https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

METEORv1.6: Spatial climate variability and integrated impact emulation

Benjamin M. Sanderson[1], Marit Sandstad[1], Maura Dewey[1], Norman Julius Steinert[1], and Johannes L. Fjeldså[2]

[1]CICERO Center for International Climate Research, Oslo 0349, Norway
[2]Section for Meteorology and Oceanography, University of Oslo, Norway

Correspondence: Benjamin M. Sanderson (benjamin.sanderson@cicero.oslo.no)

Abstract. Climate impact assessment increasingly requires spatially explicit projections with realistic temporal variability at sub-annual resolution. METEORv1.6 extends the established METEOR spatial multi-timescale, multi-forcer climate emulation framework with two major capabilities: (1) a monthly climate variability model that generates realistic sub-annual climate sequences with seasonal cycles and inter-annual variability, enabling the generation of ensemble projections and (2) a modular impact assessment framework that translates climate projections into impact metrics. The monthly climate model represents seasonal harmonics and noise, while preserving covariance structures from the source model. Seasonal cycles are represented through harmonic analysis with temperature-dependent parameterization, enabling non-stationary simulation of seasonal timing shifts under warming. Principal Component Analysis is used to decompose monthly anomalies into spatial modes, then their temporal evolution and climate variability is modeled using Vector Autoregressive with eXogenous variables (VARX) processes. The impact assessment framework provides a standardized interface for ensemble processing and uncertainty quantification through a modular system of impact calculators. The initial case implementation includes heating and cooling degree days calculations which are key drivers in estimating energy sector demand, demonstrating ensemble-based uncertainty propagation from climate projections to impact metrics. Validation against CMIP6 data demonstrates that METEORv1.6 accurately reproduces statistical properties of monthly climate variability for a range of future scenarios when trained on a single scenario from an Earth System Model (together with a CO2 quadrupling idealized experiment). The integrated impact framework enables rapid generation of probabilistic climate risk assessments suitable for sectoral applications, bridging the gap between global climate projections and local decision-making needs. The open-source implementation supports broad adoption and continued expansion to additional impact domains.

Keywords: climate emulation, pattern scaling, monthly variability, climate impacts, uncertainty quantification, VARX modeling, degree days, CMIP6

1 Introduction

Spatially resolved climate projections with realistic temporal variability are essential for robust climate impact assessment and adaptation planning (IPCC, 2021, 2022). While Earth System Models (ESMs) provide comprehensive climate simulations, their

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

computational requirements limit the range of scenarios and ensemble sizes that can be explored (Eyring et al., 2016; Nicholls et al., 2020). This is particularly problematic for sectoral impact assessments, which increasingly require large ensembles of climate projections to characterize uncertainty ranges and support probabilistic risk analysis (Tebaldi et al., 2021).

Pattern scaling has emerged as a powerful approach for generating spatially explicit climate projections with dramatically reduced computational requirements (Santer et al., 1990; Mitchell, 2003). By decomposing climate responses into spatial patterns that scale with global temperature trajectories, pattern scaling enables rapid generation of climate projections for arbitrary emission scenarios (Beusch et al., 2020; Tebaldi et al., 2021). However, traditional pattern scaling approaches face limitations when applied to impact assessment applications, particularly their lack of representation of hysteresis effects under climate reversibility experiments (Schleussner et al., 2024; Giani et al., 2025).

The METEOR climate emulation framework (Sandstad et al., 2025) addresses key limitations of conventional pattern scaling by representing multi-timescale spatial climate responses to multiple climate forcers using impulse response assumptions. Unlike traditional approaches that assume forcing-independent patterns, METEOR captures hysteresis effects and time-evolving responses essential for scenarios involving overshoots or diverse mixes of short- and long-lived climate forcers (Herger et al., 2015; Good et al., 2015; Pfleiderer et al., 2024). The framework has been validated against CMIP6 output, demonstrating accurate reproduction of spatially resolved temperature and precipitation responses across diverse emission scenarios (Sandstad et al., 2025).

Despite these advances, two critical gaps remain for sectoral impact applications. First, the annual resolution of METEORv1.0 limited its applicability to impact assessments requiring sub-annual climate information, such as energy demand modeling, agricultural applications, impacts on labour productivity and extreme event analysis (Sillmann et al., 2013; Zhao et al., 2017; Zeng et al., 2025; Dasgupta et al., 2021). Many climate impacts exhibit strong seasonal dependencies that cannot be adequately captured from annual mean projections alone. Second, the framework lacked integrated capabilities for translating climate projections into sectoral impact metrics, requiring users to implement custom post-processing workflows that may not preserve uncertainty characteristics or enable consistent ensemble analysis.

Recent developments in climate emulation have begun addressing these limitations through diverse approaches. The MESMER framework has been extended to monthly resolution (Nath et al., 2022; Schöngart et al., 2024), demonstrating the feasibility of sub-annual emulation while maintaining computational efficiency. Similarly, MESMER-X has shown capabilities for extreme event emulation (Quilcaille et al., 2022), though with simplified treatment of temporal dependencies. The RIME-X approach (Schwind et al., 2026) takes a different approach, constructing probabilistic relationships between global mean temperature and climate impacts trained on existing databases. Another approach, PRIME (Mathison et al., 2025) couples a pattern scaled driver climate to a process-resolving land surface model. However, these approaches all assume, in different ways, linear scaling relationships and may not adequately represent the multi-timescale and forcing-dependent responses that METEOR captures for annual projections.

Here we present METEORv1.6, which extends the established METEOR framework with monthly climate variability generation and integrated impact assessment capabilities. Our approach builds directly on the multi-timescale pattern scaling foundation of METEOR while adding: (1) a monthly climate and internal-variability modeling system that preserves the

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

framework's ability to represent forcing-dependent and time-evolving responses at sub-annual resolution, and (2) a modular impact framework that standardizes uncertainty quantification and ensemble analysis for sectoral applications.

The monthly climate system (hereafter METEOR-NOISE) employs a two-stage approach combining Principal Component Analysis (PCA) with Vector Autoregressive with eXogenous variables (VARX) modeling (Hamilton, 2020; Sims, 1980; Lam et al., 2010). This methodology decomposes monthly climate anomalies into dominant spatial modes, then models their temporal evolution while accounting for external forcing influences and internal variability. This approach differs from MESMER-X (Quilcaille et al., 2022), which applies first-order autoregressive processes independently at each grid point with spatially correlated innovations. Our VARX framework captures teleconnections through explicit cross-mode dependencies, allowing spatial patterns to influence each other's temporal evolution while incorporating external forcing as a predictor variable. This provides a more compact representation of spatiotemporal dynamics and enables the generation of realistic climate variability that preserves both temporal persistence and spatial covariance structures. Seasonal cycles are explicitly represented through harmonic analysis with temperature-dependent parameterization, enabling realistic simulation of seasonal timing shifts and amplitude changes under warming scenarios. Internal variability, including its spatial covariance structure, is learned from the climate anomalies and enables the generation of realistic ensembles of climate projections.

The impact assessment framework provides a standardized interface for converting climate projections into sectoral metrics through an extensible system of impact calculators. The framework enables ensemble-based uncertainty quantification that propagates climate uncertainties through to impact projections. The initial implementation focuses on degree days calculations for energy sector applications, demonstrating the framework's capability to bridge climate science and sectoral decision-making needs.

This paper describes the technical methodology, validation, and applications of METEORv1.6, with particular emphasis on how the extensions preserve the multi-timescale and forcing-dependent capabilities that distinguish METEOR from conventional pattern scaling approaches. We demonstrate that the enhanced framework maintains computational efficiency and minimal training data needs while enabling rapid climate impact applications that require monthly resolution and large ensemble output.

2 Methodology

2.1 METEOR Core Framework Overview

The METEORv1.6 extensions build directly on the core METEOR pattern scaling framework (Sandstad et al., 2025), which represents multi-timescale spatial climate responses to multiple climate forcers using spatial impulse response assumptions which are well demonstrated in a global sense to represent climate response to emissions(Hasselmann et al., 1997). The core methodology decomposes climate responses into time-evolving spatial patterns that emerge on specific timescales. (Sandstad et al., 2025) assumes that climate responses can be represented as the sum of impulse response patterns, each with its own emergence timescale. For greenhouse gas forcing, METEOR uses outputs from the *abrupt4x-CO_2* experiment to derive time-evolving patterns that capture both fast and slow climate responses. For aerosol forcing, residual signals from a concatenated historical and target scenario simulation are used to estimate separate patterns and timescales. This multi-timescale approach

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

enables METEOR to simulate non-linear climate responses and hysteresis effects that are particularly important for scenarios involving overshoot pathways or varying mixes of short- and long-lived climate forcers.

The core pattern scaling formulation represents temperature change as:

$$\Delta T(x,y,t) = \sum_{f} \sum_{i=1}^{N_f} \alpha_{f,i}(t) \cdot P_{f,i}(x,y) \tag{1}$$

where f indexes different forcing agents (greenhouse gases, aerosols), $P_{f,i}(x,y)$ are the spatial patterns for forcer f and timescale i, and $\alpha_{f,i}(t)$ are time-varying coefficients that depend on the forcing trajectory. The coefficients are computed as convolutions of the forcing timeseries with exponential impulse response functions:

$$\alpha_{f,i}(t) = \int_0^t F_f(t') \cdot \frac{1}{\tau_{f,i}} \exp\left(-\frac{t-t'}{\tau_{f,i}}\right) dt' \tag{2}$$

where $F_f(t)$ is the forcing trajectory for agent f and $\tau_{f,i}$ is the characteristic timescale for pattern i. This formulation enables METEOR to represent delayed climate responses and hysteresis effects: patterns with longer timescales respond more slowly to forcing changes, while shorter timescales track forcing more closely. The METEORv1.6 extensions preserve this multi-forcer, multi-timescale structure while extending it to monthly resolution and adding integrated impact assessment capabilities. Henceforth in the present paper, we refer to the underlying climatological model as METEOR-CORE.

2.2 METEOR-NOISE: Monthly variability model

2.2.1 Overview of the Approach

We refer to the monthly variability component of METEORv1.6 as METEOR-NOISE, which encompasses both the seasonal cycle model and the stochastic noise model, and operates in two phases: training and generation. During training, we analyze historical and scenario simulations from a CMIP6 model to learn the structure of its monthly variability. During generation, we use this learned structure to create new ensemble members for arbitrary temperature trajectories.

The noise model is trained on the raw ESM output, not the residual from the annual METEOR base model. The seasonal cycle model captures the forced climate signal through temperature dependent terms, while the PCA-VARX framework is fitted to the spatial-temporal structure of the anomalies. This approach is chosen such that the noise model learns directly from ESM behavior rather than from anomalies that might contain artifacts from imperfect pattern scaling. Temperature anomalies are computed relative to a pre-industrial baseline derived from the source model's piControl experiment, ensuring consistency with METEOR's pattern scaling framework.

Although METEOR-NOISE is trained on the full ESM output (which includes aerosol-driven signals), this does not lead to double-counting when combined with METEOR-CORE. The seasonal regression absorbs the forced annual-mean and seasonal signals through its GMT-dependent terms, while the PCA-VARX framework captures only the residual anomalies. When

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

generating ensemble members for integration with METEOR-CORE, the noise field (Eq. 11) explicitly excludes the intercept and direct temperature response, contributing only seasonal modulation and stochastic variability. The forced response, including aerosol effects, is provided entirely by METEOR-CORE.

The training phase addresses three questions. First, how does the seasonal cycle change as the climate warms? Second, after removing this seasonal pattern, what spatial structures dominate the remaining monthly fluctuations? Third, how do these fluctuations evolve over time, exhibiting persistence and teleconnections between regions?

To answer the first question, we fit a harmonic regression model at each grid point, with terms that allow seasonal amplitude and phase to be modulated by global mean temperatures. This captures state-dependent seasonal features, such as if winter warming accelerates faster than summer warming. The residuals from this fit represent deviations from expected seasonal behavior given the current global temperature state.

For the second question, we apply PCA to these residuals. This identifies a modest number of spatial patterns (by default 40) that capture a large fraction of the variance. These patterns include known modes of variability like ENSO or the North Atlantic Oscillation. Crucially, this reduces the problem from tracking thousands of grid points to tracking just a few dozen pattern amplitudes.

The third question is addressed by the VARX model. This statistical framework describes how each pattern's amplitude this month depends on previous months' amplitudes across all patterns, plus global temperature as an exogenous forcing variable. The model naturally captures both persistence through autoregressive terms (patterns tend to stay in the same phase for several months) and teleconnections through both the PCA representation and the cross-mode terms (such that activity in one pattern influences others).

