**Near-term, iterative forecasting suggests high predictability of reservoir methane ebullition at weekly time scales**

Ryan P. McClure1\*, R. Quinn Thomas2, Mary E. Lofton1, Whitney M. Woelmer1 and Cayelan C. Carey1

1Department of Biological Sciences, Virginia Tech, Blacksburg, VA, 24061, USA

2Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA, 24061, USA

\*Corresponding author; CONTACT INFO

In preparation for submission as an <ARTICLE TYPE> to *Biogeosciences*

Keywords? Authorship statement? Check what other cover page info is needed?

**Abstract**

Near-term, iterative ecological forecasting with data assimilation has shown great promise for improving our understanding of ecological dynamics, but to date has rarely been applied to predicting future biogeochemical rates. Most ecological forecasting workflows use automated, high-frequency sensor data to iteratively update predictions in near-real time as new observations become available. However, many ecosystem-level biogeochemical processes cannot be quantified reliably with automated sensors and instead depend on manually collected samples and laboratory analysis, representing a major gap in our ability to predict some ecosystem-level processes. One important biogeochemical process that near-term, iterative ecological forecasts could potentially improve estimates of – yet is challenging to both monitor and predict – is freshwater methane (CH4) ebullition (CH4 bubbling from the sediments). We developed a forecasting and data assimilation workflow that generated real-time weekly CH4 ebullition rate forecasts and partitioned forecast uncertainty among different sources (parameters, model process, initial conditions, and driver data). Our near-term, iterative forecast workflow successfully predicted an ensemble of future CH4 ebullition rates that encompassed observations, and total forecast uncertainty decreased with each iterative forecasting cycle. Forecasts with data assimilation better replicated observations throughout a 19-week summer period than forecasts without data assimilation and a persistence null model. ADD 1-2 SENTENCES MORE ON WHAT EXACTLY WAS LEARNED ABOUT EBULLITION FROM FORECASTS Despite the high variability of CH4 ebullition emissions in freshwater ecosystems, our iterative forecasting workflow demonstrates that ebullition within a reservoir can be successfully forecasted and has high predictability on a weekly time scale.

**Introduction**

Near-term, iterative ecological forecasting with data assimilation is a novel approach to improve our understanding and quantification of ecosystem processes (Dietze et al. 2018). Here, we define an ecological forecast as a prediction with specified uncertainties of the future state of an ecosystem in the near-term (days to months) that is iteratively updated as new data become available (Clark et al. 2001, Dietze 2017*a*). Similarly, we define data assimilation as iteratively incorporating observations as they are collected into a forecast model by updating model states and parameter estimates (Luo et al. 2011). The iterative process of making forecasts, collecting observations to compare with the forecast, and then updating forecast models via data assimilation before making a new forecast is a forecast cycle (Dietze 2017*a*). Iterative ecological forecasting can substantially improve our understanding of ecosystem processes because it embodies the scientific method: our hypotheses about ecosystem functioning are instantiated in a model, which generates predictions of the future that are then tested with observed data when they become available (Dietze et al. 2018). