Zhen Zhang[1,2], Sheel Bansal[3], Kuang-Yu Chang[4], Etienne Fluet-Chouinard[5], Kyle Delwiche[6], Mathias Goeckede[7], Adrian Gustafson[8], Sara Knox[9], Antti Leppänen[10], Licheng Liu[11], Jinxun Liu[12], Avni Malhotra[13], Tiina Markkanen[10], Gavin McNicol[14], Joe R. Melton[15], Paul A. Miller[8], Changhui Peng[16], Maarit Raivonen[10], William J. Riley[4], Oliver Sonnentag[17], Aalto Tuula[10], Rodrigo Vargas[18], Wenxin Zhang[8], Qing Zhu[4], Qiuan Zhu[19], Qianlai Zhuang[11], Lisamarie Windham-Myers[20], Robert B. Jackson[21], Benjamin Poulter[22]

[1]Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD 20740, USA

[2]Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, 100101, China

[3]Northern Prairie Wildlife Research Center, U.S. Geological Survey, Jamestown, ND, USA

[4]Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA

[5]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

[6]Department of Environmental Science, Policy, and Management, University of California, Berkeley, USA

[7]Department of Biogeochemical Signals, Max Planck Institute for Biogeochemistry, Jena, Germany

[8]Department of Physical Geography and Ecosystem Science, Lund University, Sweden

[9]The University of British Columbia, Vancouver, BC, Canada

[10]Finnish Meteorological Institute, Climate System Research Unit, Helsinki, Finland

[11]Department of Earth, Atmospheric, and Planetary Sciences and Department of Agronomy, Purdue University, West Lafayette IN, USA

[12]Western Geographic Science Center, U.S. Geological Survey, Moffett Field, CA 94035, USA

[13]Department of Geography, University of Zurich, Zurich, Switzerland

[14]Department of Earth and Environmental Sciences, University of Illinois Chicago, Chicago, WI, USA

[15]Environment and Climate Change Canada, Climate Research Division, Victoria, BC, Canada

[16]Department of Biology Sciences, Institute of Environment Science, University of Quebec at Montreal, Montreal, QC H3C 3P8, Canada

[17]Département de géographie, Université de Montréal, Montréal, QC, Canada

[18]Department of Plant and Soil Sciences, University of Delaware, Newark, DE, USA

[19]College of Hydrology and Water Resources, Hohai University, Nanjing, 210098, China

[20]Water Mission Area, U.S. Geological Survey, Menlo Park, CA, USA

[21]Department of Earth System Science, Stanford University, Stanford, California, CA, USA

[22]Biospheric Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20740, USA

Corresponding author: Zhen Zhang (zzhang88@umd.edu)

Key Points:

- Significant model-observation disagreements were found mainly at short time scales (< 15 days).
- Models captured variability at long time scales for boreal and Arctic tundra sites but not for temperate and tropical sites.
- Capturing flux variability at short time scales is critical to improving the performance of wetland methane models.
-

Abstract

Process-based land surface models are important tools for estimating global wetland methane (CH_4) emissions and projecting their behavior across space and time. So far there are no performance assessments of model responses to drivers at multiple time scales. In this study, we apply wavelet analysis to identify the dominant time scales contributing to model uncertainty in the frequency domain. We evaluate seven wetland models at 23 eddy covariance tower sites. Our study first characterizes site-level patterns of freshwater wetland CH_4 fluxes (FCH_4) at different time scales. A Monte Carlo approach has been developed to incorporate flux observation error to avoid misidentification of the time scales that dominate model error. Our results suggest that 1) significant model-observation disagreements are mainly at short- to intermediate time scales (< 15 days); 2) most of the models can capture the CH_4 variability at long time scales (> 32 days) for the boreal and Arctic tundra wetland sites but have limited performance for temperate and tropical/subtropical sites; 3) model error approximates pink noise patterns, indicating that biases at short time scales (< 5 days) could contribute to persistent systematic biases on longer time scales; and 4) differences in error pattern are related to model structure (e.g. proxy of CH_4 production). Our evaluation suggests the need to accurately replicate FCH_4 variability in future wetland CH_4 model developments.

Plain Language Summary

Land surface models are useful tools to estimate and predict wetland methane (CH_4) flux but so far there is no evaluation of modeled CH_4 error at different time scales. Here we use a statistical approach and observations from eddy covariance sites to evaluate the performance of seven wetland models for different wetland types. The results suggests model have captured CH_4 flux variability at monthly or longer time scales for boreal and Arctic tundra wetlands but have limited performance for temperate and tropical/subtropical wetlands. The analysis suggests that improving modeled flux at short time scale is important for future model development.

1 Introduction

Methane (CH_4) is the second most important greenhouse gas in terms of radiative forcing whose concentration in the atmosphere (~ 1900 ppb) has increased by approximately 150% since pre-industrial times (Canadell et al., 2021; IPCC, 2013). Methane emitted from wetland ecosystems is the largest natural source at ~120-180 Tg CH_4 yr^{-1} (Poulter et al., 2017; Saunois et al., 2020) and contributes to the short-term trend and interannual variability observed in atmospheric CH_4 concentration (Bousquet et al., 2006; Saunois et al., 2017; Zhen Zhang et al., 2021). Our understanding of global wetland CH_4 emissions heavily relies on process-based wetland CH_4 models, which incorporate biogeochemical mechanisms, climate forcing variables (e.g., temperature), and spatio-temporal distributions of surface inundation and wetland extent across the world (Melton et al., 2013; Wania et al., 2013; Xu et al., 2016; Zhang et al., 2021). These models play a critical role in diagnosing and forecasting terrestrial CH_4 dynamics across space and time, but their wetland CH_4 flux (FCH_4) estimates have large uncertainties due to potential biases in parameterizations, limited mechanistic characterization of known CH_4 processes, and limited integration of newly-identified processes such as thermal impact of rainfall (Neumann et al., 2019) and microbial dynamics on FCH_4 (Chadburn et al., 2020). However, it is unclear how well the current wetland models can replicate the observed FCH_4 varibility and magnitude at different time scales. Therefore, it is necessary to evaluate wetland CH_4 model performance against observations to identify model error patterns and inform future model development.

So far there has not been a major synthesis effort to evaluate multiple wetland CH_4 models against global coverage of eddy covariance (EC) observations from different biomes using a standard simulation protocol, despite a few efforts to evaluate a single wetland CH_4 model at multiple sites (Ringeval et al., 2014; Wania et al., 2010) and a model inter-comparison (Melton et al., 2013; Wania et al., 2013). Arguably, model development to represent terrestrial CH_4 processes has been hindered by 1) limited number of local-to-regional CH_4 observations to evaluate model outputs; 2) lack of understanding of the underlying processes derived from EC measurements and how well these processes represented in the models. Evaluations of wetland CH_4 models against the recently complied database FLUXNET-CH4 (Delwiche et al., 2021; Knox et al., 2019, 2021) offer

an opportunity to improve understanding of current model performance for different wetland types.

Despite previous observational synthesis studies(Chang et al., 2021; Delwiche et al., 2021; Knox et al., 2021) that have identified the major controlling factors that regulate temporal variations in freshwater wetland FCH_4 at different time scales, it is currently unknown how accurate wetland CH_4 models are in predicting FCH_4 and what factors are likely causing model biases across different time scales. Knox et al. (2021) demonstrated that dominant factors controlling the seasonality in EC-based FCH_4 vary with wetland types and the major processes that regulate FCH_4 vary at different time scales (e.g., from sub-daily to seasonal). For example, although soil temperature simulations are well established in wetland models with different thermal parameterization schemes, the representation of the modeled relationship between FCH_4 and temperature should be closely evaluated since it may affect model performance for some regions. Examples include cold regions influenced by freeze-thaw cycles where CH_4 fluxes may occur during the zero-curtain period (Tao et al., 2021; D. Zona et al., 2016). In addition, temperature hysteresis could contribute to different FCH_4 drivers across seasons (Chang et al., 2021). In contrast, models tend to use different proxies to calculate microbial CH_4 production (e.g., Gross Primary Production (GPP), Net Primary Production (NPP), ecosystem Heterotrophic Respiration (Reco), and carbon substrate concentrations), which likely influences model skill in reproducing FCH_4 at different time scales.

It is difficult to diagnose the mechanisms responsible for the lack of agreement between model and observation using conventional model-fitting approaches (Schaefer et al., 2012; Taylor, 2001) that apply statistical metrics (e.g., RMSE, r^2, standard deviation). In constrast, model-observation evaluations in the frequency domain using wavelet analysis (Figure 1) or Fourier transform can provide insights about model-observation disagreements at different temporal scales (Dietze et al., 2011; P. C. Stoy et al., 2013; Vargas et al., 2010). Wavelet analysis is especially useful for model evaluation since, compared to Fourier transform, it can identify not only the time scales that influence a signal but also inform when those time scales are significant. Previous studies have identified disagreement between models and observations for carbon dioxide (CO_2) fluxes across different ecosystems (Richardson et al., 2012; Schwalm et al., 2010; P. C. Stoy et al., 2013) but so far there is no assessment for CH_4 fluxes. Consequently, assessments of model-observation agreements using wavelet analysis are needed to identify discrepancies between observed and modeled CH_4 fluxes and provide insights for model development.

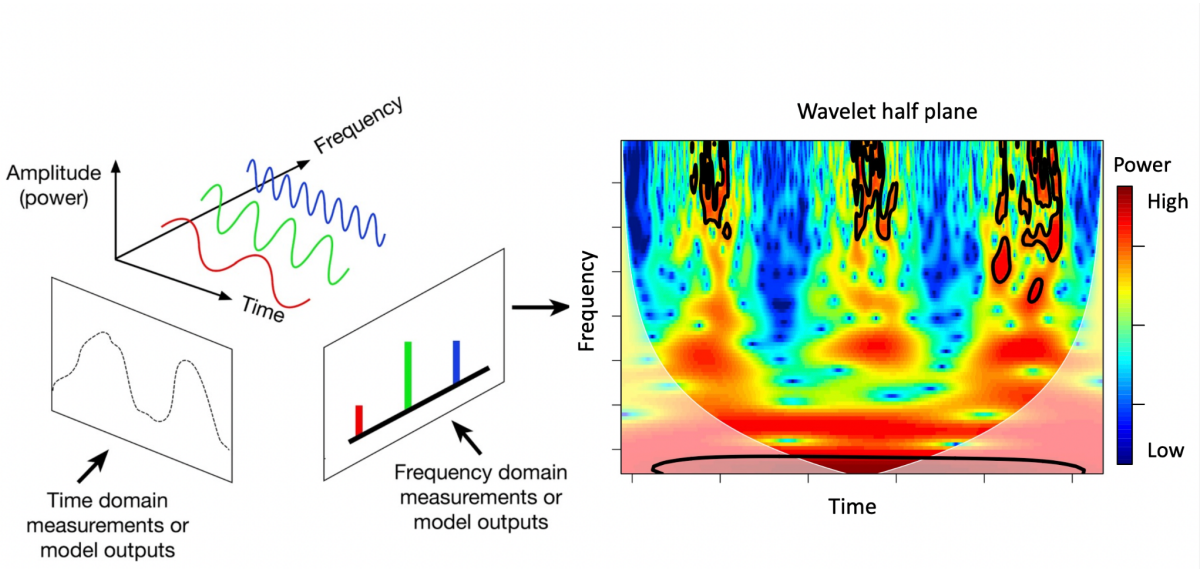

Figure 1. A conceptual description of differences between information in the time-domain and the frequency domain and an example of model-data evaluation in the frequency domain (adapted from Vargas et al., 2010). A time series can be decomposed into time and frequency (i.e. time scale) domain using the continuous wavelet transform. The resulting wavelet power spectra are plotted on what is referred to as wavelet half-plane, where time is along x axis, frequency/time scale along y axis, and spectra power indicated by color. The area of high spectra power is indicated by hot colors and vice versa. Significant frequencies are contoured with black lines. The black line is the cone of influence (COI) beyond which wavelet coefficients are unreliable (referred to as 'edge effect').

