., & Cian, E. D. (2020). Cooling demand in integrated assessment models: a methodological review. *Environmental Research Letters, 15*(11). doi:10.1088/1748-9326/abb90a

Collins, M., Knutti, R., Arblaster, J., Dufresne, J. L., Fichefet, T., Friedlingstein, P., . . . Wehner, M. (2013). Long-term Climate Change: Projections, Commitments and Irreversibility. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 1029-1136. doi:10.1017/CBO9781107415324.024

Davidson, E. A., & Janssens, I. A. (2006). Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature, 440*(7081), 165-173. doi:10.1038/nature04514

De Cian, E., Lanzi, E., & Roson, R. (2012). Seasonal temperature variations and energy demand. *Climatic Change, 116*(3-4), 805-825. doi:10.1007/s10584-012-0514-5

Dietz, S., Rising, J., Stoerk, T., & Wagner, G. (2021). Economic impacts of tipping points in the climate system. *Proceedings of the National Academy of Sciences, 118*, e2103081. doi:10.1073/pnas.2103081118

Dunne, J. P., Hewitt, H. T., Arblaster, J., Bonou, F., Boucher, O., Cavazos, T., . . . Taylor, K. E. (2024). An evolving Coupled Model Intercomparison Project phase 7 (CMIP7) and Fast Track in support of future climate assessment. *EGUsphere*, 1-51. doi:10.5194/egusphere-2024-3874

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

Forster, P., Huppmann, D., Kriegler, E., Mundaca, L., Smith, C., Rogelj, J., & R., S. (2018). Mitigation Pathways Compatible with 1.5°C in the Context of Sustainable Development In V. Masson-Delmotte, P. Zhai, H.

Pörtner, D. Roberts, J. Skea, P. R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. L. Connors, J. B. R. Matthews, Y. Chen, X. Zhou, M. I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, & T. Waterfield (Eds.), Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. Retrieved from https://www.ipcc.ch/sr15/.

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., . . . Zhang, X. (2021). The Earth's energy budget, climate feedbacks, and climate sensitivity. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, Ö. Yelekçi, R. Yu, & B. Zhou (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 923-1054). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Forster, P. M., Smith, C., Walsh, T., Lamb, W. F., Lamboll, R., Cassou, C., . . . Zhai, P. (2025). Indicators of Global Climate Change 2024: annual update of key indicators of the state of the climate system and human influence. *Earth System Science Data, 17*, 2641-2680. doi:10.5194/essd-17-2641-2025

Forster, P. M., Smith, C., Walsh, T., Lamb, W. F., Lamboll, R., Hall, B., . . . Zhai, P. (2024). Indicators of Global Climate Change 2023: annual update of key indicators of the state of the climate system and human influence. *Earth System Science Data, 16*(6), 2625-2658. doi:10.5194/essd-16-2625-2024

Freese, L. M., Giani, P., Fiore, A. M., & Selin, N. E. (2024). Spatially Resolved Temperature Response Functions to CO2 Emissions. *Geophysical Research Letters, 51*, e2024GL108788. doi:10.1029/2024GL108788

Gillett, N. P., Shiogama, H., Funke, B., Hegerl, G., Knutti, R., Matthes, K., . . . Stone, D. (2016). The Detection and Attribution Model Intercomparison Project (DAMIP v1.0) contribution to CMIP6. *Geoscientific Model Development, 9*, 3685-3697. doi:10.5194/gmd-9-3685-2016

Good, P., Lowe, J. A., Andrews, T., Wiltshire, A., Chadwick, R., Ridley, J. K., . . . Shiogama, H. (2015). Nonlinear regional warming with increasing CO2 concentrations. *Nature Climate Change, 5*. doi:10.1038/nclimate2498

Hauglustaine, D. A., Balkanski, Y., & Schulz, M. (2014). A global model simulation of present and future nitrate aerosols and their direct radiative forcing of climate. *Atmospheric Chemistry and Physics, 14*, 11031-11063. doi:10.5194/acp-14-11031-2014

Herger, N., Sanderson, B. M., & Knutti, R. (2015). Improved pattern scaling approaches for the use in climate impact studies. *Geophysical Research Letters, 42*, 3486-3494. doi:10.1002/2015GL063569

Hulme, M., Wigley, T. M. L., Barrow, E., Raper, S., Centella, A., Smith, S., & Chipanshi, A. (2000). Using a Climate Scenario Generator for Vulnerability and Adaptation Assessments. In B. Lim & J. Smith (Eds.), *MAGICC and SCENGEN Version 2.4 Workbook and CD-ROM*. East Anglia: University of East Anglia, Climatic Research Unit.

Huntingford, C. (2000). Climate sensitivity analysis of forest growth uncertainties. *Global Ecology and Biogeography, 9*, 359-374.

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

Ipcc. (2022). *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, UK and New York, NY, USA: Cambridge University Press.

Isaac, M., & van Vuuren, D. P. (2009). Modeling global residential sector energy demand for heating and air conditioning in the context of climate change. *Energy Policy, 37*(2), 507-521. doi:10.1016/j.enpol.2008.09.051

Jonko, A. K., Shell, K. M., Sanderson, B. M., & Danabasoglu, G. (2013). Climate feedbacks in CCSM3 under changing CO2 forcing. Part II: Variation of climate feedbacks and sensitivity with forcing. *Journal of Climate, 26*, 2784-2795.

Kalkuhl, M., & Wenz, L. (2020). The impact of climate conditions on economic production. Evidence from a global panel of regions. *Journal of Environmental Economics and Management, 103*. doi:10.1016/j.jeem.2020.102360

Kikstra, J. S., Waidelich, P., James, R., D, Y., Hope, C., & Brierley, C. M. (2021). The social cost of carbon dioxide under climate-economy feedbacks and temperature variability. *Environmental Research Letters, 16*. doi:10.1088/1748-9326/ac1d0b

Liao, H., & Seinfeld, J. H. (2005). Global impacts of gas-phase chemistry-aerosol interactions on direct radiative forcing by anthropogenic aerosols and ozone. *Journal of Geophysical Research: Atmospheres, 110*. doi:10.1029/2005JD005907

Liao, H., Zhang, Y., Chen, W. T., Raes, F., & Seinfeld, J. H. (2009). Effect of chemistry-aerosol-climate coupling on predictions of future climate and future levels of tropospheric ozone and aerosols. *Journal of Geophysical Research: Atmospheres, 114*. doi:10.1029/2008JD010984

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

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

Mitchell, T. D. (2003). Pattern Scaling: An Examination of the Accuracy of the Technique for Describing Future Climates. *Climatic Change, 60*(3), 217-242. doi:10.1023/A:1026035305597

Myhre, G., Forster, P. M., Samset, B. H., Hodnebrog, O., Sillmann, J., Aalbergsjo, S. G., . . . Zwiers, F. (2017). PDRMIP: A Precipitation Driver and Response Model Intercomparison Project—Protocol and Preliminary Results. *Bulletin of the American Meteorological Society, 98*, 1185-1199. doi:10.1175/BAMS-D-16-0019.1

Myhre, G., Samset, B., Forster, P. M., Hodnebrog, O., Sandstad, M., Mohr, C. W., . . . Watson-Parris, D. (2022). Scientific data from precipitation driver response model intercomparison project. *Scientific Data, 9*, 2052-4463. doi:10.1038/s41597-022-01194-9

Nath, S., Carreau, J., Kornhuber, K., Pfleiderer, P., Schleussner, C. F., & Naveau, P. (2024). MERCURY: A fast and versatile multi-resolution based global emulator of compound climate hazards. *arXiv*.

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

Nicholls, Z., Meinshausen, M., Lewis, J., Corradi, M. R., Dorheim, K., Gasser, T., . . . Woodard, D. L. (2021). Reduced Complexity Model Intercomparison Project Phase 2: Synthesizing Earth System Knowledge for Probabilistic Climate Projections. *Earth's Future, 9*. doi:10.1029/2020EF001900

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., . . . Sanderson, B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development, 9*, 3461-3482. doi:10.5194/gmd-9-3461-2016

OECD. (2024). "Feeling the heat: Working under increasing temperatures", OECD Publishing, Paris, https://www.oecd.org/en/publications/feeling-the-heat_e07ce0e6-en.html.

Penm, J. H. W., Penm, J. H., & Terrell, R. D. (1993). The Recursive Fitting of Subset Varx Models. *Journal of Time Series Analysis, 14*(6), 603-619. doi:10.1111/j.1467-9892.1993.tb00169.x

Persad, G., Samset, B. H., Wilcox, L. J., Allen, R. J., Bollasina, M. A., Booth, B. B. B., . . . Zhao, A. (2023). Rapidly evolving aerosol emissions are a dangerous omission from near-term climate risk assessments. *Environmental Research: Climate, 2*(3). doi:10.1088/2752-5295/acd6af

Peters, G. P., Al Khourdajie, A., Sognnaes, I., & Sanderson, B. M. (2023). AR6 scenarios database: an assessment of current practices and future recommendations. *npj Climate Action, 2*(1), 1-6. doi:10.1038/s44168-023-00050-9

Quilcaille, Y., Gudmundsson, L., Beusch, L., Hauser, M., & Seneviratne, S. I. (2022). Showcasing MESMER-X: Spatially Resolved Emulation of Annual Maximum Temperatures of Earth System Models. *Geophysical Research Letters, 49*. doi:10.1029/2022GL099012

Quilcaille, Y., Gudmundsson, L., & Seneviratne, S. I. (2023). Extending MESMER-X: a spatially resolved Earth system model emulator for fire weather and soil moisture. *Earth System Dynamics, 14*(6), 1333-1362. doi:10.5194/esd-14-1333-2023

Roberts, M. J., Braun, N. O., Sinclair, T. R., Lobell, D. B., & Schlenker, W. (2017). Comparing and combining process-based crop models and statistical models with some implications for climate change. *Environmental Research Letters, 12*(9). doi:10.1088/1748-9326/aa7f33

Richters, O., Kriegler, E., Anz, J., Bertram, C., Cui, R., Edmonds, J., ... & Zwerling, M. (2024). NGFS Climate Scenarios Technical Documentation: V5. 0.

Samset, B. H., Lund, M. T., Bollasina, M., Myhre, G., & Wilcox, L. (2019). Emerging Asian aerosol patterns. *Nature Geoscience, 12*(8), 582-584. doi:10.1038/s41561-019-0424-5

Samset, B. H., Sand, M., Smith, C. J., Bauer, S. E., Forster, P. M., Fuglestvedt, J. S., . . . Schleussner, C. F. (2018). Climate Impacts From a Removal of Anthropogenic Aerosol Emissions. *Geophys Res Lett, 45*(2), 1020-1029. doi:10.1002/2017GL076079

Sanderson, B., Sandstad, M., & Steinert, N. (2025). METEOR 1.5 - Climate Variability and Impacts (Version v1.5.0): Zenodo. Retrieved from https://doi.org/10.5281/zenodo.17523936

Sanderson, B. M., Brovkin, V., Fisher, R., Hohn, D., Ilyina, T., Jones, C., . . . Lawrence, D. (2024). flat10MIP: An emissions-driven experiment to diagnose the climate response to positive, zero, and negative CO2 emissions. *EGUsphere, 2024*, 1-39.