Once trained, the model enables generation of new ensemble members. Given METEOR-CORE output from a novel scenario (noting that valid METEOR-CORE output requires that the scenario's forcing characteristics remain within the range supported by the training data, as discussed in Sandstad et al. (2025)), we can simulate the pattern amplitudes forward in time using the autoregressive model with random seed inputs, then reconstruct full spatial fields, regional output or point output by combining the patterns or pattern projections onto the desired domain. Adding these noise realizations to METEOR's forced response predictions yields complete ensemble members with both the pattern-scaled climate signal and realistic monthly variability (Fig. 1)

3 Training: Learning the Structure of Variability

3.1 Preparing the Data

Training begins with monthly climate data from CMIP6, where by default we concatenate a model's historical simulation (1850–2014) with a mid-range future scenario SSP2-4.5 (2015–2100) (O'Neill et al., 2016), creating a continuous record of approximately 250 years from a single ensemble member. This extended record is used to characterize the statistical structure of internal variability (the dominant spatial patterns, persistence timescales, and teleconnections), though it represents just one

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

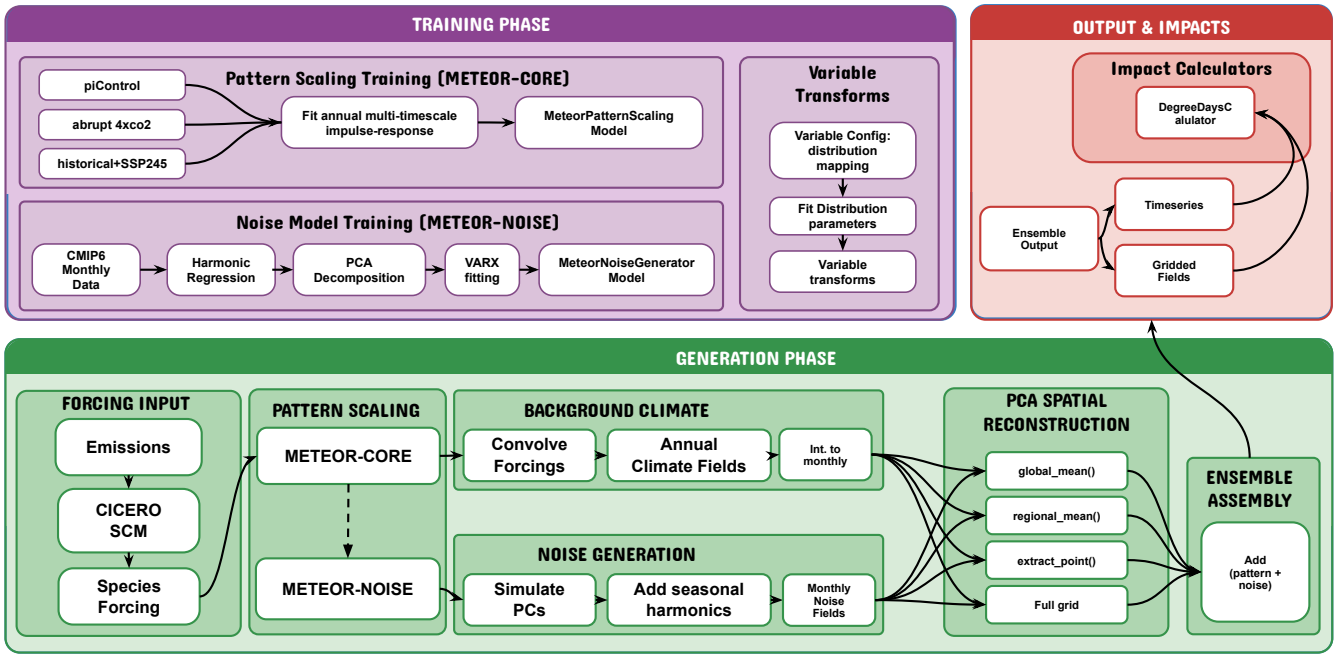

Figure 1. METEOR system workflow integrating pattern scaling, internal variability, and impacts assessment. Pattern scaling trains multi-timescale response patterns from CMIP6 annual data to predict forced climate response from emissions scenarios (purple box). Noise model training learns the structure of internal variability from CMIP6 monthly data, capturing seasonal cycles, spatial patterns (PCA), and temporal dynamics (VARX). Ensemble generation (green) combines the METEOR-predicted forced response with stochastic noise to produce large ensembles. Impacts module (red) calculates sector-relevant metrics such as heating degree days (HDD) and cooling degree days (CDD) from ensemble distributions, enabling uncertainty quantification for climate risk assessment.

realization of the climate trajectory (this keeps training data needs to a minimum and is practically found to allow defensible noise emulation, though future versions are expected to explore ensemble assimilation for noise training).

The global mean temperature (GMT) dependent seasonal model separates forced seasonality changes from internal variability, while PCA and VARX learn the structure of the residual fluctuations. We also subtract a pre-industrial baseline from the GMT using a pre-industrial control experiment, ensuring consistency with METEOR's pattern scaling framework.

Global mean temperature is computed from this data as an area-weighted spatial average:

$$T_{\text{glob}}(t) = \frac{\sum_{i=1}^{s} w_i \cdot X_i(t)}{\sum_{i=1}^{s} w_i}, \tag{3}$$

where $w_i = \cos(\text{lat}_i)$ are area weights for pixel i, with a total of s pixels. This global mean is then smoothed with a 5-year (60-month) running mean to remove high-frequency variability that would complicate seasonal cycle fitting while preserving the multi-decadal warming trend and any interannual to decadal variations that might modulate seasonality.

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

3.2 Modeling the Seasonal Cycle

The seasonal cycle can evolve as climate changes ("Harmonic Regression" in Fig. 1). For instance, sea ice loss causes Arctic regions to warm much faster in winter than summer, fundamentally altering their seasonal cycle (Sejas and Taylor, 2023). To capture such effects, we need a seasonal model that responds to warming.

We construct nine time-varying features at each monthly time step, where $T_{\text{glob}}(t)$ denotes the smoothed (5-year running mean) global mean temperature defined in Eq. (3):

$$\mathbf{X}_{\text{harm}}(t) = \begin{bmatrix} T_{\text{glob}}(t) \\ \cos\left(\frac{2\pi t}{12}\right) \\ \sin\left(\frac{2\pi t}{12}\right) \\ \cos\left(\frac{4\pi t}{12}\right) \\ \sin\left(\frac{2\pi t}{12}\right) \\ T_{\text{glob}}(t) \cdot \cos\left(\frac{2\pi t}{12}\right) \\ T_{\text{glob}}(t) \cdot \sin\left(\frac{2\pi t}{12}\right) \\ T_{\text{glob}}(t) \cdot \cos\left(\frac{4\pi t}{12}\right) \\ T_{\text{glob}}(t) \cdot \sin\left(\frac{4\pi t}{12}\right) \end{bmatrix}. \tag{4}$$

The first feature is the smoothed global temperature. The next four are standard harmonic functions representing annual and semi-annual cycles, which can reconstruct any seasonal pattern through their weighted combination. The final four features allow modulation with climate change, multiplying global temperature by the harmonic functions and allowing seasonal amplitude to grow or shrink with warming and seasonal phase to shift.

At each spatial location (x, y), we perform a linear regression with these nine features predicting the observed temperature or precipitation:

$$X(t,x,y) = \beta_0(x,y) + \sum_{j=1}^{9} \beta_j(x,y) \cdot \mathbf{X}_{\text{harm},j}(t) + \epsilon(t,x,y), \tag{5}$$

where β_0 is the intercept, β_j are the regression coefficients for predictor $j \in [1, 9]$, and ϵ represents the residual. This yields nine coefficients per location plus an intercept. The fitted values represent the expected seasonal pattern at each point in time, accounting for warming. The residuals—the difference between ESM training values and fitted values—are the anomalies:

$$A(t,x,y) = \epsilon(t,x,y) = X(t,x,y) - \left[\beta_0(x,y) + \sum_{j=1}^{9} \beta_j(x,y) \cdot \mathbf{X}_{\text{harm},j}(t)\right]. \tag{6}$$

This approach is low-dimensional (nine parameters per location) yet flexible enough to capture spatial changes in seasonality. The regression is independent across locations, and the linearity assumption is reasonable for monthly mean data over the range of warming observed in typical scenarios.

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

3.3 Extracting Spatial Patterns

After removing the seasonal cycle, we're left with monthly anomalies at each grid point over the full timeseries ("PCA decomposition" in Fig. 1). These anomalies exhibit spatial structure associated with internal variability. Our goal is to identify these dominant spatial patterns and generalize them, allowing for seasonal influence, which we achieve using Principal Component Analysis (PCA) and Vector AutoRegression with eXogenous variables (VARX).

We reshape our anomaly data into a matrix $\mathbf{A}_{\text{flat}} \in \mathbb{R}^{T \times S}$, where T is the number of monthly time steps and S is the number of spatial locations. PCA then decomposes this matrix as:

$$\mathbf{A}_{\text{flat}} = \mathbf{Z}\mathbf{E}^T + \boldsymbol{\mu}, \tag{7}$$

where $\mathbf{Z}$ contains the Principal Component (PC) timeseries, $\mathbf{E}$ contains the Empirical Orthogonal Function (EOF) spatial patterns, and $\boldsymbol{\mu}$ is the spatial mean (approximately zero after removing the seasonal cycle). The first PC and EOF capture the most variance, the second pair captures the most remaining variance orthogonal to the first, and so on. By default, we retain 40 PCs for both temperature and precipitation. For surface temperature, the seasonal model captures approximately 91% of total variance, and the 40 PCs explain 75-80% of the remaining anomaly variance, yielding a combined total of approximately 98% of variance explained. Precipitation exhibits greater stochastic variability: the seasonal model explains approximately 32% of variance, with PCs capturing 29-45% of anomaly variance, for a combined total of 51-64% (see Supplemental Table 1)

PCA representation of the anomalies has several advantages. First, it achieves massive dimensionality reduction: instead of modeling 10,000+ grid points, we model 40 timeseries. Second, the spatial orthogonality simplifies subsequent statistical modeling by ensuring that each PC can be modeled independently without concerns about spatial multicollinearity.

3.4 Modeling Monthly Temporal Dynamics

We now have 40 timeseries of PC amplitudes spanning several centuries of simulated climate ("VARX fitting" in Fig. 1). How do we characterize their temporal behavior? We need a model that captures three phenomena: persistence of anomalies, interactions between modes (e.g. tropical patterns influence extratropical circulation), and modulation by external factors (global warming might affect variability). VARX (Hamilton, 2020; Sims, 1980) is well-suited for this task. This approach models each month's PC value as a linear combination of all PCs in the previous two months, allowing for external variables and random terms:

$$\mathbf{Z}(t) = \mathbf{c} + \mathbf{A}_1\mathbf{Z}(t-1) + \mathbf{A}_2\mathbf{Z}(t-2) + \mathbf{B}\mathbf{X}_{\text{exog}}(t) + \boldsymbol{\varepsilon}(t), \tag{8}$$

where $\mathbf{Z}(t)$ is the vector of 40 PC values at time t, $\mathbf{c}$ is a constant vector, $\mathbf{A}_1$ and $\mathbf{A}_2$ are 40×40 coefficient matrices for lags 1 and 2, $\mathbf{B}$ contains coefficients for exogenous variables, and $\boldsymbol{\varepsilon}(t) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ is a multivariate Gaussian innovation with covariance matrix $\boldsymbol{\Sigma}$. The two-month lag allows the model to capture both immediate month-to-month persistence and longer memory.

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

The specific formulation we use includes three exogenous variables representing global temperature and a seasonal signal:

$$\mathbf{X}_{\text{exog}}(t) = \begin{bmatrix} T_{\text{glob}}(t) \\ \cos(2\pi t/12) \\ \sin(2\pi t/12) \end{bmatrix}. \tag{9}$$

A key feature of this framework is that the covariance matrix $\boldsymbol{\Sigma}$ of the innovation term $\boldsymbol{\varepsilon}(t)$ is learned directly from the ESM training data. This matrix encodes the ESM-specific structure of internal variability, including which modes co-vary and at what relative magnitudes. By sampling from this learned multivariate distribution, METEOR-NOISE generates ensembles whose internal variability is calibrated to the target ESM, rather than relying on generic or prescribed noise. This distinguishes METEOR from approaches that add independent noise at each grid point or use externally specified variability estimates.

