

Methane emissions from subtropical wetlands: An evaluation of the role of data filtering on annual methane budgets

C.L. Staudhammer^{a,*}, S.L. Malone^b, J. Zhao^c, Z. Yu^d, G. Starr^e, S.F. Oberbauer^f

^a Department of Biological Sciences, University of Alabama, Tuscaloosa, AL 35487 USA

^b Department of Biological Sciences, Florida International University and Institute of Environment, Miami, FL 33199 USA

^c Department of Biogeochemistry and Soil Quality, Division of Environment and Natural Resources, Norwegian Institute of Bioeconomy Research, Ås, Norway

^d Department of Biological Sciences, University of Alabama, Tuscaloosa, AL 35487 USA

^e Department of Biological Sciences, University of Alabama, Tuscaloosa, AL 35487 USA

^f Department of Biological Sciences, Florida International University and Institute of Environment, Miami, FL 33199 USA

ARTICLE INFO

Keywords:

Eddy Covariance
Random Forests
Freshwater marshes
gap-filling
U*
RSSI

ABSTRACT

Widespread adoption of eddy covariance (EC) methods for methane (CH₄) flux measurement has led to increased availability of continuous high-frequency CH₄ data. However, unreliable data frequently occur during periods of atmospheric stability, rain or instrument malfunction, requiring filtering prior to subsequent analyses. While procedures for assessing CO₂ have matured, processes to filter and gap-fill CH₄ data are less studied, as their range and controls are not as well-understood. Moreover, publications often fail to describe procedures for data processing and filtering. Our primary objective was to study effects of common filtering thresholds and provide insight on how size and timing of gaps produced by filtering affect CH₄ budgets. We utilized 4 years of data from two freshwater wetlands under the same climate regime but different hydroperiods. We applied friction velocity (U*) and signal strength filtering treatments to isolate site-specific effects and evaluate impacts of filtering on subsequent gap-filling via Random Forests (RF). We also tested sensitivity of results to predictor datasets with an “unrestricted predictors model” (using all possible predictors regardless of gaps), versus a “restricted predictors model” (using gap-filled predictors with no missing values). Depending on filtering treatment, 7 - 50% of CH₄ data were removed over the study period. Using higher signal strength thresholds introduced more small gaps. U* filtering created small gaps (mostly nighttime), and corresponding annual budget estimates were generally different from those filtered solely on signal strength but with higher uncertainty, especially at the long-hydroperiod site. Regardless of filtering method, RF models using unrestricted predictors identified 2- to 32-day average CH₄ flux as primary predictors, whereas heat and latent energy were most important when predictors were restricted. Although filtering may have less impact on CH₄ budgets than selection and pre-processing of predictor variables, it can significantly impact uncertainty and should be considered in data curation protocols.

1. Introduction

Methane (CH₄) is the third most abundant greenhouse gas in the atmosphere and it is ~34 times more effective at warming the atmosphere than carbon dioxide (CO₂) (IPCC, 2014). During 2020, CH₄ surpassed 1875 ppb (https://www.esrl.noaa.gov/gmd/ccgg/trends_ch4/), more than 2.5 times its pre-industrial concentration (Le Mer and Roger 2001; Kirschke et al. 2013). About 40% of CH₄ emissions come from natural sources (Carmichael et al. 2014, Sauniois et al. 2016). Yet, there is a great amount of uncertainty associated with estimates of

interannual variation in methane emission which needs to be further investigated (Kirschke et al. 2013; Turner et al. 2019). Within these estimates, the greatest uncertainties are thought to be associated with wetlands and inland waterways (Carmichael et al. 2014; Sauniois et al. 2020). Given that tropical and subtropical regions are the greatest biogenic contributors to rising global atmospheric CH₄ concentration, the greatest uncertainties in predicted future CH₄ budgets are therefore in the freshwater systems in these regions (Sauniois et al. 2020).

Methane in wetlands is generally produced under anaerobic conditions in saturated soils (Le Mer and Roger 2001) and apparently within

* corresponding author

E-mail address: cstaudhammer@ua.edu (C.L. Staudhammer).

<https://doi.org/10.1016/j.agrformet.2022.108972>

Received 21 October 2021; Received in revised form 17 April 2022; Accepted 19 April 2022

Available online 26 April 2022

0168-1923/© 2022 Elsevier B.V. All rights reserved.

trees (Barba et al. 2019). Methane reaches the atmosphere via transport through vascular plant aerenchyma and other vascular tissues, diffusion, and as ebullition. When the water table is below the soil surface, substantial CH₄ consumption occurs in the aerated soil zone above the anaerobic layers (Le Mer and Roger 2001; Sabrekov et al. 2016). Several factors impede understanding of the components of biogenic CH₄ cycling from local to global scales. These include, but are not limited to, variation in local micrometeorological (e.g., temperature, radiation, wind, humidity, etc.) and hydrological conditions that alter the source/sink capacity of wetlands (Conrad 2009), the lack of an observation network with continuous CH₄ measurements (e.g., eddy covariance) across regions (Delwiche et al. 2021), heterogeneity in land cover and land use (Chu et al. 2021), and human-induced changes in CH₄ budgets (i.e., climate change, land management and land-use land cover change) (Wong et al. 2020). Another increasingly important concern is the lack of standardized measurement techniques, as well as data filtering and processing criteria. The recent availability of open and closed path analyzers, which are capable of continuous high frequency methane measurements, has led to the adoption of CH₄ eddy covariance (EC) measurements and the production of ecosystem-scale CH₄ flux datasets. However, these continuous CH₄ flux datasets typically have gaps when atmospheric conditions do not meet the assumptions of EC, the wind is from an undesirable direction, sensors are fouled, rain interferes with gas analyzers and/or sonic anemometers, power is lost, etc. (Baldocchi 2003; Kim et al. 2020; Irvin et al. 2021). As a result, unsuitable data need to be identified, removed (filtered) and gap-filled. For CO₂ fluxes, the use of known processes and models to filter and fill data gaps, e.g., on the basis of relationships between fluxes and concurrently measured ancillary meteorological variables, such as light, temperature, pressure, etc. (Falge et al. 2001; Kunwor et al. 2017), is well-studied. However, for CH₄ fluxes, their range and controls are to a large extent not currently well understood, and thus robust methods for filtering and gap-filling

data are generally lacking (Dengel et al. 2013; Nemitz et al. 2018). The situation is further exacerbated by lack of clarity in publications, which often fail to describe what methods or criteria have been used for data processing, filtering and gap-filling, and how they have been applied across sites (see Table 1). In addition, many of these studies lack full-year temporal coverage (Table 1), which further hinders our ability to develop annual methane budgets with confidence. Improvement in gap filling approaches was the subject of two recent studies which compared MDS and several machine learning methods (Irvin et al. 2021; Kim et al. 2020); however, these studies were focused exclusively on the performance of gap-filling methods and relied solely on friction velocity (U*) filtering derived from relationships between U* and CO₂ fluxes.

The importance of filtering approaches that are applied to CH₄ EC data for the final estimates of fluxes seems to be greatly under-valued. The decision to select site-specific filtering criteria should be justified and presented transparently, as variation in filtering criteria and site conditions can result in a total of 20 to 83% data loss (Table 1). The removal of large amounts of data also imposes additional uncertainties in annual methane budgets that will further propagate biases in efforts to upscale and predict global methane budgets, though this uncertainty is rarely acknowledged.

In this study, we evaluate the importance of data-processing of continuous CH₄ measurements (net ecosystem exchange of CH₄; F_{CH₄}) derived via the EC method in two subtropical wetlands, with the primary objective to study the effects that data filtering has on gap-filled CH₄ in these systems. Our second objective is to provide insight on how gaps in predictor variables for gap-filling models—both their size and timing—may affect the procedures used and subsequent annual budgets derived from CH₄ data. We present a case study of two different wetlands under the same climate regime but different hydroperiods to show how filtering and gap-filling via Random Forest, using predictor datasets with an “unrestricted predictors model” (using all possible predictors

Table 1

Filtering methods utilized in studies using open-path LI-7700 infrared gas analyzers (LI-COR Biosciences, Lincoln, NE).

Ecosystem	Location	Study length	RSSI ⁺	Quality flags [†]	U* (m s ⁻¹) threshold	Weather	Feasibility threshold	Post-filter % missing [‡]	Reference
Rice field	Philippines	178 d	< 10%	0-2				20-25%	Alberto et al. (2014)
Rice field	China	7 mo	< 10%	0-2	0.12			36%	Dai et al. (2019)
Rice field	China	6 mo	< 10%	0-2					Ge et al. (2018)
Rice field	China	4 y	< 20%	1-9	0.12	Rainfall		28% -33%	Li et al. (2019)
Bog	Canada	96 d	< 20%	1-9	0.2	Wind direction	>3 SD and < -50 mg CH ₄ m ⁻² d ⁻¹		Nadeau et al. (2013)††
Bog	Japan	4 y	< 15%	/	0.15 (night)	Rainfall		15 - 21%	Ueyama et al. (2020)
Peatland	China	2 growing seasons	< 20%					39%, 21%	Yu et al. (2017)
Peatland	Finland	3 y			0.1			65%	Korrensalo et al. (2018)
Peatland	USA	5 mo		0-2	0.08		Negative values		Pypker et al. (2013)
Peatland	Canada	46 mo	< 20%		0.17		Outliers	58%	Helbig et al. (2017)
Wetland	USA	45 mo			0.2 summer, 0.15 winter				Morin et al. (2017)
Tidal salt marsh	China	22 mo	< 20%	1-9	0.15	Rainfall		66%	Li et al. (2018)
Pasture, corn, rice, restored wetlands	USA	1 y			0.15- 0.2				Knox et al. (2015); Knox et al. (2016)
Bog Forest	Germany	14 mo	< 20%		0.124			64%	Hommelntenberg et al. (2014)
Meadow	China	31 mo			0.1 (night)	Rainfall, wind dir.	>-0.1 and >0.005 μmol m ⁻² s ⁻¹	65%	Chen et al. (2020)
Forest	USA	3 growing seasons			0.1				Iwata et al. (2015)

+ RSSI Relative Signal Strength Indicator, † CH₄ quality flags as defined in EddyPro flux processing software (version 6.2.1, LI-COR Inc., Lincoln, NE, USA), ‡ The percentage of total data missing after accounting for both filtering and technical failures, †† Atmospheric stability parameter $\zeta = z/L$, where z is the measurement height and L the Obukhov length.

regardless of gaps), versus a “restricted predictors model” (using gap-filled predictors or predictors with no missing values), influence CH₄ budgets. With different hydrological conditions, the response of CH₄ fluxes to environmental factors may vary among different wetlands. For example, Yu et al (in review) found that seasonal water dynamics had a much stronger impact on CH₄ fluxes in a long hydroperiod wetland versus one with a short hydroperiod. Therefore, models for gap-filling CH₄ flux data will likely need to be site-specific. By applying the same filtering methods to sites with similar climate and contrasting hydrology, we can better isolate site-specific effects to evaluate the impact of the choice of filtering and gap-filling methods on derived CH₄ budgets to focus on the primary question: Are Everglades freshwater marshes a sink or source for CH₄? In doing so we intend to answer: 1) How does EC filtering criteria influence the size and distribution of gaps? And 2) Do filtering procedures influence annual budgets and prediction uncertainty in these systems? Our case study will provide a framework to support robust, transparent filtering criteria for subsequent research utilizing EC-derived CH₄ data.