Sandstad, M. (2025). ciceroOslo/cscm-input-data-generation: v1.0.0 (Version v1.0.0): Zenodo. Retrieved from https://doi.org/10.5281/zenodo.17195849

Sandstad, M., Aamaas, B., Johansen, A. N., Lund, M. T., Peters, G. P., Samset, B. H., . . . Skeie, R. B. (2024). CICERO Simple Climate Model (CICERO-SCM v1.1.1) – an improved simple climate model with a parameter calibration tool. *Geoscientific Model Development, 17*(17), 6589-6625. doi:10.5194/gmd-17-6589-2024

Sandstad, M., Steinert, N. J., Baur, S., & Sanderson, B. M. (2025). METEORv1.0.1: A novel framework for emulating multi-timescale regional climate responses. *EGUsphere [preprint]*. doi:10.5194/egusphere-2025-1038

Schoenberg, W., Blanz, B., Rajah, J. K., Callegari, B., Wells, C., Breier, J., . . . Mauritzen, C. (2025). An overview of FRIDA v2.1: a feedback-based, fully coupled, global integrated assessment model of climate and humans. *Geoscientific Model Development, 18*(21), 8047-8069. doi:10.5194/gmd-18-8047-2025

Schultes, A., Piontek, F., Soergel, B., Rogelj, J., Baumstark, L., Kriegler, E., . . . Luderer, G. (2021). Economic damages from on-going climate change imply deeper near-term emission cuts. *Environmental Research Letters, 16*(10). doi:10.1088/1748-9326/ac27ce

Schwaab, J., Hauser, M., Lamboll, R. D., Beusch, L., Gudmundsson, L., Quilcaille, Y., . . . Seneviratne, S. I. (2024). Spatially resolved emulated annual temperature projections for overshoot pathways. *Scientific Data, 11*. doi:10.1038/s41597-024-04122-1

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

Takakura, J. y., Fujimori, S., Hanasaki, N., Hasegawa, T., Hirabayashi, Y., Honda, Y., . . . Hijioka, Y. (2019). Dependence of economic impacts of climate change

on anthropogenically directed pathways. *Nature Climate Change, 9*(10), 737-741. doi:10.1038/s41558-019-0578-6

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

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

Watson-Parris, D. (2021). *ClimateBench*.

Watson-Parris, D., Rao, Y., Olivie, D., Seland, O., Nowack, P., Camps-Valls, G., . . . Roesch, C. (2022). ClimateBenchv1.0: A Benchmark for Data-Driven Climate Projections. *Journal of Advances in Modeling Earth Systems, 14*. doi:10.1029/2021MS002954

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

Wilcox, L. J., Allen, R. J., Samset, B. H., Bollasina, M. A., Griffiths, P. T., Keeble, J., . . . Westervelt, D. M. (2023). The Regional Aerosol Model Intercomparison Project (RAMIP). *Geoscientific Model Development, 16*, 4451-4479. doi:10.5194/gmd-16-4451-2023

Yalew, S. G., van Vliet, M. T. H., Gernaat, D. E. H. J., Ludwig, F., Miara, A., Park, C., . . . van Vuuren, D. P. (2020). Impacts of climate change on energy systems in global and regional scenarios. *Nature Energy, 5*(10), 794-802.

doi:10.1038/s41560-020-0664-z

Zapata, V., Gernaat, D. E. H. J., Yalew, S. G., Santos da Silva, S. R., Iyer, G., Hejazi, M., & van Vuuren, D. P. (2022). Climate change impacts on the energy system: a model comparison. *Environmental Research Letters, 17*(3). doi:10.1088/1748-9326/ac5141

Zelazowski, P., Huntingford, C., Mercado, L. M., & Schaller, N. (2018). 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. doi:10.5194/gmd-11-541-2018

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

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

# ANNEX

### Fitness depending on number of timescales

The METEOR model can be trained with an arbitrary number of timescales $\tau_k$. 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 an arbitrary number of timescales and corresponding spatial patterns. Figure 15 shows the effect of increasing the number of timescales on the emulation fits. Panels a and d show each model's results 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 ***Figure A 1***, show the RMSE model mean fit for temperature and precipitation with varying 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 ***Figure A 2*** (temperature) and ***Figure A 3*** (precipitation). From this, we observe that a combination of three timescales for both aerosols and GHG provides 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 timescales, 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 cannot 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 A 4*** 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 residual 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.

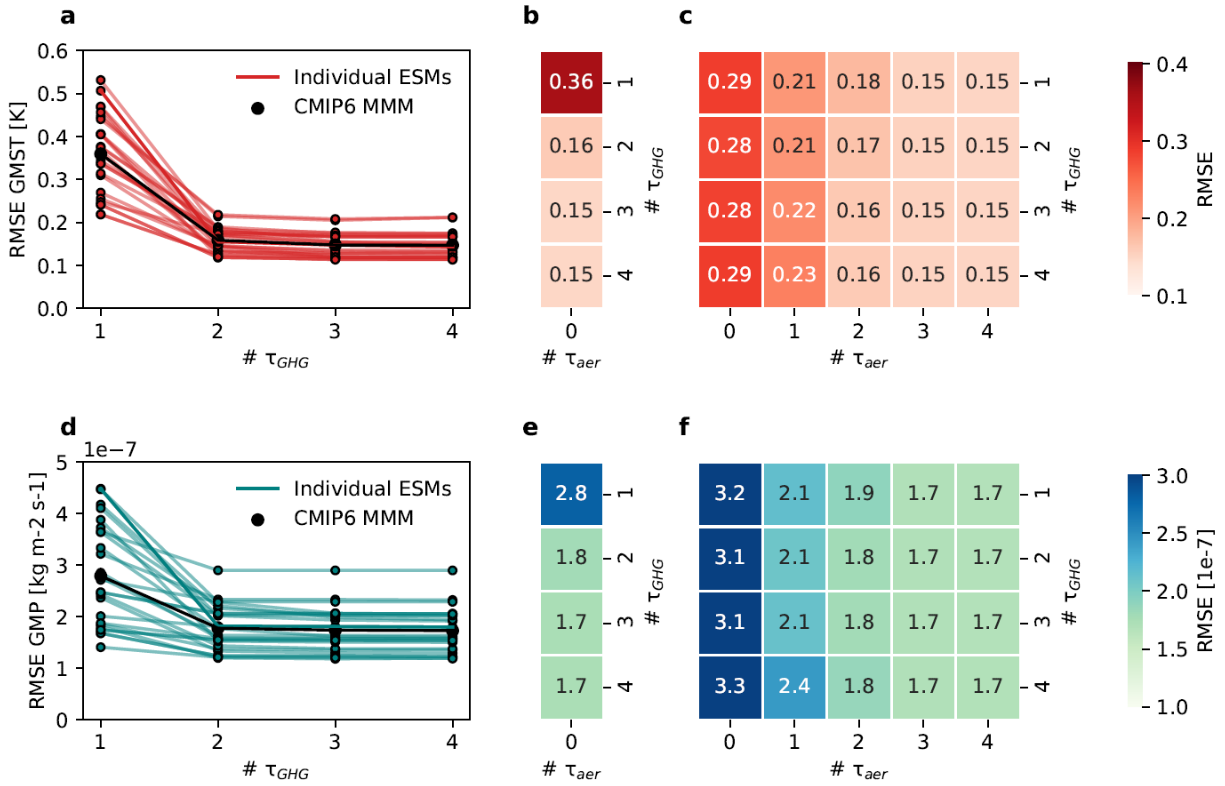



**Figure A 1:** *Evaluation of METEOR reconstruction as function of number of timescales. Panels a and d show the evolution of the RMSE parameter for temperature and precipitation respectively in the reconstruction of the abrupt4x-$CO_2$ experiment as a function of increased number of timescales for 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.*

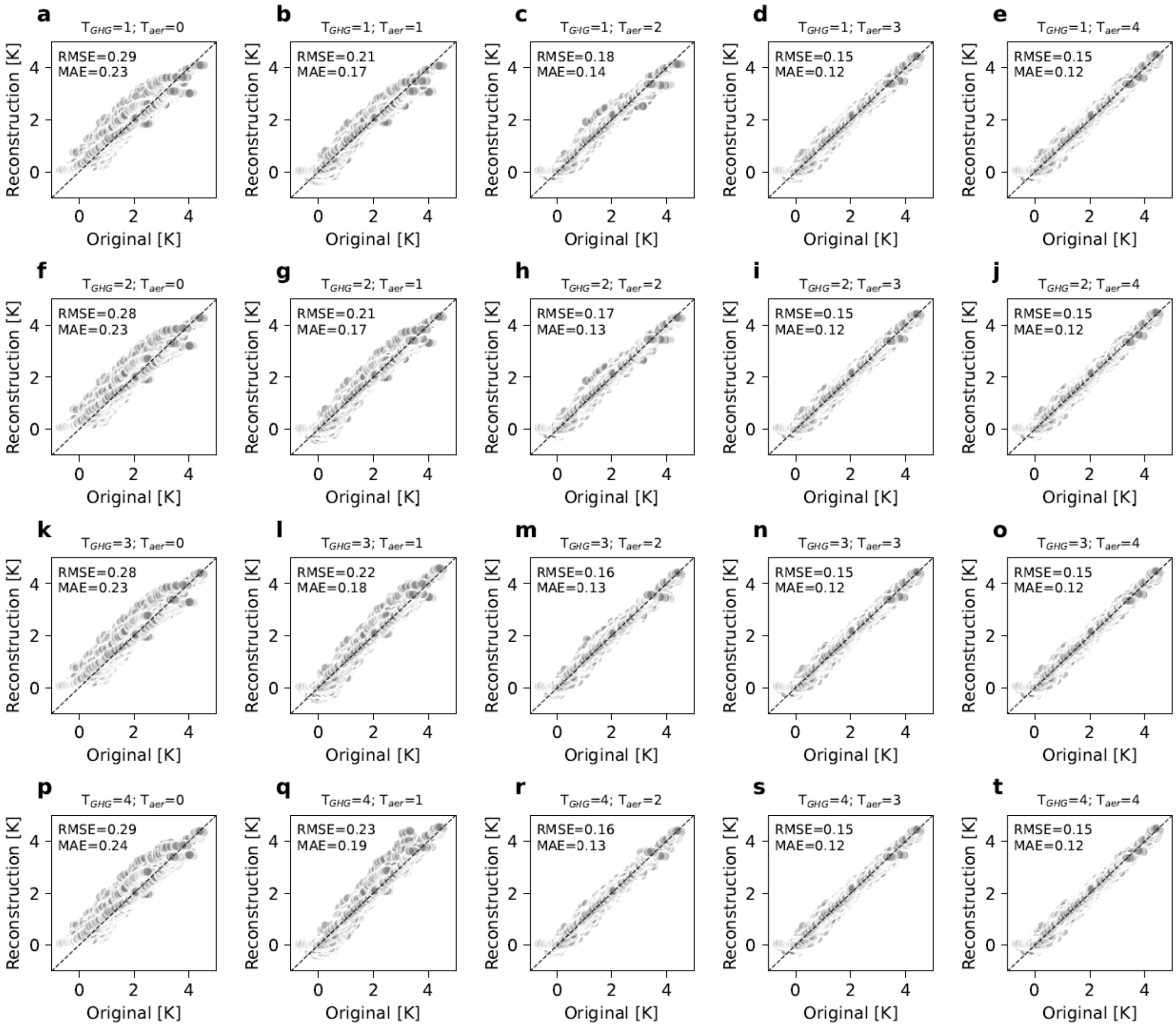



**Figure A 2:** *Evaluation of tas for varying number of timescales. Columns left to right show increasing numbers of aerosol timescales from 0 to 4, while rows from top to bottom show increasing numbers of GHG timescales from 1 to 4. Each point represents a global mean value for temperature for one model in the original CMIP6 data versus the METEOR reconstruction for SSP2-4.5. Model mean overall RMSE and mean absolute error (MAE) are displayed for each timescale combination.*