Our study aims to evaluate the performance of wetland CH_4 models in the frequency domain against a large ensemble of eddy covariance measurements of ecosystem-scale CH_4 fluxes. The goal is to quantify the most important time scales (e.g., multi-day, monthly, and seasonal) for the variability of CH_4 fluxes across wetland types and provide insights about the time periods in which models should be improved. Our specific objectives are to: i) quantify the most relevant time-scales for the variability of CH_4 fluxes in the models and observations at the site-level, ii) test the disagreement between in-situ observations and modeled CH_4 fluxes in time-frequency domain, iii) give insights into model structures responsible for model/observation mismatch. Based on previous findings for CO_2 flux (Dietze et al., 2011; P. C. Stoy et al., 2013), we hypothesize that i) models will have better model-observation agreement in terms of flux variability at longer time scales (e.g., monthly to seasonal) than short to interdediate time scale (e.g. multi-day to sub-monthly) as important biological processes regulated by seasonal variation (e.g., CH_4 production response to temperature) is properly formulated in the models; ii) models will tend to fail at

short to intermediate time scales due to forcing error propagation and limited representation of modeled plant physiology and biogeochemical processes; iii) The models have better performance over boreal and Arctic tundra sites than temperate and tropical sites, as temperature become less dominating in controlling FCH_4 variability for those wetland types.

2 Materials and Methods

We used data from 23 freshwater wetland sites included in the FLUXNET-CH4 Community Product (Delwiche et al., 2021) to evaluate seven wetland CH_4 models from the Global Carbon Project (GCP) Methane Budget (Saunois et al., 2020; Stavert et al., 2021). The model simulations follow a common simulation protocol using a gridded climate dataset from Climate Research Unit (CRU/CRU-JRA; CRU-JRA is a 6-hourly interpolated climate dataset from Japanese Reanalysis data; JRA, that is aligned with CRU on the monthly basis) as the inputs. The applied eddy covariance (EC) sites have a total of 70 site-years of data classified as boreal forest/taiga (n=25), Arctic tundra (n=15), temperate (n=25), and tropical/subtropical (n=5). We take into account the flux measurement errors in identifying model-data disagreements with observations by assessing the contribution of flux-tower observations error via a Monte Carlo approach.

2.1 FLUXNET- CH_4

Twenty-three sites from the FLUXNET-CH4 database were selected for the analysis (Table 1; Figure 2) based on three criteria: (1) the tidal, upland, and agricultural sites were excluded from the analysis as the models only simulate natural inland freshwater wetland CH_4 fluxes; (2) All seven models must have at least one complete site-year of results at the selected sites; and (3) Restored freshwater wetlands at later stages of wetland development ($>$ 10 years) were included in the analysis.

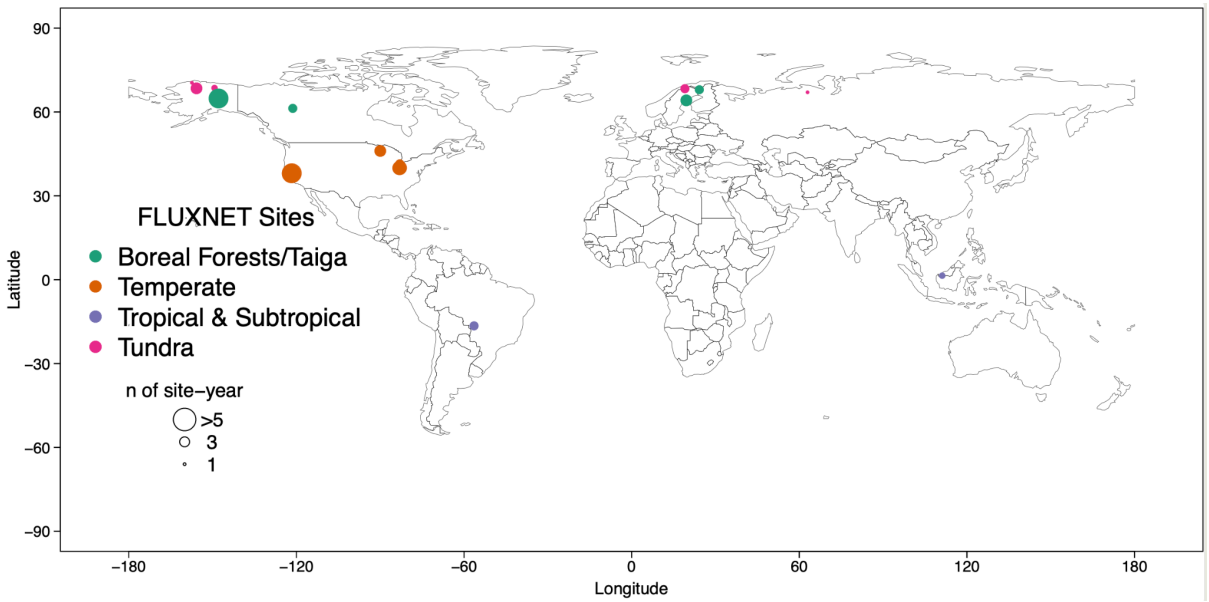

Figure 2. Locations of eddy covariance sites from FLUXNET-CH_4 in this study, with sites colored by wetland types. The variable size of dots in the map corresponds to the sample size (i.e., number of site-years) with a larger dot indicating a higher sample size. Base map used came from https://hub.arcgis.com/datasets/esri::world-countries.

In order to match the broad definition of freshwater wetlands in the models, the selection of EC sites is regrouped to represent a broad wetland/biome type along a latitudinal gradient. The original freshwater wetland types were classified into bog, fen, marsh, swamp, rice paddy, and drained wetland based on site-specific literature (Delwiche et al., 2021). The biome types (Arctic tundra, boreal forest/taiga, temperate, and tropical/subtropical), were defined based upon Olson et al. (2001) using site coordinates and vegetation types to group wetland sites. Since continuous wavelet decomposition requires a gap-free time series, we used gap-filled data from the FLUXNET-CH_4 database. Details on data standardization, gap-filling, and partitioning are described in Knox et al. (2019) and Delwiche et al. (2021). Gaps in CH_4 fluxes were filled using artificial neural networks (ANN) as described in Knox et al. (2019). An estimate of CH_4 flux observation error at every time step was generated, accounting for uncertainties associated with the gap-filling process and random measurement errors. These uncertainties are incorporated in the spectral null model, as described in the spectral analysis (Section 2.3).

2.2. Wetland FCH_4 models

Our study applies seven global wetland CH_4 models from the Global Carbon Project Methane Budget activities (Saunois et al., 2020). The details about the structure and configurations of the wetland CH_4 models can be found in

Table 2. All the models were run to steady-state using their own parameters and no site-specific tunings were done. Ancillary data such as soil texture and CH_4-related parameter sets were used as model-specific inputs (Table 2). Thus the assumptions about the local environment at each site depended on the individual model's setup. The models were run at the global scale at their native spatial resolution following a prescribed protocol to facilitate intercomparison. The models were run at the grid cell level using the CRU-JRA 6-hourly, land surface, gridded climate dataset, which was constructed by combining the Climate Research Unit (CRU) dataset and the reanalysis from Japanese Reanalysis data (JRA) produced by the Japanese Meteorological Agency (JMA). The CRU-JRA was adjusted where possible to align with the monthly climate dataset CRU (version ts3.26) data. One exception to the use of climate inputs is the LPJ-wsl model, which uses the monthly CRU dataset, and a weather generator within the model to produce precipitation events and daily temperature. Here we evaluate the wetland FCH_4 strength (Unit: $gCH_4\ m^{-2}\ day^{-1}$), which was defined as the total flux over a 24 hour period over a standardized wetland area (m^2), to excluding the effect of extent of inundation in the FCH_4 calculation.

The wetland CH_4 models can be generally described as a set of functions describing the biogeochemical processes that control CH_4 production and oxidation through methanogenesis and methanotrophy, and the biophysical processes that regulate CH_4 transport from the soil to the atmosphere (Table 2). Methanogenesis in the models is linked to different proxies (e.g., carbon substrate, heterotrophic respiration, net primary production) with a wide range of model complexity - more sophisticated models include wetland Plant Functional Types (PFTs) and explicitly simulate the processes of CH_4 production, consumption, and transport, while the simplified models use generalized empirical equations to simulate net flux without considering individual components of methane flux. The more complex model structure could offer capacity to capture the important temporal patterns of CH_4 fluxes but this invariably leads to additional extra parameter uncertainty due to the scarcity of observational constraints. The response function of CH_4 dynamics to temperature in each model is another factor that influences the simulated time series of CH_4 fluxes. For example, for high-latitude wetlands, model representations of freeze-thaw cycles influence the performance in capturing FCH_4 during the earlier thawing season and zero-curtain period (D. Zona et al., 2016).

2.3 Evaluation strategy and wavelet analyses

This analysis focused on the comparison of observed and simulated FCH_4. All analyses were conducted using daily time series. Since the modeled carbon fluxes are not directly comparable to the eddy covariance measurements due to the spatial mismatch between modeled gridded fluxes and site-level observations, we evaluate simulated FCH_4 by calculating the normalized residual error (NRE, $_{s,m,t}$) between normalized model and observation as:

$$\varepsilon_{s,m,t} = \left(\frac{Model_{s,m,t} - \overline{Model_{s,m,t}}}{\sigma_{s,m}}\right) - \left(\frac{Data_{s,t} - \overline{Data_{s,t}}}{\sigma_s}\right) \quad (1)$$

Where the subscripts denote site (s), model (m), and time (t) and the overbar denotes the average over the full length of the time series. The model and observation results were mean-centered to eliminate biases in the net flux, and divided by the standard deviation ( ) across the entire record to normalize the amplitude of variability. This NRE metric can be used to compare the synchrony of the model with the observation rather than evaluating absolute model biases.

We applied wavelet analysis to decompose the FCH_4 time series into an additive series of wave functions that have time scales of variability from 2 to 124 days. Wavelet analysis can identify the time scales that dominate a signal because wave functions that best match the fluctuations in the data will explain the most variance (i.e., power). Specifically, we used the continuous wavelet transform because of its ability to translate a time series into the frequency domain and its suitability for visual interpretation. The ability to discern small intervals of scales (i.e., spectral resolution) depends on the choice of the mother wavelet function. For this, we applied the widely Morlet wavelet, a complex non-orthogonal wavelet that has been widely used for geophysical applications (Torrence & Compo, 1998) and biometeorological measurements (Meyers et al., 1993). Following a similar definition from Knox et al., (2021), time scales of variation were classified into four bands, 'multiday scale' (2 to 5 days), 'weekly scale' (5 to 15 days), 'monthly scale' (15 to 42 days), and the 'seasonal scale ($>$ 42 days). The four bands were then summarized on both a by-site and by-model basis regarding the relative contribution of each band to the overall spectra. The continuous wavelet decomposition was computed using the Morlet wavelet basis function (function name: wt) from the R package 'biwavelet' (Gouhier et al., 2021). We use the bias-corrected wavelet power following Liu et al., (2007) to ensure a consistent definition of power in order to enable comparisons across spectral peaks. Wavelet power spectra on very long timescales ($>$ 64 days) often exceed the so-called cone-of-influence (COI) beyond which edge effects become important due to incomplete time locality across frequencies. Therefore, the power spectra outside of COI will not be interpreted here.