2. Methods

2.1. Study area

This study was conducted using data that was collected from January 2016 through December of 2019 at two freshwater marsh sites with contrasting hydrology within Everglades National Park: Taylor Slough (TS/Ph-1; 25° 26'16.5"N, 80° 35'40.68"W) and Shark River Slough (SRS-2; 25° 33'6.72"N, 80° 46'57.36"W). Both are oligotrophic marshes and are part of the Florida Coastal Everglades Long-Term Ecological Research program (FCE-LTER). The TS/Ph-1 site is a short-hydroperiod marsh that is flooded for 4–6 months each year, typically from June to November. The site is characterized by shallow marl soils (mostly fine particles of calcium carbonate as calcite, ~0.14 m deep) overlying limestone bedrock. The site is co-dominated by the C₃ macrophyte *Cladium jamaicense* Crantz (Armentano et al., 2006), commonly known as sawgrass, with the C₄ grass *Muhlenbergia capillaris* Lam (Muhly grass). Abundant periphyton mats are present during the inundation period (Gaiser et al., 2006; Gottlieb et al., 2006; Iwaniec et al., 2006). Mean canopy height (Z) and surface roughness (z₀) for this site are 0.73 and ~0.3 m, respectively. The long-hydroperiod SRS-2 site is inundated year-round in most years. Soils are deep peat (>50 cm) above limestone rock (Duever et al., 1976). The site is characterized by two different vegetation types, ridge and slough. Ridge areas are slightly elevated and dominated largely by monospecific, dense stands of robust sawgrass, and sloughs are comparatively much less vegetated areas with open water, water lilies (*Nymphaeaceae*), *Sagittaria*, and a mix of several species of spikerush (*Eliocharis*), all of which have aerenchyma tissue. For this site, Z and z₀ are 1.02 and ~0.4 m, respectively. For both sites, no notable seasonality in the canopy heights are present.

The Everglades are located in a subtropical region; however, the climate is classified as tropical with distinct annual wet and dry seasons during the summer and winter, respectively (Duever et al., 1994; Beck et al., 2006). Mean annual precipitation is 1430 mm (NCDC, 2019) with ~60% of rainfall occurring at the height of the rainy season (June to September) and ~25% in the dry season (November to April) with variability among transitional months (Duever et al., 1994). Mean annual temperature of the Everglades is 23.9°C, with the mean monthly minimum (18.1°C) in January and monthly maximum (29.4°C) in August (NCDC 2019).

2.2. Eddy Covariance

At each site, open-path infrared gas analyzers (IRGA, LI-7500 and LI-7700, LI-COR Inc., Lincoln, NE) were used to measure CO₂ (mg mol⁻¹), water vapor molar density (mg mol⁻¹), and CH₄ (μmol mol⁻¹). The IRGAs were paired with a sonic anemometer (CSAT3, Campbell

Scientific Inc., Logan, UT) to measure sonic temperature (T_{sonic}; K) and 3-dimensional wind speed (u, v and w, respectively; m s⁻¹). The sensors were installed 0.09 m apart at 3.30 and 3.24 m above ground level (a.g.l.) at TS/Ph-1 and SRS-2, respectively. Data were logged at 10 Hz, collected by the LI-COR 7550 analyzer interface unit (LI-COR Inc., Lincoln, NE) and stored on an industrial 16 GB USB drive. The LI-7500's were calibrated monthly using a trace gas standard for CO₂ in air (±1.0%), dry N₂ gas and a portable dew point generator (LI-610, LI-COR Inc.). Footprint analyses (Kljun, Rotach and Schmid, 2002; Kljun et al., 2004) indicated that 80% of measured fluxes were within 100 m of the tower during convective conditions at both sites. Other meteorological variables were measured at 1-sec and collected as half-hourly averages by a data logger (CR1000, Campbell Scientific Inc.), and included: air temperature, (T_{air}; °C) and relative humidity (RH; %) (HMP45C, Vaisala, Helsinki, Finland) mounted within an aspirated shield (43502, R.M. Young Co., Traverse City, MI), and barometric pressure (P; atm) (PTB110, Vaisala Corp., Helsinki, Finland). The T_{air}/RH sensors were installed at the same height a.g.l. as the IRGA and CSAT.

At each site, additional meteorological data were measured at 15-sec and collected as 30-min averages through a multiplexer (AM16/32A Campbell Scientific Inc.) with another data logger (CR10X Campbell Scientific Inc.). This included photosynthetically active radiation (PAR; μmol m⁻² s⁻¹) (PAR Lite, Kipp and Zonen Inc., Delft, Netherlands), incident solar radiation (R_s; W m⁻²) (LI-200SZ, LI-COR Inc.), and net radiation (R_n; W m⁻²) (CNR2-L, Kipp and Zonen). Precipitation measurements were made with tipping bucket rain gauges (mm) (TE525, Texas Electronics Inc., Dallas, TX). Soil volumetric water content (VWC; %) was calculated from equations developed for peat and marl soils using the methodology of Veldkamp & O'Brien (2000), from the dielectric constant using two soil moisture sensors (CS616, Campbell Scientific Inc.) installed at a 45° angle at the soil surface at each site. Soil temperature (T_s; °C) was measured at 5 cm, 10 cm, and 20 cm depths at two locations within each site using insulated thermocouples (Type-T, Omega Engineering Inc., Stamford, CT). When inundated at SRS-2, water temperature (T_w; °C) was measured using two pairs of insulated thermocouples (Type-T, Omega Engineering Inc.), each pair located at a fixed height 5 cm above the soil surface and another attached to shielded floats that held the thermocouples 5 cm below the water surface. At TS/Ph-1, T_w was measured using insulated thermocouples (Type-T, Omega Engineering Inc.) located at a fixed height 2 cm below the water surface. Water level (WL; m) at both sites was recorded every half-hour with a water level logger (HOBO U20-001-01, Onset Computer Corporation, Bourne, MA). Data gaps occurred during the study due to equipment failures and weather anomalies (e.g., Hurricane Irma). Gaps in WL data were filled using relationships derived from the three nearest Everglades Depth Estimation Network stations (<https://sofia.usgs.gov/eden>) and the site's water level station. PAR was gap-filled with data from nearby flux towers in the FCE-LTER for subsequent analyses.

Raw flux data were processed with EddyPro software (Advanced mode, version 6.2.1, LI-COR Inc., Lincoln, NE, USA) over 30 min time intervals. The raw CH₄ turbulence data were processed with the LI-7700 diagnostics turned off. Flux measurements for F_{CO2}, sensible heat flux (H), and latent heat flux (LE) were filtered when systematic errors were indicated, such as: (1) evidence of rainfall, condensation, or bird fouling in the sampling path of the IRGA or sonic anemometer, (2) incomplete half-hour datasets during system calibration or maintenance, (3) excessive variation from the half-hourly mean based on an analysis of standard deviation of the u wind component (std. dev. of u > 1.6), and (4) poor coupling of the canopy with the external atmospheric conditions, as defined by the friction velocity, U* (Goulden et al. 1996; Clark et al. 1999), using a threshold of 0.15 m s⁻¹. The U* threshold was determined using the CO₂ flux by site and year following Papale et al. (2006) via the EddyProc package in R (Wutzler et al. 2018). While estimated thresholds varied somewhat by year and site, the majority of 95% confidence intervals were in the range of 0.1–0.15; thus, our threshold reflects a somewhat conservative approach. Quality assurance

of the F_{CO_2} data was also maintained by examining plausibility tests (i.e., $|F_{CO_2}| < 30 \mu\text{mol m}^{-2} \text{s}^{-1}$), stationarity criteria, and integral turbulent statistics (Foken and Wichura 1996; Foken and Leclerc 2004). At TS/Ph-1, 38% and 77% of the day and nighttime data were removed, respectively. At SRS-2, 34% of daytime data and 70% of nighttime data were removed. Missing half hourly F_{CO_2} data were gap filled using separate functions for day and night. When PAR was $>10 \mu\text{mol m}^{-2} \text{s}^{-1}$, F_{CO_2} data were gap-filled using a Michaelis-Menten approach (Michaelis & Menten 1913) ($F_{CO_2 \text{ day}}$; Eq. 1), and when PAR was $\leq 10 \mu\text{mol m}^{-2} \text{s}^{-1}$, F_{CO_2} data were gap-filled using an Arrhenius approach (Arrhenius 1915) ($F_{CO_2 \text{ night}}$; Eq. 2):

$$F_{CO_2 \text{ day}} = R_{eco} - \frac{\alpha \Phi P_{max}}{\alpha \Phi + P_{max}} \quad (\text{Eq. 1})$$

where: α is the apparent quantum efficiency, Φ is PAR, R_{eco} is ecosystem respiration ($\mu\text{mol CO}_2 \text{ m}^{-2} \text{s}^{-1}$), and P_{max} is the maximum ecosystem CO_2 uptake rate ($\mu\text{mol CO}_2 \text{ m}^{-2} \text{s}^{-1}$).

$$F_{CO_2 \text{ night}} = R_{eco} = R_0 \exp^b \quad (\text{Eq. 2})$$

where: R_0 is the base respiration rate when T_{air} is 0°C and b is an empirical coefficient. In equation 1, R_{eco} is an estimated model parameter, whereas R_{eco} measurements are the dependent variable in Eq. 2. Following gap filling, gross ecosystem exchange (GEE) was calculated from half hourly F_{CO_2} and R_{eco} data (Eq. 3).

$$GEE = F_{CO_2} - R_{eco} \quad (\text{Eq. 3})$$

Missing H and LE values were gap-filled using the linear relationship between H or LE and net radiation (R_n) on a monthly, seasonal, or annual basis following Cleverly et al. (2002). When R^2 values were less than 70%, annual relationships between R_n and H or LE were used to gap fill data in that month. Less than 5% of filtered data were filled with annual equations.

2.3. CH_4 Filtering treatments

Building on the work of previous CH_4 studies, we evaluated the amount of data lost and the quality (measures of variance, outliers, feasible ranges) of gap-filling for six filtering treatments (Feasibility, RSSI 10%, RSSI 20%, U^* , RSSI 10% + U^* , and RSSI 20% + U^* ; Table 1). First, we considered that values of F_{CH_4} between -2 and $+2 \mu\text{mol m}^{-2} \text{s}^{-1}$ could be attainable in our sites by evaluating measurements under ideal conditions. Moreover, this range encompassed $\sim 95\%$ of our F_{CH_4} data. Thus our “Feasibility” filtering treatment was to limit $|F_{CH_4}| < 2 \mu\text{mol m}^{-2} \text{s}^{-1}$. While this approach may remove ebullition signals that may be accurate but sporadic, it captures the vast majority of valid data. All other filtering treatments were applied following feasibility filtering. A common parameter used for filtering is the Relative Signal Strength Indicator (RSSI). A threshold of 10% is often used, as this value is often reached when mirrors are fouled or objects are present in the IRGA sample path (McDermitt et al 2011). We also tested a threshold of 20% to encompass values commonly used in wetland CH_4 research (Table 1). Thus, our next two filtering treatments removed F_{CH_4} when RSSI was less than 10% or 20%, respectively. U^* thresholds have also been used in CH_4 flux data processing and indicate whether air samples brought to the sensor are well mixed and representative of ecosystem function at the measurement time. The U^* generally attains its lowest values at night. Thresholds are typically set based on canopy height and roughness, and based on our CO_2 studies, we set the U^* threshold at 0.15 for our fourth filtering treatment. The final two treatments combined this U^* threshold (0.15) with each of the RSSI thresholds (10%, 20%). While some previous research has used quality flags as filtering criteria (Table 1), we did not find different patterns of F_{CH_4} under different quality flags. Moreover, the use of these flags eliminated F_{CH_4} values < -0.2 , which was not appropriate based on previous studies in our sites. Therefore, we did not consider CH_4 quality flags as a filtering method for

our sites.