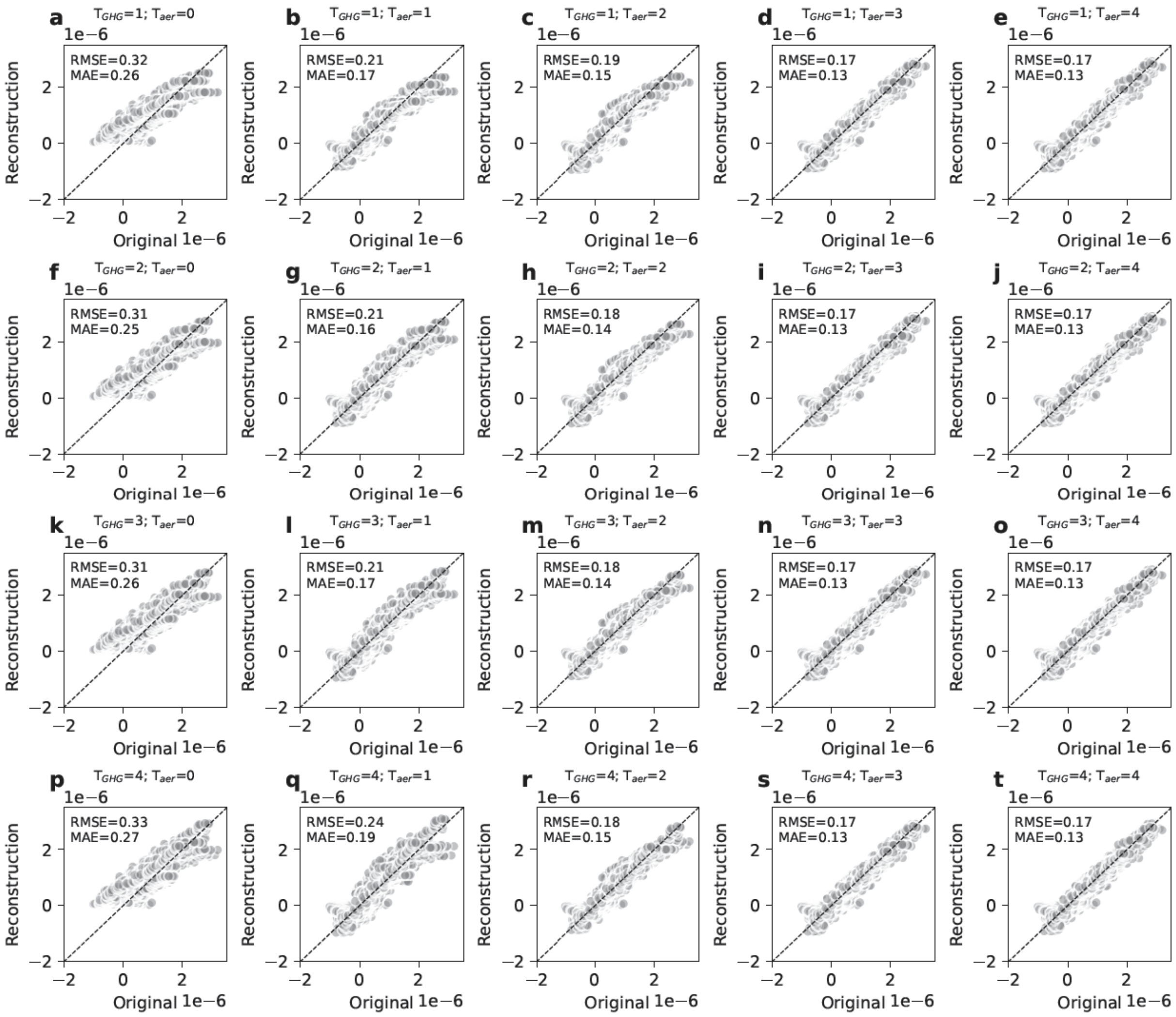



**Figure A 3:** *Evaluation of pr for varying number of timescales. Columns left to right show increasing numbers of aerosol timescales from 0 to 4, while rows from top to bottom show increasing numbers of GHG timescales from 1 to 4. Each point represents a global mean value for temperature for one model in the original CMIP6 data versus the METEOR reconstruction for SSP2-4.5. Model mean overall RMSE and mean absolute error (MAE) are displayed for each timescale combination.*

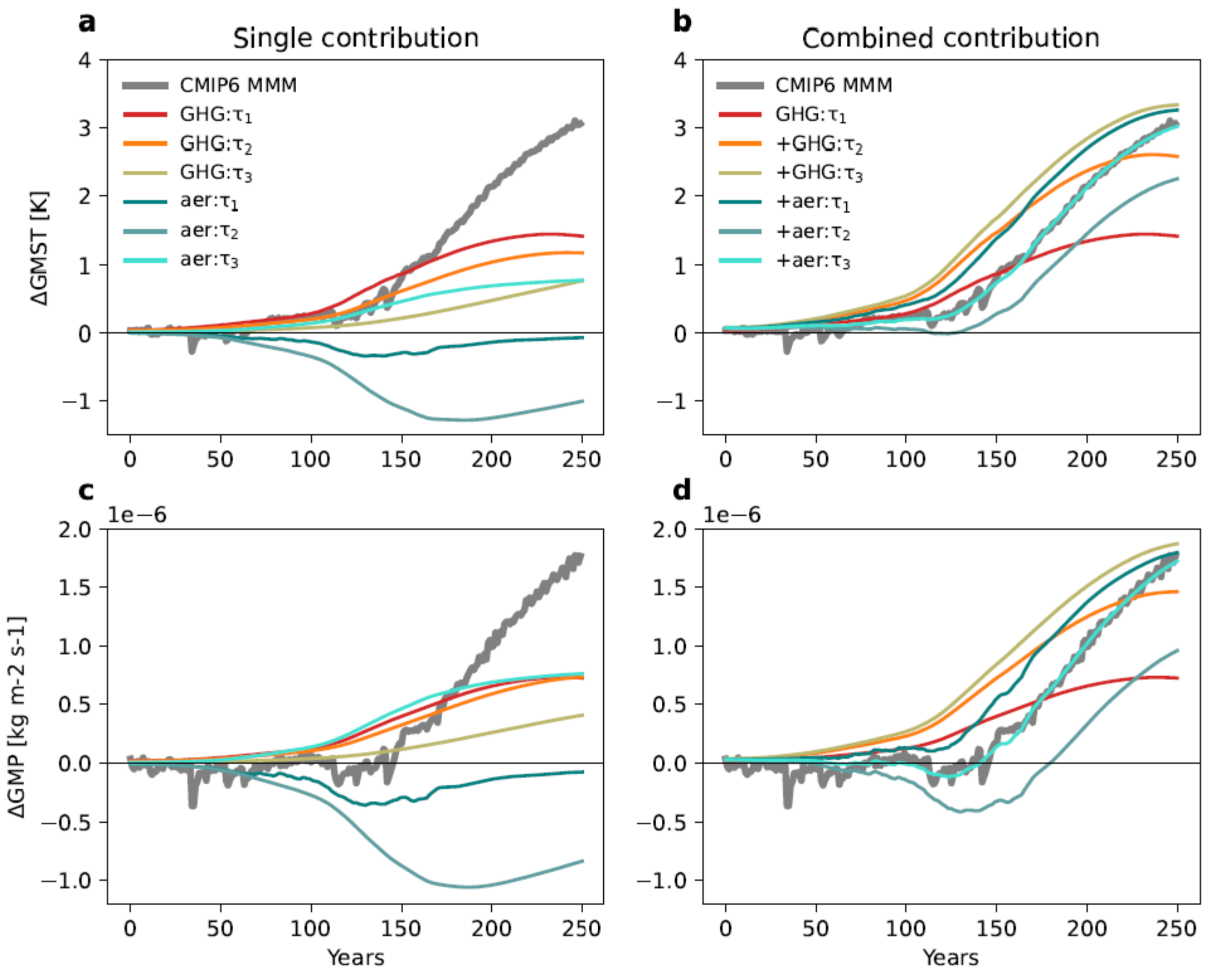



**Figure A 4:** *Multi-model mean per timescale contributions. With a model trained with three GHG and three aerosol timescales, we show the model mean global reconstruction contributions from each timescale pattern when emulating SSP2-4.5. Panels a (temperature) and c (precipitation) show the separate 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 d (precipitation) show the result of summing the patterns in this order.*

### Results for single model emulation

Here we include results per individual model emulated. Table B 1 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. ***Table B 2, Table B 3, Table B 4*** and ***Table B 5*** show ClimateBench-style error scores for all model and experiment combinations.

Global mean reconstructions per model for each of the four SSP experiments are shown in ***Figures B1-B3*** (temperature) and ***Figures B5-B7*** (precipitation). In addition, a smaller selection of models which also had data for *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 B3 (temperature) and B8 (precipitation).

We also show plots that illustrate the single model spatial performance of METEOR, by comparing end of 21st century emulation and CMIP6 data for all the standard scenarios for one model for which METEOR had some of the overall best performance (NorESM2-MM) in ***Figs. B9*** and ***B11*** and for one for which METEOR had someofthe overall worst performance (CMCC-ESM2) in ***Figs. B10*** and ***B12***. For the *SSP5-3.4-over* scenario, we display the end of century emulation map comparison for all models in ***Figs. B13*** and ***B14***, showing more of a spread in performance.

**Table B 1:** *METEOR pattern scaling timescales* $\boldsymbol{\tau}$ *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 | | |
| | $\tau_1$ | $\tau_2$ | $\tau_3$ | $\tau_1$ | $\tau_2$ | $\tau_3$ | $\tau_1$ | $\tau_2$ | $\tau_3$ | $\tau_1$ | $\tau_2$ | $\tau_3$ |
| 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 |
| 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- | 1.1 | 15.1 | 1000.0 | 1.0 | 100.0 | 1000.0 | 1.0 | 10.5 | 687.8 | 1.0 | 100.0 | 1000.0 |

Funded by the European Union

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| HR | | | | | | | | | | | | |
| 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 |

**Table B 2:** *METEOR performance for all models (A–CE) evaluated using the Climate bench evaluation metrics. Note that comparison is to single ensemble members, so variability driven errors are included.*

| | $\mathrm{NRMSE}_s$ tas | $\mathrm{NRMSE}_g$ tas | $\mathrm{NRMSE}_t$ tas | $\mathrm{NRMSE}_s$ pr | $\mathrm{NRMSE}_g$ pr | $\mathrm{NRMSE}_t$ pr |
|---|---|---|---|---|---|---|
| ACCESS-CM2 ssp126 | 0.114 | 0.037 | 0.298 | 0.079 | 0.003 | 0.096 |

| | | | | | | |
|---|---|---|---|---|---|---|
| ACCESS-CM2 ssp245 | 0.027 | 0.017 | 0.113 | 0.024 | 0.003 | 0.041 |
| ACCESS-CM2 ssp370 | 0.087 | 0.068 | 0.429 | 0.097 | 0.004 | 0.118 |
| ACCESS-CM2 ssp585 | 0.080 | 0.018 | 0.171 | 0.122 | 0.007 | 0.155 |
| ACCESS-ESM1-5 ssp126 | 0.176 | 0.090 | 0.624 | 0.076 | 0.005 | 0.101 |
| ACCESS-ESM1-5 ssp245 | 0.033 | 0.025 | 0.159 | 0.029 | 0.003 | 0.045 |
| ACCESS-ESM1-5 ssp370 | 0.080 | 0.043 | 0.294 | 0.113 | 0.003 | 0.130 |
| ACCESS-ESM1-5 ssp585 | 0.069 | 0.034 | 0.237 | 0.101 | 0.005 | 0.128 |
| AWI-CM-1-1-MR ssp126 | 0.121 | 0.054 | 0.392 | 0.057 | 0.002 | 0.069 |
| AWI-CM-1-1-MR ssp245 | 0.032 | 0.042 | 0.240 | 0.038 | 0.003 | 0.051 |
| AWI-CM-1-1-MR ssp370 | 0.058 | 0.036 | 0.239 | 0.067 | 0.005 | 0.090 |
| AWI-CM-1-1-MR ssp585 | 0.092 | 0.047 | 0.325 | 0.105 | 0.011 | 0.161 |
| BCC-CSM2-MR ssp126 | 0.133 | 0.072 | 0.492 | 0.067 | 0.003 | 0.085 |
| BCC-CSM2-MR ssp245 | 0.050 | 0.037 | 0.236 | 0.040 | 0.003 | 0.053 |
| BCC-CSM2-MR ssp370 | 0.159 | 0.114 | 0.727 | 0.065 | 0.006 | 0.094 |
| BCC-CSM2-MR ssp585 | 0.087 | 0.037 | 0.272 | 0.079 | 0.005 | 0.106 |
| CAMS-CSM1-0 ssp126 | 0.152 | 0.126 | 0.785 | 0.060 | 0.007 | 0.094 |
| CAMS-CSM1-0 ssp245 | 0.078 | 0.062 | 0.388 | 0.024 | 0.008 | 0.065 |
| CAMS-CSM1-0 ssp370 | 0.080 | 0.062 | 0.390 | 0.064 | 0.007 | 0.101 |
| CAMS-CSM1-0 | 0.153 | 0.135 | 0.830 | 0.098 | 0.016 | 0.177 |