An appropriate null model is important to determine whether the model-observation disagreement is statistically significant. We applied a similar approach to that of Dietze et al., (2011) to generate 1000 sets of 'pseudo' time series for each site using a Monte Carlo approach. The NRE between the pseudo time series and the original data and the wavelet spectra of the NRE were calculated in the same way as the model errors. The 1000 replicates of pseudo time series were generated with the uncertainties estimated by Knox et al. (2019) accounting for both uncertainties in the ANN-based gap-filling algorithm and measurement uncertainty. Systematic errors due to representativeness (Chu et al., 2021; Pallandt et al., 2021), lack of nocturnal mixing,

sub-mesoscale circulations, and other factors are not discussed here (Baldocchi, 2014; Peltola et al., 2015). We assume that the FCH_4 flux errors follow a double exponential distribution (Knox et al., 2019), which has a fatter tail than normal and is highly heteroscedastic, with error increasing linearly with the absolute magnitude of the flux, similar to CO_2 flux errors as suggested by previous studies (Hollinger & Richardson, 2005; Lasslop et al., 2008; Richardson et al., 2006, 2008). Also note that, because the uncertainty from ANN estimation was strongly linked to the sample size, the flux errors tended to be high during the non-growing season when the measurement availability was limited by local meteorological conditions such as the snow presence and soil freeze and thaw cycles.

The wavelet spectra were evaluated in the following ways:

- Significant spectra regions. The significant region was defined by counting the total number of area in the time-frequency distribution where the spectral characteristics of FCH_4 and model-data mismatch were statistically significant. It was calculated by re-coding significant power as 1 and non-significant power as 0 and then stacking all site-years to count the total number.
- Marginal distribution of power spectrum of the NRE. The disagreements in the marginal power spectra were aggregated by the four defined time bands to summarize model performance across different time scales.
- Scaling exponent for each model was calculated to explore the spectral noise properties. Scaling exponent was expressed as the slope of the log-log transformed relationship between frequency (i.e. time scale) and power. The scaling exponent with a range between 1-2 was considered as intermediate 'pink' noise between 'white' and 'red' noise. White or red noise indicated that if the modeled FCH_4 had a persistent memory effect (i.e., autocorrelation structure) that can be attributed to model error which resulted in larger and long-lived systematic biases at larger time scales.

One-way ANOVA was used to diagnose the relationship between model structure and the marginal distribution of spectra power for wetland types. The models were grouped by different structures (Table S1) to identify if there were significant differences ($p < 0.001$) between model groups.

3 Results

3.1 Wavelet decomposition of FCH_4 time series from LPJ-wsl at an example site

Figure 3 shows the time series of FCH_4 from the observations and one model (LPJ-wsl) and demonstrates its wavelet-based power spectra at one marsh site (US-WPT) in the central U.S. (Chu et al., 2015). We use this example to explain the Monte Carlo analysis with pseudo-data and discuss the model-observation disagreement in more detail. Figure 3a shows that FCH_4 simulated by the LPJ-wsl model generally captured the seasonal cycle, but with a lower magnitude at

the freshwater marsh site. The model also captured a dip in FCH_4 after the peak during the growing season, which is consistent with the observed temporal pattern. Figure 3b suggests that the temporal patterns of normalized FCH_4 between the model and observations have a good agreement ($r=0.83$, $p < 0.05$). The relatively high uncertainty in the observed FCH_4 at the beginning of 2011 is mainly due to the limited number of observations, which causes higher uncertainty in the gap-filling method. This example shows that the discrepancies between the modeled and observed FCH_4, and the NRE uncertainty range from the null model, tend to be higher during the growing season when the flux intensity is relatively high and highly variable, or when the data availability is limited (Figure 3c).

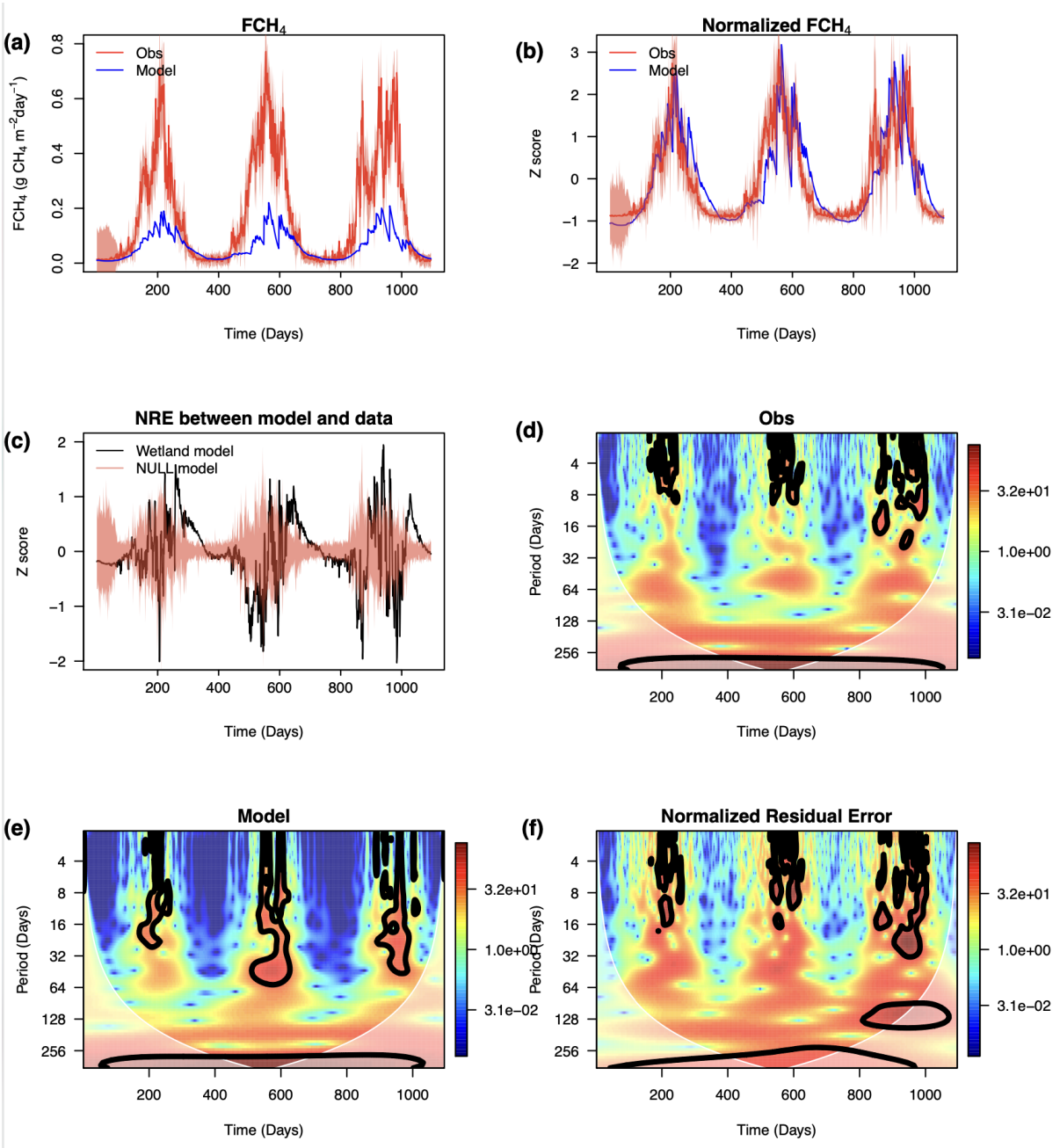

Figure 3. Example of wavelet decomposition and identification of the LPJ-wsl model error with eddy covariance (EC) observations at the US-WPT site for 2011-2013. (a) Time series of observations (Obs) methane flux (FCH_4, red line) with 1- observational uncertainty (shaded red area) and LPJ-wsl modeled

FCH_4 (Model, blue line). (b) Normalized time series of FCH_4 from model and observations; the shaded area in red represents the upper and lower range of the normalized pseudo time series from the Monte Carlo simulations. (c) Time series of normalized residual error (Z score of NRE) between wetland model and observations, with shaded area in red representing NRE between observations and normalized pseudo time series, i.e. NULL model. (d) Wavelet coefficients displayed in the wavelet half-plane for the normalized observations, (e) same as (d) but for LPJ-wsl modeled FCH_4 (Model), (f) NRE between model and the observations.

Both the observation and the model show significant power spectra during the growing seasons (Figure 3d and e). The modeled FCH_4 have a longer range of dominant time scale from 2-64 days than the observed 2-8 days. The modeled FCH_4 has weaker spectral powers (colors towards blue) during the non-growing season, indicating that the model may have less variability than the observations during the non-growing seasons (Figure 3d and e). It is important to note that the power spectra of the normalized residual error are not the difference between the wavelet coefficients displayed in Figure 3d and Figure 3e.

The wavelet plot for the NRE suggests the largest discrepancies is mostly from the growing seasons, reflected as strong spectral power in the wavelet NRE (Figure 3f). It is encouraging that there is a degree of correspondence between the model and observations: 1) the mismatch between model and observations is not significant at the long time scale (> 32 days) except for 2013 when there are strong anomaly in observed FCH_4 during the late growing season; 2) the wavelet coefficients in NRE have a low magnitude during the non-growing seasons, suggesting a less important role of the non-growing season fluxes at US-WPT. It is also worth noting that the observed FCH_4 has much higher year-to-year variations than the modeled fluxes, which is partly due to the strong influence of local environmental conditions on the measured seasonal cycle that are not captured by the model.

3.2 Evaluation of LPJ-wsl at the example site

Measurement-model discrepancies in LPJ-wsl at the US-WPT site were highest at daily to weekly scales. The significant regions (Figure 4) show that the measurements identify significant regions at high frequency (i.e., multi-day to weekly scales) while LPJ-wsl displays significant regions in the whole range of frequencies with more areas at the mid-to-low frequency (i.e., monthly to seasonal scales). LPJ-wsl tends to underestimate the time span of FCH_4 pulses at a high frequency, with strong pulse emission only occurring in late July, indicating less variability in the modeled FCH_4 during the growing season. Regarding the disagreement between model and observations, most of the significant regions are in the multi-day to weekly scales, suggesting the model failed to capture the flux variability at these time scales. The discrepancy in FCH_4 only occurs in the growing season from May to August while it is negligible during the non-growing seasons, when FCH_4 are small while uncertainty is proportionally large.