2.4. Development of CH_4 Gap Filling Model

We employed the machine learning algorithm Random Forests (RF) to gap-fill filtered F_{CH_4} for each filtering treatment. The RF algorithm was first introduced by Breiman (2001), and uses an ensemble of regression trees to predict target values. The prediction accuracy of RF has been found to be superior to other algorithms in flux upscaling studies (Bodesheim et al 2018; Tramontana et al 2015; Xu et al., 2018). While few studies have utilized RF for gap-filling, Kim et al (2020) found that RF out-performed other techniques across five sites, and Irvin et al (2021) reported that median performance was better for RF across 17 FLUXNET wetland sites. In RF, a series of bootstrapped datasets are used to generate independent regression trees; at each node, a random sample of predictor variables is selected for use. The RF prediction is the ensemble of multiple individual trees. We utilized the R package randomForest (Liaw and Wiener 2002) to create 500 trees for each year and site, using 80% of the data for model fitting and 20% for model validation. Variables included as predictors included other fluxes and the suite of micrometeorological variables, but excluded variables which were highly correlated with others (e.g., net radiation): PAR, RH, U^* , air pressure, vapor pressure deficit (VPD), net longwave radiation, T_{air} , water level, LE, H, and F_{CO_2} . Based on previous studies of F_{CO_2} , we also included the half-hourly change in water level, as well as several lagged values (12-hour, 2-day, and 8-day differences) (Malone et al., 2014). Since preliminary analyses indicated strong synoptic patterns in F_{CH_4} , we also computed 2-day, 4-day, 8-day, 16-day, and 32-day moving averages as predictors of half-hourly F_{CH_4} in RF models.

In a preliminary step, we first estimated RFs by site and year with a series of subsets of the 20 candidate independent variables using a modified backwards elimination procedure. The fit of each RF was evaluated with the out-of-bag mean square error (OOB MSE), and variable importance was computed as the amount the prediction error increased when a particular predictor was permuted. Initially, 500 RF trees per year and site were generated using all possible predictor variables. Overall model fit was evaluated with the average of the 500 OOB MSEs from the final model for each year and site, and variable importance was calculated as the average rank of each predictor variable for the 500 models. Then, the variable with the least importance was dropped as a predictor. This process was used to reduce the number of variables considered from 20 to 11 (Table 2). In our sites, this eliminated some variables which have been used for gap-filling in other sites (e.g., T_{air} , RH, air pressure, and radiation fluxes) (Irvin et al. 2021), emphasizing the need for site-specific variable selection.

Table 2

Predictor variables used in random forest (RF) models of F_{CH_4} . R=restricted model, U=unrestricted model.

Variables	Units	Description	Model
F_{CO_2}	$\mu\text{mol m}^{-2} \text{s}^{-1}$	Net ecosystem exchange of CO_2	R, U
GEE.-1d	$\mu\text{mol m}^{-2} \text{d}^{-1}$	Previous day's gross ecosystem exchange rates	R, U
F_{CH_4} -ma2d	$\mu\text{mol m}^{-2} \text{s}^{-1}$	2-day moving average of half hourly F_{CH_4} *	U
F_{CH_4} -ma8d	$\mu\text{mol m}^{-2} \text{s}^{-1}$	8-day moving average of half hourly F_{CH_4} *	U
F_{CH_4} -ma32d	$\mu\text{mol m}^{-2} \text{s}^{-1}$	32-day moving average of half hourly F_{CH_4} *	U
WL	m	Water Level	R, U
WL.12h	m	The difference in water level over 12 hours	R, U
WL.2d	m	The difference in water level over 2 days	R, U
H	W m^{-2}	Sensible heat flux	R, U
LE	W m^{-2}	Latent energy	R, U
PAR	$\mu\text{mol m}^{-2} \text{s}^{-1}$	Photosynthetically active radiation	R, U

* Calculated when $>10\%$ of observations were non-missing.

We further explored the amount of redundancy and relevance of the remaining 11 variables used to gap-fill CH_4 using the *varrank* (Kratzer and Furrer 2018) package in R. This approach provides a general implementation of the minimum redundancy maximum relevance (mRMR) model, that is based on information theory metrics. It is compatible with discrete and continuous data which are discretized using a large choice of possible rules and then a *varrank* analysis sequentially compares relevance with redundancy. The columns of the triangular matrix contain the scores at each selection step. The variable with the highest score is considered the most important and the variables are listed in order of their importance. A negative score indicates redundancy of information and a positive score is indicative of relevance. We used the R function VSURF (Genuer et al. 2010) to perform a two-stage variable selection procedure based on preliminary rankings of the explanatory variables and a stepwise forward strategy following Genuer and Poggi (2010). The procedure returns two subsets of important variables; the first is used for interpretation, but includes some redundancy, and the second eliminates redundant variables for parsimonious prediction. While gap-filling models of F_{CH_4} were estimated using only non-missing predictors, our objective was to fill F_{CH_4} data, even when predictor variables were missing. Thus, F_{CH_4} gap-filling was carried out in two steps that fit two types of models. The first type of model, hereafter referred to as the “unrestricted predictors model”, used all variables regardless of the gaps in the suite of predictors. The second model, hereafter referred to as the “restricted predictors model”, used variables that can be gap-filled with data from nearby sensors or that contain no missing values (Table 2). The latter model thus did not include variables describing synoptic patterns of F_{CH_4} , but did include filled values of (for example) PAR and WL. This latter approach allows for a two-step gap-filling process, whereby gaps in F_{CH_4} data are first filled with the strongest predictors regardless of the capacity to fill all gaps (unrestricted model), and then remaining gaps (i.e., those where one or more predictors are missing) are filled with models that include variables that were gap-filled (restricted model).

2.5. CH_4 Gap Filling Procedure

To fill gaps in F_{CH_4} , we used the *randomForestInfJack* function in the package *randomForestCI* (Wager et al. 2014). We also calculated model uncertainty for random forest predictions. Two sources of variability contribute to the variance of a RF prediction: sampling variability and Monte Carlo noise. We estimated the prediction variance using an infinitesimal jackknife estimate (Wager et al. 2014) based on bootstrapped trees using the R package *randomForestCI*. To estimate annual CH_4 budgets, we first filled gaps in F_{CH_4} with the unrestricted predictors model and any remaining gaps were filled with the restricted predictors model. To estimate annual budget prediction uncertainty, the standard deviation of each prediction was summed for missing measurements under each filtering treatment, and prediction uncertainties for half hours with a measured value were assumed 0.

3. Results

3.1. Gap distributions

The inclusion of U^* threshold filtering treatments resulted in a larger decrease in observations than that of RSSI. For example, the inclusion of a 0.15 U^* filter decreased the number of available F_{CH_4} observations by 33% and 26%, respectively at TS/Ph-1 and SRS-2, versus only using a feasibility filter (Table 3). While the inclusion of a U^* filter did not have an effect on larger gaps, its impact was to increase the number of gaps of less than one day (<48 half-hourly F_{CH_4} observations) (Table 4).

More rigorous filtering (U^* , RSSI 10% + U^* , and RSSI 20% + U^*) resulted in greater data loss (Table 3), although its effect on the number of gaps created was not distributed evenly across all gap sizes. Stricter filters created larger gaps, but not necessarily more gaps of all size

Table 3

Number of non-missing F_{CH_4} observations under each filtering treatment.

Treatment Feasibility	RSSI	U^*	TS/Ph-1	SRS-2
			58,753	64,237
(-2, 2)			53,991	59,573
(-2, 2)	$\geq 10\%$		49,221	50,634
(-2, 2)	$\geq 20\%$		43,865	41,683
(-2, 2)		≥ 0.15	34,889	42,791
(-2, 2)	$\geq 10\%$	≥ 0.15	33,152	37,774
(-2, 2)	$\geq 20\%$	≥ 0.15	30,742	32,074

classes (Table 4). Gaps that were larger than 7 days were rare and were relatively unaffected by different filters. The very large gaps (i.e., > 1 month) at both sites were a result of Hurricane Irma, which made landfall on 9 September 2017.

3.2. Redundancy between variables

When using only the feasibility, U^* , or RSSI filtering methods in TS/Ph-1, analyses showed that there was strong redundancy between F_{CO_2} and WL.2d (the two-day change in water level); however, this redundancy was greatly reduced when a combination of U^* and RSSI was used to filter half-hourly F_{CH_4} data (Supplemental Information Fig. S1). With RSSI filtering, redundancy was stronger between F_{CO_2} and all variables except PAR and the 2-day moving average of F_{CH_4} ($\text{F}_{\text{CH}_4}.\text{ma}2\text{d}$), and became stronger with greater RSSI limits; however, this redundancy was reduced when RSSI filtering was combined with the U^* filter. When considering feasibility only, redundancy was strong between LE and WL.2d. When combining RSSI at 10% with U^* , the redundancy between PAR and WL.12h was strong, but this dissipated when RSSI was increased to 20%. Under the most stringent filtering method, including both RSSI and U^* , no pairs of variables showed redundancy.

Regardless of filtering method, analyses showed that there was strong redundancy between F_{CO_2} and WL.12h at SRS-2; unlike in TS/Ph-1, this redundancy was not reduced when filtering included both U^* and RSSI (Supplemental Information Fig. S2). Like TS/Ph-1, redundancy was strong between F_{CO_2} and all variables except PAR and $\text{F}_{\text{CH}_4}.\text{ma}2\text{d}$, but it was not reduced when combined with the U^* filter. When considering feasibility only, redundancy was strong between LE and WL.12h; when combining RSSI at 10%, U^* , or their combination, redundancy was also strong.

