https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

Beyond the simple mean: an alternative way to improve multi-model bottom-up wetland CH_4 estimates

Nikola Besic[1,2], Philippe Ciais[3], Fa Li[4,5], Robert Jackson[4], Kunxiaojia Yuan[6], Shushi Peng[7], Benjamin Poulter[8], Zhen Zhang[9], and Qing Zhu[10]

[1]Université de Lorraine, Géodata Paris, IGN, LIF, Nancy, France
[2]Université Gustave Eiffel, Géodata Paris, IGN, LIF, Nancy, France
[3]LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris Saclay, Gif-sur-Yvette, France
[4]Doerr School of Sustainability, Department of Earth System Science, Stanford University, Stanford, CA, USA
[5]Department of Earth and Planetary Sciences, Jackson School of Geosciences, The University of Texas at Austin, Austin, TX, USA
[6]Department of Earth and Atmospheric Sciences, University of Houston, Houston, TX, USA
[7]Institute of Carbon Neutrality, Sino-French Institute for Earth System Science, College of Urban and Environmental Sciences, and Laboratory for Earth Surface Processes, Peking University, Beijing, China
[8]Spark Climate Solutions, San Francisco, CA, USA
[9]State Key Laboratory of Tibetan Plateau Earth System, Environment and Resources, and Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China
[10]Climate and Ecosystem Sciences Division, Climate Sciences Department, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Correspondence: Nikola Besic (n.m.besic@gmail.com, nikola.besic@ign.fr), and Philippe Ciais (philippe.ciais@lsce.ipsl.fr)

Abstract. Wetlands are the largest natural source of atmospheric methane, yet bottom-up estimates of their emissions remain highly uncertain due to structural differences, parametric uncertainties among process-based models and strong environmental heterogeneity. In addition, most wetland emission models are not systematically calibrated against methane flux measurements, limiting their ability to capture realistic spatial and temporal dynamics. The increasing availability of eddy-covariance towers measuring CH_4 fluxes across diverse wetland types now offers new opportunities to evaluate and constrain model ensembles using observational data. Multi-model ensembles are commonly used to quantify uncertainty, but the widespread use of simple model averaging implicitly assumes that all models contribute equally and optimally across sites, an assumption that is rarely justified. Here, we present a data-driven framework that moves beyond the simple mean to derive site-adaptive ensemble estimates and to characterize spatial patterns of model disagreement, and how they are reshaped by ensemble weighting, in wetland CH_4 emissions.

Using flux observations from a global network of 44 wetland sites, ranging from boreal arctic to tropical wetland ecosystems, and simulations from sixteen global wetland biogeochemistry models, we estimate site-specific optimal ensemble weights via a Bayesian model averaging framework fitted by Expectation–Maximization (EM-BMA). To improve robustness, weights are stabilized through resampling, and sites are clustered based on their stabilized weight signatures in compositional space, yielding groups of locations with similar model skill structures. Within each cluster, we perform cross-validated predictions and compare EM-BMA against simple model averaging (SMA) using standard performance metrics (R^2, normalized RMSE, and mean bias).

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

To interpret these clusters and to delineate the conditions under which different model combinations are best fitting the site measurements, we relate cluster membership to environmental predictors. We characterize cluster-specific validity domains in predictor space using low-dimensional projections and geometric and probabilistic envelopes, and we identify the most influential predictors using a machine-learning classifier with back-projected feature importance.

We show that optimal model combinations vary across sites and that Bayesian model averaging outperforms simple model averaging in cross-validation. When propagated beyond the measurement sites using environmental predictors and applied to model diagnostic outputs, the resulting global wetland CH_4 emission estimates differ only slightly from those obtained with simple averaging (less than 5%). However, substantial differences emerge at local and regional scales, highlighting the importance of accounting for spatial heterogeneity in model skill. This framework therefore provides a transparent and reproducible alternative to equal-weight ensemble means for improving bottom-up wetland CH_4 estimates across heterogeneous environments.

1 Introduction

Atmospheric methane (CH_4) is a potent greenhouse gas whose concentration continues to rise, contributing substantially to near-term climate warming (Canadell et al., 2021). Wetlands are the largest natural source of atmospheric methane, yet their emissions remain among the most uncertain components of the global methane budget (Saunois et al., 2025). This uncertainty reflects both the strong environmental heterogeneity of wetland systems and substantial structural differences among process-based models used to simulate methane production, oxidation, and transport. Reducing uncertainty in wetland methane emissions is therefore a central challenge for understanding the global methane cycle, attributing observed trends, and anticipating climate–carbon feedbacks (Zhang et al., 2023; Ciais et al., 2026).

Over the past decade, considerable progress has been made in constraining methane emissions using atmospheric observations and inverse modeling frameworks (e.g., Bloom et al., 2017; Zhang et al., 2017b; Koffi et al., 2020; Peng et al., 2022; Zhang et al., 2023). However, their ability to reduce uncertainties in natural emissions, particularly from wetlands, remains limited. This is largely due to the sparse coverage of atmospheric measurements in key regions such as the tropics and high latitudes, where wetland emissions are most uncertain. As a result, large discrepancies persist between top-down and bottom-up estimates, and wetland methane emissions continue to represent one of the dominant sources of uncertainty in the global methane budget (Saunois et al., 2025). Recent syntheses further emphasize that this uncertainty limits trend detection and attribution, especially in the tropics and high latitudes, and that progress will require better integration of observations, models, and ensemble-based uncertainty characterization (Zhu et al., 2024; Forbrich et al., 2024; Zhu et al., 2025; Zhang et al., 2025; Ciais et al., 2026).

Bottom-up estimates of wetland methane emissions are typically derived from ensembles of land surface models, dynamic global vegetation models, and ecosystem biogeochemical models (Saunois et al., 2025). Multi-model ensembles are widely used to represent structural uncertainty and to provide a more robust estimate than any single model. Large-scale intercomparison efforts, such as the Wetland and Wetland CH_4 Intercomparison of Models Project (WETCHIMP; Wania et al., 2013;

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

Melton et al., 2013), have demonstrated substantial spread among wetland methane simulations, highlighting both structural diversity and differences in hydrological and biogeochemical parameterizations across models.

In practice, ensemble means are most often computed using simple model averaging, implicitly assuming that all models contribute equally and optimally across regions and environmental conditions. This assumption is rarely justified. Models differ in their representation of wetland extent, hydrology, soil biogeochemistry, and methane transport pathways, and their relative performance can vary strongly across sites and climatic regimes. Moreover, participating models are typically not calibrated against a common and recent set of site-level methane flux observations, further contributing to ensemble spread. As a result, equal-weight ensemble means may mask systematic regional strengths and weaknesses and may fail to exploit the complementary skill of different models.

Several studies have explored alternative ways to combine model information or to benchmark model ensembles against observations, including skill-based weighting and hybrid top-down/bottom-up approaches (e.g., Bloom et al., 2017; Koffi et al., 2020; Peng et al., 2022). At the same time, the increasing availability of eddy-covariance measurements of wetland CH_4 fluxes across diverse climatic zones now offers new opportunities to evaluate ensemble performance at the process scale. Yet, most existing ensemble approaches either operate at coarse spatial scales, focus primarily on atmospheric constraints, or apply globally uniform weighting schemes. There remains a need for frameworks that (i) operate at the site scale where process understanding and evaluation are most direct, (ii) explicitly account for spatial heterogeneity in model skill, and (iii) provide transparent and reproducible uncertainty characterization that can inform both bottom-up modeling and observational network design (Forbrich et al., 2024; Ciais et al., 2026).

Here, we aim at addressing this gap by proposing a data-driven, site-adaptive ensemble framework for wetland methane emissions that moves beyond simple model averaging. Using observations from a global network of wetland sites (Delwiche et al., 2021; Kuhn et al., 2021) and simulations from sixteen process-based methane models, we estimate site-specific ensemble weights using a Bayesian model averaging approach fitted by Expectation–Maximization (EM-BMA) (Raftery et al., 2005; Besic et al., 2025; Picard et al., 2025). This allows the relative contribution of each model to vary across sites, reflecting differences in local performance and structural adequacy. To improve robustness, the inferred weights are stabilized through resampling, and sites are subsequently clustered in the resulting compositional weight space, deriving groups of locations that share similar model skill structures.

Building on this clustering, we perform cross-validated predictions within clusters and systematically compare the performance of EM-BMA against simple model averaging using standard evaluation metrics. Beyond this predictive assessment, we aim to interpret and generalize the site-level patterns of model skill identified by the clustering. To this end, we relate cluster membership to environmental predictors describing vegetation productivity, climate, and soil properties, delineate cluster-specific validity domains in predictor space using low-dimensional projections and geometric and probabilistic envelopes, and identify the most influential predictors using a supervised machine-learning classifier. This second step enables the spatial extrapolation of cluster membership beyond measurement sites and provides a pathway toward identifying where different combinations of models are expected to perform best.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

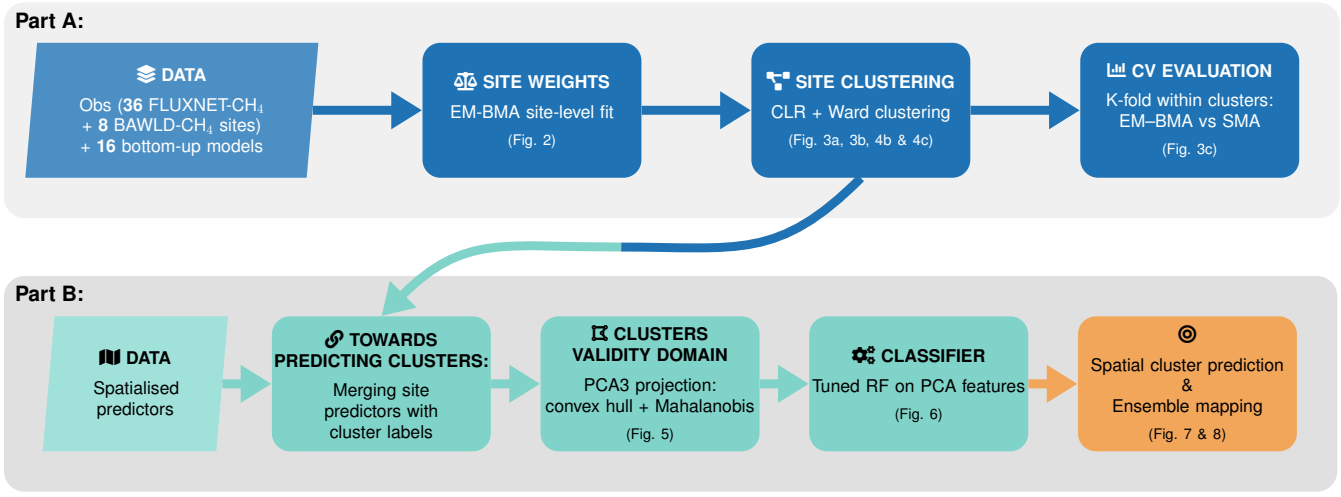


Figure 1. Overview of the study workflow and main processing steps.

The overall workflow of the study is summarized in Fig. 1. In brief, we combine observational site level flux constraints from FLUXNET-CH_4 v1 and BAWLD (Delwiche et al., 2021; Kuhn et al., 2021; Olefeldt et al., 2021; Virkkala et al., 2025) with multi-model simulations to derive site-specific ensemble weights using Bayesian model averaging, stabilize these weights through resampling, and cluster sites based on similarities in their weight signatures. The resulting clusters are then linked to environmental predictors to obtain spatially explicit validity domains and to predict cluster membership across global wetlands. Finally, the inferred cluster-specific model mixtures are applied to the diagnostic outputs of the participating models in order to map spatially constrained wetland methane emissions and to assess their implications for regional and global budgets.

By explicitly accounting for spatial heterogeneity in model performance, this framework provides a transparent and reproducible alternative to equal-weight ensemble means. More broadly, it offers a practical way to characterize structural differences among models and to support the integration of observations and process-based simulations for improved methane budget assessment and monitoring (Bloom et al., 2017; Peng et al., 2022; Zhang et al., 2023; Ciais et al., 2026).

2 Data description

2.1 Models

We analyze methane (CH_4) emissions simulated by sixteen process-based terrestrial ecosystem and land-surface models. These models, presented in Table 1 differ in their representation of wetland extent, hydrology, soil biogeochemistry, and methane production, oxidation, and transport processes, collectively spanning a broad range of structural assumptions and complexities.

Although these models share the common goal of representing wetland methane emissions, they differ substantially in their underlying process formulations, spatial and temporal resolutions, and treatment of hydrology and soil biogeochemistry, motivating the ensemble-based and site-adaptive analysis developed in this study.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

Model	Description	Reference
CH4MOD	Simulates methane emissions from wetlands by representing methanogenic substrate supply and processes of production, oxidation, and emission under varying soil and climate conditions.	Li et al. (2010).
CLASSIC	Simulates surface energy, water, carbon, and nutrient fluxes; wetland methane emissions are parameterized within its biogeochemical component.	Melton et al. (2020).
DLEM	Represents coupled carbon, nitrogen, and water cycles, including a wetland methane module for greenhouse gas fluxes.	Zhang et al. (2017a).
ELM	Includes vertically resolved soil biogeochemistry and methane processes for wetland environments.	Huang et al. (2025).
ISAM	Simulates terrestrial carbon and nitrogen dynamics with methane emissions sensitive to soil hydrology and temperature.	Shu et al. (2020).
JSBACH	Represents terrestrial carbon and water fluxes with parameterizations for wetland methane emissions.	Kaiser et al. (2017).
JULES	Simulates land–atmosphere exchanges of energy, water, and carbon, including process-based wetland methane emissions.	Best et al. (2011).
LPJ-GUESS	Represents vegetation dynamics at individual/cohort level with explicit wetland methane processes.	Kallingal et al. (2024).
LPJ-MPI	Simulates vegetation dynamics and terrestrial carbon cycling, extended to include wetland methane emissions.	Sitch et al. (2003).
LPJ-WSL	Variant of LPJ with methane module coupled to soil carbon and hydrology.	Zhang et al. (2017a).
LPX-Bern	Represents vegetation, carbon cycling, and wetland methane emissions with focus on long-term carbon–climate interactions.	Stocker et al. (2014).
ORCHIDEE	Simulates coupled energy, water, and carbon fluxes with process-based wetland methane emissions.	Salmon et al. (2022).
SDGVM	Represents vegetation dynamics and terrestrial carbon cycling, applied to wetland methane simulations.	Ringeval et al. (2013).
TEM-MDM	Couples carbon, nitrogen, and methane processes to simulate wetland methane production, oxidation, and emission.	Zhuang et al. (2004).
TRIPLEX-GHG	Simulates carbon, nitrogen, and greenhouse gas fluxes, including wetland methane emissions.	Zhu et al. (2014).
VISIT	Simulates carbon cycling and trace gas exchanges, including wetland methane emissions under varying environmental conditions.	Ito and Inatomi (2012).