Each PC's prediction then depends on lagged values of all 40 PCs and the global trend, not just itself. This allows for richer teleconnections: if tropical Pacific warming represented in one PC typically precedes North American circulation changes (represented in another PC) by one month, the model learns a non-zero coefficient connecting these PCs at lag 1. The model thus has thousands of coefficients—40 PCs times 2 lags times 40 PCs plus exogenous terms—estimated efficiently using standard least-squares regression.

3.5 Storing the Trained Model

Once training completes, we have three sets of parameters: seasonal regression coefficients (9 values per grid point), PCA components (40 spatial patterns plus their variances), and VARX parameters (coefficient and covariance matrices). Trained models can be optionally saved using Python's `pickle` format. This allows rapid loading for ensemble generation without retraining.

4 Generation: Creating New Ensemble Members

Given a trained noise model and an annual-resolution climatological trajectory from METEOR-CORE, we can generate ensemble members that combine forced response with seasonality and internal variability ("Generation Phase" in Fig. 1).

4.1 Simulating Principal Components

The first step is generating synthetic PC timeseries. We initialize the first two months (corresponding to the lag-2 VARX model) with zero PC values, representing climatological conditions. This choice ensures that ensemble members start from a neutral state and quickly diverge based on independent sampling from the fitted Gaussian noise model, rather than inheriting specific patterns from the end of the training period. For each subsequent month, we:

1. Construct exogenous variables (for temperature): the new trajectory's global temperature for this month and the month-of-year encoded as sine and cosine. For precipitation (and for any other variable specified without override), the model

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

will use a pure VAR model without exogenous variables for stability (further development will consider methods to incorporate exogenous forcing for non-temperature fields without sacrificing model stability).

2. Apply the VARX equation: compute a weighted sum of the previous two months' PC values across all modes, plus weighted exogenous variables (if applicable), plus the intercept. This gives the deterministic prediction.

3. Add random elements: draw from the fitted multivariate normal distribution with the learned covariance structure, and add to the deterministic prediction. This introduces stochasticity while preserving the inter-mode correlations observed in training.

4. Store the new PC values and advance to the next month.

This process is repeated for each ensemble member with independent random internal variability sampling. While ensemble members share the same VARX parameters and respond to the same temperature trajectory, their different random initializations cause them to diverge, exploring different realizations of internal variability.

For computational efficiency, all random innovations are pre-generated in a single batched multivariate normal draw before the time loop begins. This eliminates the overhead of thousands of separate random number generations and enables efficient vectorized operations, reducing generation time for large ensembles from minutes to seconds.

4.2 Reconstructing Spatial Fields

From simulated PCs, we reconstruct spatial climate fields in three steps. First, we compute the seasonal cycle for the new temperature trajectory. Using the nine harmonic features constructed from the new trajectory's global temperature and month-of-year, we multiply by the fitted seasonal coefficients at each location. This gives the expected seasonal pattern accounting for warming.

Second, we convert the simulated PCs back to spatial anomalies. This is the inverse of the PCA transformation:

$$A(t,x,y) = \sum_{k=1}^{40} Z_k(t) \cdot E_k(x,y), \tag{10}$$

where $Z_k(t)$ is the k-th simulated PC at time t and $E_k(x,y)$ is the k-th learned EOF pattern at location (x,y). This reconstruction exhibits the same spatial covariance structure as the training data. Finally, we combine the seasonal cycle and spatial anomalies to get the complete field.

4.3 Integration with METEOR-CORE

For integration with METEOR-CORE ("Background climate" and "Noise Generation" in Fig. 1), we generate noise-only fields that exclude the forced warming component. The noise field at each location and time is:

$$N(t,x,y) = \sum_{j=2}^{9} \beta_j(x,y) \cdot \mathbf{X}_{\text{harm},j}(t) + \sum_{k=1}^{40} Z_k(t) \cdot E_k(x,y), \tag{11}$$

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

where we omit the intercept β_0 and the direct temperature term $\beta_1 T_{\text{glob}}(t)$, keeping only the harmonic terms (including the temperature-modulated harmonics) plus the stochastic anomalies. The forced warming signal is provided by the METEOR-CORE pattern scaling model.

METEOR-CORE produces annual mean predictions $\bar{X}_{\text{METEOR}}(t_a, x, y)$ where $t_a = \lfloor (t-1)/12 \rfloor + 1$ converts monthly time index t to year index t_a. To combine with monthly noise, we simply repeat each annual value for 12 months. The final ensemble member is the sum:

$$X^{(m)}_{\text{ensemble}}(t, x, y) = \bar{X}_{\text{METEOR}}(t_a, x, y) + N^{(m)}(t, x, y), \tag{12}$$

where superscript (m) denotes ensemble member m.

This additive combination assumes forced response and internal variability are independent. The forced response provides the long-term trend and multi-decadal evolution, while the noise adds higher temporal fluctuations around that trend.

4.4 Efficient Regional Calculations

A practical feature of the PC-based approach is that we can compute regional or global means without reconstructing full spatial fields. For any region $\mathcal{R}$, we pre-compute the area-weighted regional mean of each EOF pattern:

$$\bar{E}_k^{\mathcal{R}} = \frac{\sum_{(x,y)\in\mathcal{R}} w(x) \cdot E_k(x, y)}{\sum_{(x,y)\in\mathcal{R}} w(x)}, \tag{13}$$

where $w(x) = \cos(\text{lat}(x))$ are latitude-based area weights as in Eq. (3). Then the regional mean anomaly at any time is:

$$\bar{A}^{\mathcal{R}}(t) = \sum_{k=1}^{40} Z_k(t) \cdot \bar{E}_k^{\mathcal{R}}. \tag{14}$$

This matters for computational efficiency. If we want 100 ensemble members of global mean temperature over 350 years (42,000 monthly values per member), we can compute this directly from the 40 PCs without ever calculating the full spatial fields. This vastly reduces memory requirements and enables very large ensembles for either regions or point-site estimates.

4.5 Variable-Specific Processing

Different climate variables require distinct statistical treatments. Temperature anomalies are approximately Gaussian-distributed and can be processed directly using the PCA-VARX framework described above. However, precipitation presents additional challenges: it is bounded below by zero, exhibits strong positive skewness, and has fundamentally different statistical properties than temperature.

METEORv1.6 addresses this through a variable-specific transformation framework. For precipitation, the noise model output (which is approximately Gaussian) is transformed by default via quantile mapping to a gamma distribution:

$$P_{\text{transformed}} = F_\Gamma^{-1}\left(\Phi\left(\frac{P_{\text{noise}} - \mu}{\sigma}\right); k, \theta\right), \tag{15}$$

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

where Φ is the standard normal CDF, F_Γ^{-1} is the inverse gamma CDF, and (k, θ) are the shape and scale parameters fitted to the target ESM's precipitation distribution using maximum likelihood estimation. When fitting fails (e.g., due to edge cases), the code falls back to method-of-moments estimation: $k = \mu^2/\sigma^2$ and $\theta = \sigma^2/\mu$. The quantile values are clipped to $(10^{-10}, 1 - 10^{-10})$ to ensure numerical stability at distribution tails.

Precipitation is processed as absolute values rather than anomalies. The core pattern scaling model provides the forced response, and the noise model adds realistic monthly variability, with the gamma transform ensuring the non-negativity constraint and preserving the characteristic right-skewed distribution observed in ESM output.

For temperature, anomalies are computed relative to a pre-industrial baseline derived from the pre-industrial control (piControl) experiment. The global mean piControl temperature is subtracted from the training data before computing the smoothed temperature trajectory used for seasonal cycle modulation.

The framework is extensible through a transform registry, which maps variable names to configuration objects specifying the transform type and fitting/application functions. For regional and global mean calculations, supported transforms include gamma, Weibull, log-normal, and generalized gamma distributions, as well as empirical (non-parametric) quantile mapping. For fully gridded fields, currently only the gamma distribution is implemented.

5 Computational Performance

The methodology is designed for computational efficiency at both training and generation stages.

Training a noise model for a typical CMIP6 resolution (approximately 96×144 spatial grid, 3000 monthly time steps, 40 PCs) takes 30-60 seconds on a standard laptop. The process is dominated by Principal Component Analysis, which scales roughly quadratically with spatial resolution. Models at higher resolution or with longer training periods remain tractable in several minutes at most for high-resolution ESMs.

Generation speed varies by output type. For regional or global mean timeseries (e.g., global mean temperature or precipitation for a continent), producing 100 ensemble members covering 350 years at monthly resolution takes approximately 90 seconds on a standard laptop. For point-based extraction (e.g., a single city), the same computation takes roughly 10 seconds for 100 members, enabling ensembles of 1000+ members in under two minutes. Reconstructing full spatial fields is more expensive, scaling with grid resolution. This performance comes from batch random number generation and vectorized linear algebra operations that avoid explicit loops over ensemble members.

Memory requirements depend on the use case. Full spatial fields for large ensembles can quickly exceed available memory on a conventional laptop, but storing only PCs reduces requirements by 2-3 orders of magnitude. For most applications, a hybrid approach works well: store PCs for all members, then reconstruct variability for large ensembles in regions of interest as needed.

METEORv1.6 provides a unified high-level interface that manages training, caching, and ensemble generation through a single entry point. This interface handles automatic caching of intermediate results, variable-specific transforms, and flexible

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

output aggregation (global means, IPCC Sixth Assessment Report (AR6) regional means (IPCC, 2021), or point-based extraction), enabling users to generate probabilistic climate projections with minimal configuration.

6 Impact Assessment Framework

6.1 Modular Calculator Architecture

The impact assessment framework provides a standardized interface for converting climate projections into sectoral impact metrics while preserving METEOR's uncertainty characteristics and ensemble capabilities. The system is built around an extensible `ImpactCalculator` base class that defines common methods for climate data preprocessing and validation with METEOR-specific metadata, impact metric calculation with configurable parameters, ensemble processing and uncertainty quantification, and statistical analysis and results summarization consistent with METEOR's multi-forcing framework.

This modular design enables rapid development of sector-specific impact calculators while maintaining consistency in uncertainty treatment and ensemble analysis across different applications. The framework automatically propagates uncertainty from METEOR's climate projections through to impact metrics, enabling probabilistic impact assessments that account for both climate model uncertainty and natural variability.

6.1.1 Degree Days Implementation

The initial implementation includes a degree days calculator for energy sector applications, adapting them for METEOR's monthly climate outputs. Heating degree days (HDD) and cooling degree days (CDD) are conventionally defined as:

$$\text{HDD} = \sum_d \max(T_{\text{base}} - T_d, 0) \tag{16}$$

$$\text{CDD} = \sum_d \max(T_d - T_{\text{base}}, 0) \tag{17}$$

where T_d represents daily outdoor temperature and T_{base} is the base temperature threshold (typically 15–18°C), the temperature above (below) which a building doesn't need heating (air conditioning). Since METEOR produces monthly mean temperatures rather than daily values, the implementation uses the statistical method of Isaac and Van Vuuren (2009), which estimates degree days from monthly means by assuming daily temperatures follow a normal distribution. The method uses an empirical parameterization of intra-monthly temperature variability based on Erbs et al. (1982):

$$\sigma_m = c_1 - c_2 T_m + c_3 \sigma_y, \tag{18}$$

where σ_m is the estimated standard deviation of daily temperatures within month m, T_m is the monthly mean temperature, σ_y is the annual standard deviation of monthly means, and $(c_1, c_2, c_3) = (1.45, 0.29, 0.664)$ are default empirical coefficients derived from observational data, assumed constant under climate change (a simplification that may introduce bias under substantially

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

altered variability regimes). The degree days are then computed analytically from this assumed distribution, avoiding the need for daily temperature data while still accounting for within-month variability.

6.1.2 Scaling results to a GMT timeseries

Traditional pattern scaling approaches such as MESMER (Beusch et al., 2020) and also more inventive approaches such as STITCHES (Tebaldi et al., 2022) employ global mean timeseries data, primarily global mean temperature (GMT), to scale their mean outputs. While METEOR inherently moves away from this approach to be able to model path- and forcer-dependent patterns, regional climate signals are strongly coupled to global mean temperature (Giani et al., 2025), and GMT-based scaling remains a useful approximation for many applications. This capability enables direct integration with multi-emulator comparison frameworks such as FastMIP (Seneviratne et al., 2024), where emulators are evaluated against a common GMT trajectory, and facilitates comparison with other pattern scaling approaches that use GMT as their primary scaling variable. It also allows METEOR projections to be conditioned on GMT trajectories from different ESMs or simple climate models, extending the framework's applicability beyond the training ESM.