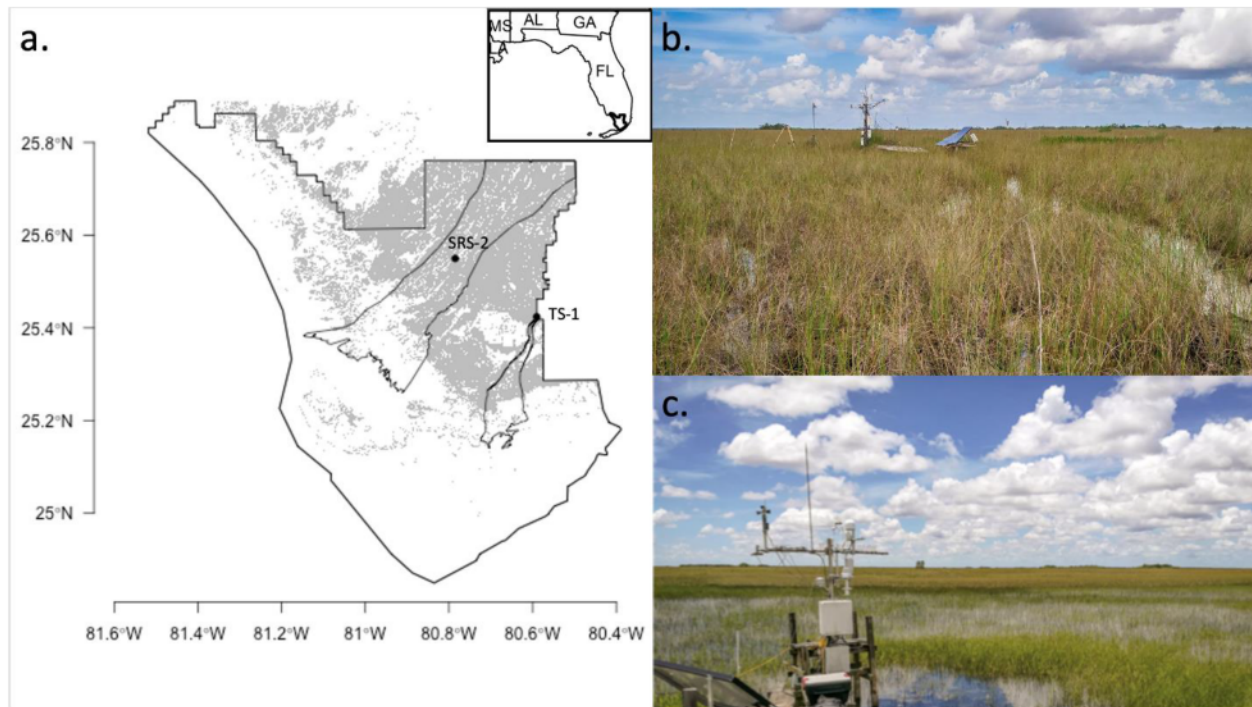
3.3. Random forest models

Variable selection procedures showed that when using the unrestricted predictors model, which included both filled and unfilled variables, synoptic patterns of F_{CH_4} , F_{CO_2} , and WL were the most important predictors of F_{CH_4} in both sites (Fig. 2 and Fig. 3). Variables that tracked multi-day synoptic patterns of F_{CH_4} (2-day, 8-day and 32-day moving averages) were the most important in predicting half-hourly F_{CH_4} ; shorter term variables were more important (2-day and 8-day) than the 32-day. The next most important variable in predicting F_{CH_4} was F_{CO_2} , followed by WL and short-term changes in water level (2-day and 12-hour). At TS/Ph-1, the order of importance for water level variables did not depend on the filtering treatment except with the strongest filter, with the synchronous value more important than the change. At SRS-2 when using the unrestricted predictors model, variables quantifying changes in water level were more important in predicting F_{CH_4} , versus that of the synchronous value in all filtering treatments except feasibility only.

When using restricted predictors, the subsets of variables selected were quite different from those of the unrestricted predictors: F_{CO_2} , GEE, H and LE, PAR, and WL variables (12-hour and 2-day change, and synchronous) (Fig. 4 and 5). Moreover, compared to restricted models estimated with unfilled and filled predictors, the two sites indicated

Table 4Number of gaps in F_{CH_4} (proportion of total number of observations) under different filtering treatments. All treatments included Feasibility Filter.

Site	Gap Length Hours (days)	Feasibility (-2, 2) only	RSSI 10%	RSSI 20%	U*	RSSI 10% + U*	RSSI 20% + U*
TS/Ph-1	0.5 - 4	1809 (5.4%)	1991 (6.7%)	2040 (7.2%)	2706 (10.4%)	2509 (9.8%)	2282 (8.9%)
	4.5 - 24	101 (2.3%)	290 (6.8%)	448 (12.1%)	943 (24.4%)	994 (26.6%)	1020 (28.9%)
	24.5 - 168 (1.01 - 7)	4 (0.9%)	10 (1.8%)	20 (3.6%)	5 (1.0%)	10 (1.9%)	22 (4.0%)
	168.5 - 672 (7.01 - 28)	4 (4.0%)	4 (4.0%)	4 (4.0%)	4 (4.0%)	4 (4.0%)	4 (4.0%)
	672.5 - 1500 (28.01 - 63)	2 (6.1%)	2 (6.1%)	2 (6.1%)	2 (6.1%)	2 (6.1%)	2 (6.1%)
SRS-2	0.5 - 4	1439 (4.5%)	1829 (6.1%)	1779 (6.4%)	3239 (12.3%)	2912 (11.1%)	2442 (9.3%)
	4.5 - 24	84 (1.7%)	318 (8.7%)	487 (14.3%)	780 (17.8%)	853 (22%)	874 (25.1%)
	24.5 - 168 (1.01 - 7)	3 (0.7%)	20 (4.4%)	30 (6%)	3 (0.7%)	20 (4.4%)	30 (6.1%)
	168.5 - 672 (7.01 - 28)	2 (1.6%)	3 (2.1%)	8 (7.3%)	2 (1.6%)	3 (2.1%)	8 (7.3%)
	672.5 - 1500 (28.01 - 63)	2 (6.5%)	2 (6.5%)	2 (6.5%)	2 (6.5%)	2 (6.5%)	2 (6.5%)

**Fig. 1.** Eddy covariance tower sites in Everglades National Park, USA: (a) location map, (b) short-hydroperiod freshwater marl prairie site, and (c) long-hydroperiod freshwater marsh site.

different orders of variable importance. For TS/Ph-1, F_{CO_2} was consistently the most important, regardless of filtering treatment (Fig. 4). Also, variables that described synchronous WL and energy (H and LE) were more important than those of water level change, PAR, and GEE. With the U* filtering treatment, GEE-1d became more important for prediction of F_{CH_4} when using restricted predictors only to estimate models, except when added to the RSSI 20% treatment.

For SRS-2, F_{CO_2} was consistently the most important predictor when using restricted predictors, except when filtering treatments were more relaxed, in which case synchronous water level was the most important (Fig. 5). As with TS/Ph-1, energy variables (H and LE) were the next most important, followed by GEE. Changes in water level and PAR, were least important for prediction of F_{CH_4} when using the set of restricted predictors in SRS-2.

3.4. Annual CH_4 budgets

At TS/Ph-1, using the feasibility filter only resulted in the lowest CH_4 budget, except in 2019 when this filter resulted in the highest CH_4 budget (Fig. 6). There was no clear pattern in terms of which filtering treatment resulted in the highest annual budget, except that it involved U* in every year except 2019. The difference between the lowest and

highest CH_4 budgets varied by filtering treatment, from 14% (2017) to 45% (2019), although there was substantial overlap among treatments when considering uncertainty. As filtering treatments became more constraining, their prediction variance increased, such that the U* + 20% RSSI filter always had the highest prediction standard errors and the feasibility only filter had the lowest standard errors in both sites. Uncertainty in TS/Ph-1 CH_4 budgets also varied by year and the type of predictor dataset used. In 2018, due to unusually good data coverage (i. e., CH_4 2-day, 8-day, and 32-day moving averages contained at least 10% data coverage), the restricted predictors model was unused in predicting CH_4 in 2018 in TS/Ph-1 (Supplemental Information Table S1), and therefore did not contribute to the annual budget (Fig. 6). In all other years, errors due to using the unrestricted predictors model in TS/Ph-1 were approximately half that of the restricted predictors model when using RSSI filtering only, whereas they were 20–40% higher when using U* or RSSI + U* filtering.

In SRS-2, CH_4 budgets were more than twice as high as that of TS/Ph-1 (Fig. 6; Supplemental Information Table S2). There was no clear pattern in annual budgets by filtering treatment; in 2017–18, the highest budgets were attained using the 10% RSSI filter, whereas the highest involved the U* filter in 2019, and U* plus 20% RSSI in 2016. In contrast, the lowest budgets were attained with the U* filter in 2016–17,

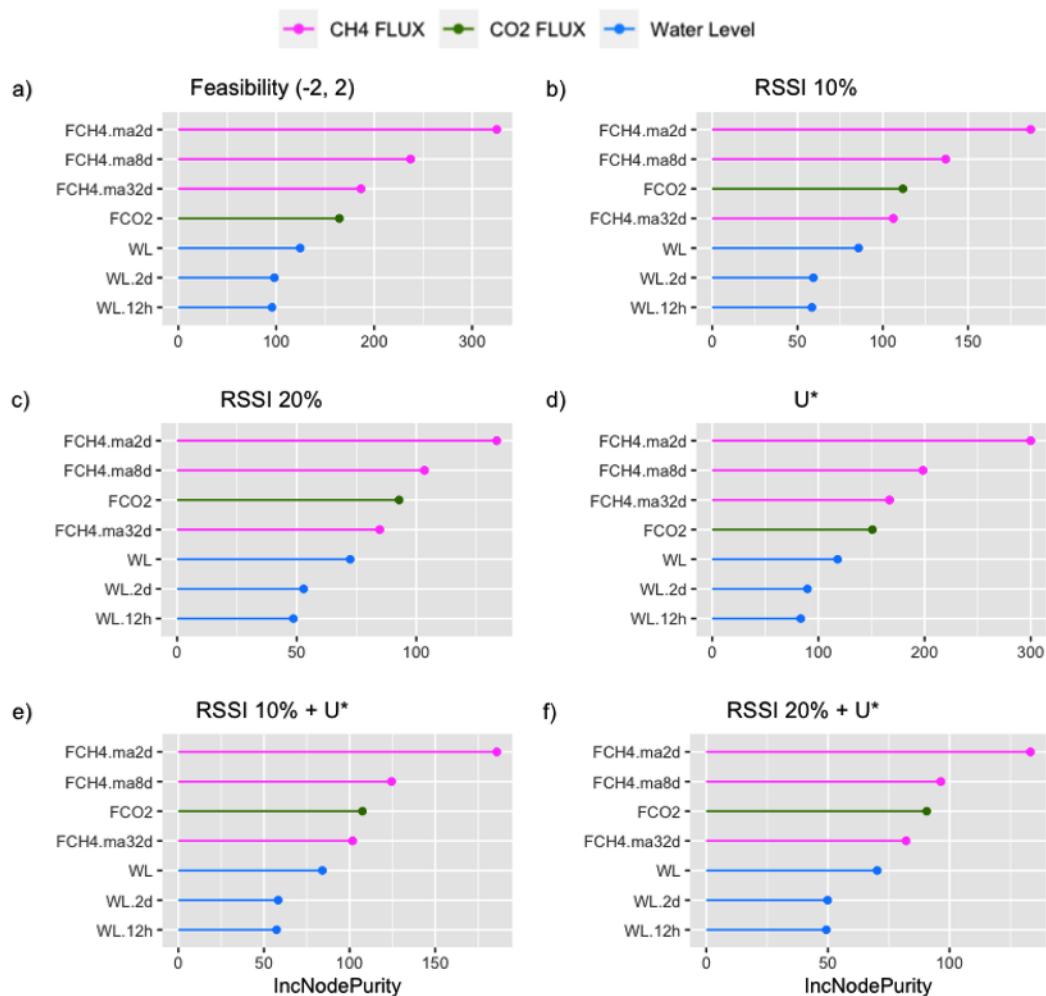


Fig. 2. Variable importance for random forest models of F_{CH_4} estimated with the unrestricted (both unfilled and filled) predictors at Taylor Slough (TS/Ph-1). IncNodePurity denotes the variable importance based on the Gini impurity index used for calculating splits in trees. All filtering treatments included the Feasibility filter.

and the 20% RSSI filter in 2018–19. The largest difference among filtering treatments occurred during 2017, when RSSI 10% filtering was 63% higher than that of the feasibility only filter. The differences were lowest the next year (2018; 7%). Unlike TS/Ph-1, errors due to using the unrestricted predictors model in SRS-2 were on average three times as high as those due to using the restrictive predictors in 2016 and 2018, and more than ten times as high using the restrictive predictors in 2019 (Supplemental Information Table S3). During 2017, both models contributed in roughly equal proportion to the total prediction error in CH_4 budget. In all other years, errors due to using the unrestricted predictors model in SRS-2 were always much higher than that of the restricted predictors model and increased when applying U^* filtering versus RSSI only filtering only.