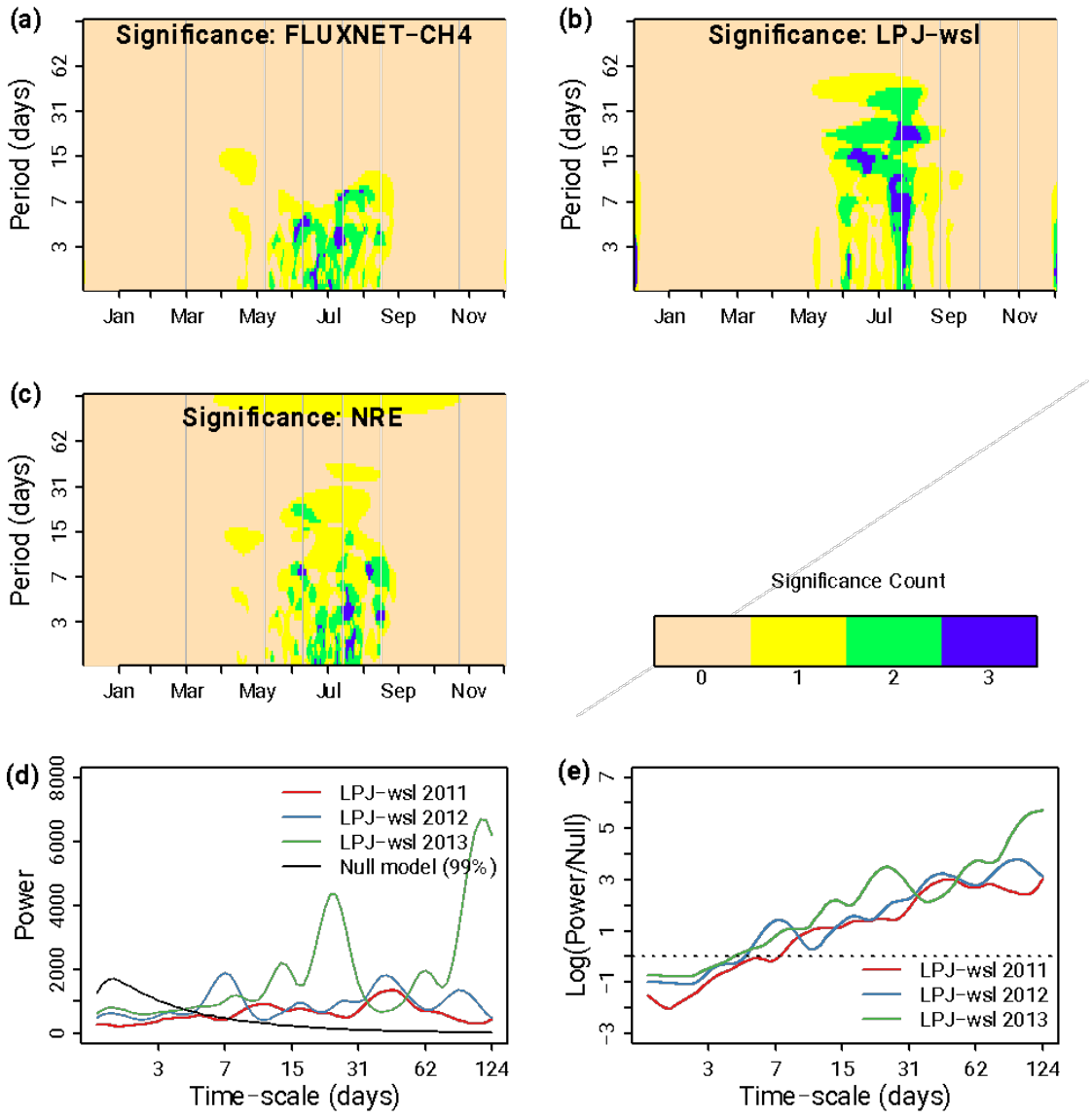

Figure 4. Wavelet evaluation of FCH_4 for the LPJ-wsl model against 3 site-year observations (2011-2013) at US-WPT site. (a) Count of significant power in the time-frequency domain for the time series of FCH_4 observations. (b) same as (a) but for LPJ-wsl modeled FCH_4. (c) Count of significant power of normalized residual errors (NRE) between model and observations. (d) Marginal distribution of power spectra of NRE as compared to the null spectra (99% of CI, solid black line) based on measurement uncertainties for each year 2011, 2012, and 2013 (red, blue, green lines, respectively). (e) The marginal distribution of power spectra of NRE divided by the maximum of the null spectra (NULL) on

a log scale. Values greater than 0 (dotted line) indicate that the model error has significantly more spectral power at those time scales than would be expected based on observation error.

Figure 4d provides an example of the model and observation mismatch in the global power spectrum for LPJ-wsl and observed FCH_4 at US-WPT in each year separately. Figure 4d is the marginal distribution of the full error spectrum by site-year in Figure 3f, in comparison to the maximum of the spectra of observation error from the Monte Carlo estimates. Here we choose a 99% confidence interval (CI) to define the criteria because, unlike CO_2, FCH_4 is highly spatially heterogeneous and has much higher year-to-year variability. To facilitate the comparison, we divided the model-data error spectra by the 99% CI of the observation error spectra for each time scale (Figure 4e). Any time scale that falls above the horizontal line (>1) indicates a model residual error that is 'significantly' higher than the uncertainty in the observations. The error of the model is constantly increasing with time-scale, while the random uncertainties in FCH_4 are declining with time-scale with the highest uncertainties at multi-day scales. Here the estimate of the scaling exponent for the LPJ-wsl model at US-WPT sites ranged from 1.5-1.7, suggesting a moderate correlation structure (i.e., pink noise).

3.3 Significant regions of NRE between the models and data

Next, we present the significant regions of model-data mismatches for all 23 sites and all seven models (Figure 5). Our results suggest that the models have diverse performance with the largest mismatch occuring at the short times scales (5-15 day). For most of the models, the significant mismatch is lower during monthly or seasonal time scales. This pattern confirms the hypothesis that the models generally have better performance in simulating the flux variability at longer time scales than at short-to-intermediate time scales. The increases in significant mismatch at the lowest frequency time scale (> 64 days) are mainly due to the edge effect, reflecting the limited length of the time series (365 days for a site–year) rather than a confirmation of model performance at capturing fluxes at the time scale. Across the wetland models there are diverse patterns of significant regions in FCH_4, most of which are different from the observation-based patterns, suggesting that there are significant discrepancies between model structures and observed process controls (Figure S1). The observation-based patterns suggest that most of the significant high power is concentrated within the time scale less than 7 days during the growing season (wet season for tropical sites), while the models tend to have relatively high power at a lower frequency (time scale larger than 14 days) at different time periods of the year dependeding on different model structure.

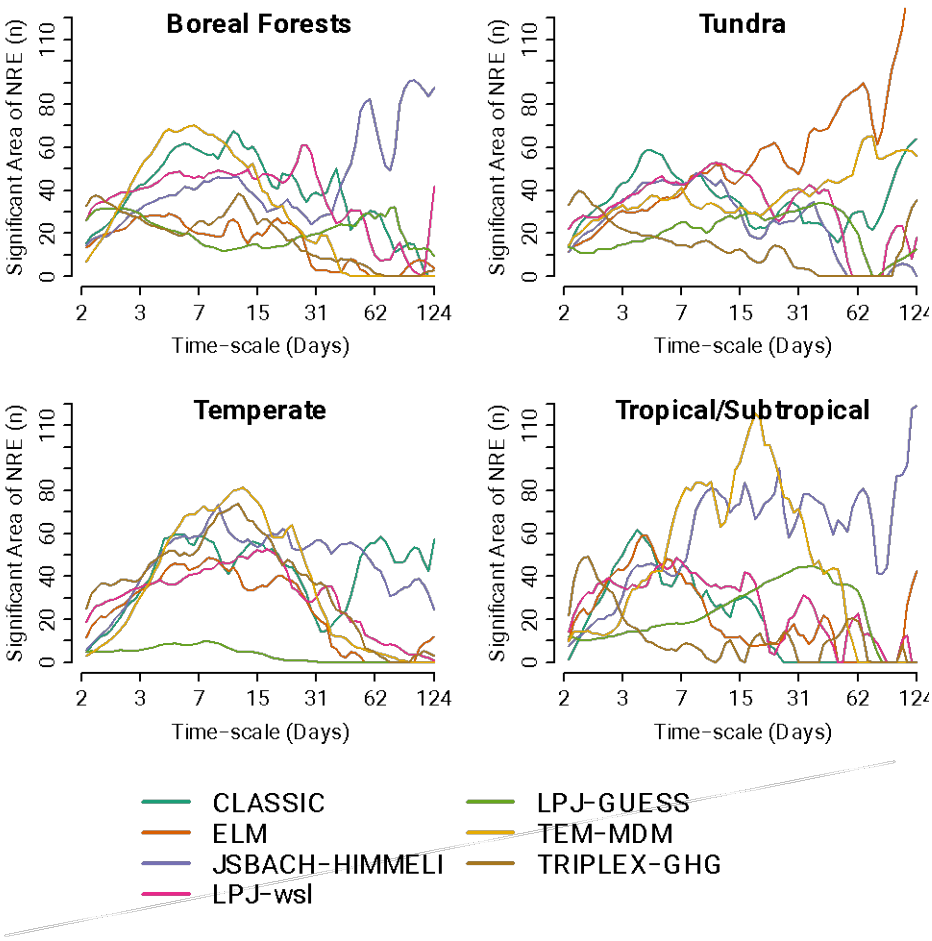

Figure 5. Significant model-observation disagreement along with time scales for all sites by biomes, represented by the marginal distribution of significant regions of NRE. High values of the significant region indicate high tendency of model-observation mismatch and vice versa. The significant region is defined as the areas where the wavelet power spectrum is statistically significant (95% CI). The marginal distribution of significant regions is then calculated by stacking all site-year to count the significant power in the time-frequency domain.

The comparison of significant regions in model-data mismatch suggests that the models have varying behavior on different wetland types (Figure 5). The majority of models show broadly consistent patterns of significant mismatch across time-scales for the boreal forest and temperate regions. In contrast, the patterns for tropical and subtropical wetland types is diverse among the wetland models. Note that the small sample size of tropical/subtropical wetlands in our study also has an impact on the representativeness of site-level observations.

The significant regions for boreal forest and Arctic tundra regions show high power during the growing seasons (Figure S2), indicating a consistent dominant control (likely temperature) in the models for these wetlands as suggested by recent studies (Irvin et al., 2021; Knox et al., 2021). For tropical/subtropical wetlands, the significant regions in NRE are spread over all time-scales with diverse patterns across the models, indicating the causes of mismatch with models differ as temperature becomes less dominant in controlling FCH_4 variability and other processes become more important.

3.4 Global Model spectra

We explored the model error patterns by calculating the scaling factor for each model. When considering the observation error in the flux data (the null model is calculated the same way as in Figure 4e), the spectral analysis of the NRE suggests the model errors approximate pink noise patterns for all the wetland models, with the mean scaling exponent of the model estimates ranging from 1.1 to 1.6 for different wetland types (Figure 6). The mean scaling exponent for the boreal forest and Arctic tundra regions (1.1-1.3) was generally lower than that for temperate and tropical regions (1.5-1.6), suggesting the wetland model performance for the temperate and tropical/subtropical regions generally has a longer memory effect (i.e., high tendency for greater persistence of model error) than wetlands in high latitudes. All the models show an increase in error at the longer time scales (monthly to seasonal) and the greatest variability across models at short time scales. The low spectral error at shorter time scales is partly due to the significant structure of the observations, which is from the noise in the data. There was a tendency for the spectral error of some models to exhibit greater persistence than other models. For example, even though the LPJ-wsl model shows relatively low error compared to the other models for boreal and temperate wetlands, the scaling exponents of LPJ-wsl (1.8 and 1.6 respectively) are higher than most of the other models, suggesting that LPJ-wsl model error tends to have a larger memory effect. For the temperate and tropical wetlands, all the models show similar scaling exponents regardless of model structure, indicating the similarity of model behaviors in environmental controls for these wetlands.