METEOR v1.6 therefore includes an option to include a GMT timeseries as input to the emulation generation, in addition to the gridded ESM training data. This GMT information is then integrated into the generation of emulated qualities described above in the following way:

First, annual patterns are created as described above, using the methodology of METEOR v1.0.1 Sandstad et al. (2025), which includes the generation of a GMT timeseries associated with those patterns, even if temperature is not a target value. These annual patterns are scaled by the ratio of the provided global mean temperature timeseries to the METEOR v1.0.1 emulated global mean timeseries. The resulting scaled annual patterns are then used to produce all downstream outputs as if they were the emulation result of the METEOR v1.0.1 part of the pipeline. Illustrative examples of this scaling approach are shown in Supplemental Figures 13 and 14.

7 Results

We validate METEORv1.6 in two stages. First, we demonstrate in-sample performance by training the emulator on SSP2-4.5 output from NorESM2-MM and evaluating the fidelity of the resulting ensemble against the training data. Second, we assess out-of-sample generalization by training on SSP2-4.5 across five CMIP6 models and evaluating predictions for scenarios not seen during training (SSP1-2.6 - a high mitigation scenario, SSP5-8.5 - an exceptionally high emissions scenario, and SSP5-3.4-Overshoot - a strong overshoot scenario) (O'Neill et al., 2016). In each case, we present an illustrative example in the main text; a more comprehensive assessment across all models and spatial scales is provided in the Supplemental Figures.

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

7.1 In-Sample Validation

7.1.1 Distributional Fidelity

Figure 2 provides an overall illustration that the METEOR emulator produces climate distributions consistent with the source ESM. The figure shows METEOR ensemble projections (100 realizations) against NorESM2-MM output for the near-term period (2015–2035), demonstrating that the emulated distributions capture both the central tendency and spread of the year-to-year variation in NorESM across multiple spatial scales (global, regional and point) and for both temperature and precipitation.

7.1.2 Long-Term Climate Trends

Figure 3 extends the evaluation to the full scenario period (2000–2100), demonstrating that long-term climate trends are preserved by METEORv1.6 at regional scales. The multi-decadal warming trajectory and associated precipitation changes of the source simulation are well captured by the emulator, indicating that the pattern scaling foundation of METEOR-CORE combined with METEOR-NOISE qualitatively captures forced responses and variability over extended time horizons.

7.1.3 Seasonal Cycle and Distributional Shifts

A critical test of the monthly extension with METEOR-NOISE is its ability to reproduce evolving seasonal patterns and distributional characteristics. Figure 4 illustrates this at four contrasting point locations: Oslo, Delhi, Cairo and Beijing, capturing evident features from the source simulation in each case. In Delhi, the emulator captures the reduction in the seasonal temperature cycle under warming and the increasing variability in precipitation. In Oslo, similar preservation of increasing precipitation variability is evident. These results validate the temperature-dependent harmonic representation of the seasonal cycle described in Section 2, which allows both seasonal amplitude and phase to evolve with the background climate state.

7.1.4 Ensemble Adequacy

We assess ensemble adequacy using rank histograms, which objectively test whether the CMIP6 output is statistically indistinguishable from a random draw of the METEOR ensemble. For each time step, we compute the normalized rank of the CMIP6 value within the METEOR ensemble; a uniform rank distribution indicates that the ensemble adequately spans the observed variability. A flat histogram indicates that the source model is indistinguishable from the ensemble, a u-shaped histogram indicated that the ensemble is underdispersive (i.e. the source model tends to fall outside the emulated ensemble more often than would be expected from a perfect emulation), while clustering towards the center of the histogram would indicate overdispersion (that the emulated ensemble exhibits too much variance).

Figure 5 shows in-sample rank histograms for NorESM2-MM across multiple regions and for both variables. Temperature ranks are broadly uniform across most regions, indicating appropriate ensemble dispersion, though slight underdispersion is evident for some regions (North South America region, Delhi gridcell). For precipitation, ranks are generally well-calibrated, with some underdispersion apparent for Northern Europe and Beijing. These results indicate that the PCA-VARX framework

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

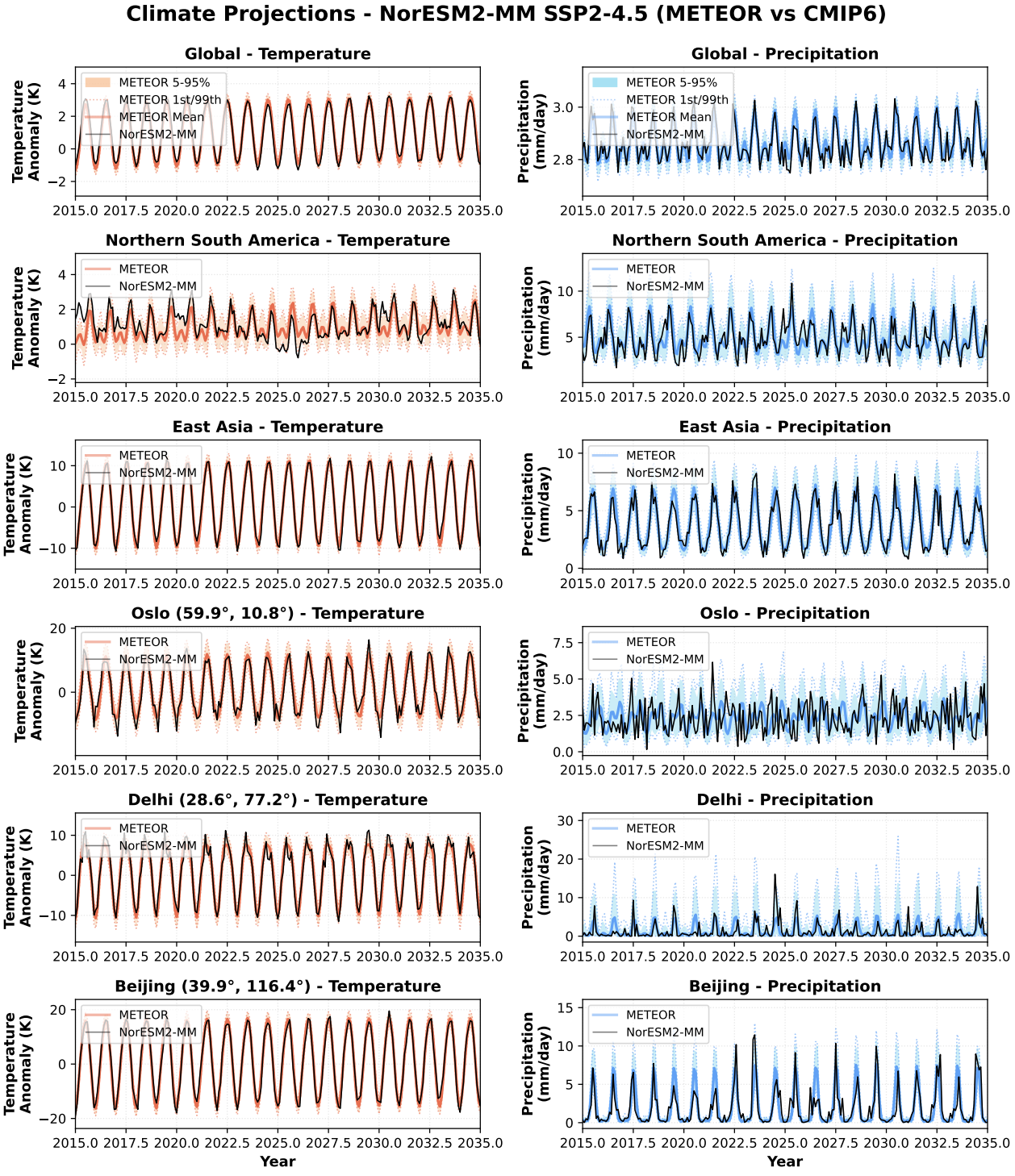

Figure 2. METEOR ensemble projections (100 realizations) compared with NorESM2-MM SSP2-4.5 output for 2015–2035. The emulated distributions capture both the forced response (emulated by METEOR-CORE) and seasonal cycle/internal variability (emulated by METEOR-NOISE) of the source model across multiple spatial scales (global, regional and point).

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

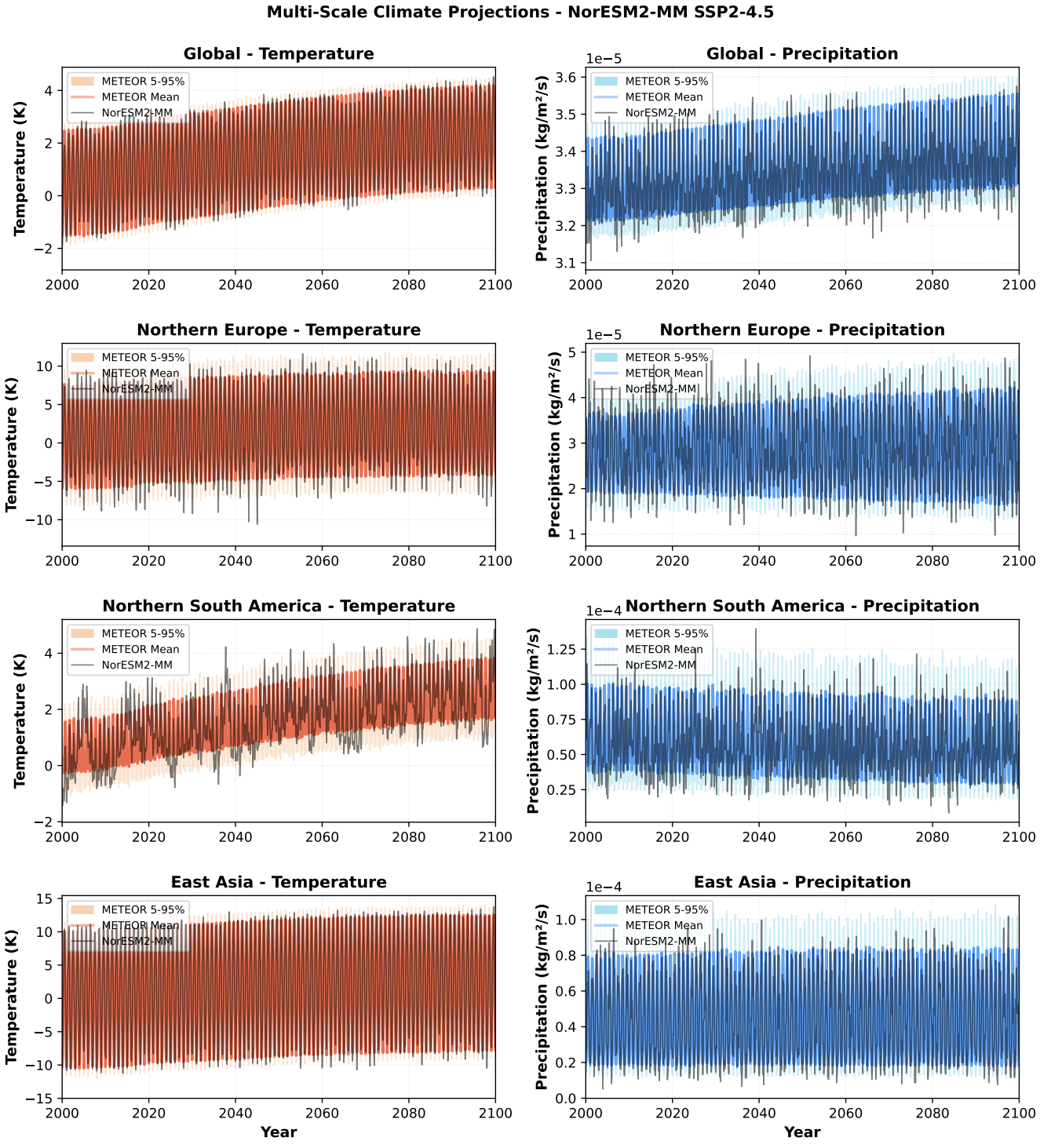

Figure 3. METEOR ensemble projections compared with NorESM2-MM SSP2-4.5 output for the extended period 2000–2100, showing that long-term regional climate trends are preserved by the emulator.