4. Discussion

While many studies have concluded that tropical/subtropical wetlands are an important natural source for CH_4 through remote sensing, modeling and chamber measurements, CH_4 budget estimates reported have been associated with large uncertainties (Saunois et al. 2016; Zhang et al. 2017; Bloom et al. 2017; Parker et al. 2018; Jeffrey et al. 2019). This study confirmed that Everglades freshwater wetlands are a source for CH_4 with budgets of 6.7–17.9 and 24.3–47.5 $g\ C\ m^{-2}\ yr^{-1}$ for the short-hydroperiod (TS/Ph-1) and long-hydroperiod wetlands (SRS-2), respectively. The annual budget of SRS-2 is close to that of a

subtropical freshwater marsh in the Mississippi River Delta (Holm et al. 2016) and also matches the flux range of temperate wetlands (27.7 and $63.4 \pm 1.0\ g\ C\ m^{-2}\ yr^{-1}$) observed by Hermes et al. (2018), but it more closely matches those of tropical wetlands ($24.5 \pm 20.7\ g\ C\ m^{-2}\ yr^{-1}$) from the FLUXNET- CH_4 database (Delwiche et al. 2021). While annual CH_4 emissions from TS/Ph-1 were also within this range, they were substantially lower. However, we also found that using different data filtering criteria in the EC technique could result in annual budget differences up to 63%. Filtering criteria vary among different studies but their effect on gap-filling results and budget estimates are hardly evaluated. Our results highlight the importance of the choice of filtering criteria in CH_4 studies using the EC approach, as it can have a significant effect on uncertainty in CH_4 budget estimates.

Among the filtering criteria we used, U^* introduced the most gaps compared to RSSI and feasibility only filters. The U^* filter is used to remove EC data collected under atmospheric conditions that do not generate sufficient turbulence to meet EC assumptions (Barr et al. 2013). These conditions occur mostly at nighttime. Therefore, the U^* filter did not introduce many gaps that were larger than one day. Nonetheless, due to the large number of gaps created by U^* filtering (Table 4), the medians of the annual budget estimates were either lower (in 2017) or higher (in 2019) versus those based on RSSI filtering alone for the long-hydroperiod site (SRS-2), but with larger uncertainties. RSSI filtering could occasionally introduce small gaps when contaminants in the optical path of the analyzer were removed by the built-in automatic

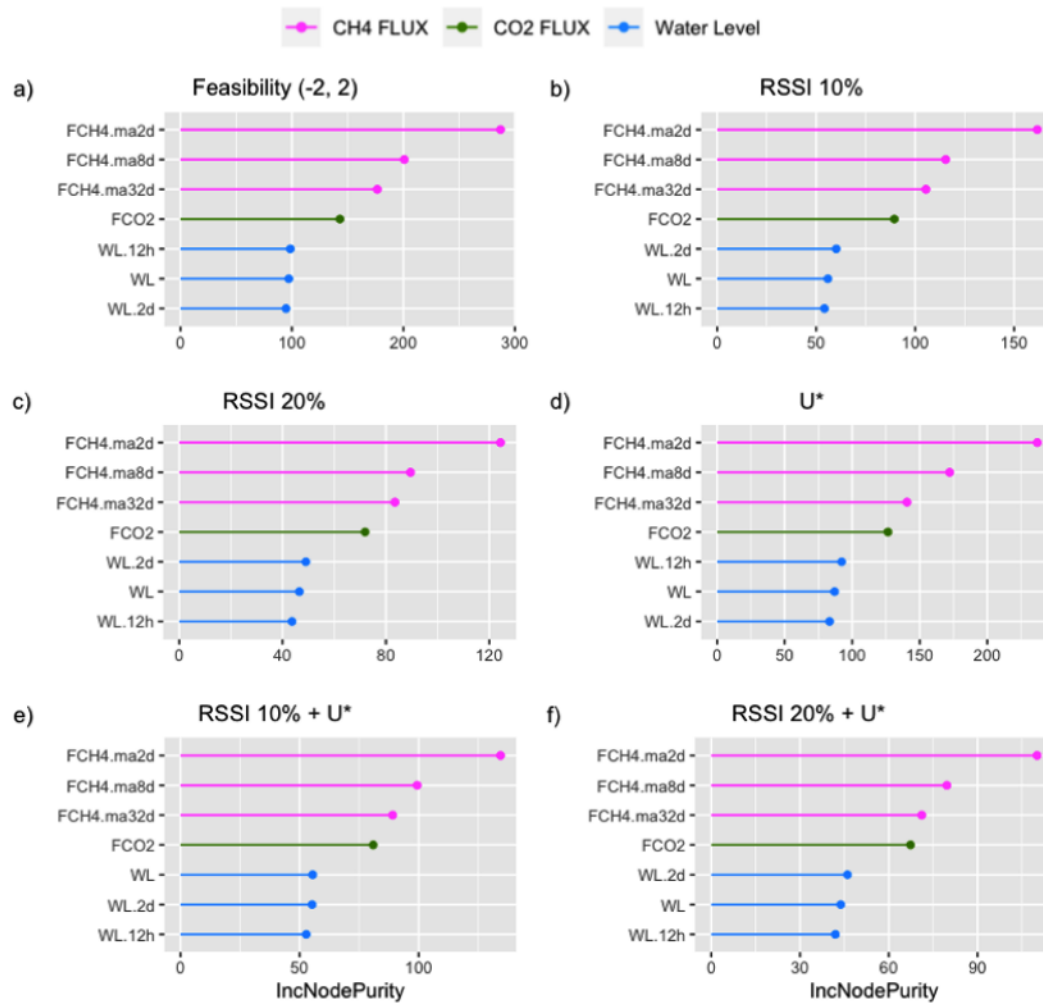


Fig. 3. Variable importance for random forest models of F_{CH_4} estimated with the unrestricted (both unfilled and filled) predictors at Shark River Slough (SRS-2). IncNodePurity denotes the variable importance based on the Gini impurity index used for calculating splits in trees. All filtering treatments included the Feasibility filter.

cleaning system at a 30-min interval (e.g., wet bird dropping, raindrops, living animals). However, when the cleaning system could not efficiently eliminate the contaminants (e.g., dry bird dropping, spiderwebs, upper mirror contamination), large gaps were generated until on-site cleaning was performed (Supplemental Information Table S4). Overall, different filtering criteria can generate gaps that can be heavily biased towards a certain length or time of day, and subsequently influence the gap-filling model and the ultimate budget estimate.

According to Kim et al. (2020) and Irvin et al. (2021), RF has the best performance in gap-filling F_{CH_4} from various ecosystems due to its efficiency in handling highly nonlinear relationships. In our RF models for gap-filling, we found that water level and F_{CO_2} were the strongest predictors, regardless of the gap size (Fig. 3-Fig. 6). Water level plays an important role in controlling the soil aerobic/anaerobic condition that determines the active levels of methanotrophs (i.e., CH_4 consumption) and methanogens (i.e., CH_4 production) (Segers 1998), and is considered as an important control for CH_4 flux (Irvin et al. 2021; Knox et al. 2019). For example, plant aerenchyma tissues are an important pathway for CH_4 emissions (Jeffrey et al., 2019); when water level is higher than mean plant canopy height, plants are fully submerged and thus gas transport is restricted. F_{CO_2} is mainly linked to F_{CH_4} in the following ways. First, carbon uptake by the ecosystem from the atmosphere can be further transported to the soil as root exudates and used by methanogens as substrates for CH_4 production (e.g., Waldo et al. 2019). Second, vascular plants mediate both CH_4 release from soil to the atmosphere

(Bhullar et al. 2013) and CO_2 uptake from the atmosphere. Hatala et al. (2012) even found that the diurnal pattern in CH_4 flux was driven by GEE. On the other hand, depending on the soil aerobic condition, the produced CH_4 can be oxidized before being released into the atmosphere and, instead, emitted in the form of CO_2 (King 1990). However, relatively few studies (~20%) included water status and F_{CO_2} (Supplemental Information Table S5). Given that our findings show that these variables are closely coupled with F_{CH_4} , we suggest that water level and F_{CO_2} should be essential factors to be included in gap-filling models for the F_{CH_4} in wetlands.

Using the restricted predictors models, energy fluxes (i.e., H and LE) became more important drivers, indicating that processes underlying F_{CH_4} are coupled with the ecosystem energy exchanges. This finding is plausible because gas movement through aerenchyma is driven by thermal processes (Dacey 1980). Interestingly, we also found that GEE of the previous day showed substantial contributions to the RF models with restricted predictors, suggesting there is a time lag for F_{CH_4} to respond to carbon substrate input through photosynthesis. This lagged effect has seldom been considered in previous studies (but see e.g., Kim et al. 2020) and should be included in F_{CH_4} models. Irvin et al. (2021) reported that the soil temperature and an indicator variable for season were the most important predictors over a wide range of wetlands at mid- and high- latitudes; however, temperatures at our sites showed very minor effects on F_{CH_4} , mainly due to the weak seasonal temperature signal in the subtropical region, highlighting the need to consider

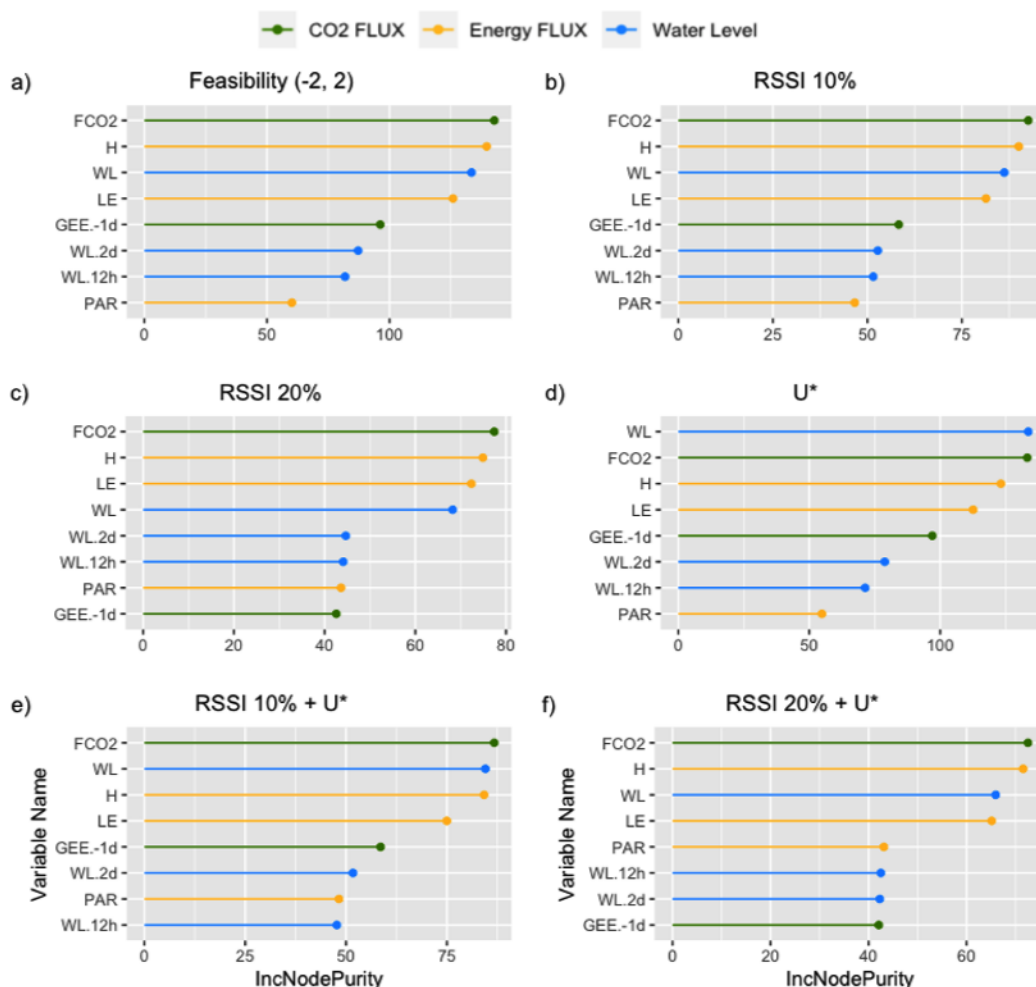


Fig. 4. Variable importance for random forest models of F_{CH_4} using restricted predictors only at Taylor Slough (TS/Ph-1). IncNodePurity denotes variable importance based on Gini impurity index used for calculating splits in trees. All filtering treatments included Feasibility filter.

site-specific variables in gap-filling models.