| | | | | | | |
|---|---|---|---|---|---|---|
| ssp585 | | | | | | |
| CAS-ESM2-0 ssp126 | 0.112 | 0.051 | 0.366 | 0.058 | 0.005 | 0.083 |
| CAS-ESM2-0 ssp245 | 0.025 | 0.031 | 0.180 | 0.020 | 0.004 | 0.040 |
| CAS-ESM2-0 ssp370 | 0.081 | 0.050 | 0.329 | 0.060 | 0.015 | 0.136 |
| CAS-ESM2-0 ssp585 | 0.065 | 0.042 | 0.273 | 0.080 | 0.010 | 0.129 |
| CESM2 ssp126 | 0.146 | 0.055 | 0.421 | 0.077 | 0.005 | 0.100 |
| CESM2 ssp245 | 0.033 | 0.040 | 0.232 | 0.026 | 0.005 | 0.051 |
| CESM2 ssp370 | 0.120 | 0.047 | 0.357 | 0.076 | 0.014 | 0.145 |
| CESM2 ssp585 | 0.087 | 0.019 | 0.182 | 0.079 | 0.005 | 0.102 |

**Table B 3:** *METEOR performance for all models (CM–Can) evaluated using the Climate bench evaluation metrics. Note that comparison is to single ensemble members, so variability driven errors are included.*

| | $\mathrm{NRMSE}_s$ tas | $\mathrm{NRMSE}_g$ tas | $\mathrm{NRMSE}_t$ tas | $\mathrm{NRMSE}_s$ pr | $\mathrm{NRMSE}_g$ pr | $\mathrm{NRMSE}_t$ pr |
|---|---|---|---|---|---|---|
| CMCC-CM2-SR5 ssp126 | 0.133 | 0.063 | 0.449 | 0.066 | 0.004 | 0.088 |
| CMCC-CM2-SR5 ssp245 | 0.039 | 0.038 | 0.229 | 0.022 | 0.005 | 0.046 |
| CMCC-CM2-SR5 ssp370 | 0.271 | 0.151 | 1.023 | 0.096 | 0.021 | 0.199 |
| CMCC-CM2-SR5 ssp585 | 0.170 | 0.081 | 0.574 | 0.078 | 0.011 | 0.132 |
| CMCC-ESM2 ssp126 | 0.159 | 0.083 | 0.574 | 0.060 | 0.007 | 0.097 |
| CMCC-ESM2 ssp245 | 0.045 | 0.033 | 0.211 | 0.024 | 0.005 | 0.049 |
| CMCC-ESM2 ssp370 | 0.264 | 0.149 | 1.007 | 0.065 | 0.020 | 0.166 |
| CMCC-ESM2 ssp585 | 0.177 | 0.085 | 0.601 | 0.083 | 0.013 | 0.146 |
| CNRM-CM6-1 ssp126 | 0.155 | 0.069 | 0.499 | 0.056 | 0.004 | 0.078 |
| CNRM-CM6-1 ssp245 | 0.052 | 0.052 | 0.311 | 0.026 | 0.005 | 0.051 |
| CNRM-CM6-1 ssp370 | 0.112 | 0.074 | 0.483 | 0.062 | 0.007 | 0.095 |
| CNRM-CM6-1 ssp585 | 0.073 | 0.048 | 0.312 | 0.061 | 0.005 | 0.089 |
| CNRM-CM6-1-HR ssp126 | 0.116 | 0.050 | 0.367 | 0.053 | 0.003 | 0.068 |
| CNRM-CM6-1-HR ssp245 | 0.048 | 0.021 | 0.153 | 0.029 | 0.003 | 0.045 |

| | | | | | | |
|---|---|---|---|---|---|---|
| CNRM-CM6-1-HR ssp370 | 0.111 | 0.039 | 0.306 | 0.051 | 0.006 | 0.080 |
| CNRM-CM6-1-HR ssp585 | 0.127 | 0.053 | 0.392 | 0.060 | 0.007 | 0.094 |
| CNRM-ESM2-1 ssp126 | 0.139 | 0.064 | 0.460 | 0.061 | 0.004 | 0.082 |
| CNRM-ESM2-1 ssp245 | 0.086 | 0.060 | 0.384 | 0.025 | 0.005 | 0.049 |
| CNRM-ESM2-1 ssp370 | 0.093 | 0.055 | 0.368 | 0.072 | 0.004 | 0.093 |
| CNRM-ESM2-1 ssp585 | 0.064 | 0.029 | 0.208 | 0.063 | 0.005 | 0.087 |
| CanESM5 ssp126 | 0.152 | 0.064 | 0.471 | 0.087 | 0.006 | 0.117 |
| CanESM5 ssp245 | 0.027 | 0.029 | 0.170 | 0.026 | 0.003 | 0.043 |
| CanESM5 ssp370 | 0.102 | 0.076 | 0.480 | 0.099 | 0.003 | 0.115 |
| CanESM5 ssp585 | 0.104 | 0.039 | 0.297 | 0.120 | 0.004 | 0.142 |

**Table B 4:** *METEOR performance for all models (E–I) evaluated using the Climate bench evaluation metrics. Note that comparison is to single ensemble members, so variability driven errors are included.*

| | $NRMSE_s$ tas | $NRMSE_g$ tas | $NRMSE_t$ tas | $NRMSE_s$ pr | $NRMSE_g$ pr | $NRMSE_t$ pr |
|---|---|---|---|---|---|---|
| EC-Earth3-Veg ssp126 | 0.997 | 0.992 | 5.958 | 0.066 | 0.009 | 0.109 |
| EC-Earth3-Veg ssp245 | 0.994 | 0.989 | 5.938 | 0.098 | 0.006 | 0.131 |
| EC-Earth3-Veg ssp370 | 0.987 | 0.986 | 5.917 | 0.080 | 0.006 | 0.111 |
| EC-Earth3-Veg ssp585 | 0.986 | 0.982 | 5.894 | 0.085 | 0.009 | 0.131 |
| FGOALS-f3-L ssp126 | 0.164 | 0.067 | 0.499 | 0.084 | 0.009 | 0.131 |
| FGOALS-f3-L ssp245 | 0.034 | 0.061 | 0.339 | 0.022 | 0.009 | 0.068 |
| FGOALS-f3-L ssp370 | 0.079 | 0.048 | 0.319 | 0.087 | 0.008 | 0.128 |
| FGOALS-f3-L ssp585 | 0.111 | 0.064 | 0.433 | 0.085 | 0.014 | 0.153 |
| GFDL-ESM4 ssp126 | 0.118 | 0.071 | 0.471 | 0.070 | 0.006 | 0.098 |
| GFDL-ESM4 ssp245 | 0.044 | 0.052 | 0.303 | 0.036 | 0.006 | 0.064 |
| GFDL-ESM4 ssp370 | 0.086 | 0.049 | 0.331 | 0.065 | 0.008 | 0.105 |
| GFDL-ESM4 ssp585 | 0.057 | 0.032 | 0.218 | 0.074 | 0.008 | 0.113 |
| GISS-E2-1-H ssp126 | 0.180 | 0.058 | 0.472 | 0.081 | 0.011 | 0.138 |

Funded by
the European Union

| | | | | | | |
|---|---|---|---|---|---|---|
| GISS-E2-1-H ssp245 | 0.067 | 0.048 | 0.308 | 0.045 | 0.003 | 0.061 |
| GISS-E2-1-H ssp370 | 0.107 | 0.050 | 0.357 | 0.111 | 0.021 | 0.215 |
| GISS-E2-1-H ssp585 | 0.102 | 0.066 | 0.430 | 0.115 | 0.009 | 0.158 |
| INM-CM4-8 ssp126 | 0.208 | 0.099 | 0.701 | 0.083 | 0.005 | 0.106 |
| INM-CM4-8 ssp245 | 0.040 | 0.053 | 0.303 | 0.024 | 0.004 | 0.045 |
| INM-CM4-8 ssp370 | 0.095 | 0.050 | 0.344 | 0.073 | 0.010 | 0.123 |
| INM-CM4-8 ssp585 | 0.114 | 0.058 | 0.407 | 0.076 | 0.009 | 0.121 |
| INM-CM5-0 ssp126 | 0.995 | 0.994 | 5.966 | 1.199 | 0.967 | 6.035 |
| INM-CM5-0 ssp245 | 0.993 | 0.992 | 5.952 | 1.184 | 0.957 | 5.971 |
| INM-CM5-0 ssp370 | 0.990 | 0.989 | 5.935 | 1.172 | 0.946 | 5.904 |
| INM-CM5-0 ssp585 | 0.988 | 0.987 | 5.924 | 1.164 | 0.940 | 5.863 |

**Table B 5:** *METEOR performance for all models (K–Z) evaluated using the Climate bench evaluation metrics. Note that comparison is to single ensemble members, so variability driven errors are included.*

| | $NRMSE_s$ tas | $NRMSE_g$ tas | $NRMSE_t$ tas | $NRMSE_s$ pr | $NRMSE_g$ pr | $NRMSE_t$ pr |
|---|---|---|---|---|---|---|
| KACE-1-0-G ssp126 | 0.283 | 0.069 | 0.628 | 0.067 | 0.003 | 0.084 |
| KACE-1-0-G ssp245 | 0.061 | 0.045 | 0.285 | 0.029 | 0.004 | 0.051 |
| KACE-1-0-G ssp370 | 0.114 | 0.056 | 0.396 | 0.074 | 0.010 | 0.126 |
| KACE-1-0-G ssp585 | 0.114 | 0.056 | 0.393 | 0.078 | 0.007 | 0.114 |
| MCM-UA-1-0 ssp126 | 0.115 | 0.072 | 0.475 | 0.064 | 0.004 | 0.083 |
| MCM-UA-1-0 ssp245 | 0.059 | 0.042 | 0.270 | 0.028 | 0.004 | 0.050 |
| MCM-UA-1-0 ssp370 | 0.114 | 0.051 | 0.369 | 0.067 | 0.006 | 0.096 |
| MCM-UA-1-0 ssp585 | 0.118 | 0.068 | 0.457 | 0.083 | 0.009 | 0.126 |
| MIROC-ES2L ssp126 | 0.158 | 0.070 | 0.505 | 0.057 | 0.005 | 0.084 |
| MIROC-ES2L ssp245 | 0.062 | 0.016 | 0.142 | 0.022 | 0.002 | 0.034 |
| MIROC-ES2L ssp370 | 0.101 | 0.069 | 0.444 | 0.056 | 0.004 | 0.075 |
| MIROC-ES2L ssp585 | 0.096 | 0.061 | 0.403 | 0.068 | 0.010 | 0.120 |

| | | | | | | |
|---|---|---|---|---|---|---|
| MPI-ESM1-2-HR ssp126 | 0.181 | 0.090 | 0.633 | 0.064 | 0.003 | 0.079 |
| MPI-ESM1-2-HR ssp245 | 0.041 | 0.030 | 0.189 | 0.022 | 0.003 | 0.037 |
| MPI-ESM1-2-HR ssp370 | 0.093 | 0.050 | 0.344 | 0.053 | 0.004 | 0.074 |
| MPI-ESM1-2-HR ssp585 | 0.075 | 0.023 | 0.188 | 0.060 | 0.005 | 0.085 |
| NorESM2-MM ssp126 | 0.996 | 0.995 | 5.970 | 1.252 | 0.985 | 6.177 |
| NorESM2-MM ssp245 | 0.993 | 0.992 | 5.953 | 1.240 | 0.976 | 6.121 |
| NorESM2-MM ssp370 | 0.990 | 0.989 | 5.934 | 1.229 | 0.969 | 6.075 |
| NorESM2-MM ssp585 | 0.987 | 0.986 | 5.916 | 1.210 | 0.955 | 5.984 |
| UKESM1-0-LL ssp126 | 0.136 | 0.051 | 0.394 | 0.070 | 0.004 | 0.091 |
| UKESM1-0-LL ssp245 | 0.061 | 0.051 | 0.314 | 0.046 | 0.003 | 0.063 |
| UKESM1-0-LL ssp370 | 0.076 | 0.028 | 0.217 | 0.102 | 0.005 | 0.125 |
| UKESM1-0-LL ssp585 | 0.075 | 0.012 | 0.137 | 0.102 | 0.006 | 0.133 |
| | | | | | | |