Table 1. Overview of wetland methane models used in this study.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

The simulations analyzed here follow a common protocol established within recent wetland methane model intercomparison efforts (Zhang et al., 2025). The ensemble provides monthly methane fluxes at the global scale over the period 2000–2020, based on harmonized environmental forcings and standardized experimental setups.

A key source of structural diversity within the ensemble arises from the representation of wetland extent. Some models simulate wetland dynamics prognostically, using internal hydrological schemes to determine inundation extent as a function of climate and soil conditions, while others rely on externally prescribed wetland extent derived from remote sensing products. This distinction contributes to differences in both the spatial distribution and temporal variability of methane emissions across models, and constitutes an important component of between-model uncertainty explored in this study.

In our analysis, we further restrict the dataset by excluding all site–month observations with zero measured CH_4 emissions. This choice ensures that the evaluation focuses on conditions where methane production is observed, and avoids penalizing models for predicting emissions in situations where no flux is detected. As a consequence, the framework emphasizes the ability of models to reproduce variability in active emitting conditions rather than their performance in predicting the occurrence of zero-flux events.

For the methodological development and evaluation presented throughout this study, we primarily use simulations driven by the CRU climate forcing dataset (Harris et al., 2020). These simulations are used to estimate model weights, perform clustering, and illustrate the proposed framework. To assess the implications of the approach for large-scale emission estimates, we additionally analyze simulations driven by both CRU and GSWP3 climate forcings (Dirmeyer et al., 2006). This allows us to examine how the regime-dependent ensemble weighting affects global wetland methane emission estimates under two commonly used meteorological forcing datasets.

We used the Wetland Area Dynamics for Methane modeling (WAD2M) (Zhang et al., 2021) dataset to derive a common wetland mask for all model outputs. Specifically, we reprojected the mean WAD2M wetland fraction over 2000–2018 onto the target $0.5°$ grid and classified as wetlands all pixels with mean fractional wetland cover $F_w \geq 0.01$. Since the model outputs already account for wetland fractional area in their annual gridded emissions, WAD2M was not used to rescale emissions. Instead, it was used solely to define a harmonized wetland domain and to replace missing annual values by zero within wetland pixels prior to temporal averaging. This ensured consistency across models when deriving mean annual emissions and when later applying the two ensemble-weighting approaches, SMA and BMA.

2.2 Measurement sites

Observational constraints on wetland methane (CH_4) emissions were compiled from two complementary sources: the FLUXNET-CH_4 network and the Boreal–Arctic Wetland and Lake Dataset (BAWLD). Eddy-covariance methane flux measurements from FLUXNET-CH_4 provide site-level, continuous observations across a range of wetland ecosystems and climatic conditions (Delwiche et al., 2021). These data are complemented by site-based and regionally representative flux information from the BAWLD synthesis, which integrates observations across boreal and Arctic wetlands and inland water systems (Kuhn et al., 2021; Olefeldt et al., 2021).

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

The site network comprises 44 wetland flux sites (36 FLUXNET-CH_4 and 8 BAWLD-CH_4 sites), including 26 located in boreal regions ($|Lat| > 50°$), 14 in temperate regions ($23.5° \leq |Lat| \leq 50°$), and 4 in tropical regions ($|Lat| < 23.5°$).

For consistency with model outputs and to reduce the influence of high-frequency variability and data gaps, all observational methane fluxes were aggregated to the monthly timescale prior to analysis. The resulting dataset provides a harmonized set of site-level monthly methane emission estimates used to evaluate model performance and to derive site-adaptive ensemble weights.

Methane fluxes were originally expressed in $\mathrm{mg\ m^{-2}\ s^{-1}}$. For numerical conditioning of the covariance estimation within the expectation–maximization framework, fluxes were linearly rescaled by a constant factor. This scaling does not affect statistical relationships or model weights.

2.3 Predictors

To interpret the spatial structure of model skill and to enable the upscaling of site clusters - defined as groups of sites with similar weight signatures - we assembled a set of environmental predictors describing vegetation productivity, climate, and soil properties. All predictors were expressed as long-term means over the period 2003–2022, in order to emphasize persistent environmental controls and to ensure consistency with the climatological nature of the cluster-based analysis.

Vegetation productivity was characterized using gross primary production (GPP) from the GOSIF dataset (Li and Xiao, 2019), which provides proxies of methane emissions linked to vegetation activities such as substrate availability and plant-mediated transport (Knox et al., 2021; Yuan et al., 2022, 2024; Li et al., 2026). Climate conditions were represented by near-surface air temperature (TA), precipitation (Prcp), and soil water content (SWC) from the ERA5 reanalysis (Hersbach et al., 2020), capturing key thermal and hydrological controls on wetland methane production, oxidation, and transport.

Soil properties were described using the Harmonized World Soil Database (HWSD) version 2.0 (Nachtergaele et al., 2023). From this dataset, we extracted soil texture information expressed as fractional proportions of the main texture classes (clay, silt, sand, and gravel) as well as soil pH, which together provide first-order constraints on soil physical and biogeochemical conditions relevant for methane cycling (Le Mer and Pierre Roger, 2001; Li et al., 2026).

All predictors were spatially collocated with the observation sites and compiled into a single predictor dataset. This predictor set was subsequently used to (i) relate cluster membership to environmental conditions, (ii) delineate cluster-specific validity domains in predictor space, and (iii) support the spatial extrapolation of cluster membership and ensemble configurations.

3 Method description

Our methodological framework is organized into two main parts (Fig. 1). Part A focuses on deriving site-adaptive ensemble estimates from multi-model simulations using Bayesian model averaging (BMA), stabilizing the resulting weights, clustering sites based on their weight signatures, and evaluating predictive performance against simple model averaging (SMA). Part B aims to interpret and generalize these clusters in environmental space by linking them to predictors, delineating cluster-specific validity domains, and training a classifier to enable spatial upscaling.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

3.1 Part A: Site-adaptive Bayesian model averaging, stabilization, clustering, and evaluation

3.1.1 Bayesian model averaging with Expectation–Maximization

Let Y denote the wetland methane flux at a given site and time i, and let $m_k(\boldsymbol{x}_i)$ be the corresponding simulation from model M_k, with $k = 1, \ldots, K$ and $K = 16$, and the input data $\boldsymbol{x}_i$. We assume that observations are generated from a Bayesian model averaging mixture over the K candidate models, such that the conditional distribution of Y given the observations $\mathcal{Y}$ can be written as

$$f(Y_i \mid \mathcal{Y}) = \sum_{k=1}^{K} w_k \, \phi(Y_i; m_k(\boldsymbol{x}_i), \sigma_k), \tag{1}$$

where $w_k = \Pr(M_k \mid \mathcal{Y})$ are non-negative model weights that sum to one, $\phi(\cdot)$ denotes the Gaussian probability density function, and σ_k^2 are model-specific error variances (Besic et al., 2025).

For each site, the weights $\mathbf{w} = (w_1, \ldots, w_K)$ and variances $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_K)$ are estimated by maximizing the log-likelihood of the observed time series using an Expectation–Maximization (EM) algorithm. In the expectation step (E), posterior responsibilities are computed for each model component given current parameters. In the maximization step (M), weights and variances are updated from these responsibilities, subject to numerical regularization floors on weights, variances, and likelihood denominators to ensure stability. Iteration continues until convergence of the log-likelihood or a maximum number of iterations is reached (Picard et al., 2012).

Given a fitted weight vector $\mathbf{w}$, the BMA prediction at time i is given by the conditional expectation of the mixture distribution,

$$\hat{Y}_i^{\mathrm{BMA}} = \mathbb{E}(Y_i \mid \mathcal{Y}) = \sum_{k=1}^{K} \tilde{w}_{ik} \, m_k(\boldsymbol{x}_i), \tag{2}$$

where $\tilde{w}_{ik}$ are renormalized weights accounting for missing model values at time i. For comparison, the simple model average (SMA) prediction is defined as the unweighted mean of the available model simulations.

3.1.2 Stabilization of site-level weights

Because site-level datasets can be limited in length and may contain gaps, raw EM-BMA weights are sensitive to sampling variability. To improve robustness, we stabilize the estimated weights using a resampling strategy. At each site, we generate multiple resampled datasets either by bootstrap resampling with replacement or by subsampling without replacement. The EM-BMA procedure is applied independently to each replicate, yielding an ensemble of weight vectors.

From this ensemble, we compute the mean stabilized weight vector $\bar{\mathbf{w}}$, which is subsequently used as the site's representative weight signature, and the standard deviation of weights, which provides a diagnostic of weight stability. If too few valid data points are available in a resample to fit the EM model, that replicate is discarded. If no valid replicate can be fitted, weights are estimated from the full dataset.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

3.1.3 Clustering of sites in weight space

To identify groups of sites that share similar model skill structures, we cluster sites based on their stabilized weight signatures. Because the weights are compositional data (non-negative and summing to one), we first apply a centered log-ratio (CLR) transform:

$$\mathrm{CLR}(w_k) = \log(w_k) - \frac{1}{K}\sum_{j=1}^{K}\log(w_j), \tag{3}$$

with a small numerical offset applied to avoid taking the logarithm of zero (Egozcue et al., 2003).

The transformed weight vectors are then clustered using agglomerative hierarchical clustering with Ward's linkage (Kaufman and Rousseeuw, 1990). The number of clusters is chosen using the elbow criterion, subject to the constraint that no cluster contains only a single site. Cluster centroids are computed in the original weight space for interpretability, and cluster-level summaries of weight stability are derived from the corresponding standard deviations.

3.1.4 Cross-validation and performance metrics

To evaluate predictive performance, we perform a K-fold cross-validation within each cluster. For each fold and each cluster, the EM-BMA model is fitted on the training subset of fluxes, and predictions are generated for the validation subset using both BMA and SMA. This ensures that model weights are always estimated independently of the validation data.

Performance is assessed using the coefficient of determination (R^2), the normalized root-mean-square error (NRMSE, normalized by the standard deviation of observations), and the mean bias error (MBE). In addition, bootstrap resampling stratified by cluster is used to derive confidence intervals for global performance metrics.

3.1.5 Variance decomposition

Beyond the mean prediction, the predictive variance of the BMA ensemble can be decomposed following Raftery (1993) as

$$\begin{aligned}\mathrm{Var}(Y_i \mid \mathcal{Y}) &= \sigma_{bm}^2 + \sigma_{wm}^2 = \\ &= \sum_{k=1}^{K} w_k \left(m_k(\boldsymbol{x}_i) - \sum_{l=1}^{K} w_l\, m_l(\boldsymbol{x}_i) \right)^2 \\ &+ \sum_{k=1}^{K} w_k\, \sigma_k^2,\end{aligned} \tag{4}$$

where the first term represents the between-model variance (σ_{bm}^2), i.e. the weighted spread of individual model predictions around the BMA mean, and the second term represents the within-model variance (σ_{wm}^2), i.e. the weighted average of the model-specific error variances.

The between-model component quantifies structural disagreement among models, and becomes large when models with comparable weights produce contrasted predictions. The within-model component reflects the typical model–data mismatch captured by the Gaussian error terms. This decomposition provides a natural way to distinguish uncertainty arising from model

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

structural spread from uncertainty associated with residual errors (the need for such a separation is discussed in Ciais et al. (2026)), and forms the basis for the cluster-level variance analysis presented in Sect. 4.

3.2 Part B: Linking clusters of sites in weight space to environmental predictors and spatial generalization

3.2.1 Preparation of predictor matrix

As evoked in Sec. 2.3, environmental predictors describing vegetation productivity, climate, and soil properties are assembled at each site and expressed as long-term means over 2003–2022. The predictor matrix is merged with cluster labels derived in Part A. Missing predictor values are imputed using median imputation, and all predictors are standardized to zero mean and unit variance prior to multivariate analysis.

3.2.2 Dimensionality reduction and validity domains

To visualize and characterize the distribution of clusters in predictor space, we apply principal component analysis (PCA) to the standardized predictor matrix and retain the first three components. These components are used to construct pairwise two-dimensional projections.

For each cluster and each projection, we delineate validity domains using two complementary approaches: (i) the convex hull of the projected points, which provides a geometric envelope of occupied predictor space (Preparata and Shamos, 1985), and (ii) a Mahalanobis-distance-based ellipse derived from the cluster covariance matrix, corresponding to a prescribed confidence level (Johnson and Wichern, 2007). These envelopes provide an interpretable representation of the environmental conditions under which each cluster is supported by observations.

3.2.3 Supervised classification of cluster membership

To enable prediction of cluster membership from environmental predictors, we train a supervised classifier using a pipeline consisting of median imputation, standardization, PCA with a tunable number of components, and a random forest classifier (Pedregosa et al., 2011). Hyperparameters, including the number of retained PCA components and random forest settings, are optimized using randomized search with stratified cross-validation. To avoid unstable behavior, clusters with fewer than a minimum number of sites can be excluded from the training set.

Model performance is evaluated using cross-validated predictions and standard classification metrics, including overall accuracy, balanced accuracy, macro-averaged F1 score, and the confusion matrix. Feature importance scores from the trained random forest are back-projected onto the original predictor space using the PCA loadings to identify the most influential environmental controls on cluster membership.

3.2.4 Towards spatial upscaling

The trained classifier provides a practical pathway to predict cluster membership in regions without direct flux observations, based solely on environmental predictors. Combined with the cluster-specific ensemble configurations derived in Part A, this

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

enables spatial mapping of dominant model combinations and associated validity domains, thereby supporting the generalization of site-level results to larger spatial scales.

4 Results

4.1 Site-level EM-BMA behavior and comparison with simple averaging

We first examine the behavior of the EM-BMA framework at the site level using representative examples spanning contrasting environmental conditions and flux regimes (Fig. 2). For each site, EM-BMA yields a distinct set of model weights, reflecting differences in local model skill, and produces ensemble predictions that track observed variability more closely than the simple model average (SMA) in most cases.

Across the illustrated sites, the inferred weights are typically sparse, with a limited subset of models receiving most of the weight, while others contribute marginally. This behavior highlights that model skill is strongly site dependent and that uniform weighting is rarely optimal. In the time series and scatter comparisons, BMA predictions generally show reduced bias and improved alignment with observations compared to SMA, although the magnitude of improvement varies across sites and regimes.

4.2 Cluster structure in weight space and global performance

Clustering the stabilized site-level weight signatures reveals four distinct clusters of sites characterized by different dominant model combinations (Fig. 3a). The cluster centroids show distinct weight patterns, indicating that different subsets of models tend to perform best in different groups of sites. The corresponding mean weight standard deviations (Fig. 3b) are generally small, indicating that the stabilization procedure yields robust cluster-level weight structures.