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

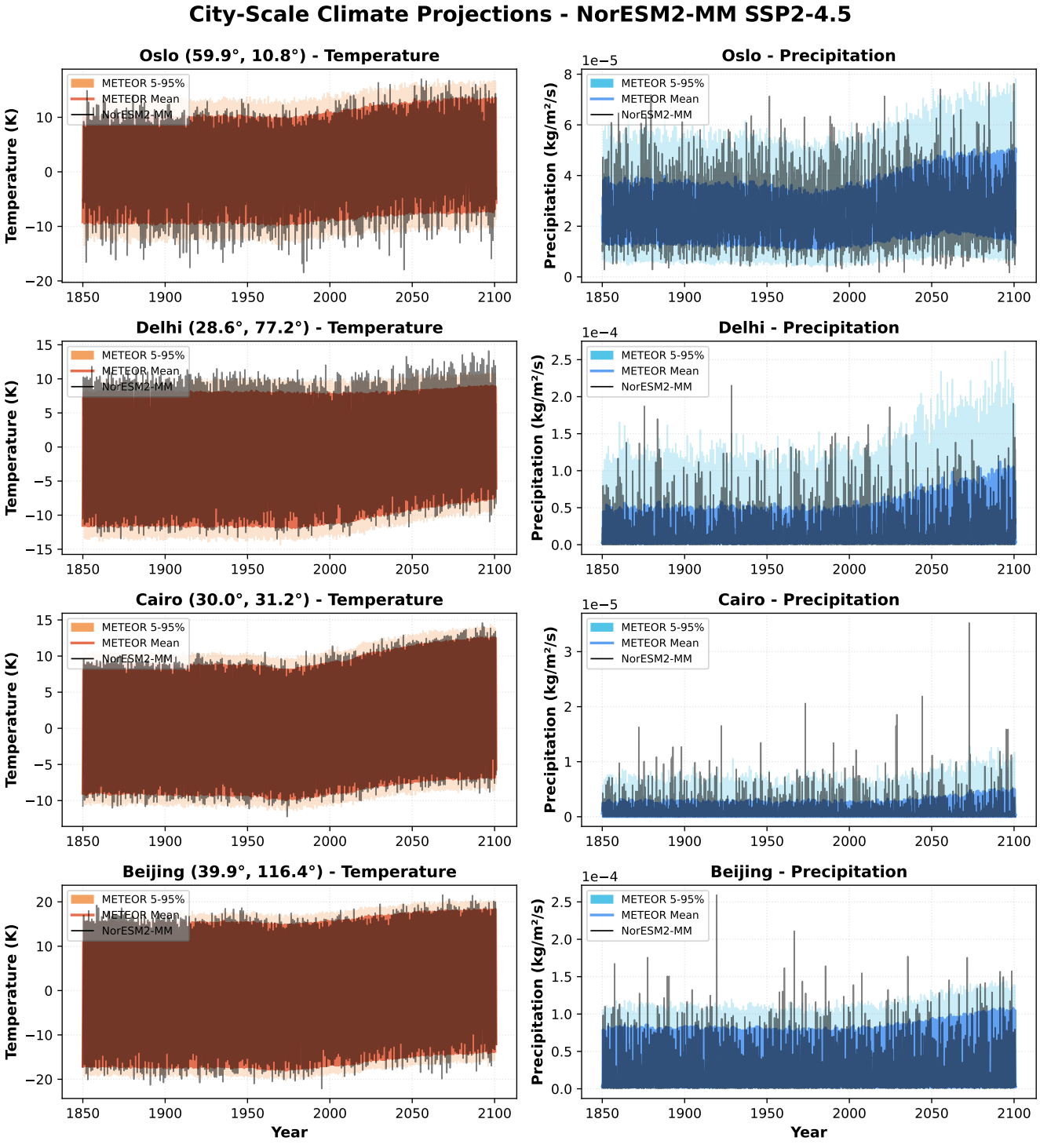

Figure 4. City-scale projections (1850–2100) for a number of cities, demonstrating that METEORv1.6 preserves distributional trends and seasonal cycle shifts.

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

captures the dominant modes of internal variability, though the Gaussian noise assumption may slightly underestimate the tails of the distribution in certain regions. Global projections for both temperature and precipitation are slightly underdispersive.

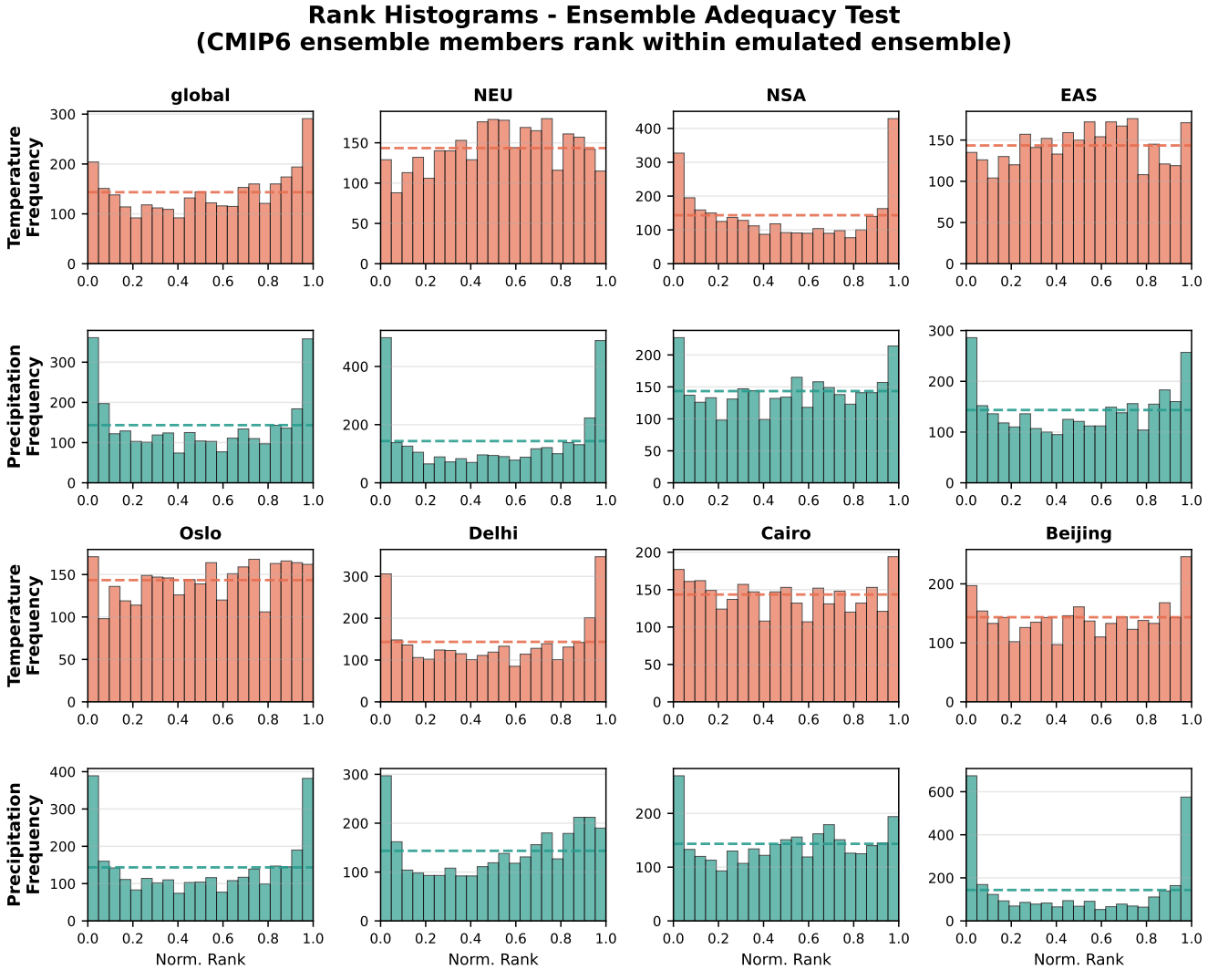

Figure 5. In-sample rank histograms for NorESM2-MM (SSP2-4.5) across multiple regions and variables. A uniform distribution (dashed line) indicates a perfectly representative ensemble.

7.1.5 Impact Metrics: Degree Days

Figure 6 demonstrates the integrated impact assessment capabilities of METEORv1.6 through heating degree day (HDD) and cooling degree day (CDD) calculations. The emulator shows strong in-sample performance for these evolving impact metrics, capturing several notable features. In Beijing, the emulator reproduces the increase in HDD demands during the late 20th

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

century associated with aerosol-driven regional cooling, demonstrating that METEOR's multi-forcer pattern scaling correctly propagates aerosol effects through to impact metrics at monthly resolution. In Oslo, cooling demands are essentially absent until the mid-21st century—a feature faithfully preserved by METEOR, which correctly represents this threshold behavior despite never being explicitly trained on degree day targets. These results demonstrate end-to-end uncertainty propagation from climate projections through to sectoral impact metrics.

Here, the target is the Isaac and Van Vuuren (2009) algorithm applied directly to NorESM output. Future work will look into daily and subdaily temperature emulation using high temporal frequency output from NorESM to test this algorithm's performance, but this requires significant additional processing of high frequency output and is beyond the scope for this paper.

7.2 Out-of-Sample Validation

To assess generalization beyond the training scenario, we train METEORv1.6 on SSP2-4.5 and evaluate predictions for SSP1-2.6, SSP5-8.5, and SSP5-3.4-Overshoot. Here, we demonstrate model performance for five models which performed simulations in CMIP6 for this entire set of scenarios: CanESM5, CESM2-WACCM, CNRM-ESM2-1, IPSL-CM6A-LR, and UKESM1-0-LL.

7.2.1 Multi-Scale Timeseries

In the main paper, we show CNRM-ESM2-1 as an out-of-sample example; equivalent results for all five models are provided in the Supplemental Figures. Figure 7 shows temperature anomaly projections across five spatial scales (global mean, East Asia, Northern Europe, Mumbai, and Oslo) for all four scenarios. The METEOR ensemble captures the CMIP6 climate signal across all spatial scales and scenarios. Both mean signals and variability characteristics are well captured in diverse out-of-sample scenarios, including under strong mitigation in SSP1-2.6 and for the overshoot scenario SSP5-3.4-OS, where CNRM-CM6's non-monotonic temperature trajectory is well reproduced. This highlights that METEOR-CORE's multi-timescale impulse response representation for scenarios, combined with the METEOR-NOISE model can represent the outcome of reversing forcing trends.

Figure 8 shows the corresponding precipitation projections. The forced precipitation response is generally well captured, though precipitation changes under the high-emissions SSP5-8.5 are somewhat underestimated at some scales. This is consistent with the known challenge that precipitation responses can exhibit stronger nonlinearity with forcing magnitude than temperature responses, and that the pattern scaling relationship trained on a moderate forcing scenario (SSP2-4.5) may not fully extrapolate to the strongest forcing case.

7.2.2 Out-of-Sample Ensemble Adequacy

Figure 9 shows rank histograms for an example gridcell (Oslo) across all four validation scenarios and five models, assessing whether the METEOR ensemble adequately spans the CMIP6 variability in out-of-sample conditions. Temperature distributions are well emulated for all scenarios, with rank histograms either uniform or slightly overdispersive across all models. Precipitation