To gap-fill for annual budgets, we used two types of models (i.e., unrestricted and restricted predictors models). This hybrid approach takes advantage of the observed values of predictors while still having the flexibility to fill gaps even when predictors are associated with gaps themselves. This approach is practical especially when F_{CH_4} is also included as a predictor in the model, which we found accounted for synoptic patterns in CH_4 . While we found some studies that used averages of F_{CH_4} in gap-filling, to our knowledge, this is the first study that implemented a hybrid approach to gap-fill when those averages could not be calculated due to missingness. We did not find particularly large uncertainties associated with either of the two model types in the final budget estimates, although generally the unrestricted predictors model was a larger source of uncertainty than the restricted predictors model, and the uncertainties depended largely on the amount of gaps that were filled by them. Overall, this hybrid approach enabled us to fill data gaps with predictors that themselves had gaps, which commonly occur in most EC studies of F_{CH_4} .

In this study, we set the feasibility filter for F_{CH_4} to $(-2, 2) \mu\text{mol m}^{-2} \text{s}^{-1}$. It is noteworthy that this filter will exclude emissions from ebullition that are $>2 \mu\text{mol m}^{-2} \text{s}^{-1}$. Ebullition is a pathway of releasing a large amount of the produced CH_4 in the soil to the atmosphere, which can account for $\sim 20\%$ of the annual CH_4 budget in a subtropical wetland (Jeffrey et al. 2019). Because ebullition occurs episodically, no approach is currently available to partition the ebullition fluxes from outliers caused by technical limitations of the EC method. Since the ebullition

fluxes are most likely removed and filled with relatively small values, the annual CH_4 budget can be underestimated to a certain degree. Despite this, our site-level estimates were similar to those of tropical sites in the FLUXNET- CH_4 network (Delwiche et al 2021). Future studies to develop approaches to compensate for this underestimation are needed.

Given the effect of data filtering criteria on final CH_4 budget estimates and the fact that both the RSSI and U^* filters address the limitations of the EC technique, we recommend that both filters should be applied in the standard workflow for F_{CH_4} data processing to increase the confidence of CH_4 budget estimates. Rather than ignoring data filtering procedures, publications should clearly and transparently describe protocols for data processing. Furthermore, since different filters can generate gaps of different sizes, affecting gap-filling models as well as final budget estimates, we also encourage studies to document the size and frequency of gaps introduced during data processing.

5. Conclusions

Although EC data filtering is important in determining annual budgets of F_{CH_4} , the filtering approaches for CH_4 flux are greatly understudied. The decision to select site-specific filtering criteria should be justified and presented transparently, as variation in filtering criteria and site conditions leads to the removal of a substantial amount of collected data (Table 1). The removal of large amounts of data also imposes additional uncertainties on annual CH_4 budgets that will further

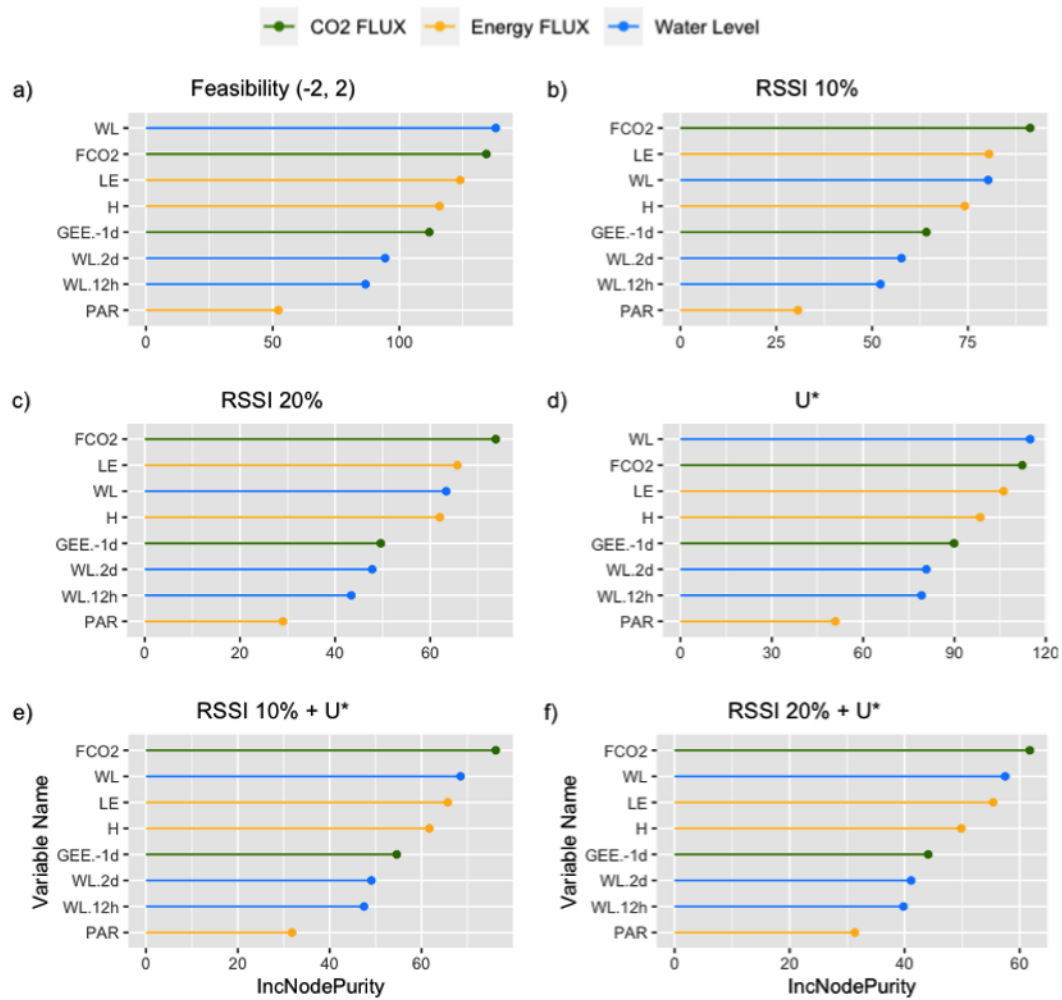


Fig. 5. Variable importance for random forest models of F_{CH_4} using restricted predictors only at Shark River Slough (SRS-2). IncNodePurity denotes variable importance based on Gini impurity index used for calculating splits in trees. All filtering treatments included Feasibility filter.

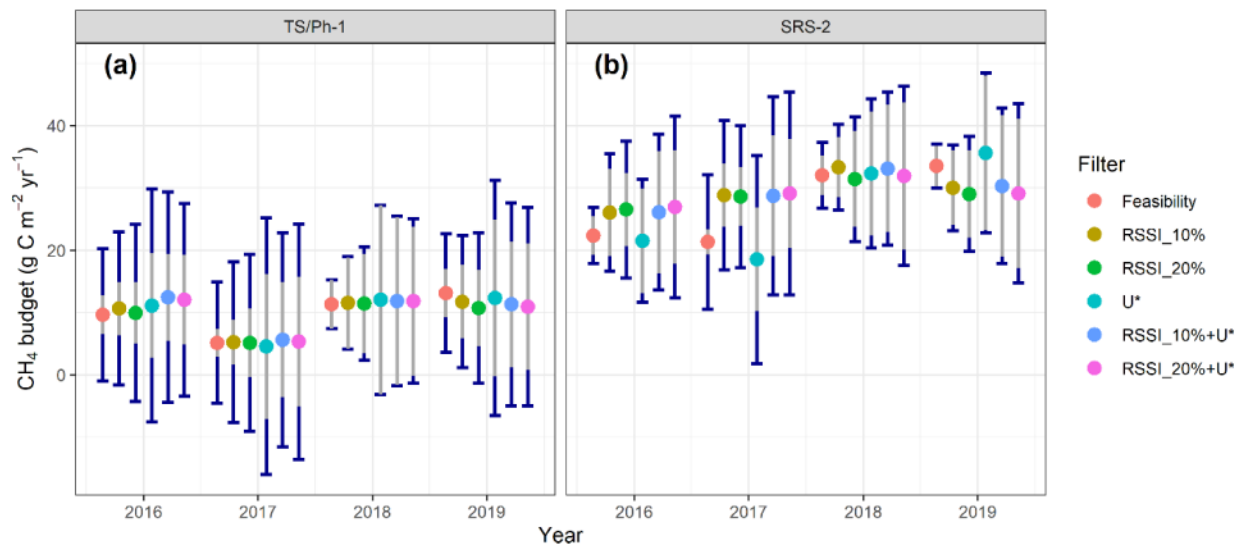


Fig. 6. CH_4 budget for Everglades freshwater wetlands sites at (a) TS/Ph-1 and (b) SRS-2 in $g\ CH_4\ m^{-2}\ yr^{-1}$ by filtering treatment. The circle denotes the annual budget and the error bars show the prediction interval for the annual budget (± 1 standard error; gray section = proportion of error from unrestricted predictors model; dark blue section = proportion of error from restricted predictors model).

propagate biases in efforts to upscale and predict global CH₄ budgets, though this uncertainty is rarely acknowledged. In this study, we evaluated the effects that data filtering has on gap-filling of F_{CH₄} from EC measurements in short- and long-hydroperiod freshwater wetlands in the subtropical Florida Everglades. We show how filtering criteria influence annual CH₄ budget estimates through their impact on the size and timing of the data gaps produced. How these gaps are filled is further a function of the size of gaps in the predictor variables. In all, we show that F_{CO₂}, water level, and energy fluxes are effective at gap filling F_{CH₄}, and that while annual budgets do vary, Everglades freshwater wetlands are a small source of CH₄.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

All research was performed under permits issued by Everglades National Park (EVER-2009-SCI-0070 and EVER-2013-SCI-0058). The authors would like to acknowledge the excellent support provided by the Florida Coastal Everglades Long Term Ecological Research Program (FCE-LTER) and the Southeast Environmental Research Center in the Institute of Water and Environment at Florida International University. This research was also supported by the Everglades Foundation. This manuscript is partly based upon work supported by the National Science Foundation (NSF) through the FCE-LTER program under Grant DEB-1237517. This research is also based in part on support from the Department of Energy's (DOE) National Institute for Climate Change Research (NICCR) through grant 07-SC-NICCR-1059 and the National Science Foundation Division of Atmospheric & Geospace Sciences Atmospheric Chemistry Program awards 1561139, 1233006, 1801310 and 1807533. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of FCE-LTER, NSF, or DOE. This is contribution #1428 from the Institute of Environment at Florida International University.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.agrformet.2022.108972.