When aggregating predictions across all sites, EM-BMA outperforms SMA (Fig. 3c). The global scatter plot shows that BMA achieves a higher coefficient of determination and a lower normalized RMSE and bias than SMA, which exhibits both larger dispersion and stronger systematic deviations from the one-to-one line. This confirms that allowing site-adaptive weights leads to an overall improvement compared to uniform model averaging.

4.3 Within- and between-model variance across clusters

The cross-validated variance decomposition further highlights structural differences between clusters (Fig. 4a). For some clusters, the within-model component (i.e., the weighted sum of model-specific error variances) dominates, indicating that uncertainty is primarily driven by model–data mismatch. In other clusters, the between-model component (i.e., the ensemble spread around the weighted mean) is comparable to or larger than the within-model term, pointing to substantial structural disagreement among models in those regimes.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

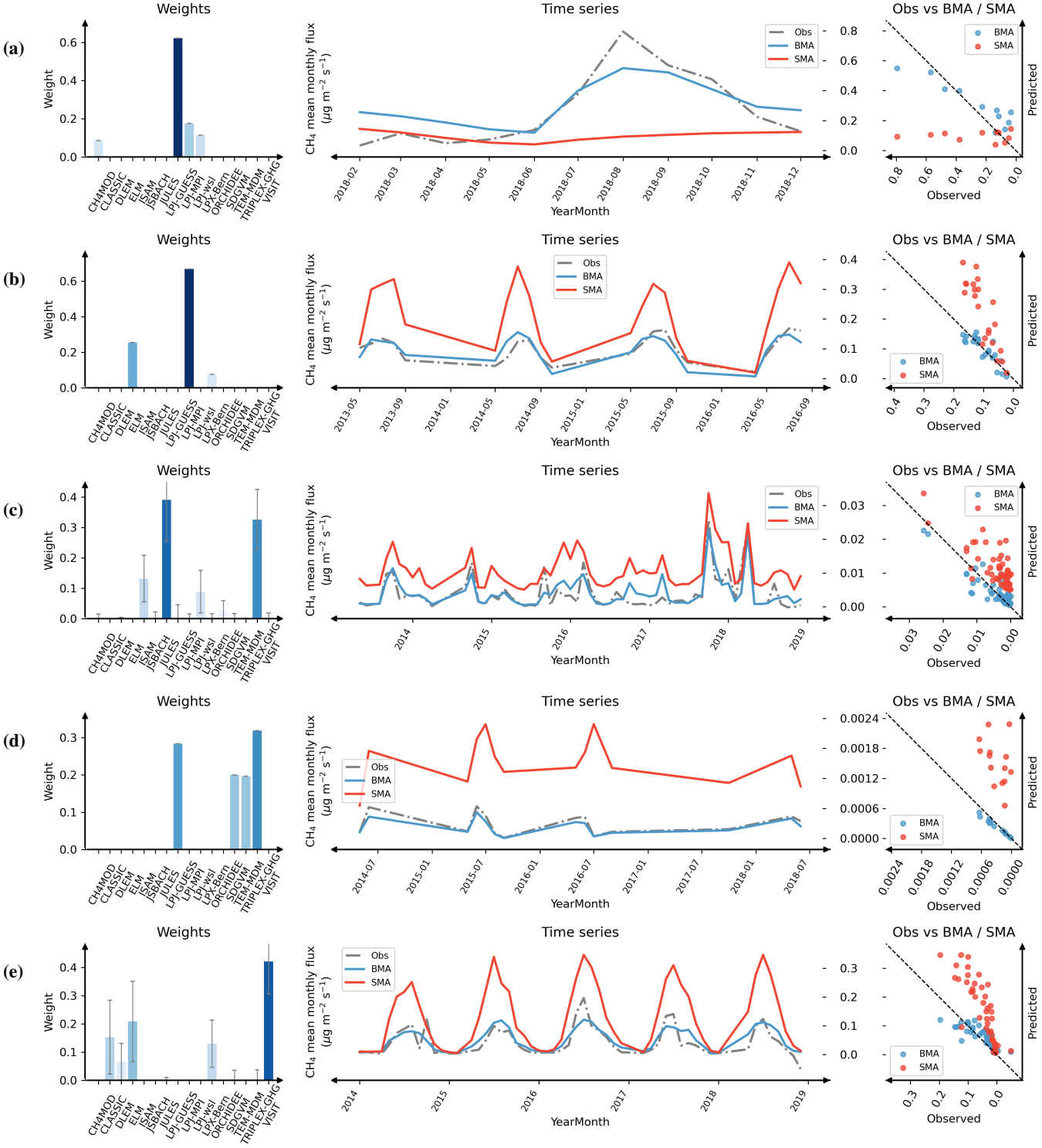


Figure 2. Examples of site-level EM-BMA fits at wetland sites: (a) BWNXR (Botswana), (b) CASCC (Canada), (c) DEZRK (Germany), (d) SEDEG (Sweden), and (e) USLOS (USA). For each site, the left panel shows the inferred model weights (with stabilization uncertainty), the middle panel presents the corresponding time series, and the right panel displays the observation–prediction comparison for BMA and SMA.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

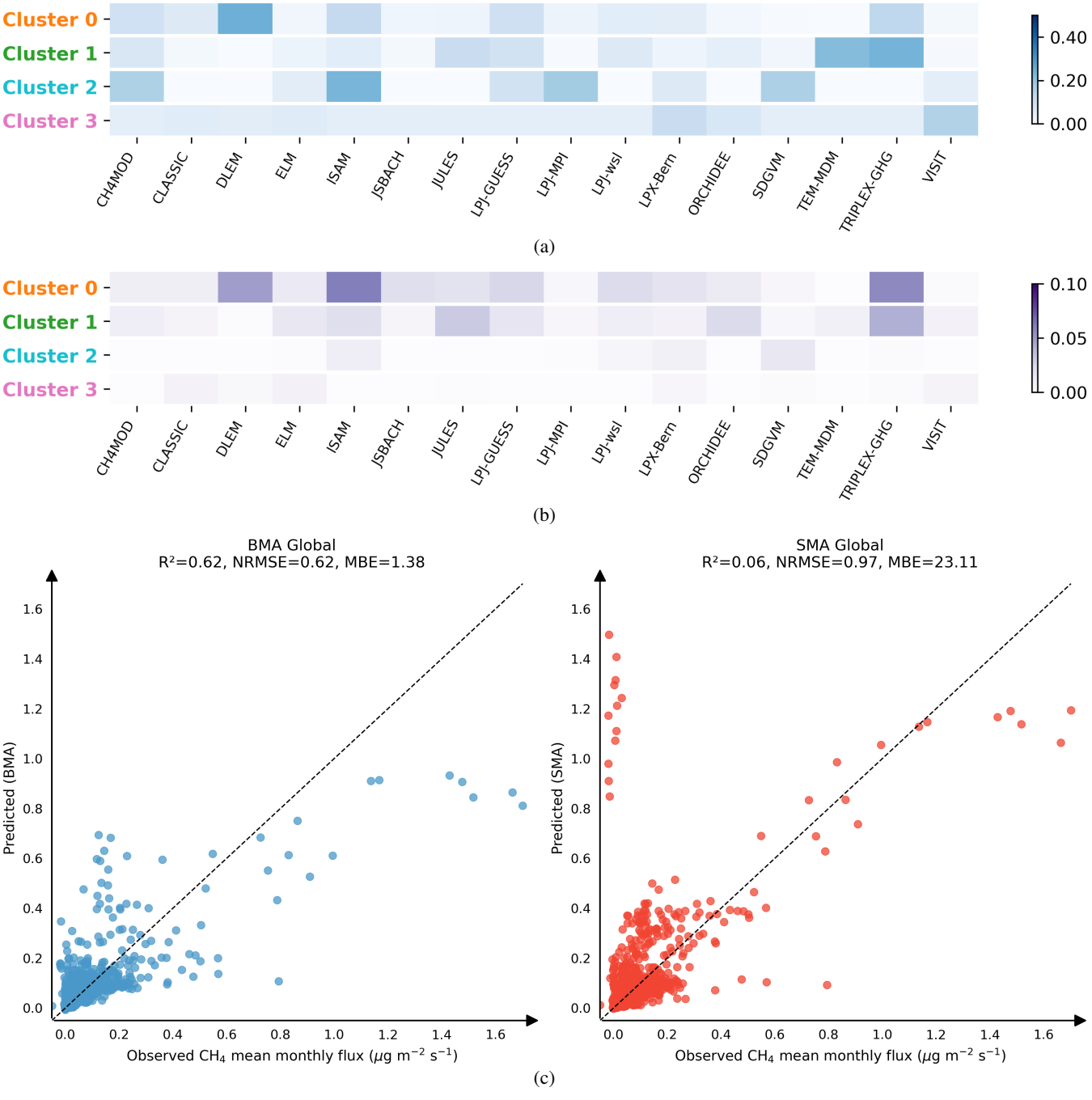


Figure 3. (a) Cluster centroid weights in the original model space, (b) mean standard deviation of stabilized weights per cluster (as a measure of weight stability), and (c) global scatter comparison between observations and predictions from EM-BMA and simple model averaging (SMA). Reported metrics summarize overall performance across all sites.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

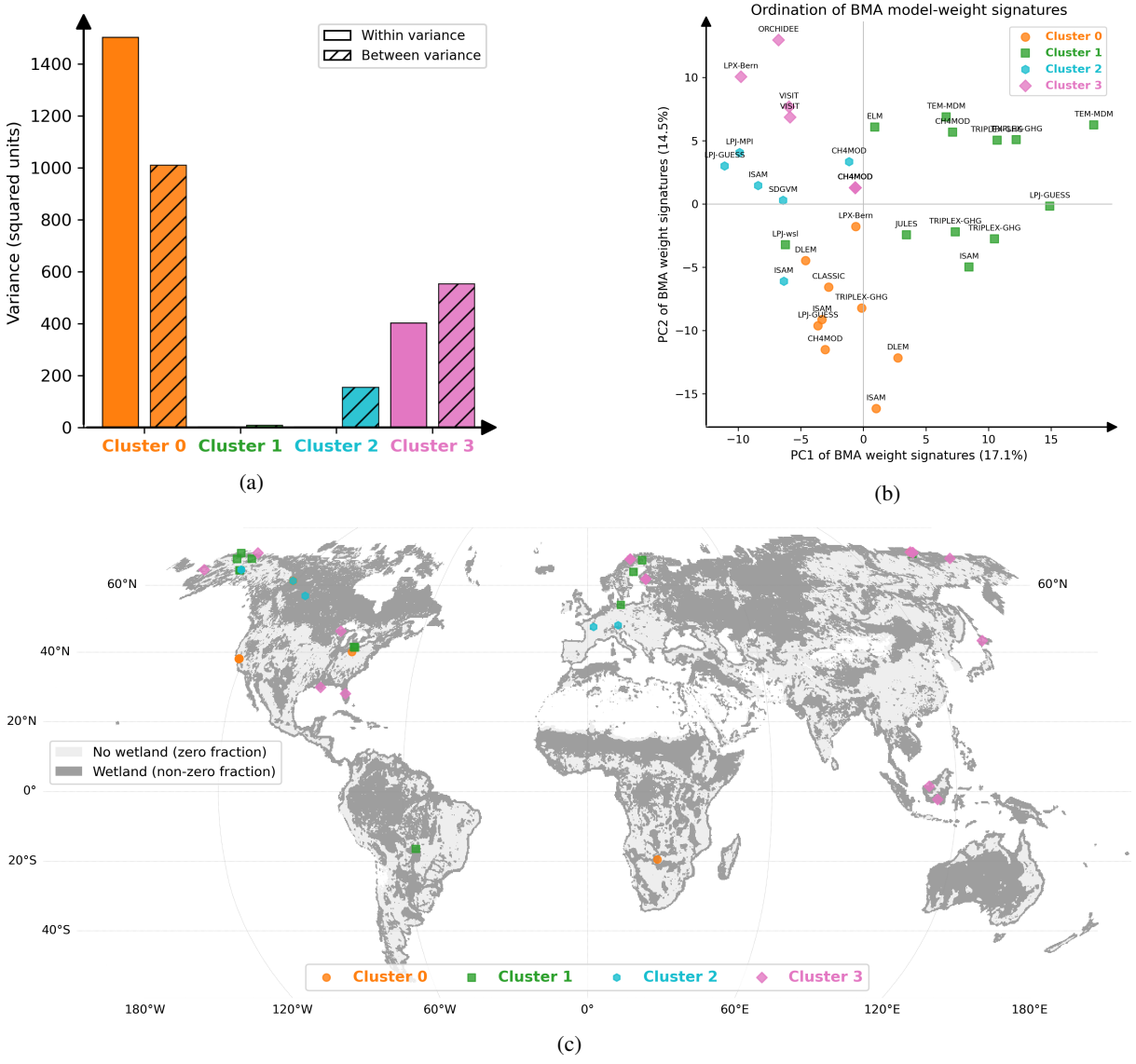


Figure 4. (a) Cross-validated decomposition of BMA predictive variance by cluster into within-model (weighted error variance) and between-model (ensemble spread on validation data) components. (b) PCA of stabilized BMA weight signatures across sites - points represent sites, coloured by cluster and labelled by their dominant model. (c) Geographic distribution of measurement sites colored by cluster membership, showing the spatial organization of model-skill regimes. Darker gray areas in the background indicate wetlands according to the Wetland Area Dynamics for Methane modeling (WAD2M) (Zhang et al., 2021).

These contrasts show that uncertainty is not only heterogeneous in magnitude across sites, but also in nature: some environ-
285 ments are characterized by relatively consistent but biased models, whereas others exhibit large structural spread among model simulations.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

4.4 Ordination of model-weight signatures

To investigate whether the clusters reflect distinct regimes of model performance, we performed a principal component analysis (PCA) on the stabilized BMA weight signatures across sites. The first two principal components explain 17.1% and 14.5% of the variance, respectively (Fig. 4b).

The ordination reveals a clear separation of sites into four groups consistent with the clusters identified previously. In particular, the first axis primarily separates Cluster 1 from Cluster 2, while the second axis distinguishes Cluster 3 from Cluster 0, indicating that clusters correspond to distinct configurations of model weights rather than arbitrary partitions.

When labeling sites by their dominant model (i.e. the model receiving the highest BMA weight), distinct regions of the ordination space are associated with different models. While no single model dominates globally, specific models or model families tend to prevail within particular regions of the space. This confirmes that model performance varies systematically across sites, with different models receiving stronger support depending on the site-specific conditions.

4.5 Spatial distribution of clusters

Projecting cluster membership onto the global map reveals a geographically structured pattern (Fig. 4c). The four clusters are not randomly distributed but tend to group in broad regions, with certain clusters dominating in specific climatic and biogeographical zones. This spatial organization suggests that the differences in model skill captured by the clustering are linked to large-scale environmental controls rather than being purely site-specific artifacts.

4.6 Validity domains in environmental predictor space

To interpret these spatial patterns, we analyze the distribution of clusters in environmental predictor space using a PCA-based representation (Fig. 5 and A1). In the first three principal components, clusters occupy partially overlapping but distinct regions. The convex hulls and Mahalanobis envelopes provide complementary views of these domains: the hulls delineate the outer bounds of observed conditions for each cluster, while the ellipses summarize their core distribution at a given confidence level.

Although overlap exists, especially in transitional regions of predictor space, each cluster exhibits a characteristic environmental signature, supporting the idea that different model combinations are preferentially suited to different combinations of productivity, climate, and soil conditions.

4.7 Predictability of cluster membership and dominant controls

Finally, we assess to what extent cluster membership can be predicted from environmental predictors alone. The random forest classifier trained on PCA-reduced predictors achieves a moderate but meaningful performance (Fig. 6a), with an overall accuracy of about 0.70 and a balanced accuracy of about 0.66 (macro f1-score: 0.68, k=5). Most misclassifications occur between neighboring clusters in predictor space, consistent with the partial overlap seen in the validity domains.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

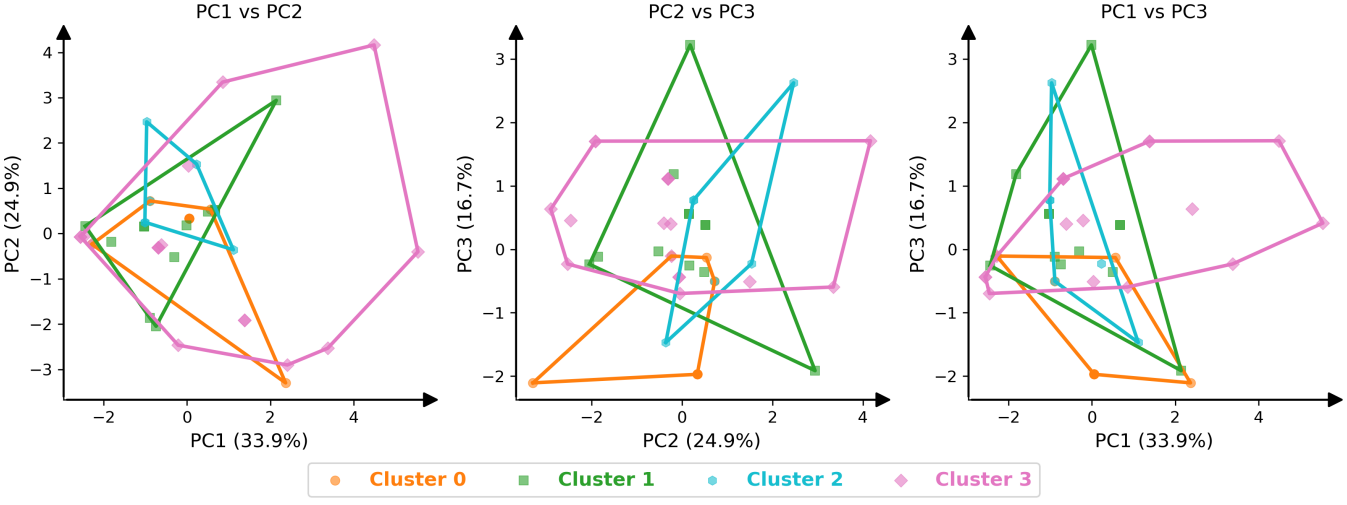
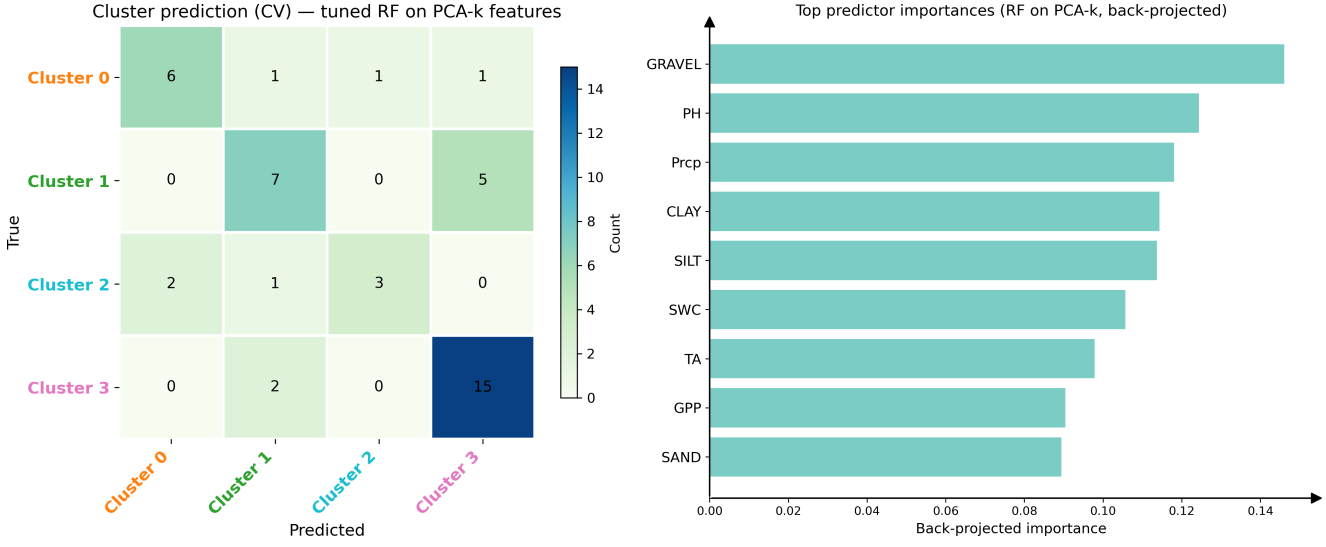


Figure 5. Validity domains of clusters in environmental predictor space using the first three principal components (PCA3) - Convex hulls.



(a) (b)

Figure 6. (a) Confusion matrix of the random forest classifier predicting cluster membership from environmental predictors (trained on PCA-reduced features), and (b) back-projected feature importances identifying the most influential predictors.

The back-projected feature importance analysis (Fig. 6b) indicates that soil properties (e.g., clay content, bulk density, and soil pH), and key climatic and hydrological variables (including precipitation, air temperature, and soil water content) are among the most influential predictors of cluster membership. This confirms that both biotic and abiotic controls jointly structure the regimes in which different model combinations perform best.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

4.8 Synthesis and spatial prediction

Taken together, these results show that (i) optimal model combinations vary across sites, (ii) accounting for this heterogeneity through EM-BMA improves predictive performance relative to simple model averaging, and (iii) the resulting clusters of model weights exhibit structure in both geographic and environmental space. This provides the basis for extending the site-level analysis to a spatially explicit framework.

Using the environmental predictors identified in the previous section, cluster membership was predicted across global wetlands (Fig. 7 and B1). The resulting maps reveal spatially coherent patterns in the predicted clusters. Cluster 3, which corresponds to weight configurations close to simple averaging, dominates large portions of boreal and tropical wetlands. Other clusters occur in more spatially restricted regions, including parts of tropical floodplains and temperate wetlands.

The predicted clusters were then used to propagate cluster-specific ensemble weights to the outputs of the participating models. Figures 8 (CRU climate forcing) and C1 (GSWP3 climate forcing) show the resulting differences between emissions estimated using EM-BMA and those obtained using simple model averaging. At the global scale, the total wetland methane emissions derived from the two approaches differ only slightly, with changes below 5%. However, the spatial distribution of emissions differs locally, with relative differences exceeding several tens of percent in some regions.

Panel (b) of Figs. 8 and C1 shows the relative difference in between-model variance between the BMA and SMA ensembles, providing a comparison of the ensemble spread associated with the two approaches. These variances represent the spread among model simulations only and therefore do not include the residual noise or internal uncertainty associated with individual models. Under both climate forcings, the between-model spread associated with the BMA ensemble is comparable to that of the SMA ensemble, with no systematic reduction in ensemble variance. However, substantial local differences are observed, indicating that the BMA framework primarily redistributes structural uncertainty across regions rather than reducing it.

Regional time series extracted from selected wetland systems further illustrate these differences (Figs. 8c–d and C1c–d). While the temporal variability of emissions remains broadly consistent between the two approaches, the magnitude of emissions differs between BMA and SMA in several regions.

5 Discussion

5.1 Spatial regimes of model skill and implications for emission estimates

The spatial extrapolation of cluster membership (Fig. 7) reveals that the regimes of model skill identified at flux tower sites are not randomly distributed but exhibit coherent geographic structure across global wetlands. In particular, the dominant cluster occupies large portions of boreal and tropical wetlands, whereas other clusters occur in more spatially confined regions such as tropical floodplains and parts of temperate wetlands. This spatial organization suggests that similarities in model performance are linked to broad environmental gradients captured by the predictor variables.

Interestingly, the dominant cluster corresponds to weight configurations close to simple model averaging. This indicates that, for a large fraction of wetland environments, the optimal ensemble mixture does not depart strongly from equal weighting. In

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

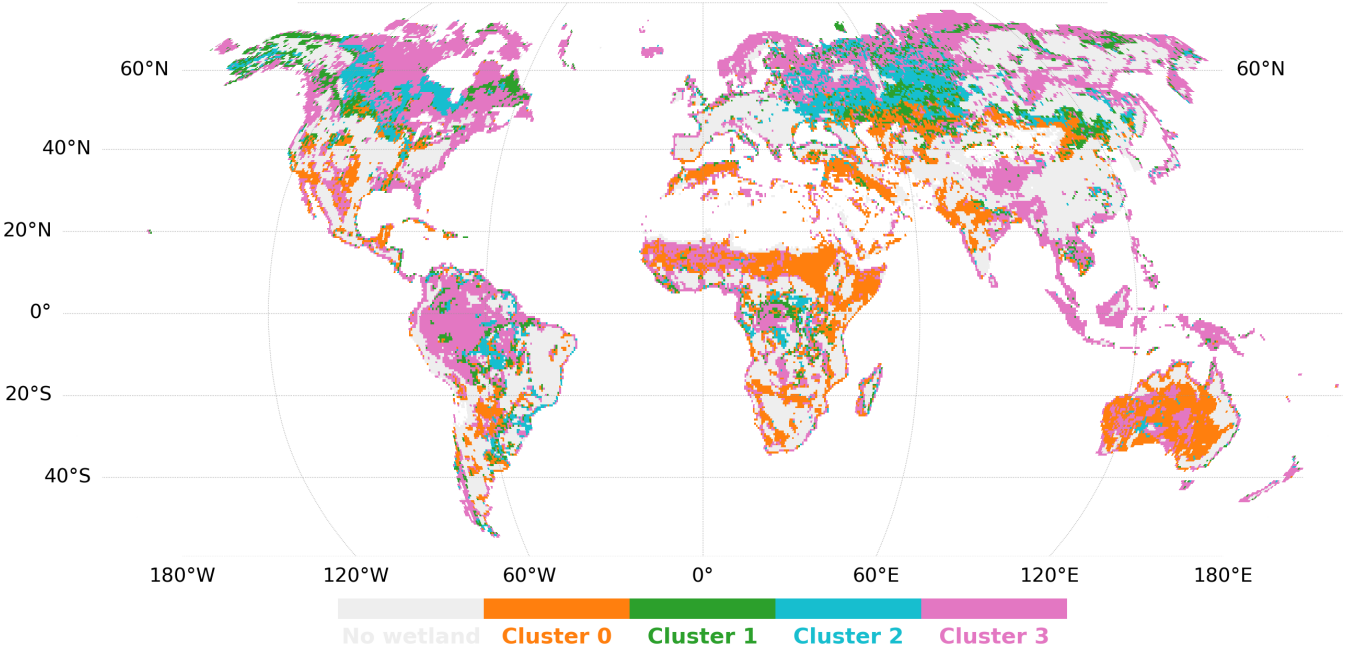


Figure 7. Predicted clusters. Lighter gray areas in the background indicate no wetlands according to the Wetland Area Dynamics for Methane modeling (WAD2M) (Zhang et al., 2021).

this sense, the proposed framework identifies where simple averaging is an adequate approximation and where alternative model combinations are statistically supported by the observations.

When these cluster-specific mixtures are applied to the outputs of the participating models (Figs. 8 and C1), the resulting global wetland methane emissions differ only marginally from those obtained using simple model averaging. The total emissions change by less than five percent, indicating that the proposed framework preserves the overall magnitude of global bottom-up estimates. However, the spatial distribution of emissions differs substantially in some regions, with localized deviations reaching several tens of percent. These differences exhibit a clear dependence on cluster membership (Figs. 7 and 8), with distinct clusters associated with characteristic ranges and signs of relative deviations, reflecting differences in the degree of model agreement across cluster-defined regimes.

The between-model variance shown in panel (b) of Figs. 8 and C1 provides an additional perspective on ensemble uncertainty. Because this quantity represents only the spread among model simulations, it isolates the structural variability of the ensemble without including the residual noise associated with individual models. In our results, the BMA-based ensemble does not substantially reduce the overall magnitude of between-model variability, but instead redistributes it across regions in a manner consistent with the identified regimes of model performance.

While the overall magnitude of ensemble spread remains comparable to that obtained with simple model averaging, the framework represents a structural advancement in how uncertainty is characterized. Rather than relying on simple min–max envelopes across models, as commonly done in global methane budget assessments (Saunois et al., 2025; Ciais et al., 2026),

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

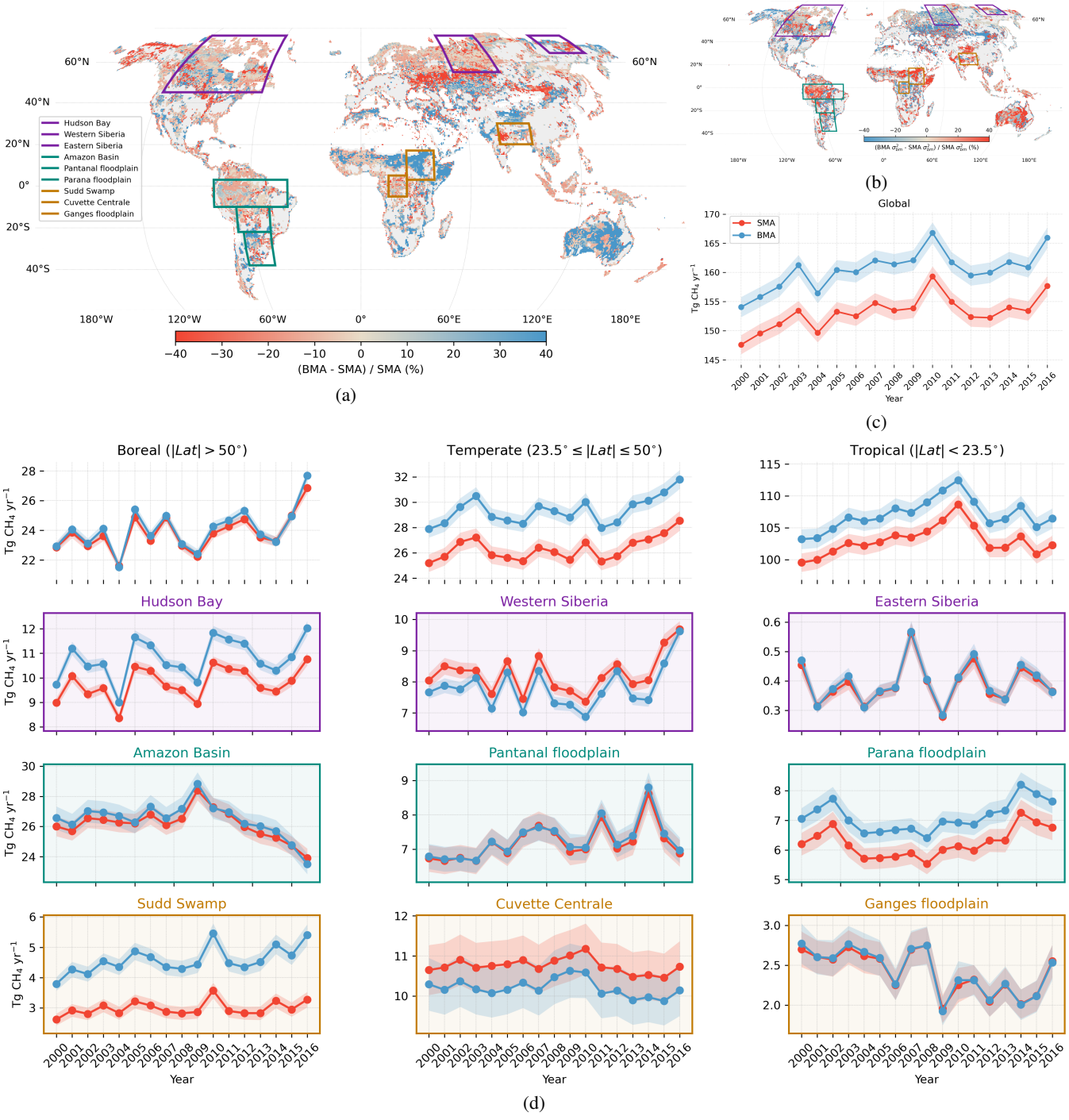


Figure 8. (a) - Relative difference between BMA and SMA wetland methane emissions. (b) - Relative difference between BMA and SMA between variances. (c) and (d) - Comparison of yearly emissions from SMA and BMA, with envelopes showing the between-model variance. Models are driven by CRU climate forcing. Employed regional rectangles as defined in Lin et al. (2024).

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

the proposed approach provides spatially explicit ensemble mixtures constrained by site-level observations. This enables uncertainty to be interpreted in terms of regimes of model performance, rather than solely through the range of model outputs.

At the same time, the spatial prediction of clusters relies on a limited set of environmental predictors and on the current distribution of observational sites. Although the resulting spatial patterns are coherent, uncertainties remain in regions where observational coverage is sparse, particularly in tropical wetlands. Improving the spatial representativeness of flux measurements and incorporating additional predictors describing wetland hydrology or wetland type may further refine the identification of model skill regimes and the extrapolation of ensemble weights.

Overall, the spatial cluster prediction and ensemble mapping illustrate how site-level observational constraints can be translated into spatially explicit information about model performance and emission estimates. This provides a complementary perspective to traditional ensemble averaging approaches and contributes to ongoing efforts to reduce uncertainty in bottom-up wetland methane budgets.

5.2 Between- versus within-model uncertainty and structural spread

The variance decomposition further provides insight into the nature of uncertainty across clusters. In particular, we observe that, for several clusters - and especially for C3 - the between-model variance exceeds the within-model variance. This implies that structural disagreement among models contributes more to predictive uncertainty than model-data mismatch alone.

In practical terms, this suggests that in regimes where SMA-like behavior dominates, the ensemble may consist of models that are structurally too dissimilar to be meaningfully averaged without careful weighting (Mo et al., 2023). The dominance of between-model variance in the principal cluster therefore raises the question of whether simple averaging implicitly blends models that respond differently to environmental drivers, potentially obscuring regime-specific process controls.

5.3 Implications for model performance regimes

The ordination of BMA weight signatures provides a complementary perspective on the clustering by revealing how model performance organizes across sites. The clear separation of clusters in the ordination space suggests that the inferred groups correspond to distinct regimes of model agreement and disagreement.

Importantly, these regimes are not driven by a single model but rather by shifts in the relative contribution of multiple models. This indicates that differences in model performance arise from structural differences in how models represent key processes, rather than from isolated model behaviour.

The fact that different models dominate in different regions of the ordination space suggests that no single model performs best across all conditions, but instead that model skill is context-dependent. This supports the use of adaptive ensemble approaches such as BMA, which allow the contribution of individual models to vary across space.

At the same time, the absence of a simple one-to-one correspondence between individual models and the main axes of variability highlights the difficulty of attributing performance differences to specific processes. Instead, the results point toward complex interactions between model structure and environmental conditions, which are only partially captured by the available predictors.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

5.4 Influence of heterogeneous calibration datasets

An additional factor contributing to structural spread is the fact that the participating models were not calibrated using identical observational datasets or site selections. Differences in calibration sites, temporal coverage, and process parameterizations inevitably propagate into ensemble spread. As a result, part of the between-model variance likely reflects differences in training data rather than purely structural contrasts. This highlights a broader issue in multi-model intercomparison: ensemble diversity combines both epistemic uncertainty (structural assumptions) and differences in calibration strategy.

5.5 Methodological considerations and potential dependence

The EM-BMA framework is effectively used twice in this study: first, to derive site-level weight signatures that define clusters, and second, to evaluate predictive skill within those clusters through cross-validation. While the cross-validation procedure ensures that predictions are generated from weights estimated independently of validation data, the clustering itself is derived from the same underlying weight structures.

This introduces a potential, albeit limited, dependence between cluster definition and within-cluster evaluation. The clusters represent higher-order statistical summaries of model performance, and the subsequent predictive assessment tests the consistency of these summaries. Nevertheless, future extensions could explore fully nested or external clustering strategies to further decouple structural identification from predictive evaluation.

5.6 Data limitations and the need for expanded observations

The robustness of any site-adaptive ensemble framework ultimately depends on the availability of high-quality observations. Although the FLUXNET and BAWLD datasets provide valuable constraints, spatial coverage remains uneven, with limited representation in several tropical and high-latitude regions (Knox et al., 2019). Temporal coverage is also heterogeneous, with varying record lengths across sites.

Expanding both the spatial and temporal extent of methane flux observations is therefore essential for improving ensemble calibration and for more reliably identifying regime-dependent model mixtures. This need is consistent with broader calls for enhanced observing systems to reduce uncertainty in natural methane sources and improve trend detection at regional scales (Ciais et al., 2026).

5.7 Toward improved environmental predictors

The environmental predictors used to interpret and upscale cluster membership capture broad controls related to productivity, climate, and soil properties. However, the moderate classification performance indicates that these variables do not fully explain the observed heterogeneity in model skill. Incorporating additional predictors, such as explicit wetland type classifications, hydrogeomorphic setting, inundation dynamics, or more explicit substrate availability proxies, may provide a more process-informed basis for distinguishing regimes.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

In particular, wetland functional types or dynamic wetland extent products could help bridge the gap between structural model assumptions and observed flux variability, thereby improving the interpretability and transferability of cluster-based ensemble configurations.

5.8 Proof of concept and operational perspectives

The framework presented here should be viewed as a proof of concept demonstrating that site-adaptive ensemble weighting and regime identification can improve predictive skill relative to simple averaging. While the improvements are statistically robust, operational deployment for large-scale methane budgeting would require additional data, systematic tuning, and potentially harmonized model calibration protocols.

Future work should aim to integrate expanded observational constraints, refine predictor selection, and test the framework under fully independent validation datasets. Such developments would strengthen the applicability of site-adaptive ensemble approaches for regional and global methane budget assessments.

6 Conclusions

We presented a site-adaptive ensemble framework for wetland methane (CH_4) emissions that moves beyond simple model averaging by estimating model weights using Bayesian model averaging (EM-BMA), stabilizing these weights through resampling, and clustering sites based on their weight signatures. This approach allows ensemble configurations to vary across environmental regimes rather than assuming uniform model contributions everywhere.

Our results show that optimal model combinations differ across sites and that EM-BMA improves predictive performance relative to simple model averaging. At the same time, the principal cluster exhibits near-uniform weights, suggesting that equal-weight averaging can emerge as a limiting case when models perform comparably. The variance decomposition further highlights that structural spread among models often dominates residual error, underscoring the importance of explicitly accounting for between-model uncertainty.

By linking cluster membership to environmental predictors, we demonstrate a pathway toward the spatial generalization of ensemble regimes. The predicted clusters reveal coherent geographic patterns across global wetlands and allow the propagation of cluster-specific model mixtures to the diagnostic outputs of the participating models. When applied globally, the resulting Bayesian model averaged emissions differ only marginally from simple model averaging at the global scale, while producing more pronounced regional differences in the spatial distribution of emissions.

Overall, this study provides a transparent and reproducible proof of concept for regime-dependent ensemble weighting in bottom-up wetland methane modeling. With additional data, refined environmental predictors, and improved observational coverage across wetland types, such approaches could contribute to more robust regional and global methane budget assessments and complement ongoing efforts to reduce uncertainty in natural methane sources.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

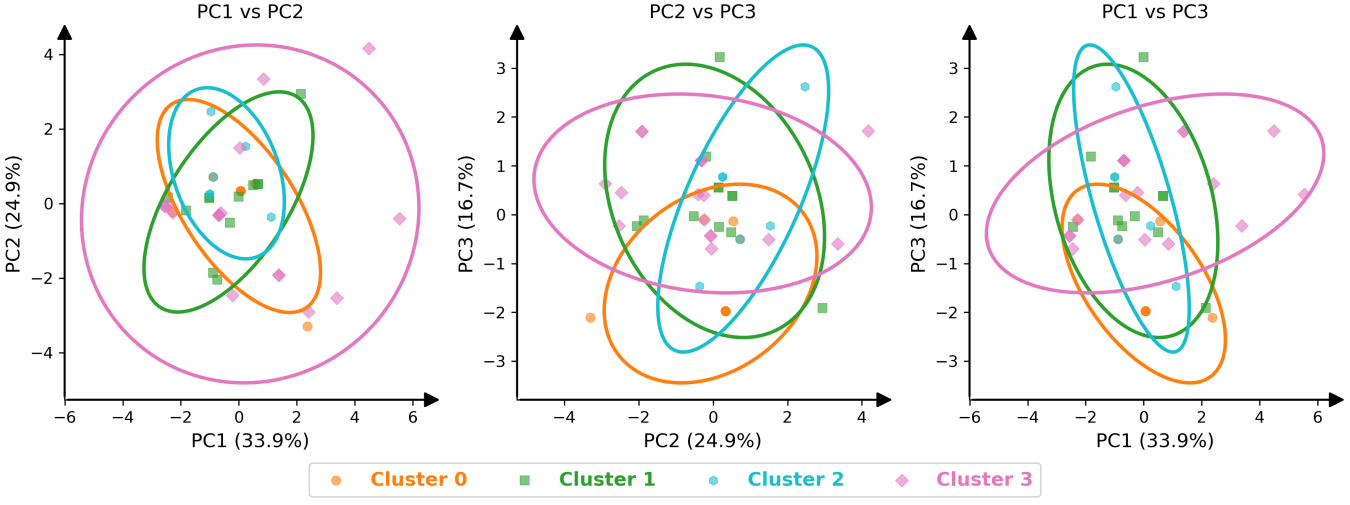


Figure A1. Validity domains of clusters in environmental predictor space using the first three principal components (PCA3) - Mahalanobis envelopes (95% level) shown for pairwise PCA projections.

Code and data availability. The code and the data allowing to reproduce all the results presented in the article are publicly accessible at the following link: https://doi.org/10.5281/zenodo.20043618 (Besic and Ciais, 2026).

Appendix A: Validity domains of clusters

This appendix primarily complements Sec. 3.2.2 and 4.6.

Appendix B: Predicted clusters probability

This appendix primarily complements Sec. 4.8.

Appendix C: Results with GSWP3 climate forcing

This appendix primarily complements Sec. 4.8 and 5.1.

Author contributions. NB and PC conceived the study. NB performed the analysis and wrote the manuscript. FL, RJ, KY, SP, BP, ZZ, and QZ provided the datasets and contributed to the study design. All authors contributed to the interpretation and discussion of the results, and reviewed and edited the manuscript.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

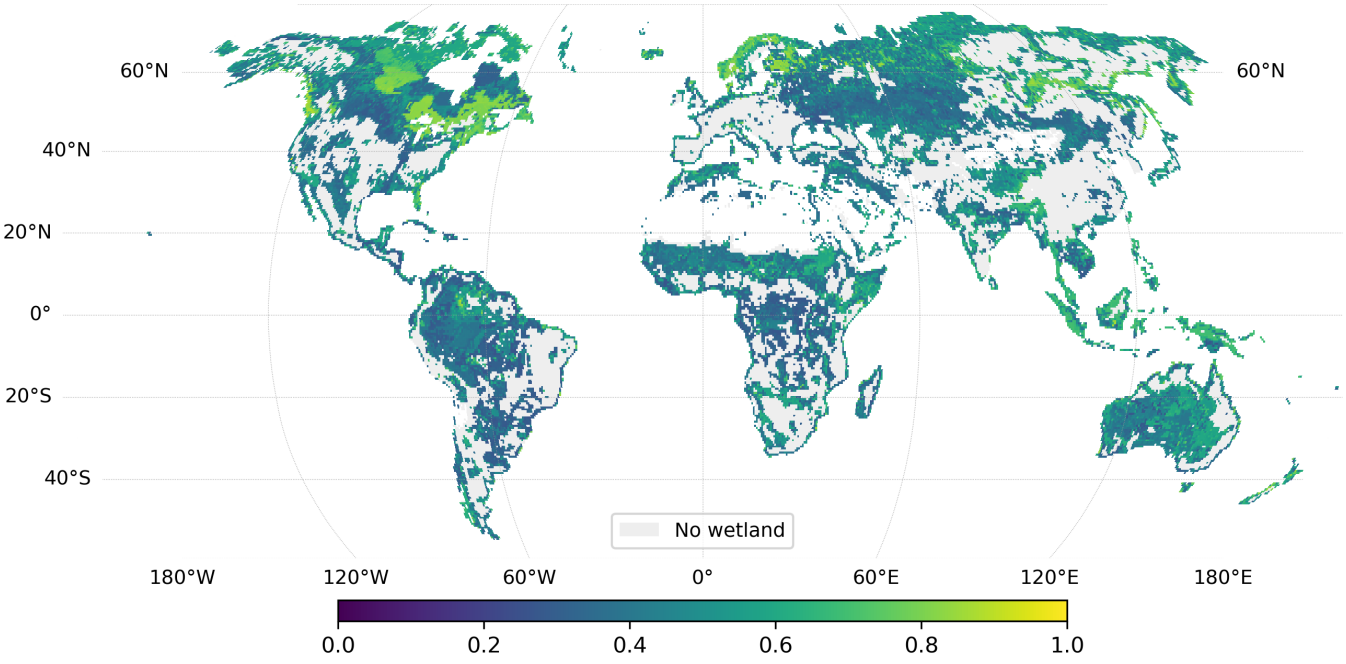


Figure B1. Predicted clusters probability (confidence). Lighter gray areas in the background indicate no wetlands according to the Wetland Area Dynamics for Methane modeling (WAD2M) (Zhang et al., 2021).

Competing interests. The contact author has declared that none of the authors has any competing interests.

Acknowledgements. NB and PC are supported by the French France 2030 program "Initiative d'Excellence Lorraine (LUE)", reference ANR-15-IDEX-04-LUE. They are also supported by project ALAMOD of the exploratory research program FairCarboN and received government funding managed by the Agence Nationale de la Recherche under the France 2030 program, reference ANR-22-PEXF-0002. QZ is supported by the biological and environmental research program of the US Department of Energy under contract number DE-AC02-05CH11231 to Lawrence Berkeley National Laboratory. FL and RBJ are supported by the Gordon and Betty Moore Foundation through grant GBMF11519 "Advancing the Understanding of Methane Emissions from Tropical Wetlands" and by Stanford University's Global Methane Office of the Global Carbon Project. FL acknowledges support from startup funds provided by The University of Texas at Austin. The authors acknowledge the use of generative artificial intelligence tools for language editing and improvement of the manuscript text.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

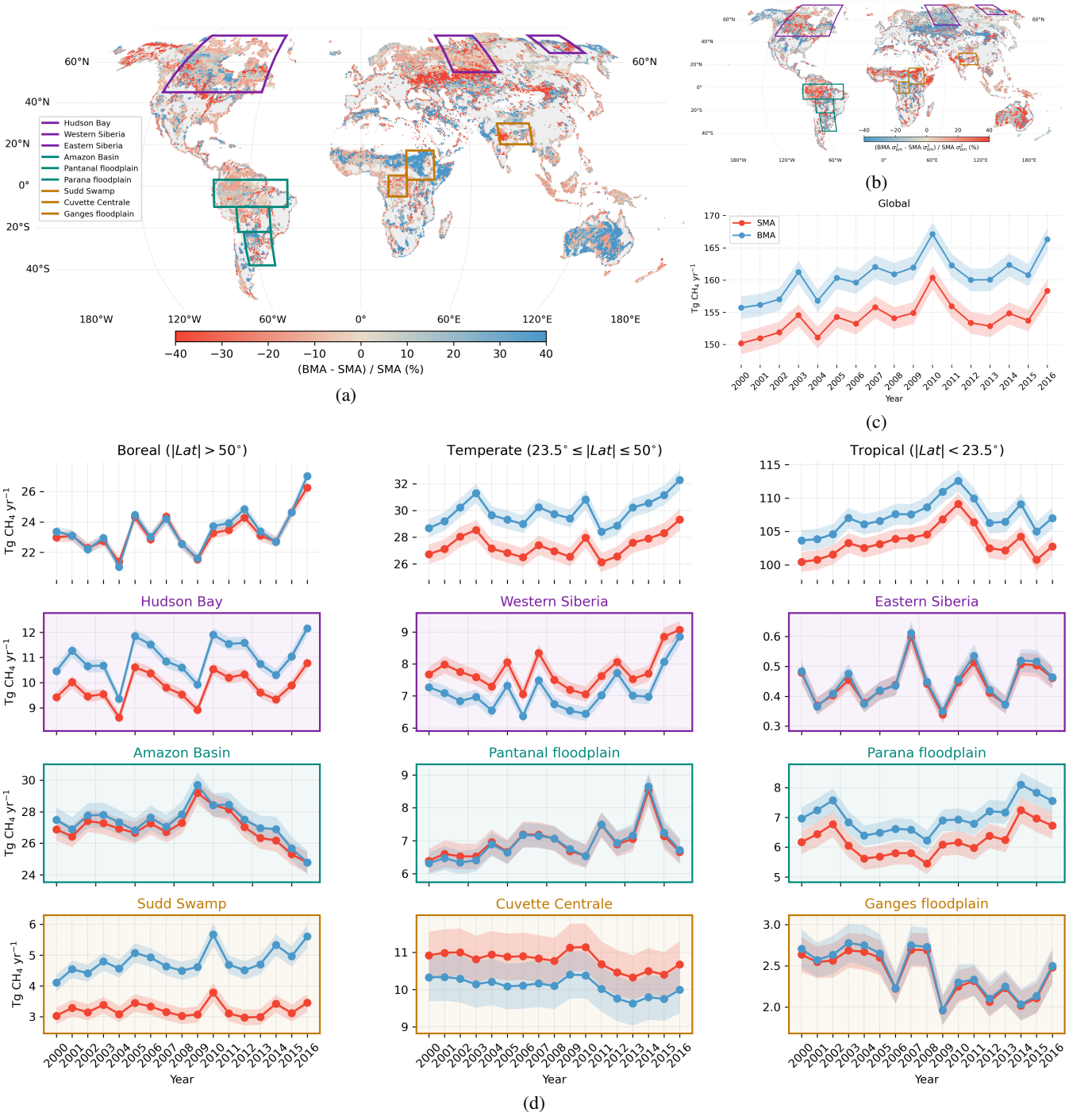


Figure C1. (a) - Relative difference between BMA and SMA wetland methane emissions. (b) - Relative difference between BMA and SMA between variances. (c) and (d) - Comparison of yearly emissions from SMA and BMA, with envelopes showing the between-model variance. Models are driven by GSWP3 climate forcing. Employed regional rectangles as defined in Lin et al. (2024).

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

References

Besic, N. and Ciais, P.: Code and data associated to the manuscript "Beyond the simple mean: an alternative way to improve multi-model bottom-up wetland CH_4 estimates", https://doi.org/10.5281/zenodo.20043618, 2026.

Besic, N., Picard, N., Vega, C., Bontemps, J.-D., Hertzog, L., Renaud, J.-P., Fogel, F., Schwartz, M., Pellissier-Tanon, A., Destouet, G., Mortier, F., Planells-Rodriguez, M., and Ciais, P.: Remote-sensing-based forest canopy height mapping: some models are useful, but might they provide us with even more insights when combined?, Geoscientific Model Development, 18, 337–359, https://doi.org/10.5194/gmd-18-337-2025, 2025.

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, Geoscientific Model Development, 4, 677–699, https://doi.org/10.5194/gmd-4-677-2011, 2011.

Bloom, A. A., Bowman, K. W., Lee, M., Turner, A. J., Schroeder, R., Worden, J. R., Weidner, R., McDonald, K. C., and Jacob, D. J.: A global wetland methane emissions and uncertainty dataset for atmospheric chemical transport models (WetCHARTs version 1.0), Geoscientific Model Development, 10, 2141–2156, https://doi.org/10.5194/gmd-10-2141-2017, 2017.

Canadell, J. G., Monteiro, P. M. S., Costa, M. H., et al.: Global Carbon and other Biogeochemical Cycles, in: Climate Change 2021: The Physical Science Basis, edited by Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/9781009157896.007, 2021.

Ciais, P., Peng, S., Chang, J., Li, F., Zhu, Q., Yuan, K., Hugelius, G., Li, H., Cai, Y., Chevallier, F., Tibrewal, K., Kort, E. A., Arndt, K., Watts, J., Buma, B., Besic, N., Palmer, P. I., Cadillo-Quiroz, H., Euskirchen, E., Gondwem, M., Hoyt, A., Jackson, R., Malone, S., Monteverde, D., Natali, S., Ramonet, M., Rey-Sanchez, C., Sagang, L., Schuur, E., Vargas, R., Varner, R., and Poulter, B.: A global methane observation system to reduce uncertainty for anthropogenic and natural sources and sinks for detecting and attributing climate feedbacks, Advanced Science, undergoing review, 2026.

Delwiche, K. B., Knox, S. H., Malhotra, A., Fluet-Chouinard, E., McNicol, G., Feron, S., Ouyang, Z., Papale, D., Trotta, C., Canfora, E., Cheah, Y.-W., Christianson, D., Alberto, M. C. R., Alekseychik, P., Aurela, M., Baldocchi, D., Bansal, S., Billesbach, D. P., Bohrer, G., Bracho, R., Buchmann, N., Campbell, D. I., Celis, G., Chen, J., Chen, W., Chu, H., Dalmagro, H. J., Dengel, S., Desai, A. R., Detto, M., Dolman, H., Eichelmann, E., Euskirchen, E., Famulari, D., Fuchs, K., Goeckede, M., Gogo, S., Gondwe, M. J., Goodrich, J. P., Gottschalk, P., Graham, S. L., Heimann, M., Helbig, M., Helfter, C., Hemes, K. S., Hirano, T., Hollinger, D., Hörtnagl, L., Iwata, H., Jacotot, A., Jurasinski, G., Kang, M., Kasak, K., King, J., Klatt, J., Koebsch, F., Krauss, K. W., Lai, D. Y. F., Lohila, A., Mammarella, I., Belelli Marchesini, L., Manca, G., Matthes, J. H., Maximov, T., Merbold, L., Mitra, B., Morin, T. H., Nemitz, E., Nilsson, M. B., Niu, S., Oechel, W. C., Oikawa, P. Y., Ono, K., Peichl, M., Peltola, O., Reba, M. L., Richardson, A. D., Riley, W., Runkle, B. R. K., Ryu, Y., Sachs, T., Sakabe, A., Sanchez, C. R., Schuur, E. A., Schäfer, K. V. R., Sonnentag, O., Sparks, J. P., Stuart-Haëntjens, E., Sturtevant, C., Sullivan, R. C., Szutu, D. J., Thom, J. E., Torn, M. S., Tuittila, E.-S., Turner, J., Ueyama, M., Valach, A. C., Vargas, R., Varlagin, A., Vazquez-Lule, A., Verfaillie, J. G., Vesala, T., Vourlitis, G. L., Ward, E. J., Wille, C., Wohlfahrt, G., Wong, G. X., Zhang, Z., Zona, D., Windham-Myers, L., Poulter, B., and Jackson, R. B.: FLUXNET-CH_4: a global, multi-ecosystem dataset and analysis of methane seasonality from freshwater wetlands, Earth System Science Data, 13, 3607–3689, https://doi.org/10.5194/essd-13-3607-2021, 2021.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T., and Hanasaki, N.: GSWP-2: Multimodel Analysis and Implications for Our Perception of the Land Surface, Bulletin of the American Meteorological Society, 87, 1381 – 1398, https://doi.org/10.1175/BAMS-87-10-1381, 2006.

Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., and Barceló-Vidal, C.: Isometric Logratio Transformations for Compositional Data Analysis, Mathematical Geology, 35, 279–300, https://doi.org/10.1023/A:1023818214614, 2003.

Forbrich, I., Yazbeck, T., Sulman, B., Morin, T. H., Tang, A. C. I., and Bohrer, G.: Three Decades of Wetland Methane Surface Flux Modeling by Earth System Models-Advances, Applications, and Challenges, Journal of Geophysical Research: Biogeosciences, 129, e2023JG007 915, https://doi.org/https://doi.org/10.1029/2023JG007915, 2024.

Harris, I., Osborn, T. J., Jones, P., and Lister, D.: Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset, Scientific Data, 7, 109, https://doi.org/10.1038/s41597-020-0453-3, 2020.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/https://doi.org/10.1002/qj.3803, 2020.

Huang, H., Qian, Y., Bisht, G., Wang, J., Chakraborty, T., Hao, D., Li, J., Thurber, T., Singh, B., Yang, Z., Liu, Y., Xue, P., Sacks, W. J., Coon, E., and Hetland, R.: WRF-ELM v1.0: a regional climate model to study land–atmosphere interactions over heterogeneous land use regions, Geoscientific Model Development, 18, 1427–1443, https://doi.org/10.5194/gmd-18-1427-2025, 2025.

Ito, A. and Inatomi, M.: Use of a process-based model for assessing the methane budgets of global terrestrial ecosystems and evaluation of uncertainty, Biogeosciences, 9, 759–773, https://doi.org/10.5194/bg-9-759-2012, 2012.

Johnson, R. A. and Wichern, D. W.: Applied Multivariate Statistical Analysis, Prentice Hall, 6 edn., 2007.

Kaiser, S., Göckede, M., Castro-Morales, K., Knoblauch, C., Ekici, A., Kleinen, T., Zubrzycki, S., Sachs, T., Wille, C., and Beer, C.: Process-based modelling of the methane balance in periglacial landscapes (JSBACH-methane), Geoscientific Model Development, 10, 333–358, https://doi.org/10.5194/gmd-10-333-2017, 2017.

Kallingal, J. T., Lindström, J., Miller, P. A., Rinne, J., Raivonen, M., and Scholze, M.: Optimising CH_4 simulations from the LPJ-GUESS model v4.1 using an adaptive Markov chain Monte Carlo algorithm, Geoscientific Model Development, 17, 2299–2324, https://doi.org/10.5194/gmd-17-2299-2024, 2024.

Kaufman, L. and Rousseeuw, P. J.: Finding Groups in Data: An Introduction to Cluster Analysis, Wiley, New York, NY, USA, 1990.

Knox, S. H., Jackson, R. B., Poulter, B., McNicol, G., Fluet-Chouinard, E., Zhang, Z., Hugelius, G., Bousquet, P., Canadell, J. G., Saunois, M., Papale, D., Chu, H., Keenan, T. F., Baldocchi, D., Torn, M. S., Mammarella, I., Trotta, C., Aurela, M., Bohrer, G., Campbell, D. I., Cescatti, A., Chamberlain, S., Chen, J., Chen, W., Dengel, S., Desai, A. R., Euskirchen, E., Friborg, T., Gasbarra, D., Goded, I., Goeckede, M., Heimann, M., Helbig, M., Hirano, T., Hollinger, D. Y., Iwata, H., Kang, M., Klatt, J., Krauss, K. W., Kutzbach, L., Lohila, A., Mitra, B., Morin, T. H., Nilsson, M. B., Niu, S., Noormets, A., Oechel, W. C., Peichl, M., Peltola, O., Reba, M. L., Richardson, A. D., Runkle, B. R. K., Ryu, Y., Sachs, T., Schäfer, K. V. R., Schmid, H. P., Shurpali, N., Sonnentag, O., Tang, A. C. I., Ueyama, M., Vargas, R., Vesala, T., Ward, E. J., Windham-Myers, L., Wohlfahrt, G., and Zona, D.: FLUXNET-CH4 Synthesis Activity: Objectives, Observations, and Future Directions, Bulletin of the American Meteorological Society, 100, 2607 – 2632, https://doi.org/10.1175/BAMS-D-18-0268.1, 2019.

Knox, S. H., Bansal, S., McNicol, G., Schafer, K., Sturtevant, C., Ueyama, M., Valach, A. C., Baldocchi, D., Delwiche, K., Desai, A. R., Euskirchen, E., Liu, J., Lohila, A., Malhotra, A., Melling, L., Riley, W., Runkle, B. R. K., Turner, J., Vargas, R., Zhu, Q., Alto, T.,

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

Fluet-Chouinard, E., Goeckede, M., Melton, J. R., Sonnentag, O., Vesala, T., Ward, E., Zhang, Z., Feron, S., Ouyang, Z., Alekseychik, P., Aurela, M., Bohrer, G., Campbell, D. I., Chen, J., Chu, H., Dalmagro, H. J., Goodrich, J. P., Gottschalk, P., Hirano, T., Iwata, H., Jurasinski, G., Kang, M., Koebsch, F., Mammarella, I., Nilsson, M. B., Ono, K., Peichl, M., Peltola, O., Ryu, Y., Sachs, T., Sakabe, A., Sparks, J. P., Tuittila, E.-S., Vourlitis, G, L., Wong, G. X., Windham-Myers, L., Poulter, B., and Jackson, R, B.: Identifying dominant environmental predictors of freshwater wetland methane fluxes across diurnal to seasonal time scales, Global Change Biology, 27, 3582–3604, https://doi.org/https://doi.org/10.1111/gcb.15661, 2021.

Koffi, E. N., Bergamaschi, P., Alkama, R., and Cescatti, A.: An observation-constrained assessment of the climate sensitivity and future trajectories of wetland methane emissions, Science Advances, 6, https://doi.org/10.1126/sciadv.aay4444, 2020.

Kuhn, M. A., Varner, R. K., Bastviken, D., Crill, P., MacIntyre, S., Turetsky, M., Walter Anthony, K., McGuire, A. D., and Olefeldt, D.: BAWLD-CH_4: a comprehensive dataset of methane fluxes from boreal and arctic ecosystems, Earth System Science Data, 13, 5151–5189, https://doi.org/10.5194/essd-13-5151-2021, 2021.

Le Mer, J. and Pierre Roger, P.: Production, oxidation, emission and consumption of methane by soils: A review, European Journal of Soil Biology, 37, 25–50, https://doi.org/https://doi.org/10.1016/S1164-5563(01)01067-6, 2001.

Li, F., Zhu, Q., Yuan, K., Fluet-Chouinard, E., Zhang, X., Wang, J., Knox, S. H., You, H., Chen, M., Li, M., Stern, R., Hoyt, A. M., McNicol, G., Riley, W. J., Peng, S., Poulter, B., Malhotra, A., Cooley, S., Zhang, Z., Hong, S., Chen, Z., Zhu, Z., Raymond, P. A., Ciais, P., and Jackson, R. B.: The underappreciated importance of small wetlands in global methane emissions, Nature Climate Change, https://doi.org/10.1038/s41558-026-02609-w, 2026.

Li, T., Huang, Y., Zhang, W., and Song, C.: CH4MODwetland: A biogeophysical model for simulating methane emissions from natural wetlands, Ecological Modelling, 221, 666–680, https://doi.org/https://doi.org/10.1016/j.ecolmodel.2009.05.017, 2010.

Li, X. and Xiao, J.: A Global, 0.05-Degree Product of Solar-Induced Chlorophyll Fluorescence Derived from OCO-2, MODIS, and Reanalysis Data, Remote Sensing, 11, https://doi.org/10.3390/rs11050517, 2019.

Lin, X., Peng, S., Ciais, P., Hauglustaine, D., Lan, X., Liu, G., Ramonet, M., Xi, Y., Yin, Y., Zhang, Z., Bösch, H., Bousquet, P., Chevallier, F., Dong, B., Gerlein-Safdi, C., Halder, S., Parker, R. J., Poulter, B., Pu, T., Remaud, M., Runge, A., Saunois, M., Thompson, R. L., Yoshida, Y., and Zheng, B.: Recent methane surges reveal heightened emissions from tropical inundated areas, Nature Communications, 15, 10 894, https://doi.org/10.1038/s41467-024-55266-y, 2024.

Melton, J. R., Wania, R., Hodson, E. L., Poulter, B., Ringeval, B., Spahni, R., Bohn, T., Avis, C. A., Beerling, D. J., Chen, G., Eliseev, A. V., Denisov, S. N., Hopcroft, P. O., Lettenmaier, D. P., Riley, W. J., Singarayer, J. S., Subin, Z. M., Tian, H., Zürcher, S., Brovkin, V., van Bodegom, P. M., Kleinen, T., Yu, Z. C., and Kaplan, J. O.: Present state of global wetland extent and wetland methane modelling: conclusions from a model inter-comparison project (WETCHIMP), Biogeosciences, 10, 753–788, https://doi.org/10.5194/bg-10-753-2013, 2013.

Melton, J. R., Arora, V. K., Wisernig-Cojoc, E., Seiler, C., Fortier, M., Chan, E., and Teckentrup, L.: CLASSIC v1.0: the open-source community successor to the Canadian Land Surface Scheme (CLASS) and the Canadian Terrestrial Ecosystem Model (CTEM) – Part 1: Model framework and site-level performance, Geoscientific Model Development, 13, 2825–2850, https://doi.org/10.5194/gmd-13-2825-2020, 2020.

Mo, L., Zohner, C. M., Reich, P. B., Liang, J., de Miguel, S., Nabuurs, G.-J., Renner, S. S., van den Hoogen, J., Araza, A., Herold, M., Mirzagholi, L., Ma, H., Averill, C., Phillips, O. L., Gamarra, J. G. P., Hordijk, I., Routh, D., Abegg, M., Adou Yao, Y. C., Alberti, G., Almeyda Zambrano, A. M., Alvarado, B. V., Alvarez-Dávila, E., Alvarez-Loayza, P., Alves, L. F., Amaral, I., Ammer, C., Antón-Fernández, C., Araujo-Murakami, A., Arroyo, L., Avitabile, V., Aymard, G. A., Baker, T. R., Bałazy, R., Banki, O., Barroso, J. G.,

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

Bastian, M. L., Bastin, J.-F., Birigazzi, L., Birnbaum, P., Bitariho, R., Boeckx, P., Bongers, F., Bouriaud, O., Brancalion, P. H. S., Brandl, S., Brearley, F. Q., Brienen, R., Broadbent, E. N., Bruelheide, H., Bussotti, F., Cazzolla Gatti, R., César, R. G., Cesljar, G., Chazdon, R. L., Chen, H. Y. H., Chisholm, C., Cho, H., Cienciala, E., Clark, C., Clark, D., Colletta, G. D., Coomes, D. A., Cornejo Valverde, F., Corral-Rivas, J. J., Crim, P. M., Cumming, J. R., Dayanandan, S., de Gasper, A. L., Decuyper, M., Derroire, G., DeVries, B., Djordjevic, I., Dolezal, J., Dourdain, A., Engone Obiang, N. L., Enquist, B. J., Eyre, T. J., Fandohan, A. B., Fayle, T. M., Feldpausch, T. R., Ferreira, L. V., Finér, L., Fischer, M., Fletcher, C., Frizzera, L., Gianelle, D., Glick, H. B., Harris, D. J., Hector, A., Hemp, A., Hengeveld, G., Hérault, B., Herbohn, J. L., Hillers, A., Honorio Coronado, E. N., Hui, C., Ibanez, T., Imai, N., Jagodziński, A. M., Jaroszewicz, B., Johannsen, V. K., Joly, C. A., Jucker, T., Jung, I., Karminov, V., Kartawinata, K., Kearsley, E., Kenfack, D., Kennard, D. K., Kepfer-Rojas, S., Keppel, G., Khan, M. L., Killeen, T. J., Kim, H. S., Kitayama, K., Köhl, M., Korjus, H., Kraxner, F., Kucher, D., Laarmann, D., Lang, M., Lu, H., Lukina, N. V., Maitner, B. S., Malhi, Y., Marcon, E., Marimon, B. S., Marimon-Junior, B. H., Marshall, A. R., Martin, E. H., Meave, J. A., Melo-Cruz, O., Mendoza, C., Mendoza-Polo, I., Miscicki, S., Merow, C., Monteagudo Mendoza, A., Moreno, V. S., Mukul, S. A., Mundhenk, P., Nava-Miranda, M. G., Neill, D., Neldner, V. J., Nevenic, R. V., Ngugi, M. R., Niklaus, P. A., Oleksyn, J., Ontikov, P., Ortiz-Malavasi, E., Pan, Y., Paquette, A., Parada-Gutierrez, A., Parfenova, E. I., Park, M., Parren, M., Parthasarathy, N., Peri, P. L., Pfautsch, S., Picard, N., Piedade, M. T. F., Piotto, D., Pitman, N. C. A., Poulsen, A. D., Poulsen, J. R., Pretzsch, H., Ramirez Arevalo, F., Restrepo-Correa, Z., Rodeghiero, M., Rolim, S. G., Roopsind, A., Rovero, F., Rutishauser, E., Saikia, P., Salas-Eljatib, C., Saner, P., Schall, P., Schelhaas, M.-J., Schepaschenko, D., Scherer-Lorenzen, M., Schmid, B., Schöngart, J., Searle, E. B., Seben, V., Serra-Diaz, J. M., Sheil, D., Shvidenko, A. Z., Silva-Espejo, J. E., Silveira, M., Singh, J., Sist, P., Slik, F., Sonké, B., Souza, A. F., Stereńczak, K. J., Svenning, J.-C., Svoboda, M., Swanepoel, B., Targhetta, N., Tchebakova, N., ter Steege, H., Thomas, R., Tikhonova, E., Umunay, P. M., Usoltsev, V. A., Valencia, R., Valladares, F., van der Plas, F., Van Do, T., van Nuland, M. E., Vasquez, R. M., Verbeeck, H., Viana, H., Vibrans, A. C., Vieira, S., von Gadow, K., Wang, H.-F., Watson, J. V., Werner, G. D. A., Wiser, S. K., Wittmann, F., Woell, H., Wortel, V., Zagt, R., Zawiła-Niedźwiecki, T., Zhang, C., Zhao, X., Zhou, M., Zhu, Z.-X., Zo-Bi, I. C., Gann, G. D., and Crowther, T. W.: Integrated global assessment of the natural forest carbon potential, Nature, 624, 92–101, https://doi.org/10.1038/s41586-023-06723-z, 2023.

Nachtergaele, F., Velthuizen, H., Verelst, L., Wiberg, D., Henry, M., Chiozza, F., Yigini, Y., Fischer, G., Tramberend, S., Batjes, N., Montanarella, L., Jones, A., Aksoy, E., Boateng, E., and Shi, X.: Harmonized World Soil Database version 2.0, https://doi.org/10.4060/cc3823en, 2023.

Olefeldt, D., Hovemyr, M., Kuhn, M. A., Bastviken, D., Bohn, T. J., Connolly, J., Crill, P., Euskirchen, E. S., Finkelstein, S. A., Genet, H., Grosse, G., Harris, L. I., Heffernan, L., Helbig, M., Hugelius, G., Hutchins, R., Juutinen, S., Lara, M. J., Malhotra, A., Manies, K., McGuire, A. D., Natali, S. M., O'Donnell, J. A., Parmentier, F.-J. W., Räsänen, A., Schädel, C., Sonnentag, O., Strack, M., Tank, S. E., Treat, C., Varner, R. K., Virtanen, T., Warren, R. K., and Watts, J. D.: The Boreal–Arctic Wetland and Lake Dataset (BAWLD), Earth System Science Data, 13, 5127–5149, https://doi.org/10.5194/essd-13-5127-2021, 2021.

Pedregosa, F., Varoquaux, G., Gramfort, A., et al.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Peng, S., Lin, X., Thompson, R. L., Xi, Y., Liu, G., Hauglustaine, D., Lan, X., Poulter, B., Ramonet, M., Saunois, M., Yin, Y., Zhang, Z., Zheng, B., and Ciais, P.: Wetland emission and atmospheric sink changes explain methane growth in 2020, Nature, 612, 477–482, https://doi.org/10.1038/s41586-022-05447-w, 2022.

Picard, N., Henry, M., Mortier, F., Trotta, C., and Saint-André, L.: Using Bayesian Model Averaging to Predict Tree Aboveground Biomass in Tropical Moist Forests, Forest Science, 58, 15–23, https://doi.org/10.5849/forsci.10-083, 2012.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

Picard, N., Besic, N., Meliho, M., Sainte-Marie, J., Mortier, F., and Legay, M.: Bayesian model averaging of climate-dependent forest models using Expectation–Maximization, Ecological Modelling, 510, 111 355, https://doi.org/https://doi.org/10.1016/j.ecolmodel.2025.111355, 2025.

Preparata, F. P. and Shamos, M. I.: Computational Geometry: An Introduction, Springer, New York, NY, USA, 1985.

Raftery, A. E.: Bayesian model selection in structural equation models, in: Testing structural equation models, edited by Bollen, K. A. and Log, J. S., pp. 163–180, 1993.

Raftery, A. E., Gneiting, T. Fadoua Balabdaoui, F., and Polakowski, M.: Using Bayesian Model Averaging to Calibrate Forecast Ensembles, Monthly Weather Review, 133, 1155 – 1174, https://doi.org/10.1175/MWR2906.1, 2005.

Ringeval, B., Hopcroft, P. O., Valdes, P. J., Ciais, P., Ramstein, G., Dolman, A. J., and Kageyama, M.: Response of methane emissions from wetlands to the Last Glacial Maximum and an idealized Dansgaard–Oeschger climate event: insights from two models of different complexity, Climate of the Past, 9, 149–171, https://doi.org/10.5194/cp-9-149-2013, 2013.