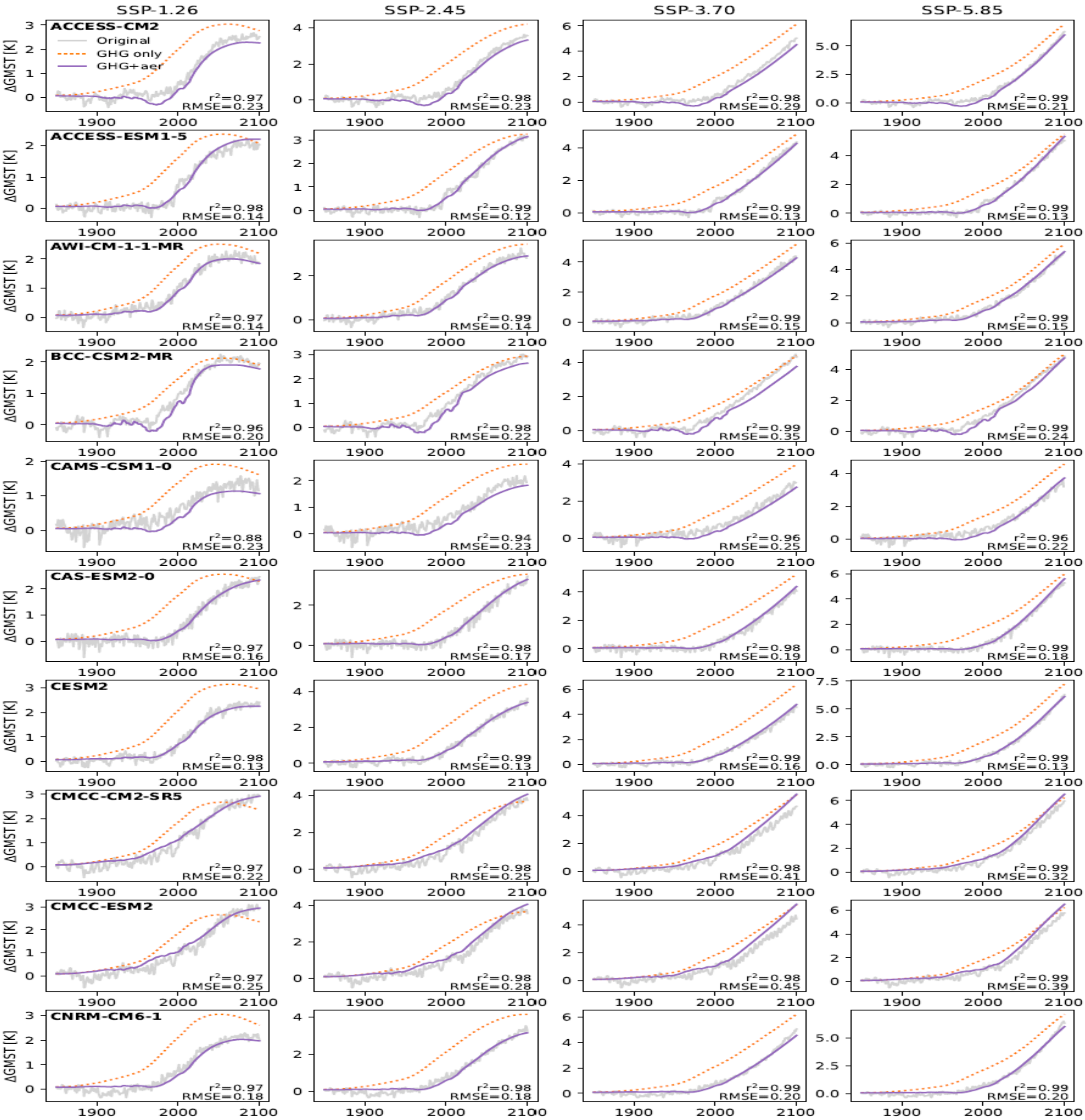



**Figure B 1:** *GMST original data (grey), fit reconstruction for only GHG (orange) and full reconstruction (purple) for single models, one model per row. The results were obtained using SSP2-4.5 to fit the aerosol sulfate residual. From left to right the columns show SSP1-2.6 (out-of-sample), SSP2-4.5 (in-sample), SSP3-7.0 (out-of-sample) and SSP5-8.5 (out-of-sample) fits.*

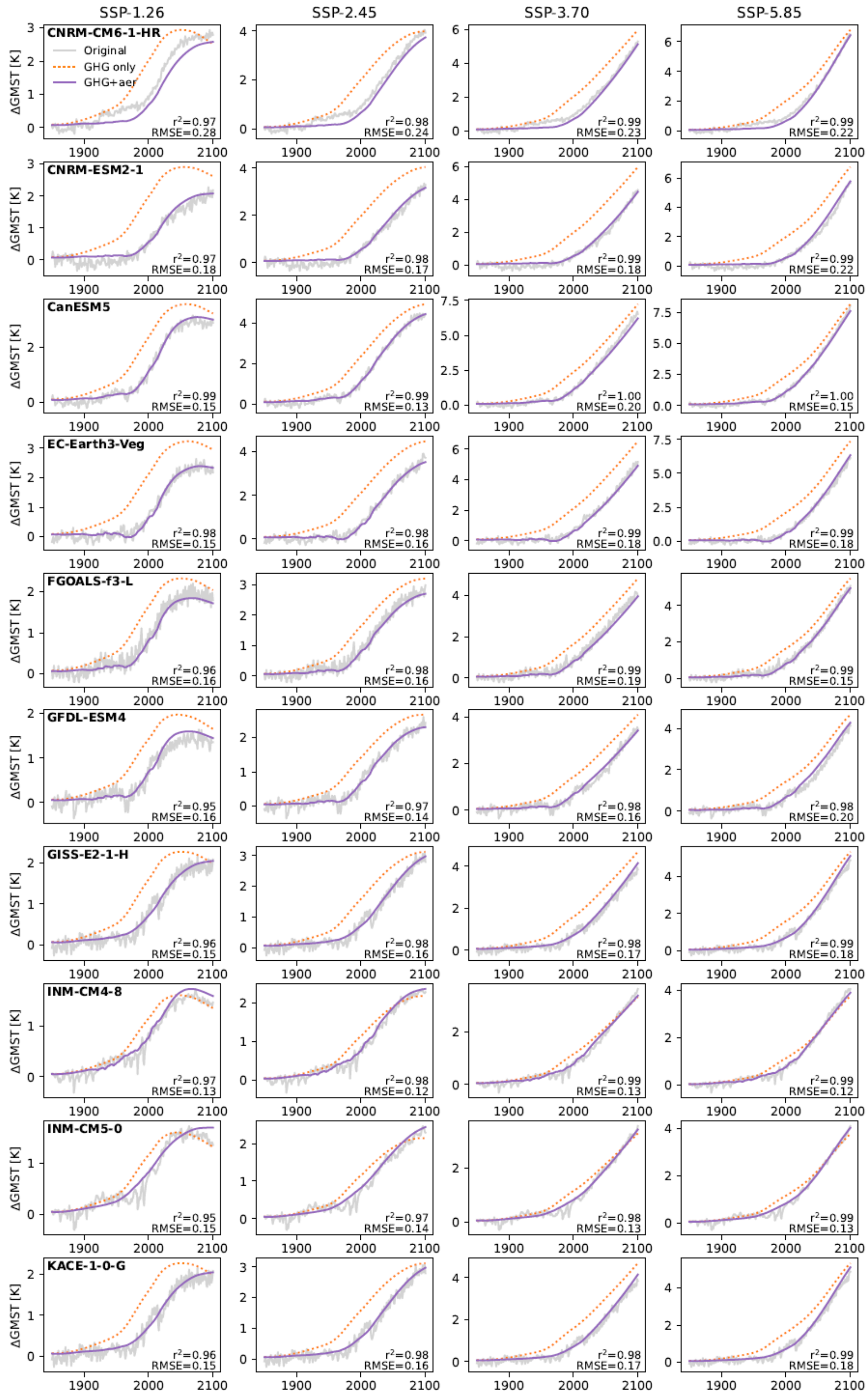



**Figure B 2:** *GMST original data (grey), fit reconstruction for only GHG (orange) and full reconstruction (purple) for single models, one model per row. The results were obtained using SSP2-4.5 to fit the aerosol sulfate residual. From left to right the columns show SSP1-2.6 (out-of-sample), SSP2-4.5 (in-sample), SSP3-7.0 (out-of-sample) and SSP5-8.5 (out-of-sample) fits.*

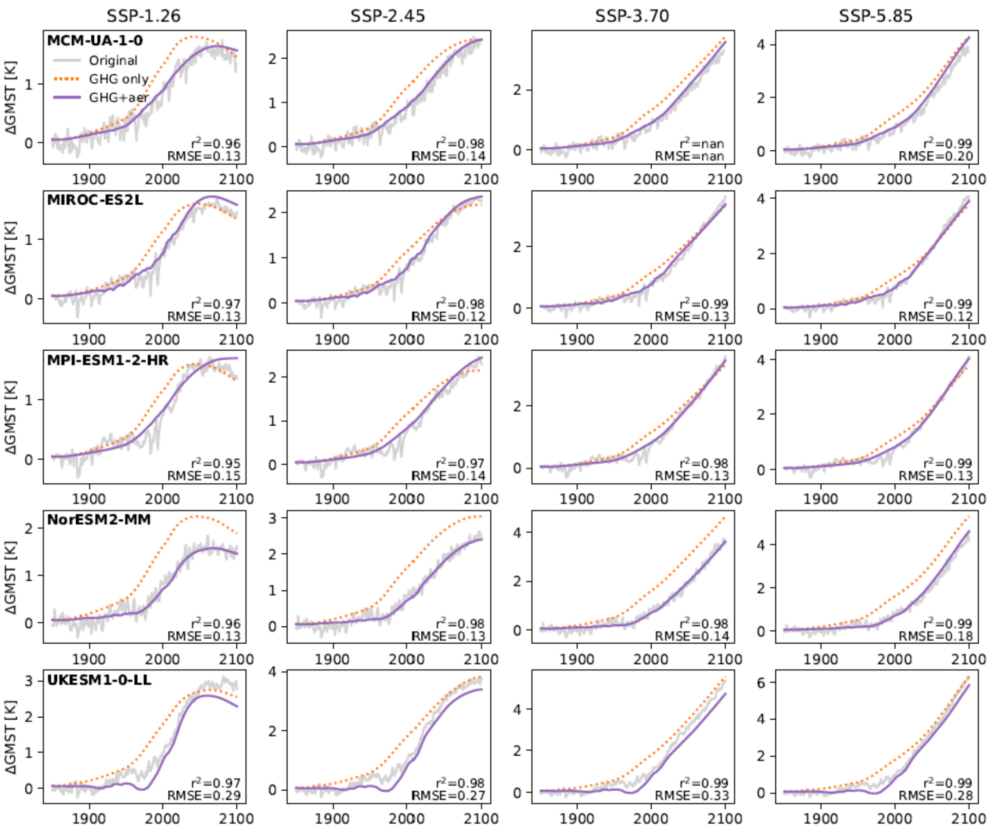
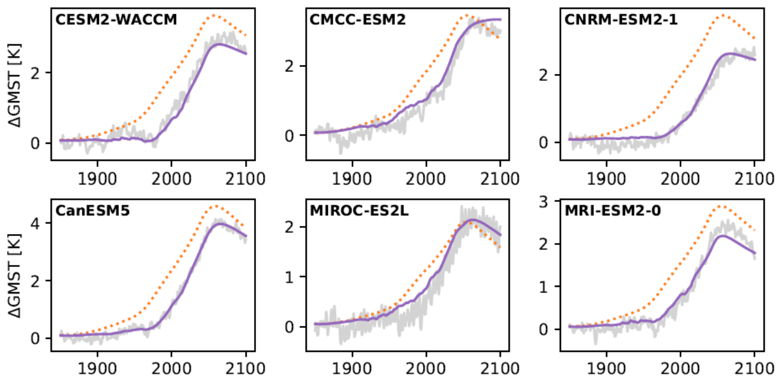



**Figure B 3:** *GMST original data (grey), fit reconstruction for only GHG (orange) and full reconstruction (purple) for single models, one model per row. The results were obtained using SSP2-4.5 to fit the aerosol sulfate residual. From left to right the columns show SSP1-2.6 (out-of-sample), SSP2-4.5 (in-sample), SSP3-7.0 (out-of-sample) and SSP5-8.5 (out-of-sample) fits*




.