References

- Alberto, M.C.R., Wassmann, R., Buresh, R.J., Quilty, J.R., Correa, T.Q., Sandro, J.M., Centeno, C.A.R., 2014. Measuring methane flux from irrigated rice fields by eddy covariance method using open-path gas analyzer. *Field Crops Research* 160, 12–21.
- Armentano, T.V., Sah, J.P., Ross, M.S., Jones, D.T., Cooley, H.C., Smith, C.S., 2006. Rapid responses of vegetation to hydrological changes in Taylor Slough, Everglades National Park, Florida, USA. *Hydrobiologia* 569, 293–309. <https://doi.org/10.1007/s10750-006-0138-8>.
- Arrhenius, S., 1915. *Quantitative laws in biological chemistry*. Bell, London.
- Baldocchi, D.D., 2003. Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. *Global Change Biology* 9, 479–492.
- Barba, J., Bradford, M.A., Brewer, P.E., Bruhn, D., Covey, K., van Haren, J., Megonigal, J. P., Mikkelsen, T.N., Pangala, S.R., Pihlatie, M., Poulter, B., Rivas-Ubach, A., Schadt, C.W., Terazawa, K., Warner, D.L., Zhang, Z., Vargas, R., 2019. Methane emissions from tree stems: a new frontier in the global carbon cycle. *The New Phytologist* 222 (1), 18–28.
- Barr, A.G., Richardson, A.D., Hollinger, D.Y., Papale, D., Arain, M.A., Black, T.A., Bohrer, G., Dragoni, D., Fischer, M.L., Gu, L., Law, B.E., Margolis, H.A., McCaughey, J.H., Munger, J.W., Oechel, W., Schaeffer, K., 2013. Use of change-point detection for friction-velocity threshold evaluation in eddy-covariance studies. *Agricultural and Forest Meteorology* 171–172, 31–45.
- Beck, C., Grieser, M., Kottek, M., Rubel, F., Rudolf, B., 2006. Characterizing global climate change by means of Köppen Climate Classification. *Klimastatusbericht* 2005, 139–149.
- Bhullar, G.S., Edwards, P.J., Olde Venterink, H., 2013. Variation in the plant-mediated methane transport and its importance for methane emission from intact wetland peat mesocosms. *Journal of Plant Ecology* 6 (4), 298–304.
- Bloom, A.A., Bowman, K.W., Lee, M., Turner, A.J., Schroeder, R., Worden, J.R., Weidner, R., McDonald, K.C., Jacob, D.J., 2017. A global wetland methane emissions and uncertainty dataset for atmospheric chemical transport models (WetCHARTs version 1.0). *Geoscientific Model Development* 10 (6), 2141–2156.
- Bodesheim, P., Jung, M., Gans, F., Mahecha, M.D., Reichstein, M., 2018. Upscaled diurnal cycles of land-atmosphere fluxes: a new global half-hourly data product. *Earth Systems Science Data* 10, 1327–1365. <https://doi.org/10.5194/essd-10-1327-2018>.
- Breiman, L., 2001. Random Forests. *Machine Learning* 45 (1), 5–32.
- Carmichael, M.J., Bernhardt, E.S., Bräuer, S.L., Smith, W.K., 2014. The role of vegetation in methane flux to the atmosphere: should vegetation be included as a distinct category in the global methane budget? *Biogeochemistry* 119 (1), 1–24.
- Chen, W., Wang, B., Zhang, F., Li, Z., Wang, J., Yu, G., Wen, X., Niu, S., 2020. Hysteric relationship between plant productivity and methane uptake in an alpine meadow. *Agricultural and Forest Meteorology* 288–289, 107982.
- Chu, H., Luo, X., Ouyang, Z., Chan, W.S., Dengel, S., Biraud, S.C., Torn, M.S., Metzger, S., Kumar, J., Arain, M.A., Arkebauer, T.J., Baldocchi, D., Bernacchi, C., Billesbach, D., Black, T.A., Blanken, P.D., Bohrer, G., Bracho, R., Brown, S., Zona, D., 2021. Representativeness of Eddy-Covariance flux footprints for areas surrounding AmeriFlux sites. *Agricultural and Forest Meteorology* 301–302, 108350.
- Clark, K.L., Gholz, H.L., Moncrieff, J.B., Cropley, F., Loescher, H.W., 1999. Environmental Controls over Net Exchanges of Carbon Dioxide from Contrasting Florida Ecosystems. *Ecological Applications* 9 (3), 936–948.
- Cleverly, J., Dahm, C., Thibault, J., Gilroy, D., Coonrod, J., 2002. Seasonal estimates of actual evapotranspiration from *Tamarix ramosissima* stands using three-dimensional eddy covariance. *Journal of Arid Environments* 52, 181–197.
- Conrad, R., 2009. The global methane cycle: recent advances in understanding the microbial processes involved. *Environmental Microbiology Reports* 1 (5), 285–292.
- Dacey, J.W., 1980. Internal winds in water lilies: an adaptation for life in anaerobic sediments. *Science* 210 (4473), 1017–1019.
- Dai, S., Ju, W., Zhang, Y., He, Q., Song, L., Li, J., 2019. Variations and drivers of methane fluxes from a rice-wheat rotation agroecosystem in eastern China at seasonal and diurnal scales. *Science of the Total Environment* 690, 973–990. <https://doi.org/10.1016/j.scitotenv.2019.07.012>.
- 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., Jackson, R.B., 2021. FLUXNET-CH₄: a global, multi-ecosystem dataset and analysis of methane seasonality from freshwater wetlands. *Earth System Science Data* 13 (7), 3607–3689.
- Dengel, S., Zona, D., Sachs, T., Aurela, M., Jammot, M., Parmentier, F.-J.W., Oechel, W., Vesala, T., Others, 2013. Testing the applicability of neural networks as a gap-filling method using CH₄ flux data from high latitude wetlands. *Biogeosciences* 10, 8185–8200. https://helda.helsinki.fi/bitstream/handle/10138/162381/bg_10_8185_2013.pdf?sequence=1.
- Duever, M.J., Meeder, J.F., Meeder, L.C., McCollom, J.M., 1994. The climate of south Florida and its role in shaping the Everglades ecosystem. In: Davis, S.M., Ogden, J.C. (Eds.), *Everglades: the Ecosystem and its Restoration*. St. Lucie Press, Delray Beach, FL, pp. 225–248.
- Duever, M.J., Carlson, J.E., Riopelle, L.A., Gunderson, L.H., Duever, L.C., 1976. Ecosystem analyses at corkscrew swamp. Third Annual Report on Cypress Wetlands, Florida University, Center for Wetlands, Gainesville P 707–737.
- Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger, D., Jensen, N.-O., Katul, G., Keronen, P., Kowalski, A., Lai, C.T., Wofsy, S., 2001. Gap filling strategies for defensible annual sums of net ecosystem exchange. *Agricultural and Forest Meteorology* 107, 43–69.
- Foken, T., Leclerc, M.Y., 2004. Methods and limitations in validation of footprint models. *Agricultural and Forest Meteorology* 127 (3), 223–234.
- Foken, T., Wichura, B., 1996. Tools for quality assessment of surface-based flux measurements. *Agricultural and Forest Meteorology* 78 (1), 83–105.
- Gaiser, E.E., Childers, D.L., Jones, R.D., Richards, J.H., Scinto, L.J., Trexler, J.C., 2006. Periphyton responses to eutrophication in the Florida Everglades: Cross-system patterns of structural and compositional change. *Limnology and Oceanography* 51 (1part2), 617–630.
- Ge, H.-X., Zhang, H.-S., Zhang, H., Cai, X.-H., Song, Y., Kang, L., 2018. The characteristics of methane flux from an irrigated rice farm in East China measured using the eddy covariance method. *Agricultural and Forest Meteorology* 249, 228–238.
- Genier, R., Poggi, J.M., Tuleau-Malot, C., 2010. Variable selection using random forests. *Pattern Recognition Letters* 31, 2225. <https://doi.org/10.1016/j.patrec.14>.
- Gottlieb, A.D., Richards, J.H., Gaiser, E.E., 2006. Comparative study of periphyton community structure in long and short-hydroperiod Everglades marshes. *Hydrobiologia* 569, 195–207. <https://doi.org/10.1007/s10750-006-0132-1>.
- Goulden, M.L., Munger, J.W., Fan, S.-M., Daube, B.C., Wofsy, S.C., 1996. Measurements of carbon sequestration by long-term eddy covariance: Methods and a critical evaluation of accuracy. *Global Change Biology* 2 (3), 169–182.
- Hatala, J.A., Detto, M., Baldocchi, D.D., 2012. Gross ecosystem photosynthesis causes a diurnal pattern in methane emission from rice. *Geophysical Research Letters* 39, L06409. <https://doi.org/10.1029/2012GL051303>.
- Helbig, M., Chasmer, L.E., Kljun, N., Quinton, W.L., Treat, C.C., Sonnentag, O., 2017. The positive net radiative greenhouse gas forcing of increasing methane emissions from a thawing boreal forest-wetland landscape. *Global Change Biology* 23 (6), 2413–2427.