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

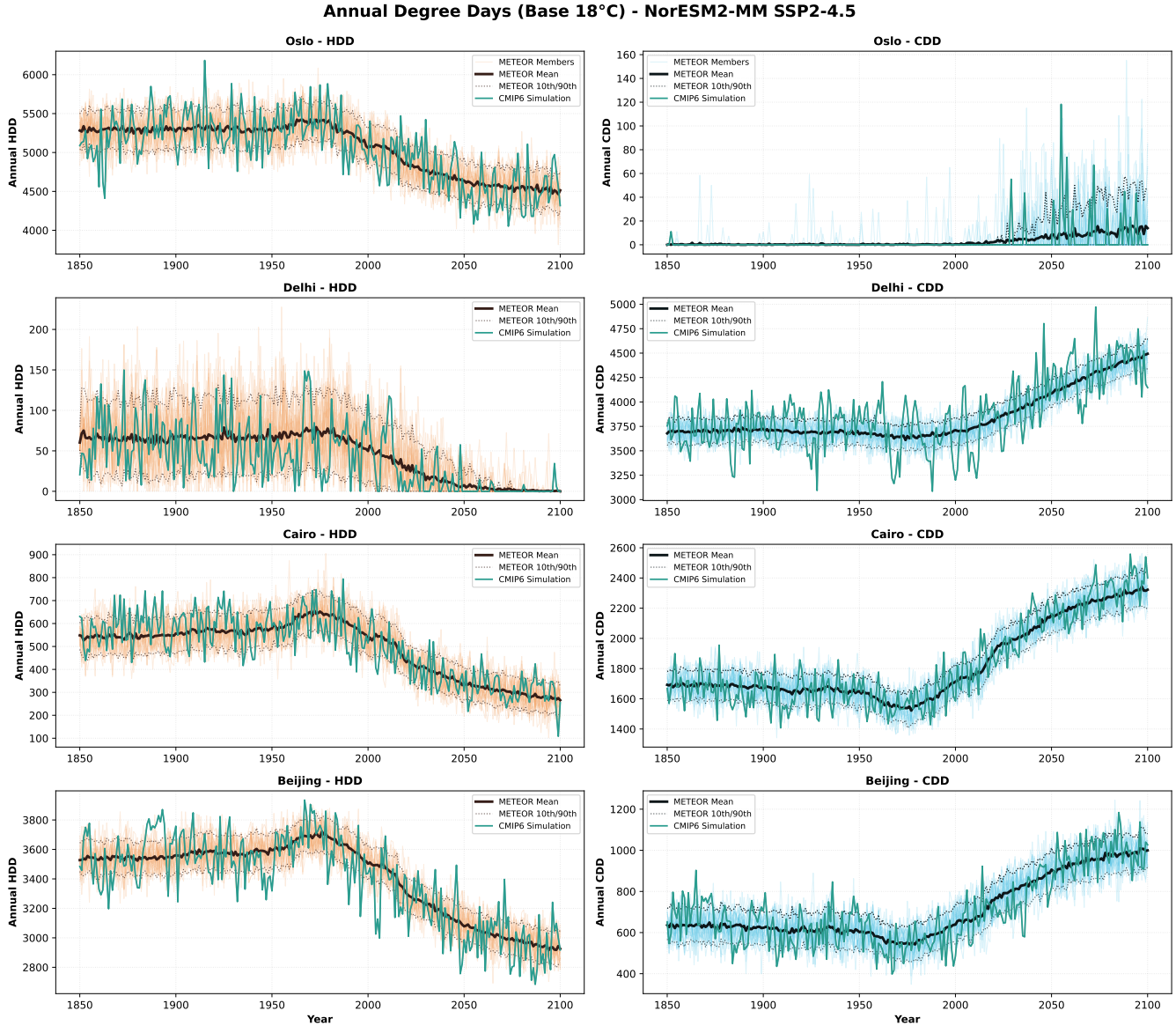

Figure 6. Heating degree days (HDD) and cooling degree days (CDD) for selected cities, comparing METEOR ensemble projections with NorESM2-MM. Notable features include the aerosol-related increase in HDD in Beijing during the late 20th century and the absence of cooling demands in Oslo until mid-21st century, both faithfully reproduced by the emulator.

shows underdispersion to different degrees in different models - IPSL-CM6A-LR is well represented, for example but CanESM5 is notably underdispersive). Interestingly, global-scale precipitation rank histograms (Supplemental Figure 9) show more pronounced underdispersion than point-scale results, suggesting that the precipitation distributional transform may be less appropriate for large-scale spatial aggregations where individual grid-cell biases are averaged rather than cancelling. Coordinate

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

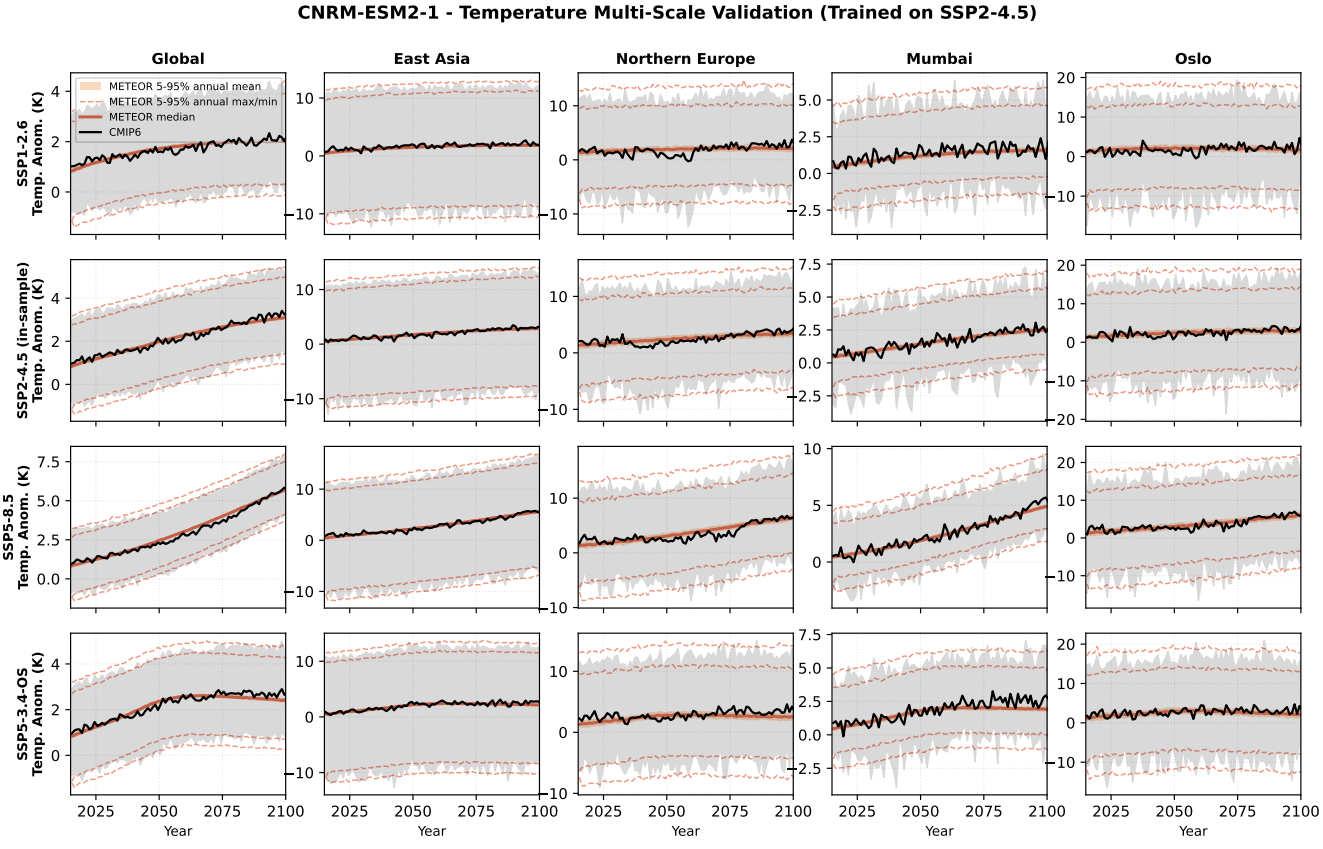

Figure 7. Out-of-sample temperature validation for CNRM-ESM2-1 trained on SSP2-4.5. Each row shows a validation scenario (SSP1-2.6, SSP2-4.5, SSP5-8.5, SSP5-3.4-OS); each column shows a spatial scale (global, East Asia, Northern Europe, Mumbai, Oslo). Orange shading: METEOR 5–95% ensemble range of annual means. Orange dashed lines: 5–95% range of annual maximum and minimum values. Orange solid line: ensemble median of annual means. Black line and grey shading: CMIP6 annual mean and monthly min/max envelope. Results for other models are shown in Supplemental Figures 1–4.

transforms as a function of spatial scale are a development goal for future versions. Rank histograms for additional spatial scales are provided in the Supplemental Figures.

7.2.3 Quantitative Validation Metrics

For a stochastic emulator, the appropriate validation question is not whether individual ensemble members match the CMIP6 realization (they cannot, since stochastic sequences are inherently unpredictable) but whether the CMIP6 realization is statistically indistinguishable from an ensemble member. We assess this using *skill ratios* that compare member-CMIP6 metrics against member-member metrics. For both correlation and RMSE, we define the ratio such that values greater than 1 indicate poorer

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

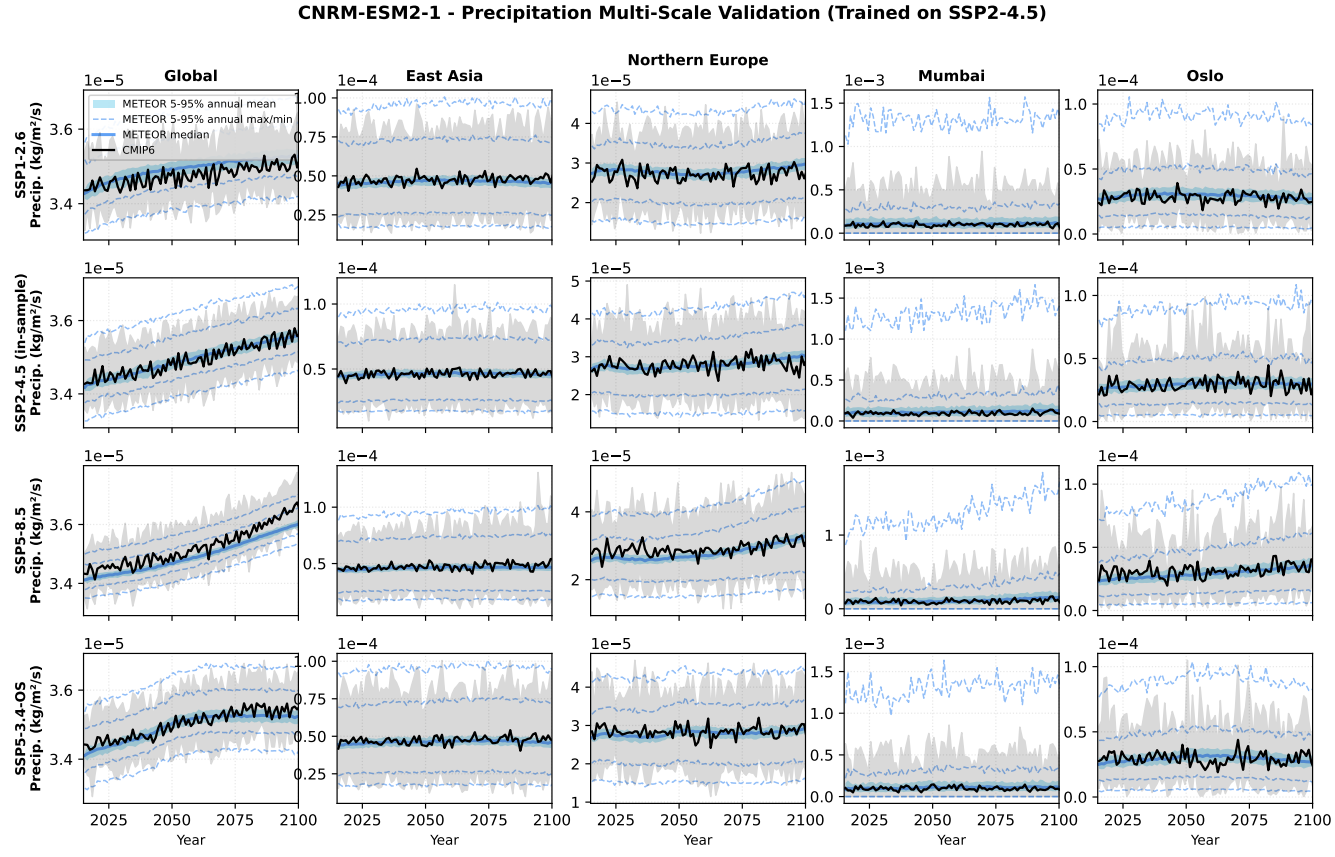

Figure 8. Out-of-sample precipitation validation for CNRM-ESM2-1 trained on SSP2-4.5. Layout and visual encoding as in Figure 7: blue shading shows the 5–95% range of annual means, dashed lines show the 5–95% range of annual extrema, and the solid line shows the ensemble median. Precipitation changes are generally well captured, though SSP5-8.5 changes are somewhat underestimated. Results for other models are shown in Supplemental Figures 5–8.

460 emulation:

$$\text{Correlation Skill Ratio} = \frac{\text{Member-Member correlation}}{\text{Member-CMIP6 correlation}} \tag{19}$$

$$\text{RMSE Skill Ratio} = \frac{\text{Member-CMIP6 RMSE}}{\text{Member-Member RMSE}} \tag{20}$$

If the emulator is well-calibrated, the CMIP6 output should be no more or less similar to ensemble members than members are to each other, yielding skill ratios near unity. For both metrics, ratios greater than 1 indicate poorer emulation: either
465 systematic bias in the forced response (for RMSE) or lower-than-expected correlation (for correlation). Ratios below 1 indicate the ensemble may be underdispersive-members are too similar to each other, consistent with U-shaped rank histograms. This can be considered 'hedging' results by underestimating variability with a simulation closer to the ensemble mean.