**Figure B 4:** *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 shows results for a single model.*

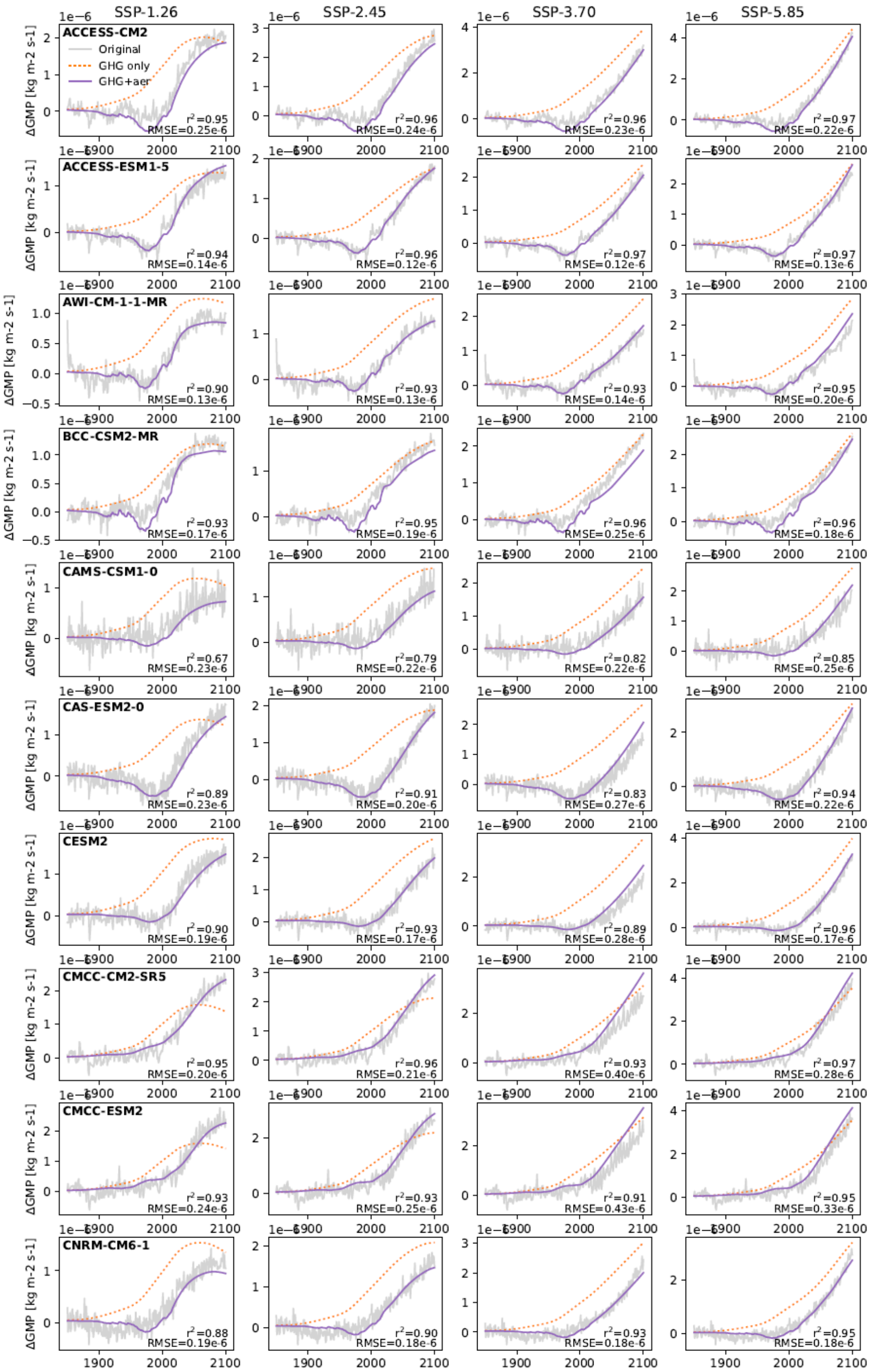



**Figure B 5:** *GMP original data (grey), fit reconstruction for only GHG (orange) and full reconstruction (purple) for single models, one model per row. The results were obtained using SSP2-4.5 to fit the aerosol sulfate residual. From left to right the columns show SSP1-2.6 (out-of-sample), SSP2-4.5 (in-sample), SSP3-7.0 (out-of-sample) and SSP5 8.5 (out-f-sample) fits.*

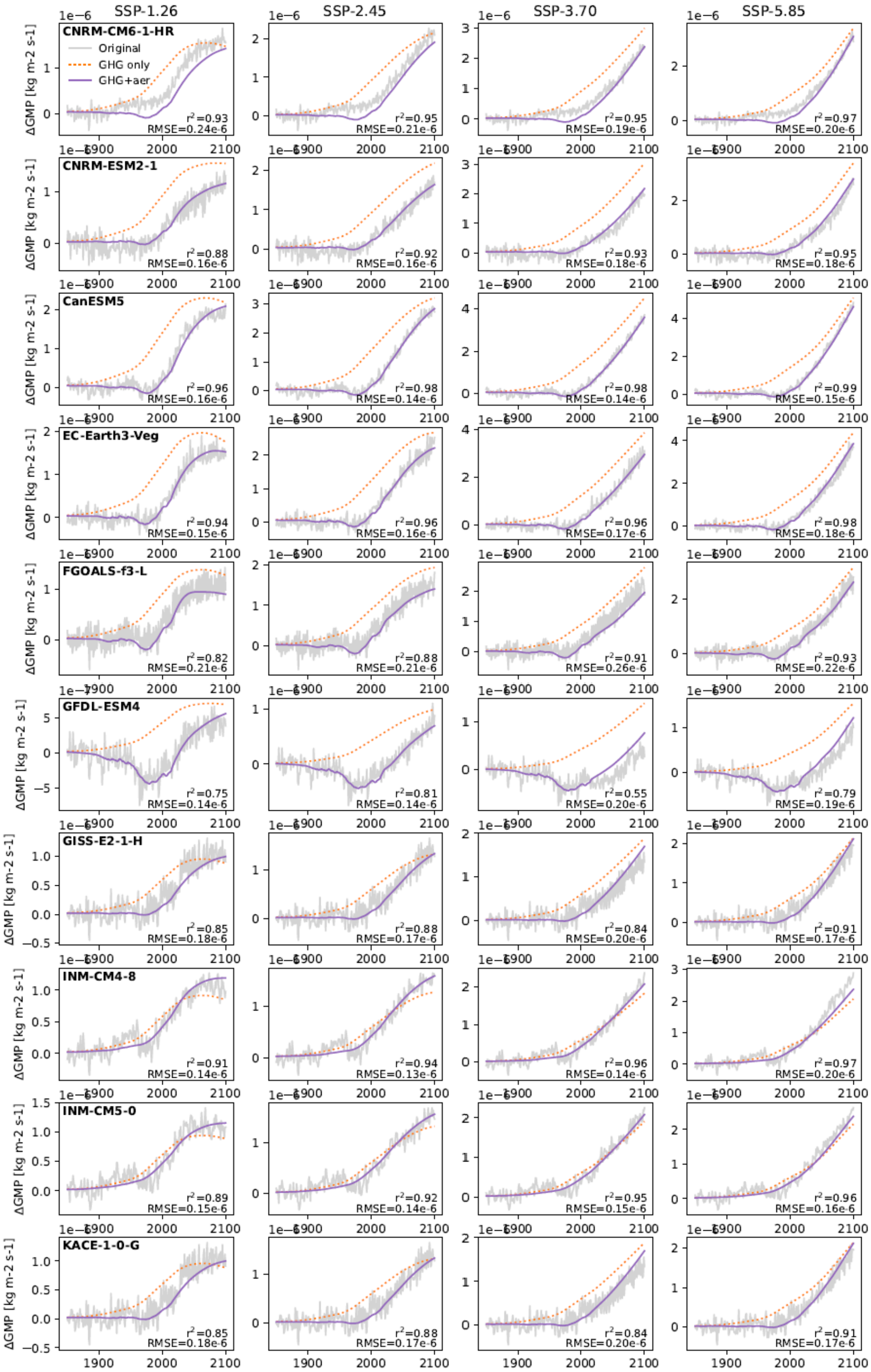



**Figure B 6:** *GMP original data (grey), fit reconstruction for only GHG (orange) and full reconstruction (purple) for single models, one model per row. The results were obtained using SSP2-4.5 to fit the aerosol sulfate residual. From left to right the columns show SSP1-2.6 (out-of-sample), SSP2-4.5 (in-sample), SSP3-7.0 (out-of-sample) and SSP5-8.5 (out-of-sample) fits.*