- Holm, G., Perez, B., McWhorter, D., Krauss, K., Johnson, D., Raynie, R., Killebrew, C., 2016. Ecosystem Level Methane Fluxes from Tidal Freshwater and Brackish Marshes of the Mississippi River Delta: Implications for Coastal Wetland Carbon Projects. *Wetlands* 36, 401–413.
- Hemes, K.S., Chamberlain, S.D., Eichelmann, E., Knox, S.H., Baldocchi, D.D., 2018. A biogeochemical compromise: The high methane cost of sequestering carbon in restored wetlands. *Geophysical Research Letters* 45 (12), 6081–6091.
- Hommelberg, J., Mauder, M., Drösler, M., Heidbach, K., Werle, P., Schmid, H.P., 2014. Ecosystem scale methane fluxes in a natural temperate bog-pine forest in southern Germany. *Agricultural and Forest Meteorology* 198–199, 273–284.
- Intergovernmental Panel on Climate Change, 2014. *Climate Change 2013: The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.
- Irvin, J., Zhou, S., McNicol, G., Lu, F., Liu, V., Fluet-Chouinard, E., Ouyang, Z., Knox, S. H., Lucas-Moffat, A., Trotta, C., Papale, D., Vitale, D., Mammarella, I., Alekseychik, P., Aurela, M., Avati, A., Baldocchi, D., Bansal, S., Bohrer, G., Jackson, R.B., 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.
- Iwaniec, D.M., Childers, D.L., Rondeau, D., Madden, C.J., Saunders, C., 2006. Effects of hydrologic and water quality drivers on periphyton dynamics in the southern Everglades. *Hydrobiologia* 569, 223–235. <https://doi.org/10.1007/s10750-006-0134-z>.
- Iwata, H., Harazono, Y., Ueyama, M., Sakabe, A., Nagano, H., Kosugi, Y., Takahashi, K., Kim, Y., 2015. Methane exchange in a poorly-drained black spruce forest over permafrost observed using the eddy covariance technique. *Agricultural and Forest Meteorology* 214–215, 157–168.
- Jeffrey, L.C., Maher, D.T., Johnston, S.G., Kelaher, B.P., Steven, A., Tait, D.R., 2019. Wetland methane emissions dominated by plant-mediated fluxes: Contrasting emissions pathways and seasons within a shallow freshwater subtropical wetland. *Limnology and Oceanography* 64 (5), 1895–1912.
- Kim, Y., Johnson, M.S., Knox, S.H., Black, T.A., Dalmagro, H.J., Kang, M., Kim, J., Baldocchi, D., 2020. Gap-filling approaches for eddy covariance methane fluxes: A comparison of three machine learning algorithms and a traditional method with principal component analysis. *Global Change Biology* 26 (3), 1499–1518.
- King, G.M., 1990. Dynamics and controls of methane oxidation in a Danish wetland sediment. *FEMS Microbiology Letters* 74 (4), 309–323.
- Kirschke, S., Bousquet, P., Ciais, P., Saunio, M., Canadell, J.G., Dlugokencky, E.J., Bergamaschi, P., Bergmann, D., Blake, D.R., Bruhwiler, L., Cameron-Smith, P., Castaldi, S., Chevallier, F., Feng, L., Fraser, A., Heimann, M., Hodson, E.L., Houweling, S., Josse, B., Zeng, G., 2013. Three decades of global methane sources and sinks. *Nature Geoscience* 6, 813.
- Kljun, N., Calanca, P., Rotach, M.W., Schmid, H.P., 2004. A simple parameterisation for flux footprint predictions. *Boundary-Layer Meteorology* 112 (3), 503–523.
- Kljun, N., Rotach, M.W., Schmid, H.P., 2002. A three-dimensional backward Lagrangian footprint model for a wide range of boundary-layer stratifications. *Boundary-Layer Meteorology* 103 (2), 205–226.
- Knox, S.H., Matthes, J.H., Sturtevant, C., Oikawa, P.Y., Verfaillie, J., Baldocchi, D., 2016. Biophysical controls on interannual variability in ecosystem-scale CO₂ and CH₄ exchange in a California rice paddy. *Journal of Geophysical Research*. *Biogeosciences* 121 (3), 978–1001.
- Knox, S.H., Sturtevant, C., Matthes, J.H., Koteen, L., Verfaillie, J., Baldocchi, D., 2015. Agricultural peatland restoration: effects of land-use change on greenhouse gas (CO₂ and CH₄) fluxes in the Sacramento-San Joaquin Delta. *Global Change Biology* 21 (2), 750–765.
- Knox, S.H., Jackson, R.B., Poulter, B., McNicol, G., Fluet-Chouinard, E., Zhang, Z., Hugelius, G., Bousquet, P., Canadell, J.G., Saunio, M., Zona, D., 2019. FLUXNET-CH₄ Synthesis Activity: Objectives, Observations, and Future Directions. *Bulletin of the American Meteorological Society* 100 (12), 2607–2632.
- Korrensalo, A., Männistö, E., Alekseychik, P., Mammarella, I., Rinne, J., Vesala, T., Tuittila, E.-S., 2018. Small spatial variability in methane emission measured from a wet patterned boreal bog. *Biogeosciences* 15 (6), 1749–1761.
- Kratzer, G., Furrer, R., 2018. *varrank: An R Package for Variable Ranking Based on Mutual Information with Applications to Systems Epidemiology*. R package version 0.1. <https://CRAN.R-project.org/package=varrank>.
- Kunwor, S., Starr, G., Loescher, H.W., Staudhammer, C.L., 2017. Preserving the variance in imputed eddy-covariance measurements: Alternative methods for defensible gap filling. *Agricultural and Forest Meteorology* 232, 635–649.
- Le Mer, J., Roger, P., 2001. Production, oxidation, emission and consumption of methane by soils: A review. *European Journal of Soil Biology* 37 (1), 25–50.
- Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. *R News* 2 (3), 18–22.
- Li, H., Dai, S., Ouyang, Z., Xie, X., Guo, H., Gu, C., Xiao, X., Ge, Z., Peng, C., Zhao, B., 2018. Multi-scale temporal variation of methane flux and its controls in a subtropical tidal salt marsh in eastern China. *Biogeochemistry* 137 (1), 163–179.
- Li, H., Guo, H.-Q., Helbig, M., Dai, S.-Q., Zhang, M.-S., Zhao, M., Peng, C.-H., Xiao, X.-M., Zhao, B., 2019. Does direct-seeded rice decrease ecosystem-scale methane emissions?—A case study from a rice paddy in southeast China. *Agricultural and Forest Meteorology* 272–273, 118–127.
- Malone, S.L., Staudhammer, C.L., Oberbauer, S.F., Olivas, P., Ryan, M.G., Schedlbauer, J. L., Loescher, H.W., Starr, G., 2014. El Niño Southern Oscillation (ENSO) enhances CO₂ exchange rates in freshwater marsh ecosystems in the Florida Everglades. *PLoS One* 9 (12), e115058.
- McDermitt, D., Burba, G., Xu, L., Anderson, T., Komissarov, A., Riensche, B., Schedlbauer, J., Starr, G., Zona, D., Oechel, W., Oberbauer, S., Hastings, S., 2011. A new low-power, open-path instrument for measuring methane flux by eddy covariance. *Applied Physics, B, Lasers and Optics*, 102 (2), 391–405.
- Michaelis, L., Menten, M.L., 1913. Die Kinetik der Invertinwirkung. *Biochemische Zeitschrift* 49, 333–369.
- Morin, T.H., Bohrer, G., Stefanik, K.C., Rey-Sanchez, A.C., Matheny, A.M., Mitsch, W.J., 2017. Combining eddy-covariance and chamber measurements to determine the methane budget from a small, heterogeneous urban floodplain wetland park. *Agricultural and Forest Meteorology* 237–238, 160–170.
- Nadeau, D.F., Rousseau, A.N., Coursolle, C., Margolis, H.A., Parlange, M.B., 2013. Summer methane fluxes from a boreal bog in northern Quebec, Canada, using eddy covariance measurements. *Atmospheric Environment* 81, 464–474.
- National Climate Data Center (NCDC), 2019. Royal Palm Rs daily surface data. NCDC, Asheville. <http://www.ncdc.noaa.gov/oa/ncdc.html>.
- Nemitz, E., Mammarella, I., Ibrom, A., Aurela, M., 2018. Standardisation of eddy-covariance flux measurements of methane and nitrous oxide. *International Agrophysics* 32 (4), 517–549. <http://agro.icm.edu.pl/agro/element/bwmeta1.element.agro-2aceedee-1d65-4e20-ab07-b2309782b67a/c/s.517.pdf>.
- Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Yakir, D., 2006. Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. *Biogeosciences* 3, 571–583. <https://doi.org/10.5194/bg-3-571-2006>.
- Parker, R.J., Boesch, H., McNorton, J., Conyn-Platt, E., Gloor, M., Wilson, C., Chipperfield, M.P., Hayman, G.D., Bloom, A.A., 2018. Evaluating year-to-year anomalies in tropical wetland methane emissions using satellite CH₄ observations. *Remote Sensing of Environment* 211, 261–275.
- Pypker, T.G., Moore, P.A., Waddington, J.M., Hribljan, J.A., Chimner, R.C., 2013. Shifting environmental controls on CH₄ fluxes in a sub-boreal peatland. *Biogeosciences* 10 (12), 7971–7981.
- Sabrekov, A.F., Glagolev, M.V., Alekseychik, P.K., Smolentsev, B.A., Terentiev, I.E., Krivenok, L.A., Maksyutov, S.S., 2016. A process-based model of methane consumption by upland soils. *Environmental Research Letters* 11 (7), 075001.
- Saunio, M., Bousquet, P., Poulter, B., Peregon, A., Ciais, P., Canadell, J.G., Dlugokencky, E.J., Etiope, G., Bastviken, D., Houweling, S., Janssens-Maenhout, G., Tubiello, F.N., Castaldi, S., Jackson, R.B., Alexe, M., Arora, V.K., Beerling, D.J., Bergamaschi, P., Blake, D.R., Zhu, Q., 2016. The Global Methane Budget: 2000–2012. *Earth System Science Data Discussions* 1–79.
- Saunio, M., Staver, A.R., Poulter, B., Bousquet, P., Canadell, J.G., Jackson, R.B., Raymond, P.A., Dlugokencky, E.J., Houweling, S., Patra, P.K., Ciais, P., Arora, V.K., Bastviken, D., Bergamaschi, P., Blake, D.R., Brailsford, G., Bruhwiler, L., Carlson, K. M., Carrol, M., Zhuang, Q., 2020. The global methane budget 2000–2017. *Earth System Science Data* 12 (3), 1561–1623.
- Segers, R., 1998. Methane production and methane consumption: a review of processes underlying wetland methane fluxes. *Biogeochemistry* 41, 23–51.
- Tramontana, G., Ichii, K., Camps-Valls, G., Tomelleri, E., Papale, D., 2015. Uncertainty analysis of gross primary production upscaling using random forests, remote sensing and eddy covariance data. *Remote Sensing of the Environment* 168, 360–373.
- Turner, A.J., Frankenberg, C., Kort, E.A., 2019. Interpreting contemporary trends in atmospheric methane. *Proceedings of the National Academy of Sciences of the United States of America* 116 (8), 2805–2813.
- Ueyama, M., Yazaki, T., Hirano, T., Futakuchi, Y., Okamura, M., 2020. Environmental controls on methane fluxes in a cool temperate bog. *Agricultural and Forest Meteorology* 281, 107852.
- Veldkamp, E., O'Brien, J.J., 2000. Calibration of a Frequency Domain Reflectometry Sensor for Humid Tropical Soils of Volcanic Origin. *Soil Science Society of America Journal* 64 (5), 1549–1553.
- Wager, S., Hastie, T., Efron, B., 2014. Confidence Intervals for Random Forests: The Jackknife and the Infinitesimal Jackknife. *Journal of Machine Learning Research* 15 (1), 1625–1651.
- Waldo, N.B., Hunt, B.K., Fadely, E.C., Moran, J.J., Neumann, R.B., 2019. Plant root exudates increase methane emissions through direct and indirect pathways. *Biogeochemistry* 145 (1), 213–234.
- Wong, G.X., Hirata, R., Hirano, T., Kiew, R., Aeries, E.B., Musin, K.K., Waili, J.W., Lo, K. S., Melling, L., 2020. How do land use practices affect methane emissions from tropical peat ecosystems? *Agricultural and Forest Meteorology* 282–283, 107869.
- Wutzler, T., Lucas-Moffat, A., Migliavacca, M., Knauer, J., Sickel, K., Sigut, L., Menzer, O., Reichstein, M., 2018. Basic and extensive post-processing of eddy covariance flux data with REdDyProc. *Biogeosciences*, Copernicus, 15. <https://doi.org/10.5194/bg-15-5015-2018>.
- Xu, T., Guo, Z., Liu, S., He, X., Meng, Y., Xu, Z., Xia, Y., Xiao, J., Zhang, Y., Ma, Y., Song, L., 2018. Evaluating different machine learning methods for upscaling evapotranspiration from flux towers to the regional scale. *Journal of Geophysical Research: Atmospheres* 123, 8674–8690. <https://doi.org/10.1029/2018JD028447>.
- Yu, X., Song, C., Sun, L., Wang, X., Shi, F., Cui, Q., Tan, W., 2017. Growing season methane emissions from a permafrost peatland of northeast China: Observations using open-path eddy covariance method. *Atmospheric Environment* 153, 135–149.
- Yu, Z., Staudhammer, C.L., Malone, S.L., Oberbauer, S.F., Zhao, J., Starr, G., 2022. Biophysical Factors and Water Dynamics Impact Methane Fluxes in Everglades Freshwater Marshes. *Ecosystems*. In review.
- Zhang, Z., Zimmermann, N.E., Stenke, A., Li, X., Hodson, E.L., Zhu, G., Huang, C., Poulter, B., 2017. Emerging role of wetland methane emissions in driving 21st century climate change. In: *Proceedings of the National Academy of Sciences of the United States of America*, 114, pp. 9647–9652.