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

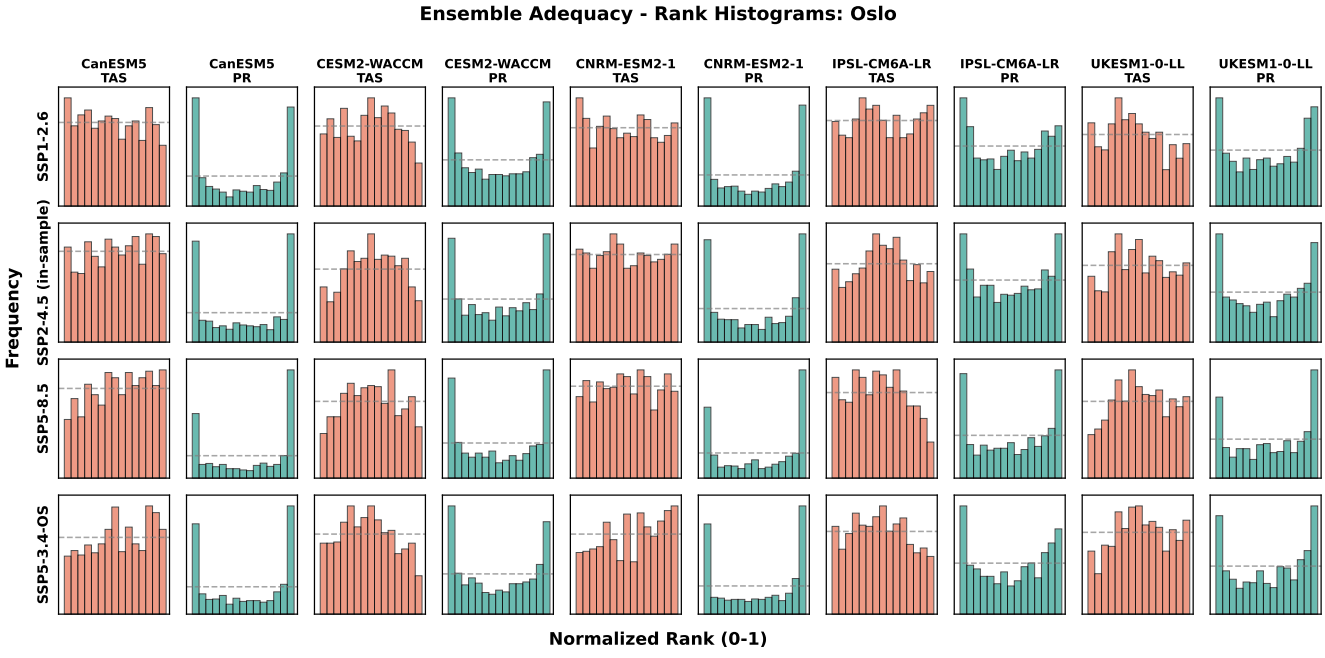

Figure 9. Out-of-sample rank histograms for Oslo across all scenarios and models. Each row shows a scenario; each pair of columns shows temperature and precipitation for one model. Temperature ranks are approximately uniform, indicating good ensemble calibration. Precipitation shows slight underdispersion. Rank histograms for other spatial scales are shown in Supplemental Figures 9–12.

Figure 10 shows skill ratios across all spatial scales, models, and scenarios. Temperature skill ratios are consistently near 1.0 for both correlation (range 0.98–1.12) and RMSE (0.88–1.84), indicating that the emulator faithfully reproduces temperature variability characteristics across all scales with minimal systematic bias.

Precipitation skill varies more with scale. At global scales, precipitation correlation ratios are near unity (1.02–1.18) but RMSE ratios are elevated (1.4–7.0), suggesting pattern scaling does not fully capture the forced global precipitation response. At regional scales (East Asia, Northern Europe), precipitation correlation ratios remain near unity (1.02–1.43) with RMSE ratios of 1.1–1.8. At point scales, precipitation correlation ratios show greater variability (0.88–3.4 across Oslo and Mumbai), with Oslo showing the largest values, reflecting that point-scale precipitation correlation is dominated by stochastic variability where even small absolute differences produce large ratios. Mumbai shows ratios closer to unity, suggesting better distributional matching for monsoon-dominated precipitation.

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

8 Discussion

8.1 Advantages and Integration with METEOR Core Framework

METEOR-NOISE extends the METEOR-CORE framework to monthly resolution, enabling the generation of realistic sub-annual climate variability and integrated impact metrics while preserving the multi-timescale and forcing-dependent characteristics that distinguish METEOR from conventional pattern scaling approaches.

The framework's primary strength is computational efficiency combined with interpretable statistical realism. It generates ensembles far faster than ESMs while preserving many aspects of spatial coherence, temporal persistence, and teleconnections. The decomposition into seasonal and anomaly components aids interpretability, and the PC-based approach enables calculations at multiple spatial scales without redundancy.

The temperature-modulated seasonal cycle can represent warming effects on annual cycles. This is particularly important for high-latitude regions undergoing rapid change. METEOR-NOISE is fully integrated with METEOR-CORE in a combined interface, enabling rapid generation of large climate ensembles (100+ members in under two minutes) for any scenario METEOR can emulate. Each ensemble member shares the same forced response but explores a distinct realization of internal variability, providing a direct characterization of the irreducible uncertainty in regional climate projections.

METEOR-NOISE maintains the framework's ability to represent hysteresis effects and forcing-dependent responses, which are critical for scenarios involving emission overshoots or varying forcer mixes (Good et al., 2015; Herger et al., 2015). Monthly variability generation adds very modest computational overhead while maintaining METEOR's performance advantages over ESM simulations, enabling large ensemble generation for uncertainty quantification. The temporal resolution supports sectoral applications requiring sub-annual climate information, including energy demand modeling, agricultural impact assessment, and extreme event analysis (Sillmann et al., 2013).

The framework propagates uncertainty from climate projections through to impact metrics, enabling probabilistic risk assessment that accounts for both forced climate change and natural variability. The impact assessment here is just illustrated with a simple example, but the model structure provides standardized interfaces that facilitate rapid development of sector-specific applications while maintaining consistency in uncertainty treatment across different sectors and applications.

8.2 Limitations and Future Developments

Current limitations of METEORv1.6 include dependencies on ESM training data quality, simplified treatment of compound events, and assumptions of stationarity in some statistical relationships. These limitations are inherited from the core METEOR framework and represent areas for continued development.

The linear seasonal model works well for temperature, but application to precipitation is more challenging due to its non-Gaussian distribution and stronger seasonality - leading to the choice to represent precipitation as a static VAR model in the default configuration. Future versions should explore more robust approaches to representing seasonality across variables, including nonlinear and non-parametric methods, noting that the linear model cannot capture highly nonlinear changes in seasonality, though the temperature-interaction terms provide substantial flexibility.

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

The Gaussian assumption for noise is well-justified for monthly means but likely underestimates extreme event probabilities. Extensions using fat-tailed distributions could address this for applications focused on extremes. Similarly, the current impact assessment framework focuses on a limited number of validated outputs, requiring expansion to precipitation-dependent and compound event applications.

The VARX model assumes some stationarity, for example that the lag dependence of a particular PC on others does not change over time. Under very strong forcing, teleconnection strengths or persistence timescales might evolve, requiring time-varying or state-dependent parameters. Similarly, the additive combination of forced response and variability assumes independence, whereas the real climate system exhibits some nonlinearity wherein forcing can modulate variability amplitude.

These limitations are important to recognize but do not undermine the model's utility for most applications. Monthly mean climate on seasonal to decadal timescales, under moderate forcing scenarios, is generally well-represented. For studies requiring daily extremes or investigating potential regime shifts, more sophisticated approaches may be needed.

Future versions will explore higher temporal sampling, including daily and sub-daily climate generation, to support applications requiring extreme event statistics. Machine learning methods for temporal (Kajbaf et al., 2022; Bassetti et al., 2024) and spatial (Wang and Tian, 2022; Kheir et al., 2023; Harder et al., 2024) downscaling represent a promising avenue for enhancing resolution while preserving physical consistency with the underlying pattern scaling framework. Additionally, improvements to forcer treatment in METEOR-CORE are planned, including incorporation of the AeroGP Gaussian process regional aerosol model (Dewey et al., 2025), which provides more sophisticated representation of spatially heterogeneous aerosol forcing responses than the current linear pattern scaling approach.

8.3 Relationship to Other Climate Emulation Approaches

METEORv1.6 complements other existing climate emulation frameworks while addressing specific limitations. Compared to MESMER's monthly extensions (Nath et al., 2022; Schöngart et al., 2024), METEORv1.6 preserves forcing-dependent and time-evolving response capabilities present in METEOR v1.0.1 Sandstad et al. (2025) while also preserving spatial covariance structures of the target model. Other recent approaches also employ impulse response formulations for spatially resolved emulation (Freese et al., 2024; Womack et al., 2025). METEOR shares the impulse response philosophy but differs in its multi-forcer decomposition and its integration of monthly variability generation with the forced response, enabling ensemble-based uncertainty quantification alongside the pattern-scaled climate signal. Relative to machine learning approaches (Watson-Parris et al., 2021; Kendon et al., 2025; Kitsios et al., 2023), METEOR provides physically grounded and fully explainable pattern scaling with interpretable structure and significantly lower training data requirements. Further, the integration of impact assessment capabilities distinguishes METEORv1.6 from purely climate-focused emulators, providing end-to-end workflows from emissions to sectoral impacts. This integration enables consistent uncertainty propagation and ensemble analysis that may be difficult to achieve with separate climate and impact modeling systems.

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

9 Conclusions

METEORv1.6 occupies a distinctive niche in the climate modeling landscape: lightweight enough for interactive use, yet grounded in the physical dynamics that govern Earth system model responses. By distilling ESM behavior into interpretable pattern scaling relationships while preserving forcing-dependent and time-evolving characteristics, METEOR bridges the gap between computationally intensive process models and purely statistical approaches that sacrifice physical realism for speed.

The monthly extension presented here adds realistic sub-annual variability and integrated impact assessment while maintaining this balance. Validation against CMIP6 demonstrates that METEOR reproduces key statistical properties of climate variability across spatial scales, from global means to individual grid cells. The framework captures forced responses, seasonal cycles, and interannual variability with sufficient fidelity for many impact assessment applications, while generating 100-member ensembles in under two minutes on a standard laptop.

This computational efficiency opens new possibilities for climate information delivery. METEOR could serve as an interactive climate component within integrated assessment models (IAMs), providing spatially resolved climate feedbacks and probabilistic impact information in real time as users explore emission pathways. End-users (planners, policymakers, or researchers) can use METEOR to rapidly assess climate impacts for novel scenarios without waiting for ESM simulations, enabling truly interactive exploration of the relationship between emission choices and their consequences.

Important challenges remain beyond the current framework's scope. Nonlinear climate responses (such as potential Atlantic Meridional Overturning Circulation (AMOC) collapse or ice sheet instabilities) cannot be captured by pattern scaling approaches that assume linear superposition of forcing responses. Novel forcing scenarios, including solar radiation management interventions, require dedicated training data and potentially new model structures to represent their distinct spatial and temporal signatures (Farley et al., 2025). These current limitations represent desirable targets for future development, to increase the applicability of the framework to a wider range of questions and scenarios.

The value of emulators like METEOR lies not in reproducing the vast data volumes of initiatives like CMIP, but in providing rapid turnaround and interactive tools that illuminate the relationship between emission choices and impact outcomes. As climate-informed decision-making becomes increasingly urgent across sectors, frameworks that can deliver probabilistic, spatially resolved climate information on demand, while remaining true to process-resolving models, will play an essential role in translating climate science into actionable guidance for adaptation and mitigation planning.

Data and Code Availability

METEORv1.6 is freely available as open-source software at https://doi.org/10.5281/zenodo.18977718 (Sanderson et al., 2026). The software includes comprehensive documentation, tutorials, and example applications demonstrating both the monthly climate generation and impact assessment capabilities. All code used in this study is available in the repository with version control and continuous integration testing.

The framework builds on the established METEORv1.0.1 codebase (Sandstad et al., 2025) with backwards compatibility maintained for existing applications. Version 1.6 introduces a simplified high-level programming interface, improved precipitation

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

handling with variable-specific distributional transforms, and enhanced caching for computational efficiency. Training data requirements follow the same CMIP6 specifications as the core METEOR framework, with additional monthly resolution data used for the variability system validation.

Author Contributions

BMS wrote the manuscript and developed the initial METEORv1.6 variability and impact framework, extending METEOR-CORE model developed by MS. Code development and testing were performed by BMS, MS, JF and MD. All co-authors reviewed and commented on the manuscript.

Competing Interests

The authors declare that they have no conflict of interest.

Acknowledgments

BMS, MS, NS and MD acknowledge the Norwegian Funding Council project "TRIFECTA" (grant number 334811) and the Horizon-Europe Project "DIAMOND" (Grant agreement ID: 101081179). AI tools (the Anthropic Claude family of models) were used to assist with initial code structuring and in preparing the first draft of the paper, though all text and code has been extensively manually reviewed, edited and validated by the authors. We thank the CMIP6 modeling groups for making their output available, which was essential for this work.

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

References

Bassetti, S., Hutchinson, B., Tebaldi, C., and Kravitz, B.: DiffESM: Conditional Emulation of Temperature and Precipitation in Earth System Models With 3D Diffusion Models, Journal of Advances in Modeling Earth Systems, 16, e2023MS004 194, https://doi.org/https://doi.org/10.1029/2023MS004194, 2024.

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

Dasgupta, S., Van Maanen, N., Gosling, S. N., Piontek, F., Otto, C., and Schleussner, C.-F.: Effects of climate change on combined labour productivity and supply: an empirical, multi-model study, The Lancet Planetary Health, 5, e455–e465, 2021.

Dewey, M., Hansson, H.-C., Watson-Parris, D., Samset, B. H., Wilcox, L. J., Lewinschal, A., Sand, M., Seland, Ø., Krishnan, S., and Ekman, A. M.: AeroGP: machine learning how aerosols impact regional climate, Journal of Geophysical Research: Machine learning and computation, 2, e2025JH000 741, 2025.

Erbs, D., Klein, S., and Duffie, J.: Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation, Solar energy, 28, 293–302, 1982.

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

Farley, J., MacMartin, D. G., Visioni, D., Kravitz, B., Bednarz, E., Duffey, A., and Henry, M.: A Climate Intervention Dynamical Emulator (CIDER) for Scenario Space Exploration, EGUsphere, 2025, 1–23, https://doi.org/10.5194/egusphere-2025-1830, 2025.

Freese, L. M., Giani, P., Fiore, A. M., and Selin, N. E.: Spatially Resolved Temperature Response Functions to CO2 Emissions, Geophysical Research Letters, 51, e2024GL108 788, https://doi.org/https://doi.org/10.1029/2024GL108788, e2024GL108788 2024GL108788, 2024.

Giani, P., Fiore, A. M., Flierl, G., Ferrari, R., and Selin, N. E.: Origin and Limits of Invariant Warming Patterns in Climate Models, Journal of Climate, p. e240683, https://doi.org/10.1175/JCLI-D-24-0683.1, 2025.

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

Hamilton, J. D.: Time series analysis, Princeton university press, 2020.

Harder, P., Hernandez-Garcia, A., Ramesh, V., Yang, Q., Sattegeri, P., Szwarcman, D., Watson, C., and Rolnick, D.: Generating physically-consistent high-resolution climate data with hard-constrained neural networks, ArXiv e-prints, arXiv:2208.05424, 2024.

Hasselmann, K., Hasselmann, S., Giering, R., Ocana, V., and Storch, H.: Sensitivity study of optimal CO2 emission paths using a simplified structural integrated assessment model (SIAM), Climatic Change, 37, 345–386, 1997.

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

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

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

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

Isaac, M. and Van Vuuren, D. P.: Modeling global residential sector energy demand for heating and air conditioning in the context of climate change, Energy policy, 37, 507–521, 2009.

Kajbaf, A. A., Bensi, M., and Brubaker, K. L.: Temporal downscaling of precipitation from climate model projections using machine learning, Stochastic Environmental Research and Risk Assessment, 36, 2173–2194, 2022.

Kendon, E. J., Addison, H., Doury, A., Somot, S., Watson, P. A. G., Booth, B. B. B., Coppola, E., Gutiérrez, J. M., Murphy, J., and Scullion, C.: Potential for Machine Learning Emulators to Augment Regional Climate Simulations in Provision of Local Climate Change Information, Bulletin of the American Meteorological Society, 106, E1175 – E1203, https://doi.org/10.1175/BAMS-D-24-0114.1, 2025.

Kheir, A. M., Elnashar, A., Mosad, A., and Govind, A.: An improved deep learning procedure for statistical downscaling of climate data, Heliyon, 9, 2023.

Kitsios, V., O'Kane, T. J., and Newth, D.: A machine learning approach to rapidly project climate responses under a multitude of net-zero emission pathways, Communications Earth & Environment, 4, 355, 2023.

Lam, T. N., Wan, K. K., Wong, S., and Lam, J. C.: Impact of climate change on commercial sector air conditioning energy consumption in subtropical Hong Kong, Applied Energy, 87, 2321–2327, 2010.

Mathison, C., Burke, E. J., Munday, G., Jones, C. D., Smith, C. J., Steinert, N. J., Wiltshire, A. J., Huntingford, C., Kovacs, E., Gohar, L. K., et al.: A rapid-application emissions-to-impacts tool for scenario assessment: Probabilistic Regional Impacts from Model patterns and Emissions (PRIME), Geoscientific Model Development, 18, 1785–1808, 2025.

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

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

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

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, Geoscientific Model Development, 9, 3461–3482, https://doi.org/10.5194/gmd-9-3461-2016, 2016.

Pfleiderer, P., Schleussner, C.-F., and Sillmann, J.: Limited reversal of regional climate signals in overshoot scenarios, Environmental Research: Climate, 3, 015 005, https://doi.org/10.1088/2752-5295/ad1c45, 2024.

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

Sanderson, B., Sandstad, M., Steinert, N. J., Dewey, M., and Fjeldså, J. L.: benmsanderson/METEOR: v1.6.0, https://doi.org/10.5281/zenodo.18977718, 2026.

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

Sandstad, M., Steinert, N. J., Baur, S., and Sanderson, B. M.: METEORv1.0.1: a novel framework for emulating multi-timescale regional climate responses, Geoscientific Model Development, 18, 8269–8312, https://doi.org/10.5194/gmd-18-8269-2025, 2025.

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

Schleussner, C.-F., Ganti, G., Lejeune, Q., Zhu, B., Pfleiderer, P., Prütz, R., Ciais, P., Frölicher, T. L., Fuss, S., Gasser, T., et al.: Overconfidence in climate overshoot, Nature, 634, 366–373, 2024.

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

Schwind, N., Perrette, M., Byers, E., Högner, A., Lejeune, Q., Möller, T., Nicholls, Z., Pfleiderer, P., Schöngart, S., Werning, M., et al.: RIME-X v1. 0: Combining Simple Climate Models, Earth System Models, and Climate Impact Models into a Unified Statistical Emulator for Regional Climate Indicators, EGUsphere, 2026, 1–27, 2026.

Sejas, S. A. and Taylor, P. C.: The role of sea ice in establishing the seasonal Arctic warming pattern, Environmental Research: Climate, 2, 035 008, 2023.

Seneviratne, S., Quilcaille, Y., Windisch, M., Gudmundsson, L., Biess, B., Jaeger, F., Hauser, M., and Hirschi, M.: Using regional ESM emulators to assess climate feedbacks to IAMs: The" FASTMIP" experimental protocol, European Geosciences Union General Assembly 2024 (EGU24), p. 18731, 2024.

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

Sims, C. A.: Macroeconomics and reality, Econometrica: journal of the Econometric Society, pp. 1–48, 1980.

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

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

Wang, F. and Tian, D.: On deep learning-based bias correction and downscaling of multiple climate models simulations, Climate dynamics, 59, 3451–3468, 2022.

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

Womack, C. B., Giani, P., Eastham, S. D., and Selin, N. E.: Rapid Emulation of Spatially Resolved Temperature Response to Effective Radiative Forcing, Journal of Advances in Modeling Earth Systems, 17, e2024MS004 523, https://doi.org/https://doi.org/10.1029/2024MS004523, e2024MS004523 2024MS004523, 2025.

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

Zeng, Z., Kim, J.-H. J., Tan, H., Hu, Y., Cameron-Rastogi, P., Villa, D., New, J., Wang, J., and Muehleisen, R. T.: A review of future weather data for assessing climate change impacts on buildings and energy systems, Renewable and Sustainable Energy Reviews, 212, 115 213, https://doi.org/https://doi.org/10.1016/j.rser.2024.115213, 2025.

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

https://doi.org/10.5194/egusphere-2026-2029
Preprint. Discussion started: 18 May 2026

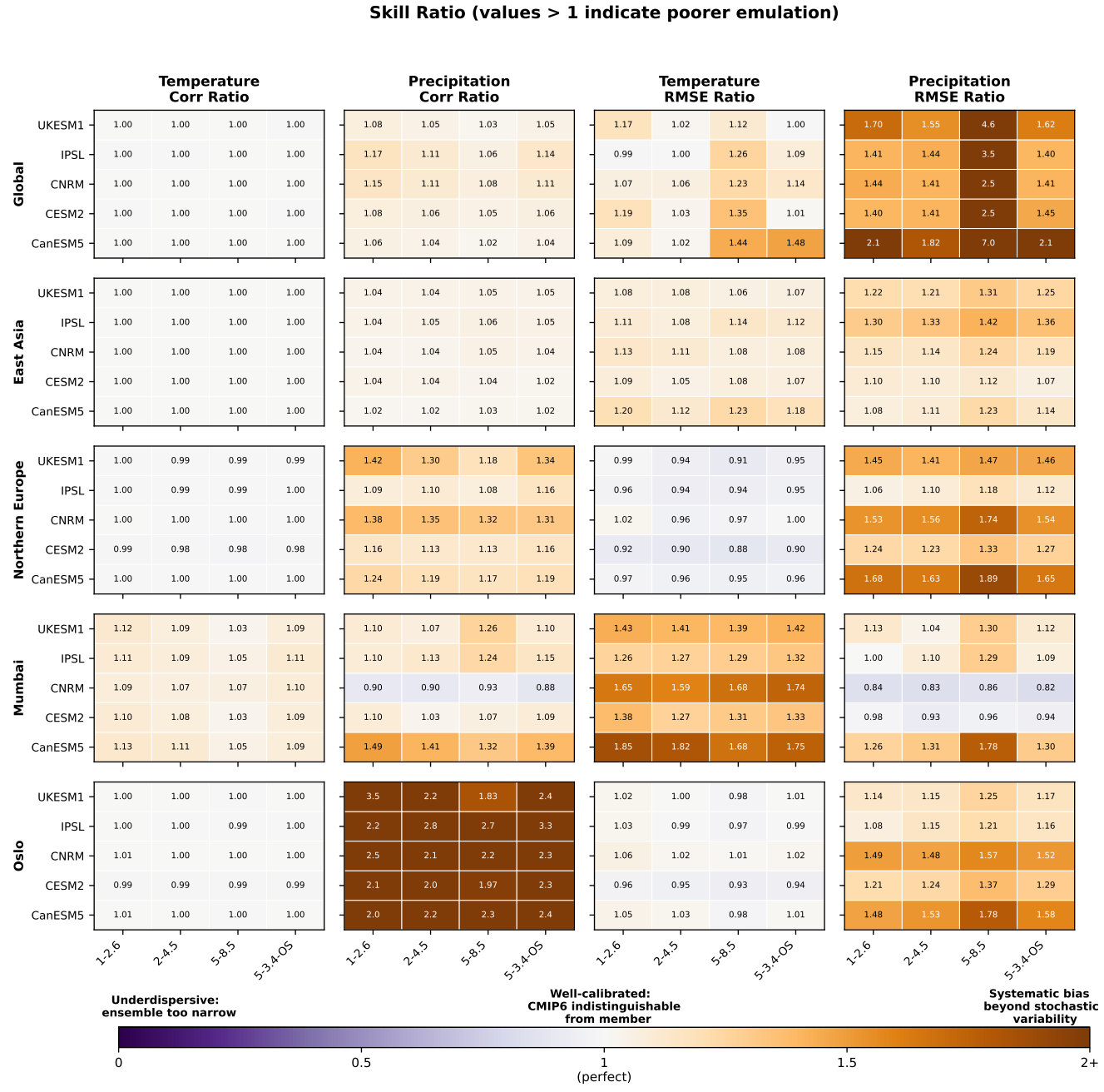

Figure 10. Skill ratio analysis for out-of-sample validation across all spatial scales (rows), models, and scenarios. Each panel shows a grid of model-scenario combinations. Skill ratios are defined such that values >1 (orange) indicate poorer emulation for both metrics: for correlation, ratio = Member-Member / Member-CMIP6; for RMSE, ratio = Member-CMIP6 / Member-Member. Values near 1 (white) indicate CMIP6 is statistically indistinguishable from a METEOR ensemble member; values <1 (purple) suggest ensemble underdispersion. Temperature ratios are consistently near unity; precipitation shows more variability, particularly at global scales where forced response bias is evident.