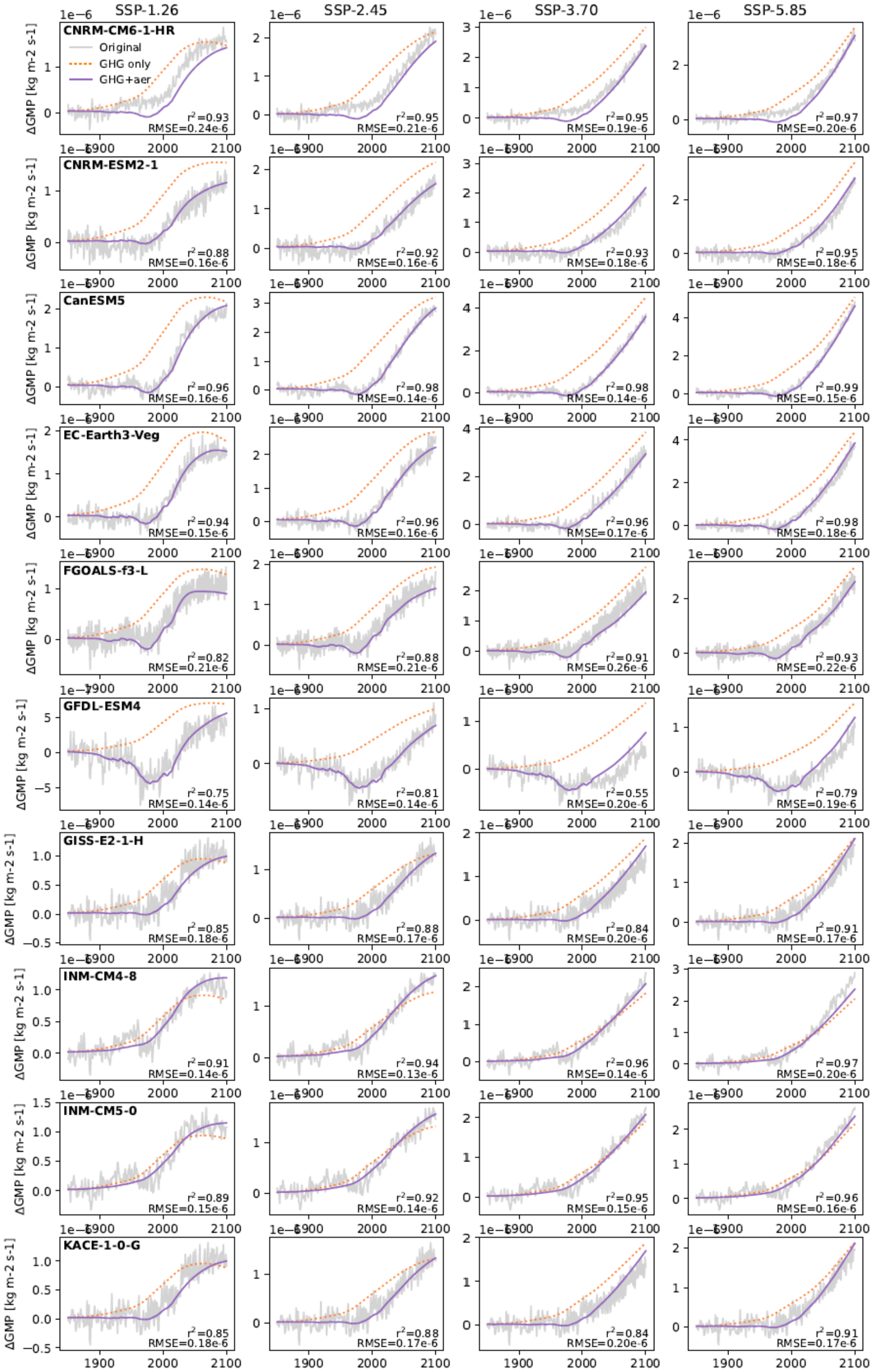



**Figure B 7:** *GMP original data (grey), fit reconstruction for only GHG (orange) and full reconstruction (purple) for single models, one model per row. The results were obtained using SSP2-4.5 to fit the aerosol sulfate residual. From left to right the columns show SSP1-2.6 (out-of-sample), SSP2-4.5 (in-sample), SSP3-7.0 (out-of-sample) and SSP5-8.5 (out-of-sample) fits.*

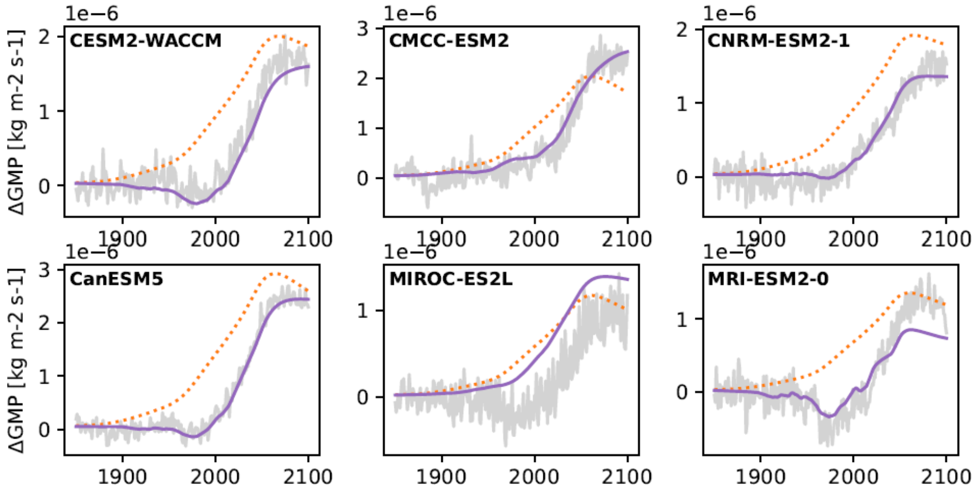



**Figure B 8** *GMP original data (grey), fit reconstruction for only GHG (orange) and full reconstruction (purple) for single models that had reasonable data quality for the overshoot scenario SSP5-3.4-over. The results were obtained using SSP2-4.5 to fit the aerosol sulfate residual. Each sub-plot shows results for a single model.*

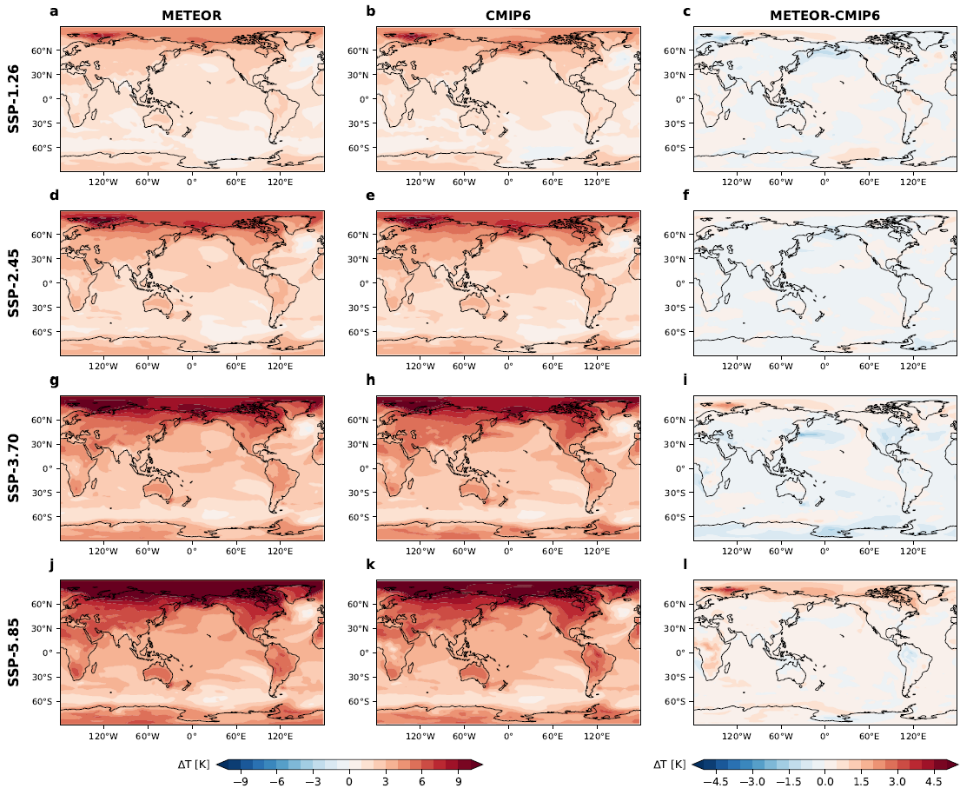


**Figure B 9:** *NorESM2-MM temperature emulation comparison. NorESM2-MM was found to be one of the models for which METEOR has the best overall fit. Here we display the emulated patterns (a, d, g, j) compared to the CMIP6 data (b, e, h, k) and the difference between them (c, f, i, l) for the temperature emulation of the standard SSP scenarios at the end of the century; SSP1-2.6 (a, b, c), SSP2-4.5 (d, e, f), SSP3-7.0 (g, h, i) and SSP5-8.5 (j, k, l) for this model.*

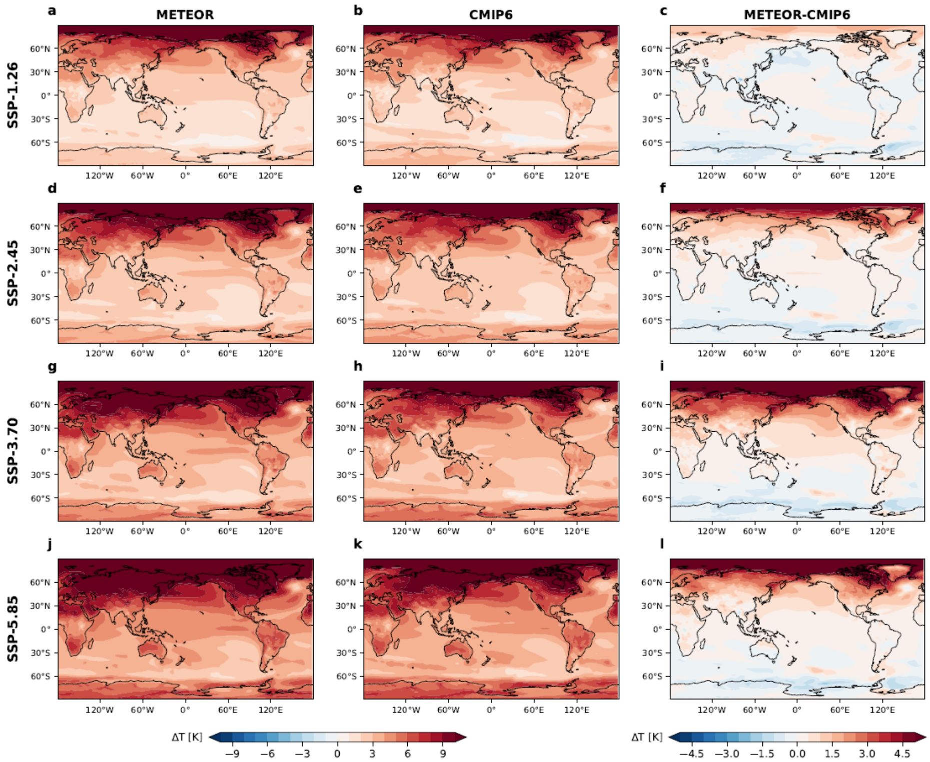



**Figure B 10:** *CMCC-ESM2 temperature emulation comparison. CMCC-ESM2 was found to be one of the models for which METEOR has the worst overall fit. Here we display the emulation patterns (a, d, g, j) compared to the CMIP6 data (b, e, h, k) and the difference between them (c, f, i, l) for the temperature emulation of the standard SSP scenarios at the end of the century; SSP1-2.6 (a, b, c), SSP2-4.5 (d, e, f), SSP3-7.0 (g, h, i) and SSP5-8.5 (j, k, l) for this model.*

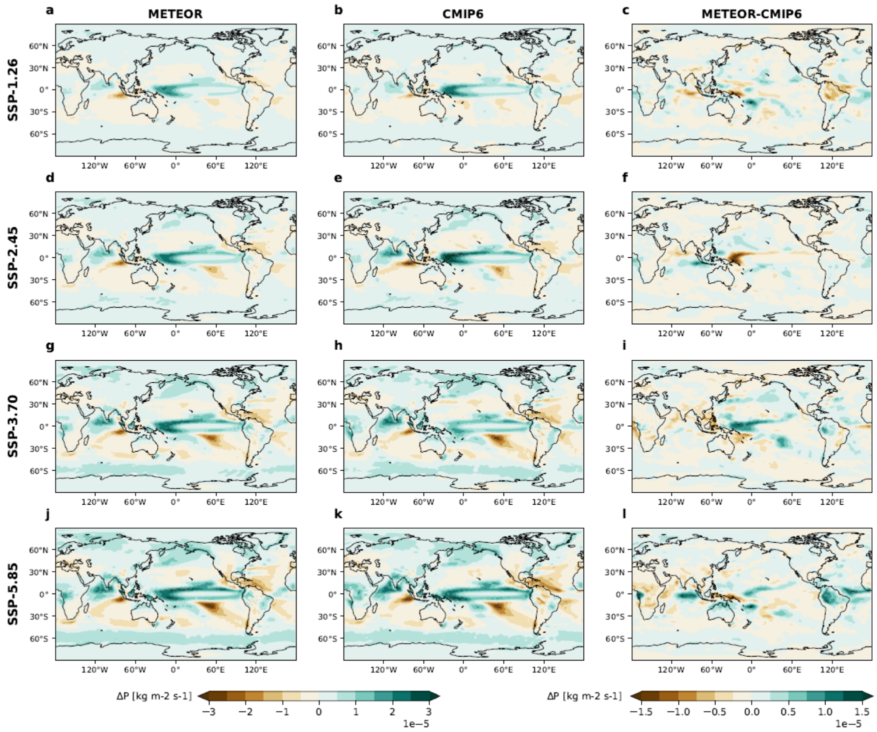



**Figure B 11:** *NorESM2-MM precipitation emulation comparison. NorESM2-MM was found to be one of the models for which METEOR has the best overall fit. Here we display the emulation patterns (a, d, g, j) compared to the CMIP6 data (b, e, h, k) and the difference between them (c, f, i, l) for the precipitation emulation of the standard SSP scenarios at the end of the century; SSP1-2.6 (a, b, c), SSP2-4.5 (d, e, f), SSP3-7.0 (g, h, i) and SSP5-8.5 (j, k, l) for this model.*

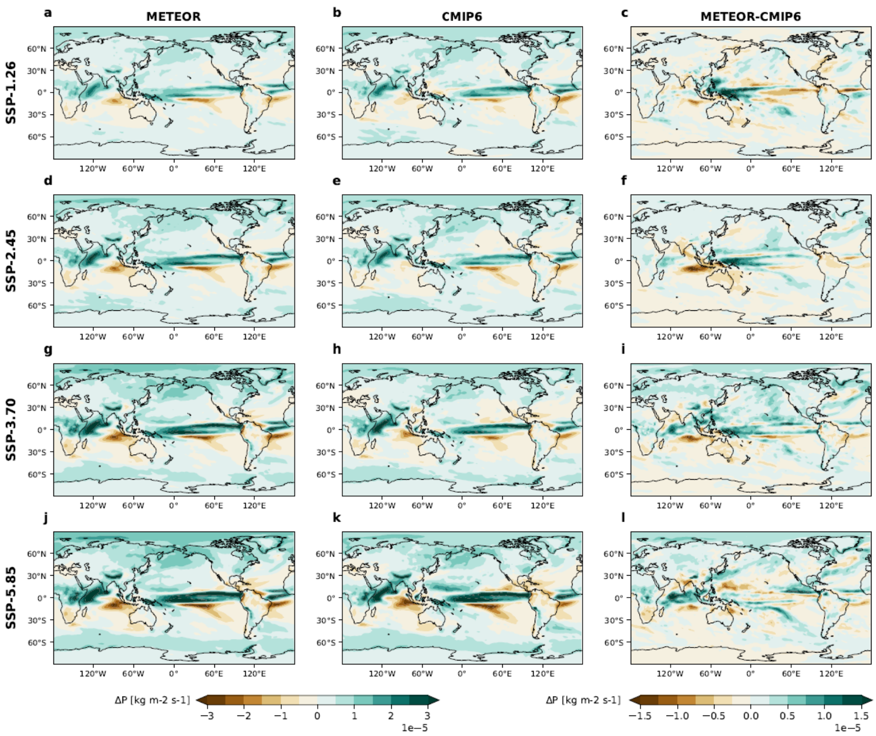



**Figure B 12:** *CMCC-ESM2 precipitation emulation comparison. CMCC-ESM2 was found to be one of the models for which METEOR has the worst overall fit. Here we display the emulation patterns (a, d, g, j) compared to the CMIP6 data (b, e, h, k) and the* *difference between them (c, f, i, l) for the precipitation emulation of the standard SSP scenarios at the end of the century; SSP1-2.6 (a, b, c), SSP2-4.5 (d, e, f), SSP3-7.0 (g, h, i) and SSP5-8.5 (j, k, l) for this model.*

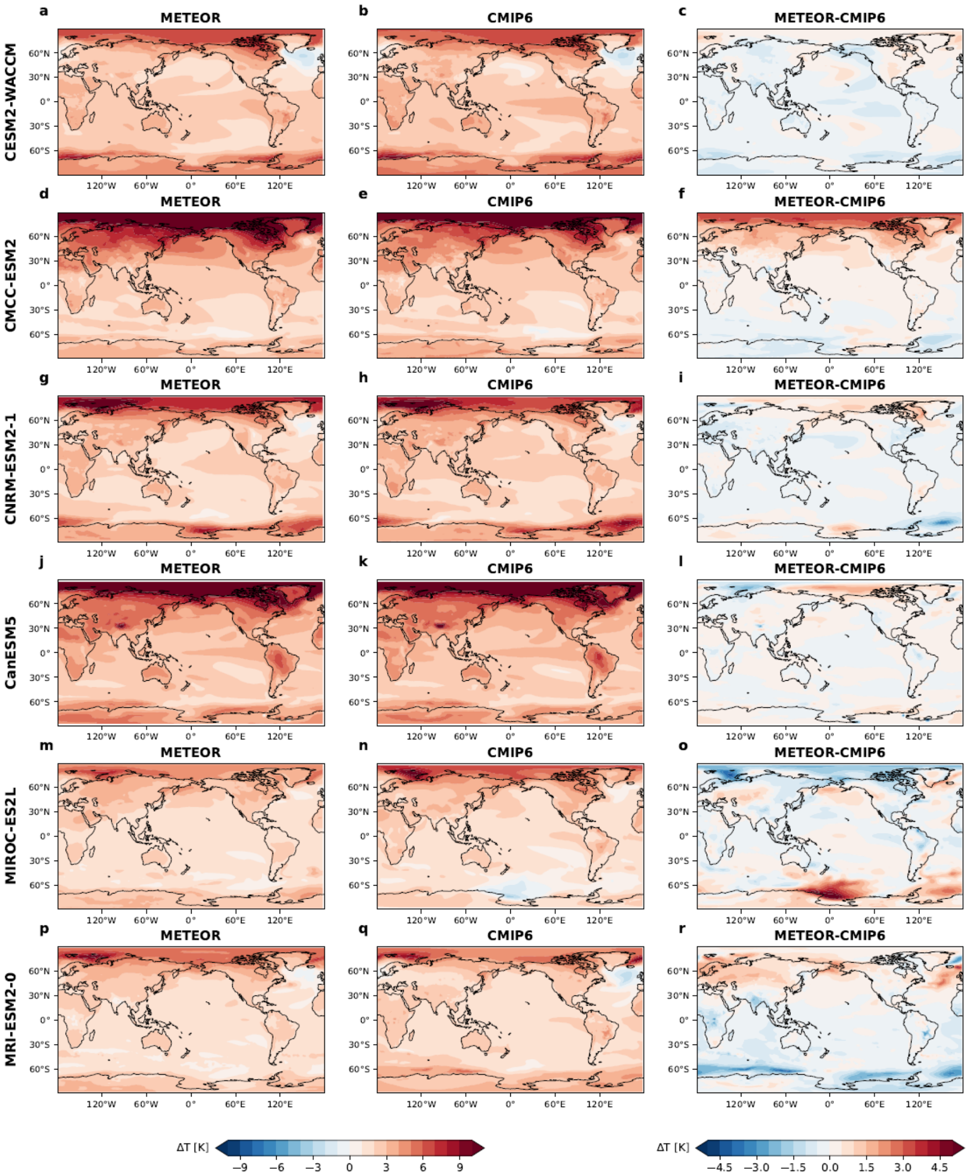



**Figure B 13:** *SSP5-3.4-over temperature emulation comparison. SSP5-3.4-over spatial temperature emulation patterns (a, d, g, j, m, p) compared to the CMIP6 data (b, e, h, k, n, q) and the difference between them (c, f, i, l, o, r, u) at the end of the century for each of the models that had sufficient SSP5-3.4-over data available.*

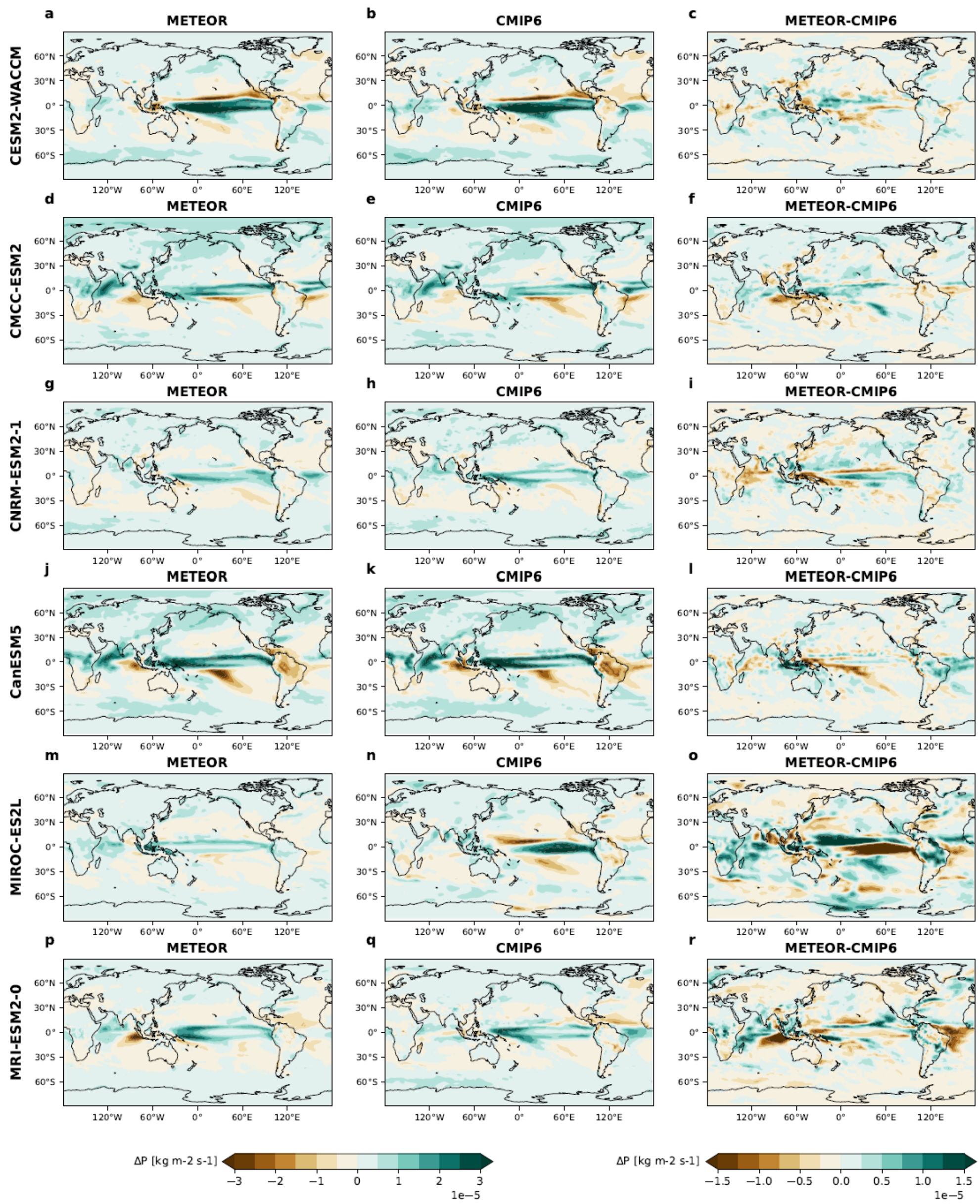



**Figure B 14:** *SSP5-3.4-over precipitation emulation comparison. SSP5-3.4-over spatial precipitation emulation patterns (a, d, g, j, m, p) compared to the CMIP6 data (b, e, h, k, n, q) and the difference between them (c, f, i, l, o, r, u) at the end of the century for each of the models that had sufficient SSP5-3.4-over data available.*