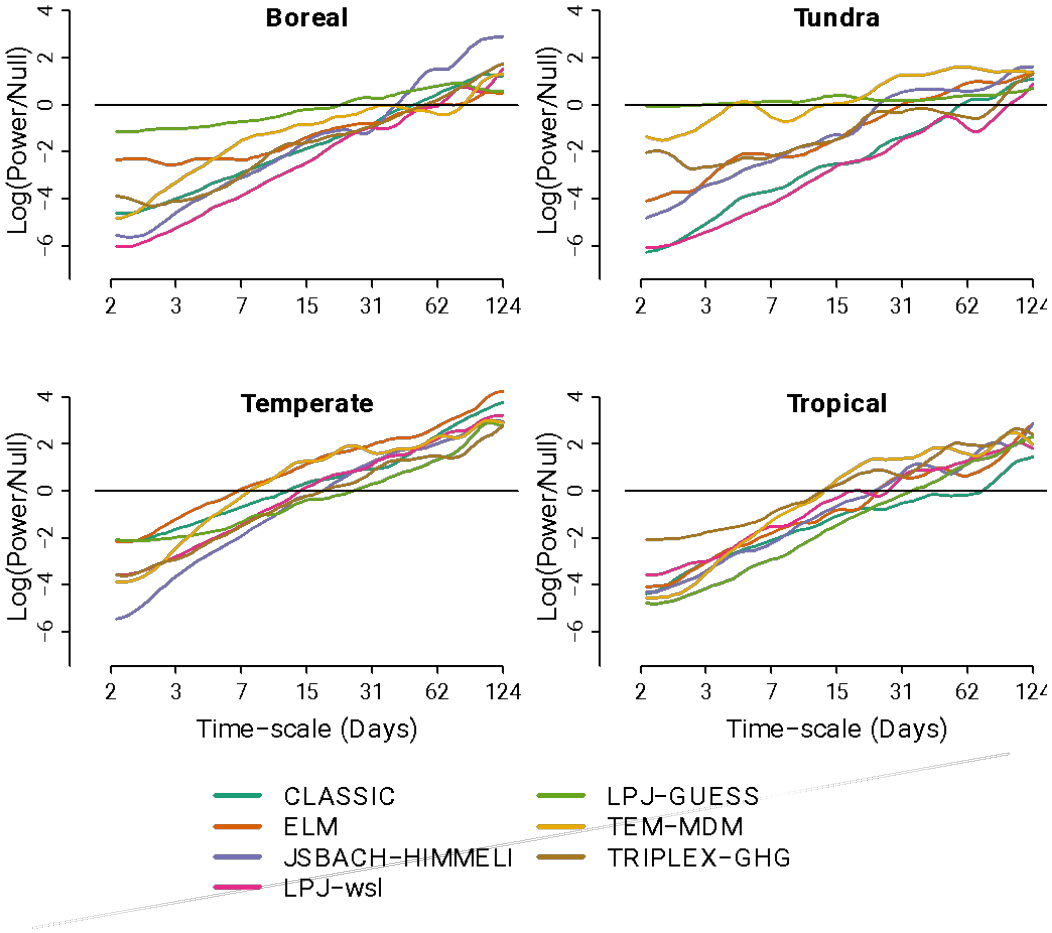

Figure 6. Model performance along time scales with a consideration of uncertainty in observations, reflected by the comparison of model error spectra to the null spectra. The power spectra (Power) are divided by the upper confidence interval of the null (NULL) model (99% quantile) based on logged observations, following the same calculation as Fig. 4e. A model error spectrum greater than 0 (horizontal black line) indicates more significant spectral power at these time scales.

4 Discussion and Conclusions

Our initial hypothesis was that models would perform well at monthly and seasonal time scales because the biogeochemical processes at these time scales are largely driven by solar radiation cycles and corresponding changes in soil temperature. Our results support this hypothesis for Arctic tundra and boreal wetland types where the variations of temperature are the dominant control of FCH_4 (Knox et al., 2021). However, in contrast to our expectations, the mod-

els have difficulty capturing variability at monthly and seasonal time scales for temperate and tropical wetlands, where other environmental controls emerge. Considering that the precipitation-driven variables such as water table depth are significantly correlated with the seasonal cycle of FCH_4 at the site level for temperate and tropical sites (Knox et al., 2021), the lower agreements between model and data may be partly caused by discrepancies in precipitation between gridded climate datasets and site-level meteorological conditions. The models also lack representation of hydrological processes at a scale fine enough to reflect the lateral flow from uplands to lowlands and its influence on the water dynamics. The distribution of model wavelet spectra (Figure S1) on visual inspection appears very different from the site-level measurements, indicating that the models' structures need to better capture variability at short to intermediate time scales (e.g., multi-day to weekly). This finding indicates that current models may have a biased seasonal cycle over temperate and tropical wetlands, as suggested by a few recent regional studies (Lunt et al., 2019; Maasakkers et al., 2021; Yu et al., 2021).

Our analysis further reveals important characteristics in the time series of model errors, which indicates that the errors at short time scales have a memory effect on biases at long time scales. These results suggest that further model development should focus first on correctly replicating flux variability and magnitude at short time scales. Investigations into modeled FCH_4 spectra (Figure S1) suggest that in general models are not variable enough over the year and tend to smooth over multi-day scale variability. One reason is likely that other environmental variables (e.g., vapor pressure deficit, atmospheric pressure) that regulate FCH_4 variability at short time scales (Stoy et al., 2005) are not included in the model inputs. Despite of this, many of the models predict a strong pulse in variability during a short time period, especially for the growing season, which causes significant errors at long time scales (Figure S2). This pattern has not been observed by the EC measurements, indicating shared model errors due to the meteorological forcing among models and/or due to missing processes arising from limited understanding of wetland ecosystem dynamics (Neumann et al., 2019; D. Zona et al., 2016).

The spectral properties of the model errors along with time scales (Figure 6) indicates that the model structure has an impact on FCH_4 variability, and different groups of models that share similar structure tend to have lower errors propagated from short time scales to high time scales. The ANOVA analysis (Table 3) suggests that the explicit representation of wetland plant functional types (PFTs), CH_4 component fluxes, and wetland production proxies is significantly associated with variance for boreal and Arctic tundra wetland FCH_4 prediction, with a modest and inconsistent effect for temperate and tropical wetlands. The effects of including the nitrogen cycle, fire, and spatial resolution of grid cells were non-significant for most of the time scales. In addition, CH_4 transport through aerenchyma and stomata, which is linked to photosynthesis, and other processes such as ventilation in aerenchymatous vegetation with influence from latent heat are critical for models to capture the variability at the

diel scale (Knox et al., 2021). Unfortunately, we did not have sub-daily FCH_4 model predictions nor were they driven by site-level meteorological forcings, so we could not evaluate whether representation of processes at the diel scale has an impact on model performance at intermediate scales.

The ranking of model performance across different time scales suggests that no model outperforms others at all time scales (Figure 7). Given different biogeochemical structures and parameterizations, the analysis suggests inclusion of representation of some key processes in wetland models and proper parameterizations are the basis for improving model performance. However, complex model structure does not guarantee superior model performance, which highlights the importance of properly parameterizing processes at a certain time scale. For instance, models with explicit CH_4 components and multiple wetland PFTs could perform worse than simple models at some time scales, which is likely due to increased uncertainty from parameterization due to poor observational constraints. A further diagnosis of what environmental and biotic parameters impact the agreement with EC measurements is needed for a better choice of parameter values in representing the realistic temporal variability of FCH_4.

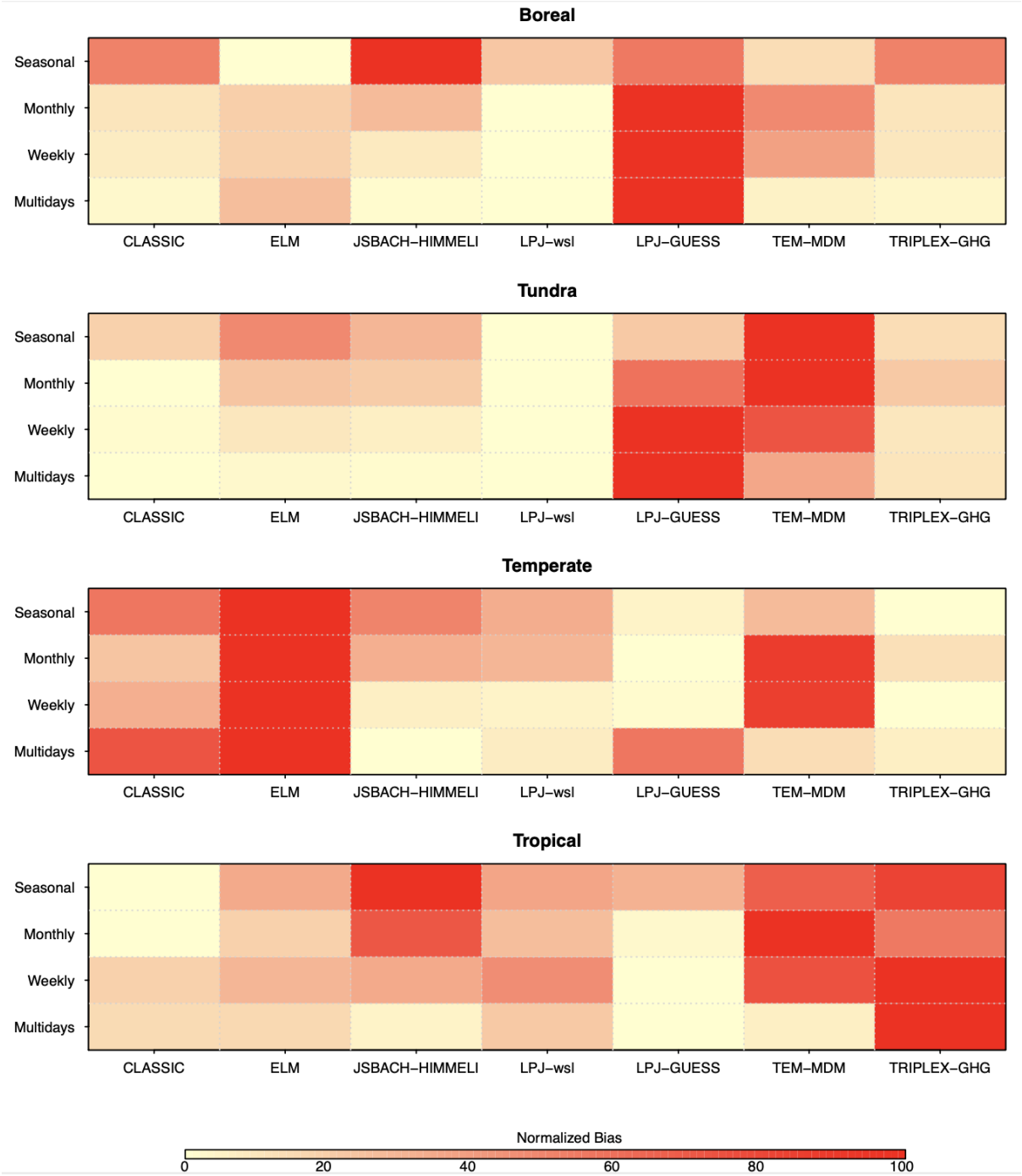

Figure 7. Heat map showing model error by time scales for different wetland

types. All of the model errors per time scale are normalized to 1-100 with the value of highest model error equal to 100 and lowest to 1. Light yellow and red represent the lowest and highest errors, respectively. The time scales are defined as 'Multiday ' (2 to 5 days), 'Weekly' (5 to 15 days), 'Monthly' (15 to 42 days), and 'Seasonal scale (> 42 days).