Salmon, E., Jégou, F., Guenet, B., Jourdain, L., Qiu, C., Bastrikov, V., Guimbaud, C., Zhu, D., Ciais, P., Peylin, P., Gogo, S., Laggoun-Défarge, F., Aurela, M., Bret-Harte, M. S., Chen, J., Chojnicki, B. H., Chu, H., Edgar, C. W., Euskirchen, E. S., Flanagan, L. B., Fortuniak, K., Holl, D., Klatt, J., Kolle, O., Kowalska, N., Kutzbach, L., Lohila, A., Merbold, L., Pawlak, W., Sachs, T., and Ziemblińska, K.: Assessing methane emissions for northern peatlands in ORCHIDEE-PEAT revision 7020, Geoscientific Model Development, 15, 2813–2838, https://doi.org/10.5194/gmd-15-2813-2022, 2022.

Saunois, M., Martinez, A., Poulter, B., Zhang, Z., Raymond, P. A., Regnier, P., Canadell, J. G., Jackson, R. B., Patra, P. K., Bousquet, P., Ciais, P., Dlugokencky, E. J., Lan, X., Allen, G. H., Bastviken, D., Beerling, D. J., Belikov, D. A., Blake, D. R., Castaldi, S., Crippa, M., Deemer, B. R., Dennison, F., Etiope, G., Gedney, N., Höglund-Isaksson, L., Holgerson, M. A., Hopcroft, P. O., Hugelius, G., Ito, A., Jain, A. K., Janardanan, R., Johnson, M. S., Kleinen, T., Krummel, P. B., Lauerwald, R., Li, T., Liu, X., McDonald, K. C., Melton, J. R., Mühle, J., Müller, J., Murguia-Flores, F., Niwa, Y., Noce, S., Pan, S., Parker, R. J., Peng, C., Ramonet, M., Riley, W. J., Rocher-Ros, G., Rosentreter, J. A., Sasakawa, M., Segers, A., Smith, S. J., Stanley, E. H., Thanwerdas, J., Tian, H., Tsuruta, A., Tubiello, F. N., Weber, T. S., van der Werf, G. R., Worthy, D. E. J., Xi, Y., Yoshida, Y., Zhang, W., Zheng, B., Zhu, Q., Zhu, Q., and Zhuang, Q.: Global Methane Budget 2000–2020, Earth System Science Data, 17, 1873–1958, https://doi.org/10.5194/essd-17-1873-2025, 2025.

Shu, S., Jain, A. K., and Kheshgi, H. S.: Investigating Wetland and Nonwetland Soil Methane Emissions and Sinks Across the Contiguous United States Using a Land Surface Model, Global Biogeochemical Cycles, 34, e2019GB006 251, https://doi.org/https://doi.org/10.1029/2019GB006251, 2020.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K., and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, Global Change Biology, 9, 161–185, https://doi.org/https://doi.org/10.1046/j.1365-2486.2003.00569.x, 2003.

Stocker, B. D., Spahni, R., and Joos, F.: DYPTOP: a cost-efficient TOPMODEL implementation to simulate sub-grid spatio-temporal dynamics of global wetlands and peatlands, Geoscientific Model Development, 7, 3089–3110, https://doi.org/10.5194/gmd-7-3089-2014, 2014.

Virkkala, A.-M., Wargowsky, I., Vogt, J., Kuhn, M. A., Madaan, S., O'Keefe, R., Windholz, T., Arndt, K. A., Rogers, B. M., Watts, J. D., Kent, K., Göckede, M., Olefeldt, D., Rocher-Ros, G., Schuur, E. A. G., Bastviken, D., Aalstad, K., Aho, K., Ala-Könni, J., Alcock, H., Althuizen, I., Arp, C. D., Asanuma, J., Attermeyer, K., Aurela, M., Balathandayuthabani, S., Barr, A., Barret, M., Batkhishig, O., Biasi, C., Björkman, M. P., Black, A., Blanc-Betes, E., Bodmer, P., Boike, J., Bolek, A., Bouchard, F., Bussmann, I., Cabrol, L., Canfora, E., Carey, S., Castro-Morales, K., Chae, N., Christen, A., Christensen, T. R., Christiansen, C. T., Chu, H., Clark, G., Clayer, F., Crill, P., Cunada, C.,

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

Davidson, S. J., Dean, J. F., Dengel, S., Detto, M., Dieleman, C., Domine, F., Dyukarev, E., Edgar, C., Elberling, B., Emmerton, C. A., Euskirchen, E., Falvo, G., Friborg, T., Garneau, M., Giamberini, M., Glagolev, M. V., Gonzalez-Meler, M. A., Granath, G., Guðmundsson, J., Happonen, K., Harazono, Y., Harris, L., Hashemi, J., Hasson, N., Heerah, J., Heffernan, L., Helbig, M., Helgason, W., Heliasz, M., Henry, G., Hensgens, G., Hiyama, T., Hock, M., Holl, D., Holmes, B., Holst, J., Holst, T., Hould-Gosselin, G., Humphreys, E., Hung, J., Huotari, J., Ikawa, H., Ilyasov, D. V., Ishikawa, M., Iwahana, G., Iwata, H., Jackowicz-Korczynski, M. A., Jansen, J., Järveoja, J., Jassey, V. E. J., Jensen, R., Jentzsch, K., Jespersen, R. G., Johannesson, C.-F., Jones, C. P., Jonsson, A., Jung, J. Y., Juutinen, S., Kane, E., Karlsson, J., Karsanaev, S., Kasak, K., Kelly, J., Kempton, K., Klaus, M., Kling, G. W., Kljun, N., Knutson, J., Kobayashi, H., Kochendorfer, J., Kohonen, K.-M., Kolari, P., Korkiakoski, M., Korrensalo, A., Kortelainen, P., Koster, E., Koster, K., Kotani, A., Krishnan, P., Kurbatova, J., Kutzbach, L., Kwon, M. J., Kyzivat, E. D., Lagroix, J., Langhorst, T., Lapshina, E., Larmola, T., Larsen, K. S., Laurion, I., Ledman, J., Lee, H., Leffler, A. J., Lesack, L., Lindroth, A., Lipson, D., Lohila, A., López-Blanco, E., St. Louis, V. L., Lundin, E., Luoto, M., Machimura, T., Magnani, M., Malhotra, A., Maljanen, M., Mammarella, I., Männistö, E., Marchesini, L. B., Marsh, P., Martkainen, P. J., Marushchak, M. E., Mastepanov, M., Mavrovic, A., Maximov, T., Minions, C., Montemayor, M., Morishita, T., Murphy, P., Nadeau, D. F., Nicholls, E., Nilsson, M. B., Niyazova, A., Nordén, J., Noumonvi, K. D., Nykanen, H., Oechel, W., Ojala, A., Okadera, T., Pal, S., Panov, A. V., Papakyriakou, T., Papale, D., Park, S.-J., Parmentier, F.-J. W., Pastorello, G., Peacock, M., Peichl, M., Petrov, R., St. Pierre, K., Pirk, N., Plein, J., Preskienis, V., Prokushkin, A., Pumpanen, J., Rains, H. A., Rakos, N., Räsänen, A., Rautakoski, H., Rinnan, R., Rinne, J., Rocha, A., Roulet, N., Roy, A., Rutgersson, A., Sabrekov, A. F., Sachs, T., Sahlée, E., Salazar, A., Sawakuchi, H. O., Schulze, C., Seco, R., Sepulveda-Jauregui, A., Serikova, S., Serrone, A., Silvennoinen, H. M., Sjogersten, S., Skeeter, J., Snöälv, J., Sobek, S., Sonnentag, O., Stanley, E. H., Strack, M., Strom, L., Sullivan, P., Sullivan, R., Sytiuk, A., Tagesson, T., Taillardat, P., Talbot, J., Tank, S. E., Tenuta, M., Terenteva, I., Thalasso, F., Thiboult, A., Thorgeirsson, H., Garcia Tigreros, F., Torn, M., Townsend-Small, A., Treat, C., Tremblay, A., Trotta, C., Tuittila, E.-S., Turetsky, M., Ueyama, M., Umair, M., Vähä, A., van Delden, L., van Hardenbroek, M., Varlagin, A., Varner, R. K., Veretennikova, E., Vesala, T., Virtanen, T., Voigt, C., Vonk, J. E., Wagner, R., Walter Anthony, K., Wang, Q., Watanabe, M., Webb, H., Welker, J. M., Westergaard-Nielsen, A., Westermann, S., White, J. R., Wille, C., Williamson, S. N., Zolkos, S., Zona, D., and Natali, S. M.: ABCFlux v2: Arctic–boreal CO_2 and CH_4 monthly flux observations and ancillary information across terrestrial and freshwater ecosystems, Earth System Science Data Discussions, 2025, 1–86, https://doi.org/10.5194/essd-2025-585, 2025.

Wania, R., Melton, J. R., Hodson, E. L., Poulter, B., Ringeval, B., Spahni, R., Bohn, T., Avis, C. A., Chen, G., Eliseev, A. V., Hopcroft, P. O., Riley, W. J., Subin, Z. M., Tian, H., van Bodegom, P. M., Kleinen, T., Yu, Z. C., Singarayer, J. S., Zürcher, S., Lettenmaier, D. P., Beerling, D. J., Denisov, S. N., Prigent, C., Papa, F., and Kaplan, J. O.: Present state of global wetland extent and wetland methane modelling: methodology of a model inter-comparison project (WETCHIMP), Geoscientific Model Development, 6, 617–641, https://doi.org/10.5194/gmd-6-617-2013, 2013.

Yuan, K., Zhu, Q., Li, F., Riley, W. J., Torn, M., Chu, H., McNicol, G., Chen, M., Knox, S., Delwiche, K., Wu, H., Baldocchi, D., Ma, H., Desai, A. R., Chen, J., Sachs, T., Ueyama, M., Sonnentag, O., Helbig, M., Tuittila, E.-S., Jurasinski, G., Koebsch, F., Campbell, D., Schmid, H. P., Lohila, A., Goeckede, M., Nilsson, M. B., Friborg, T., Jansen, J., Zona, D., Euskirchen, E., Ward, E. J., Bohrer, G., Jin, Z., Liu, L., Iwata, H., Goodrich, J., and Jackson, R.: Causality guided machine learning model on wetland CH4 emissions across global wetlands, Agricultural and Forest Meteorology, 324, 109 115, https://doi.org/https://doi.org/10.1016/j.agrformet.2022.109115, 2022.

Yuan, K., Li, F., McNicol, G., Chen, M., Hoyt, A., Knox, S., Riley, W. J., Jackson, R., and Zhu, Q.: Boreal–Arctic wetland methane emissions modulated by warming and vegetation activity, Nature Climate Change, 14, 282–288, https://doi.org/10.1038/s41558-024-01933-3, 2024.

https://doi.org/10.5194/egusphere-2026-2604
Preprint. Discussion started: 29 June 2026

Zhang, B., Tian, H., Lu, C., Chen, G., Pan, S., Anderson, C., and Poulter, B.: Methane emissions from global wetlands: An assessment of the uncertainty associated with various wetland extent data sets, Atmospheric Environment, 165, 310–321, https://doi.org/https://doi.org/10.1016/j.atmosenv.2017.07.001, 2017a.

Zhang, Z., Zimmermann, N. E., Stenke, A., Li, X., Hodson, E. L., Zhu, G., Huang, C., and Poulter, B.: Emerging role of wetland methane emissions in driving 21st century climate change, Proceedings of the National Academy of Sciences, 114, 9647–9652, https://doi.org/10.1073/pnas.1618765114, 2017b.

Zhang, Z., Fluet-Chouinard, E., Jensen, K., McDonald, K., Hugelius, G., Gumbricht, T., Carroll, M., Prigent, C., Bartsch, A., and Poulter, B.: Development of the global dataset of Wetland Area and Dynamics for Methane Modeling (WAD2M), Earth System Science Data, 13, 2001–2023, https://doi.org/10.5194/essd-13-2001-2021, 2021.

Zhang, Z., Poulter, B., Feldman, A. F., Ying, Q., Ciais, P., Peng, S., and Li, .: Recent intensification of wetland methane feedback, Nature Climate Change, 13, 430–433, https://doi.org/10.1038/s41558-023-01629-0, 2023.

Zhang, Z., Poulter, B., Melton, J. R., Riley, W. J., Allen, G. H., Beerling, D. J., Bousquet, P., Canadell, J. G., Fluet-Chouinard, E., Ciais, P., Gedney, N., Hopcroft, P. O., Ito, A., Jackson, R. B., Jain, A. K., Jensen, K., Joos, F., Kleinen, T., Knox, S. H., Li, T., Li, X., Liu, X., McDonald, K., McNicol, G., Miller, P. A., Müller, J., Patra, P. K., Peng, C., Peng, S., Qin, Z., Riggs, R. M., Saunois, M., Sun, Q., Tian, H., Xu, X., Yao, Y., Xi, Y., Zhang, W., Zhu, Q., Zhu, Q., and Zhuang, Q.: Ensemble estimates of global wetland methane emissions over 2000–2020, Biogeosciences, 22, 305–321, https://doi.org/10.5194/bg-22-305-2025, 2025.

Zhu, Q., Liu, J., Peng, C., Chen, H., Fang, X., Jiang, H., Yang, G., Zhu, D., Wang, W., and Zhou, X.: Modelling methane emissions from natural wetlands by development and application of the TRIPLEX-GHG model, Geoscientific Model Development, 7, 981–999, https://doi.org/10.5194/gmd-7-981-2014, 2014.

Zhu, Q., Yuan, K., Li, F., Riley, W. J., Hoyt, A., Jackson, R., McNicol, G., Chen, M., Knox, S. H., Briner, O., Beerling, D., Gedney, N., Hopcroft, P. O., Ito, A., Jain, A. K., Jensen, K., Kleinen, T., Li, T., Liu, X., McDonald, K. C., Melton, J. R., Miller, P. A., Müller, J., Peng, C., Poulter, B., Qin, Z., Peng, S., Tian, H., Xu, X., Yao, Y., Xi, Y., Zhang, Z., Zhang, W., Zhu, Q., and Zhuang, Q.: Critical needs to close monitoring gaps in pan-tropical wetland CH4 emissions, Environmental Research Letters, 19, 114 046, https://doi.org/10.1088/1748-9326/ad8019, 2024.

Zhu, Q., Jacob, D. J., Yuan, K., Li, F., Runkle, B. R. K., Chen, M., Bloom, A. A., Poulter, B., East, J. D., Riley, W. J., McNicol, G., Worden, J., Frankenberg, C., and Halabisky, M.: Advancements and opportunities to improve bottom–up estimates of global wetland methane emissions, Environmental Research Letters, 20, 023 001, https://doi.org/10.1088/1748-9326/adad02, 2025.

Zhuang, Q., Melillo, J. M., Kicklighter, D. W., Prinn, R. G., McGuire, A. D., Steudler, P. A., Felzer, B. S., and Hu, S.: Methane fluxes between terrestrial ecosystems and the atmosphere at northern high latitudes during the past century: A retrospective analysis with a process-based biogeochemistry model, Global Biogeochemical Cycles, 18, https://doi.org/https://doi.org/10.1029/2004GB002239, 2004.