There are a few limitations in the observations affecting our model evaluation. First, the length of observed time series is limited across sites with few sites having more than 5-year records. Unlike CO_2, measurements of CH_4 are only beginning to cover multiple-year records and thus the EC tower records are not long enough to assess the model's performance in capturing annual and interannual variability. For spectral methods, the short records are particularly problematic for longer sub-annual time scales (e.g., seasonal) due to edge effects on the amount of usable data. Given that the wetland model results at annual and interannual time scales are particularly of interest to the global methane budget, having long-term records of measurements is importan for an evaluation of model performance at longer time scales. Second, the model-site comparisons are statistically challenging as the model-site-year combinations are not randomly distributed but rather depend on the performance at a few sites given the reality of unevenly distributed EC wetland sites. Both undoubtedly have the potential to introduce biases in statistical interpretation and thus influence model score. For instance, the evaluation of model performance for temperate wetlands is strongly affected by model simulations at one Marsh site US-TW1 in the United States, which is a restored wetland that contributes ~ 28% (n=7 site-years) of the total site-years for temperate wetlands. US-TW1 has a water table height managed at ~ 25 cm above the soil (Oikawa et al., 2017), which influences the temporal pattern of FCH_4 via hydrological control and thus model evaluations. All of the limitations indicate a critical need for more detailed evaluation of model performance at site-level and long-term measurements for underrepresented regions.

One of the important aspects of this analysis is that it is possible that the model performance was underestimated due to the limitation in estimating observation uncertainty and due to potential spatial mismatch between models and EC observations. Although we calculated the spectral uncertainty with the inclusion of observational errors in the evaluation across time scales, the interpretation of whether model errors falls outside the acceptable range is strongly influenced by the uncertainty of FCH_4 observations. The default gap-filling methods such as ANN-based estimates for observational uncertainty appear to be overly tight across all sites as suggested by a recent study (Irvin et al., 2021), indicating that actual observation error is higher than the estimates in our study. In addition, on top of the uncertainty of all the measurements, there is uncertainty originating from a mismatch between the footprints of the individual towers that are usually <1 km^2 and the size of gridded pixels that are often 0.5 degrees or larger. This footprint mismatch introduces additional noise.

Our study evaluated seven global-scale wetland models from the Global Carbon

Project Methane Assessment against eddy covariance CH_4 flux measurements from the FLUXNET-CH4 dataset in the time-frequency domain. This analysis helped to identify model errors in variability across different time scales and provided guidelines for further wetland model developments. Further detailed intercomparison of model structure and parameterizations is needed to diagnose model structural and parameterization errors. In particular, a more advanced intercomparison protocol would help distinguish structural and parameterization limitations by 1) testing multiple parmeterization schemes for major wetland processes (e.g. methane production rate and transport); 2) running the models with FLUXNET-CH4 meteorological forcing inputs and local site information such as slope,drainage, and vegetation characteristics; and 3) including longer-term records and spatially representative observations with full uncertainty characterization from EC tower measurements. Future intercomparison of wetland methane models would improve understanding of the role of wetland emissions in the variations of atmospheric CH_4 concentration during the past decades and future projections.

Acknowledgments

We acknowledge primary support from the Gordon and Betty Moore Foundation (Grant GBMF5439) and from the John Wesley Powell Center for Analysis and Synthesis of the U.S. Geological Survey ("Wetland FLUXNET Synthesis for Methane"). W.J.R., Q.Z, and K.Y.C were supported by the Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Computation (RUBISCO) Scientific Focus Area which is sponsored by the Earth and Environmental Systems Modeling (EESM) Program under the Office of Biological and Environmental Research of the U.S. Department of Energy Office of Science to Lawrence Berkeley National Laboratory (LBNL) under contract no. DEAC02-05CH11231. O.S. acknowledges funding by the Canada Research Chairs, Canada Foundation for Innovation Leaders Opportunity Fund, and Natural Sciences and Engineering Research Council Discovery Grant Programs. SB was funded by the U.S. Geological Survey, Ecosystem Mission Area, Land Change Science Climate R&D Program. W.Z. acknowledged the support by the Swedish Research Council (Vetenskapsrådet) start grant (2020-05338). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Data availability statement

The observational data that support the findings of this study are available in the FLUXNET-CH4 Community Product, available at https://fluxnet.org/data/fluxnet-ch4-community-product/. The modeled results are available at https://doi.org/10.5281/zenodo.7246403.

References

Arora, V. K., Melton, J. R., & Plummer, D. (2018). An assessment of natural methane fluxes simulated by the CLASS-CTEM model. *Biogeosciences*, *15*(15), 4683–4709. https://doi.org/10.5194/bg-15-4683-2018Baldocchi, D.

(2014). Measuring fluxes of trace gases and energy between ecosystems and the atmosphere – the state and future of the eddy covariance method. *Global Change Biology*, *20*(12), 3600–3609. https://doi.org/10.1111/gcb.12649Bohrer, G. [The O. S. U. (ORCID:0000000292099540), & Morin, T. H. [The O. S. U. (ORCID:0000000265046937). (2020, January 1). FLUXNET-CH4 US-ORv Olentangy River Wetland Research Park. United States. https://doi.org/10.18140/FLX/1669689Bousquet, P., Ciais, P., Miller, J. B., Dlugokencky, E. J., Hauglustaine, D. A., Prigent, C., et al. (2006). Contribution of anthropogenic and natural sources to atmospheric methane variability. *Nature*, *443*(7110), 439–443. https://doi.org/10.1038/nature05132Canadell, J. G., Monteiro, P. M. S., Costa, M. H., Cunha, L. C. da, Cox, P. M., Eliseev, A. V., et al. (2021). Global Carbon and other Biogeochemical Cycles and Feedbacks. In *IPCC AR6 WGI, Final Government Distribution* (p. chapter 5). Retrieved from https://hal.archives-ouvertes.fr/hal-03336145Chadburn, S. E., Aalto, T., Aurela, M., Baldocchi, D., Biasi, C., Boike, J., et al. (2020). Modeled Microbial Dynamics Explain the Apparent Temperature Sensitivity of Wetland Methane Emissions. *Global Biogeochemical Cycles*, *34*(11), e2020GB006678. https://doi.org/10.1029/2020GB006678Chang, K.-Y., Riley, W. J., Knox, S. H., Jackson, R. B., McNicol, G., Poulter, B., et al. (2021). Substantial hysteresis in emergent temperature sensitivity of global wetland CH4 emissions. *Nature Communications*, *12*(1), 2266. https://doi.org/10.1038/s41467-021-22452-1Chen, J. [Michigan S. U., & Chu, H. [Lawrence B. L. (ORCID:0000000281314938). (2020, January 1). FLUXNET-CH4 US-WPT Winous Point North Marsh. United States. https://doi.org/10.18140/FLX/1669702Chu, H., Gottgens, J. F., Chen, J., Sun, G., Desai, A. R., Ouyang, Z., et al. (2015). Climatic variability, hydrologic anomaly, and methane emission can turn productive freshwater marshes into net carbon sources. *Global Change Biology*, *21*(3), 1165–1181. https://doi.org/10.1111/gcb.12760Chu, H., Luo, X., Ouyang, Z., Chan, W. S., Dengel, S., Biraud, S. C., et al. (2021). Representativeness of Eddy-Covariance flux footprints for areas surrounding AmeriFlux sites. *Agricultural and Forest Meteorology*, *301–302*, 108350. https://doi.org/10.1016/j.agrformet.2021.108350Dalmagro, H. J., Zanella de Arruda, P. H., Vourlitis, G. L., Lathuillière, M. J., de S. Nogueira, J., Couto, E. G., & Johnson, M. S. (2019). Radiative forcing of methane fluxes offsets net carbon dioxide uptake for a tropical flooded forest. *Global Change Biology*, *25*(6), 1967–1981. https://doi.org/10.1111/gcb.14615Delwiche, K. B., Knox, S. H., Malhotra, A., Fluet-Chouinard, E., McNicol, G., Feron, S., et al. (2021). FLUXNET-CH : a global, multi-ecosystem dataset and analysis of methane seasonality from freshwater wetlands. *Earth System Science Data*, *13*(7), 3607–3689. https://doi.org/10.5194/essd-13-3607-2021Desai, A. R. [University of W.-M. (ORCID:0000000252266041), & Thom, J. [University of W.-M. (2020, January 1). FLUXNET-CH4 US-Los Lost Creek. United States. https://doi.org/10.18140/FLX/1669682Dietze, M. C., Vargas, R., Richardson, A. D., Stoy, P. C., Barr, A. G., Anderson, R. S., et al. (2011). Characterizing the performance of ecosystem models across time scales: A spectral analysis of the North American Carbon Pro-

gram site-level synthesis. *Journal of Geophysical Research: Biogeosciences*, *116*(G4), G04029s. https://doi.org/10.1029/2011JG001661Euskirchen, E. [University of A. F., Institute of Arctic Biology]. (2022a, January 1). AmeriFlux FLUXNET-1F US-BZF Bonanza Creek Rich Fen. United States. https://doi.org/10.17190/AMF/1881570Euskirchen, E. [University of A. F., Institute of Arctic Biology]. (2022b, January 1). AmeriFlux FLUXNET-1F US-BZS Bonanza Creek Black Spruce. United States. https://doi.org/10.17190/AMF/1881572Euskirchen, E. [University of A. F., Institute of Arctic Biology] (ORCID:0000000208484295), & Edgar, C. [University of A. F., Institute of Arctic Biology] (ORCID:0000000270268358). (2020, January 1). FLUXNET-CH4 US-BZB Bonanza Creek Thermokarst Bog. United States. https://doi.org/10.18140/FLX/1669668Euskirchen, E. [University of A. F., Institute of Arctic Biology] (ORCID:0000000208484295), Bret-Harte, M. [University of A. F., Institute of Arctic Biology] (ORCID:0000000151513947), & Edgar, C. [University of A. F., Institute of Arctic Biology] (ORCID:0000000270268358). (2020, January 1). FLUXNET-CH4 US-ICs Imnavait Creek Watershed Wet Sedge Tundra. United States. https://doi.org/10.18140/FLX/1669678Friborg, T. [University of C. (ORCID:0000000156336097), & Shurpali, N. [University of E. F., Finland] (ORCID:0000000310524396). (2020, January 1). FLUXNET-CH4 RU-Vrk Seida/Vorkuta. Russian Federation. https://doi.org/10.18140/FLX/1669658Gouhier, T. C., Grinsted, A., & Simko, V. (2021). *R package biwavelet: Conduct Univariate and Bivariate Wavelet Analyses*. Retrieved from https://github.com/tgouhier/biwaveletHollinger, D. Y., & Richardson, A. D. (2005). Uncertainty in eddy covariance measurements and its application to physiological models. *Tree Physiology*, *25*(7), 873–885. https://doi.org/10.1093/treephys/25.7.873IPCC. (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. Retrieved from https://www.ipcc.ch/report/ar5/wg1/Irvin, J., Zhou, S., McNicol, G., Lu, F., Liu, V., Fluet-Chouinard, E., et al. (2021). Gap-filling eddy covariance methane fluxes: Comparison of machine learning model predictions and uncertainties at FLUXNET-CH4 wetlands. *Agricultural and Forest Meteorology*, *308–309*, 108528. https://doi.org/10.1016/j.agrformet.2021.108528Iwata, H. [Shinshu U. (ORCID:0000000289628982), Ueyama, M. [Osaka P. U. (ORCID:0000000240004888), & Harazono, Y. [Osaka P. U. (2020, January 1). FLUXNET-CH4 US-Uaf University of Alaska, Fairbanks. United States. https://doi.org/10.18140/FLX/1669701Jansen, J. [Stockholm U. (ORCID:0000000159657662), Friborg, T. [University of C. (ORCID:0000000156336097), Jammet, M. [University of C. (ORCID:0000000305479842), & Crill, P. [Stockholm U. (ORCID:0000000311103059). (2020, January 1). FLUXNET-CH4 SE-St1 Stordalen grassland. Sweden. https://doi.org/10.18140/FLX/1669660Knox, S. H., Jackson, R. B., Poulter, B., McNicol, G., Fluet-Chouinard, E., Zhang, Z., et al. (2019). FLUXNET-CH4 Synthesis Activity: Objectives, Observations, and Future Directions. *Bulletin of the American Meteorological Society*,

100(12), 2607–2632. https://doi.org/10.1175/BAMS-D-18-0268.1Knox, S. H., Bansal, S., McNicol, G., Schafer, K., Sturtevant, C., Ueyama, M., et al. (2021). Identifying dominant environmental predictors of freshwater wetland methane fluxes across diurnal to seasonal time scales. *Global Change Biology*, *27*(15), 3582–3604. https://doi.org/10.1111/gcb.15661Lasslop, G., Reichstein, M., Kattge, J., & Papale, D. (2008). Influences of observation errors in eddy flux data on inverse model parameter estimation. *Biogeosciences*, *5*(5), 1311–1324. https://doi.org/10.5194/bg-5-1311-2008Liu, L., Zhuang, Q., Oh, Y., Shurpali, N. J., Kim, S., & Poulter, B. (2020). Uncertainty Quantification of Global Net Methane Emissions From Terrestrial Ecosystems Using a Mechanistically Based Biogeochemistry Model. *Journal of Geophysical Research: Biogeosciences*, *125*(6), e2019JG005428. https://doi.org/10.1029/2019JG005428Liu, Y., Liang, X. S., & Weisberg, R. H. (2007). Rectification of the Bias in the Wavelet Power Spectrum. *Journal of Atmospheric and Oceanic Technology*, *24*(12), 2093–2102. https://doi.org/10.1175/2007JTECHO511.1Lohila, A., Aurela, M., Tuovinen, J.-P., Laurila, T., Hatakka, J., Rainne, J., & Mäkelä, T. (2020). FLUXNET-CH4 FI-Lom Lompolojankka [Data set]. https://doi.org/10.18140/FLX/1669638. Retrieved from https://doi.org/10.18140/FLX/1669638Lunt, M. F., Palmer, P. I., Feng, L., Taylor, C. M., Boesch, H., & Parker, R. J. (2019). An increase in methane emissions from tropical Africa between 2010 and 2016 inferred from satellite data. *Atmospheric Chemistry and Physics*, *19*(23), 14721–14740. https://doi.org/10.5194/acp-19-14721-2019Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Scarpelli, T. R., Nesser, H., Sheng, J., et al. (2021). 2010–2015 North American methane emissions, sectoral contributions, and trends: a high-resolution inversion of GOSAT observations of atmospheric methane. *Atmospheric Chemistry and Physics*, *21*(6), 4339–4356. https://doi.org/10.5194/acp-21-4339-2021McGuire, A. D., Christensen, T. R., Hayes, D., Heroult, A., Euskirchen, E., Kimball, J. S., et al. (2012). An assessment of the carbon balance of Arctic tundra: comparisons among observations, process models, and atmospheric inversions. *Biogeosciences*, *9*(8), 3185–3204. https://doi.org/10.5194/bg-9-3185-2012Melton, J. R., & Arora, V. K. (2016). Competition between plant functional types in the Canadian Terrestrial Ecosystem Model (CTEM) v. 2.0. *Geoscientific Model Development*, *9*(1), 323–361. https://doi.org/10.5194/gmd-9-323-2016Melton, J. R., Wania, R., Hodson, E. L., Poulter, B., Ringeval, B., Spahni, R., et al. (2013). Present state of global wetland extent and wetland methane modelling: conclusions from a model inter-comparison project (WETCHIMP). *Biogeosciences*, *10*(2), 753–788. https://doi.org/10.5194/bg-10-753-2013Meyers, S. D., Kelly, B. G., & O'Brien, J. J. (1993). An Introduction to Wavelet Analysis in Oceanography and Meteorology: With Application to the Dispersion of Yanai Waves. *Monthly Weather Review*, *121*(10), 2858–2866. https://doi.org/10.1175/1520-0493(1993)121<2858:AITWAI>2.0.CO;2Neumann, R. B., Moorberg, C. J., Lundquist, J. D., Turner, J. C., Waldrop, M. P., McFarland, J. W., et al. (2019). Warming Effects of Spring Rainfall Increase Methane Emissions From Thawing Permafrost. *Geophysical Research Letters*, *46*(3), 1393–1401.

https://doi.org/etiopeNilsson, M. B. [Swedish U. of A. S., Department of Forest Ecology and Management] (ORCID:0000000337656399), & Peichl, M. [Swedish U. of A. S., Department of Forest Ecology and Management] (ORCID:0000000299405846). (2020, January 1). FLUXNET-CH4 SE-Deg Degero. Sweden. https://doi.org/10.18140/FLX/1669659Oikawa, P. Y., Jenerette, G. D., Knox, S. H., Sturtevant, C., Verfaillie, J., Dronova, I., et al. (2017). Evaluation of a hierarchy of models reveals importance of substrate limitation for predicting carbon dioxide and methane exchange in restored wetlands. *Journal of Geophysical Research: Biogeosciences*, *122*(1), 145–167. https://doi.org/10.1002/2016JG003438Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., et al. (2001). Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*, *51*(11), 933–938. https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2Pallandt, M., Kumar, J., Mauritz, M., Schuur, E., Virkkala, A.-M., Celis, G., et al. (2021). Representativeness assessment of the pan-Arctic eddy-covariance site network, and optimized future enhancements. *Biogeosciences Discussions*, 1–42. https://doi.org/10.5194/bg-2021-133Peltola, O., Hensen, A., Belelli Marchesini, L., Helfter, C., Bosveld, F. C., van den Bulk, W. C. M., et al. (2015). Studying the spatial variability of methane flux with five eddy covariance towers of varying height. *Agricultural and Forest Meteorology*, *214–215*, 456–472. https://doi.org/10.1016/j.agrformet.2015.09.007Poulter, B., Bousquet, P., Canadell, J. G., Ciais, P., Peregon, A., Saunois, M., et al. (2017). Global wetland contribution to 2000–2012 atmospheric methane growth rate dynamics. *Environmental Research Letters*, *12*(9), 094013. https://doi.org/10.1088/1748-9326/aa8391Raivonen, M., Smolander, S., Backman, L., Susiluoto, J., Aalto, T., Markkanen, T., et al. (2017). HIMMELI v1.0: HelsinkI Model of MEthane buiLd-up and emIssionfor peatlands. *Geoscientific Model Development Discussions*, 1–45. https://doi.org/10.5194/gmd-2017-52Richardson, A. D., Hollinger, D. Y., Burba, G. G., Davis, K. J., Flanagan, L. B., Katul, G. G., et al. (2006). A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes. *Agricultural and Forest Meteorology*, *136*(1), 1–18. https://doi.org/10.1016/j.agrformet.2006.01.007Richardson, A. D., Mahecha, M. D., Falge, E., Kattge, J., Moffat, A. M., Papale, D., et al. (2008). Statistical properties of random CO2 flux measurement uncertainty inferred from model residuals. *Agricultural and Forest Meteorology*, *148*(1), 38–50. https://doi.org/10.1016/j.agrformet.2007.09.001Richardson, A. D., Anderson, R. S., Arain, M. A., Barr, A. G., Bohrer, G., Chen, G., et al. (2012). Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis. *Global Change Biology*, *18*(2), 566–584. https://doi.org/10.1111/j.1365-2486.2011.02562.xRiley, W. J., Subin, Z. M., Lawrence, D. M., Swenson, S. C., Torn, M. S., Meng, L., et al. (2011). Barriers to predicting changes in global terrestrial methane fluxes: analyses using CLM4Me, a methane biogeochemistry model integrated in CESM. *Biogeosciences*, *8*(7), 1925–1953. https://doi.org/10.5194/bg-8-1925-

2011Ringeval, B., Houweling, S., van Bodegom, P. M., Spahni, R., van Beek, R., Joos, F., & Röckmann, T. (2014). Methane emissions from floodplains in the Amazon Basin: challenges in developing a process-based model for global applications. *Biogeosciences*, *11*(6), 1519–1558. https://doi.org/10.5194/bg-11-1519-2014Saunois, M., Bousquet, P., Poulter, B., Peregon, A., Ciais, P., Canadell, J. G., et al. (2017). Variability and quasi-decadal changes in the methane budget over the period 2000–2012. *Atmospheric Chemistry and Physics*, *17*(18), 11135–11161. https://doi.org/10.5194/acp-17-11135-2017Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., et al. (2020). The Global Methane Budget 2000–2017. *Earth System Science Data*, *12*(3), 1561–1623. https://doi.org/10.5194/essd-12-1561-2020Schaefer, K., Schwalm, C. R., Williams, C., Arain, M. A., Barr, A., Chen, J. M., et al. (2012). A model-data comparison of gross primary productivity: Results from the North American Carbon Program site synthesis. *Journal of Geophysical Research: Biogeosciences*, *117*(G3). https://doi.org/10.1029/2012JG001960Schuur, E. A. [Northern A. U. (ORCID:0000000210962436). (2020, January 1). FLUXNET-CH4 US-EML Eight Mile Lake Permafrost thaw gradient, Healy Alaska. United States. https://doi.org/10.18140/FLX/1669674Schwalm, C. R., Williams, C. A., Schaefer, K., Anderson, R., Arain, M. A., Baker, I., et al. (2010). A model-data intercomparison of CO2 exchange across North America: Results from the North American Carbon Program site synthesis. *Journal of Geophysical Research: Biogeosciences*, *115*(G3). https://doi.org/10.1029/2009JG001229Shortt, R. [University of C., Berkeley] (ORCID:0000000156904656), Hemes, K. [University of C., Berkeley] (ORCID:0000000150901083), Szutu, D. [University of C., Berkeley] (ORCID:0000000176980461), Verfaillie, J. [University of C., Berkeley] (ORCID:0000000270098942), & Baldocchi, D. [University of C., Berkeley] (ORCID:0000000150901083). (2020, January 1). FLUXNET-CH4 US-Sne Sherman Island Restored Wetland. United States. https://doi.org/10.18140/FLX/1669693Sonnentag, O., & Helbig, M. (2020). FLUXNET-CH4 CA-SCB Scotty Creek Bog (2014-2017) [Data set]. https://doi.org/10.18140/FLX/1669613. Retrieved from https://doi.org/10.18140/FLX/1669613Stavert, A. R., Saunois, M., Canadell, J. G., Poulter, B., Jackson, R. B., Regnier, P., et al. (2021). Regional trends and drivers of the global methane budget. *Global Change Biology*, *n/a*(n/a). https://doi.org/10.1111/gcb.15901Stoy, P. C., Dietze, M. C., Richardson, A. D., Vargas, R., Barr, A. G., Anderson, R. S., et al. (2013). Evaluating the agreement between measurements and models of net ecosystem exchange at different times and timescales using wavelet coherence: an example using data from the North American Carbon Program Site-Level Interim Synthesis. *Biogeosciences*, *10*(11), 6893–6909. https://doi.org/10.5194/bg-10-6893-2013Stoy, Paul C., Katul, G. G., Siqueira, M. B. S., Juang, J.-Y., McCarthy, H. R., Kim, H.-S., et al. (2005). Variability in net ecosystem exchange from hourly to inter-annual time scales at adjacent pine and hardwood forests: a wavelet analysis. *Tree Physiology*, *25*(7), 887–902. https://doi.org/10.1093/treephys/25.7.887Tao, J., Zhu, Q., Riley, W. J., &

Neumann, R. B. (2021). Improved ELMv1-ECA simulations of zero-curtain periods and cold-season CH_4 and CO_2 emissions at Alaskan Arctic tundra sites. *The Cryosphere*, *15*(12), 5281–5307. https://doi.org/10.5194/tc-15-5281-2021Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, *106*(D7), 7183–7192. https://doi.org/10.1029/2000JD900719Torrence, C., & Compo, G. P. (1998). A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, *79*(1), 61–78. https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2Valach, A. [University of C., Berkeley] (ORCID:0000000347825766), Szutu, D. [University of C., Berkeley] (ORCID:0000000176980461), Eichelmann, E. [University of C., Berkeley] (ORCID:0000000195167951), Knox, S. [University of C., Berkeley], Verfaillie, J. [University of C., Berkeley] (ORCID:0000000270098942), & Baldocchi, D. [University of C., Berkeley] (ORCID:0000000334964919). (2020, January 1). FLUXNET-CH4 US-Tw1 Twitchell Wetland West Pond. United States. https://doi.org/10.18140/FLX/1669696Vargas, R., Detto, M., Baldocchi, D. D., & Allen, M. F. (2010). Multiscale analysis of temporal variability of soil CO2 production as influenced by weather and vegetation. *Global Change Biology*, *16*(5), 1589–1605. https://doi.org/10.1111/j.1365-2486.2009.02111.xWania, R., Ross, I., & Prentice, I. C. (2009). Integrating peatlands and permafrost into a dynamic global vegetation model: 2. Evaluation and sensitivity of vegetation and carbon cycle processes: PEATLANDS AND PERMAFROST IN LPJ, 2. *Global Biogeochemical Cycles*, *23*(3), GB3015. https://doi.org/10.1029/2008GB003413Wania, R., Ross, I., & Prentice, I. C. (2010). Implementation and evaluation of a new methane model within a dynamic global vegetation model: LPJ-WHyMe v1.3.1. *Geoscientific Model Development*, *3*(2), 565–584. https://doi.org/10.5194/gmd-3-565-2010Wania, R., Melton, J. R., Hodson, E. L., Poulter, B., Ringeval, B., Spahni, R., et al. (2013). Present state of global wetland extent and wetland methane modelling: methodology of a model inter-comparison project (WETCHIMP). *Geoscientific Model Development*, *6*(3), 617–641. https://doi.org/10.5194/gmd-6-617-2013Windham-Myers, L. [United S. G. S. (ORCID:0000000302819581), Stuart-Haëntjens, E. [United S. G. S. (ORCID:0000000199017643), Bergamaschi, B. [United S. G. S. (ORCID:0000000296105581), Knox, S. [University of B. C. (ORCID:0000000322555835), Anderson, F. [Land I. (ORCID:0000000214184678), & Nakatsuka, K. [United S. G. S. (2020, January 1). FLUXNET-CH4 US-Srr Suisun marsh - Rush Ranch. United States. https://doi.org/10.18140/FLX/1669694Wong, G., Melling, L., Tang, A., Aeries, E., Waili, J., Musin, K., et al. (2020). FLUXNET-CH4 MY-MLM Maludam National Park [Data set].Xu, Xiaofeng, Yuan, F., Hanson, P. J., Wullschleger, S. D., Thornton, P. E., Riley, W. J., et al. (2016). Reviews and syntheses: Four Decades of Modeling Methane Cycling in Terrestrial Ecosystems. *Biogeosciences Discussions*, 1–56. https://doi.org/10.5194/bg-2016-37Xu, Xiyan, Riley, W. J., Koven, C. D., Billesbach, D. P., Chang, R. Y.-W., Commane, R., et al. (2016). A multi-scale comparison of modeled and observed seasonal methane emissions in northern wetlands. *Biogeo-*

sciences, *13*(17), 5043–5056. https://doi.org/10.5194/bg-13-5043-2016Yu, X., Millet, D. B., Wells, K. C., Henze, D. K., Cao, H., Griffis, T. J., et al. (2021). Aircraft-based inversions quantify the importance of wetlands and livestock for Upper Midwest methane emissions. *Atmospheric Chemistry and Physics*, *21*(2), 951–971. https://doi.org/10.5194/acp-21-951-2021Zhang, Z., Fluet-Chouinard, E., Jensen, K., McDonald, K., Hugelius, G., Gumbricht, T., et al. (2021). Development of the global dataset of Wetland Area and Dynamics for Methane Modeling (WAD2M). *Earth System Science Data*, *13*(5), 2001–2023. https://doi.org/10.5194/essd-13-2001-2021Zhang, Zhen, Zimmermann, N. E., Kaplan, J. O., & Poulter, B. (2016). Modeling spatiotemporal dynamics of global wetlands: comprehensive evaluation of a new sub-grid TOPMODEL parameterization and uncertainties. *Biogeosciences*, *13*(5), 1387–1408. https://doi.org/10.5194/bg-13-1387-2016Zhang, Zhen, Zimmermann, N. E., Calle, L., Hurtt, G., Chatterjee, A., & Poulter, B. (2018). Enhanced response of global wetland methane emissions to the 2015–2016 El Niño-Southern Oscillation event. *Environmental Research Letters*, *13*(7), 074009. https://doi.org/10.1088/1748-9326/aac939Zhang, Zhen, Poulter, B., Knox, S., Stavert, A., McNicol, G., Fluet-Chouinard, E., et al. (2021). Anthropogenic emission is the main contributor to the rise of atmospheric methane during 1993–2017. *National Science Review*, nwab200. https://doi.org/10.1093/nsr/nwab200Zhu, Q., Peng, C., Chen, H., Fang, X., Liu, J., Jiang, H., et al. (2015). Estimating global natural wetland methane emissions using process modelling: spatio-temporal patterns and contributions to atmospheric methane fluctuations: Global natural wetland methane emissions. *Global Ecology and Biogeography*, *24*(8), 959–972. https://doi.org/10.1111/geb.12307Zhuang, Q., Melillo, J. M., Kicklighter, D. W., Prinn, R. G., McGuire, A. D., Steudler, P. A., et al. (2004). 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*(3). https://doi.org/10.1029/2004GB002239Zhuang, Qianlai, Chen, M., Xu, K., Tang, J., Saikawa, E., Lu, Y., et al. (2013). Response of global soil consumption of atmospheric methane to changes in atmospheric climate and nitrogen deposition: GLOBAL SOIL CONSUMPTION OF METHANE. *Global Biogeochemical Cycles*, *27*(3), 650–663. https://doi.org/10.1002/gbc.20057Zona, D., Gioli, B., Commane, R., Lindaas, J., Wofsy, S. C., Miller, C. E., et al. (2016). Cold season emissions dominate the Arctic tundra methane budget. *Proceedings of the National Academy of Sciences*, *113*(1), 40–45.https://doi.org/10.1073/pnas.1516017113Zona, D. [San D. S. U., & Oechel, W. C. [San D. S. U. (2020a, January 1). FLUXNET-CH4 US-Atq Atqasuk. United States. https://doi.org/10.18140/FLX/1669663Zona, D. [San D. S. U., & Oechel, W. C. [San D. S. U. (2020b, January 1). FLUXNET-CH4 US-Ivo Ivotuk. United States. https://doi.org/10.18140/FLX/1669679

Tables

Table 1. Summary of site characteristics.

Site ID	Country	Latitude	Longtiude	Biome type	Wetland type	Start year	End year	Data reference
BR-NPW	Brazil			Tropical/Subtropical	Seasonal			(Dalmagro et al., 2019)
CA-SCB	Canada			Boreal forest	Bog			(Sonnentag & Helbig, 2020)
FI-LOM	Finland			Boreal forest	Fen			(Lohila et al., 2020)
MY-MLM	Malaysia			Tropical/Subtropical	Swamp			(Wong et al., 2020)
RU-VRK	Russia			Arctic tundra	Wet Tundra			(Friborg & Shurpali, 2020)
SE-DEG	Sweden			Boreal forest	Fen			(Nilsson & Peichl, 2020)
SE-ST1	Sweden			Arctic tundra	Fen			(Jansen et al., 2020)
SE-STO	Sweden			Arctic tundra	Bog			(Jansen et al., 2020)
US-ATQ	USA			Arctic tundra	Wet Tundra			(Zona & Oechel, 2020)
US-BZB	USA			Boreal forest	Bog			(Euskirchen & Edgar, 2020)
US-BZF	USA			Boreal forest	Fen			(Euskirchen, 2022a)
US-BZS	USA			Boreal forest	Swamp			(Euskirchen, 2022b)

US-EML	USA	Arctic tundra	Bog	(Schuur, 2020)
US-ICS	USA	Arctic tundra	Wet Tundra	(Euskirchen et al., 2020)
US-IVO	USA	Arctic tundra	Wet Tundra	(Zona & Oechel, 2020b)
US-LOS	USA	Temperate	Fen	(Desai & Thom, 2020)
US-ORV	USA	Temperate	Marsh	(Bohrer & Morin, 2020)
US-OWC	USA	Temperate	Marsh	(Bohrer & Morin, 2020)
US-SNE	USA	Temperate	Marsh	(Shortt et al., 2020)
US-SRR	USA	Temperate	Salt Marsh	(Windham-Myers et al., 2020)
US-TW1	USA	Temperate	Marsh	(Valach et al., 2020)
US-UAF	USA	Boreal forest	Bog	(Iwata et al., 2020)
US-WPT	USA	Temperate	Marsh	(Chen & Chu, 2020)

Table 2. Summary of Model Characteristics

Model	Wetland PFT
CLASSIC	No wetland-specific PFTs

ELM	No wetland-specific PFTs
JSBACH-HIMMELI	Generic wetland PFT with C3 grass parameters for vegetation
LPJ-wsl	No wetland-specific PFTs
LPJ-GUESS	High-latitude (> 40°N): Wetland grass, cushion forbs, lichens, sphagnum moss. South
TEM-MDM	Five primary types of wetlands are considered in boreal, temperate and tropical region
TRIPLEX-GHG	a general wetland PFT was added without considering specific wetland plants type

Table 3. The p values of ANOVA analysis for the impact of model structure on the spectral power for different wetland types within each of the four-time scales. ns: non significant.

Wetland Type	Time scale	Wetland PFT	Component of CH_4 flux	CH_4 production proxy	Incor
Boreal forest	Multiday	<0.001	<0.001	<0.001	<0.0
	Weekly	<0.001	<0.001	<0.001	<0.0
	Monthly	<0.001	<0.001	<0.001	<0.0
	Seasonal	<0.001	<0.001	ns	ns
Arctic tundra	Multiday	<0.001	<0.001	<0.001	<0.0
	Weekly	<0.001	<0.001	<0.001	<0.0
	Monthly	<0.001	<0.001	<0.001	<0.0
	Seasonal	<0.001	<0.001	<0.001	ns
Temperate	Multiday	<0.001	<0.001	ns	<0.0
	Weekly	<0.001	<0.001	ns	ns
	Monthly	<0.001	0.002	ns	0.018
	Seasonal	ns	ns	0.001	0.007
Tropical/subtropical	Multiday	<0.001	<0.001	<0.001	ns
	Weekly	<0.001	ns	ns	ns
	Monthly	<0.001	ns	ns	ns
	Seasonal	<0.001	ns	ns	ns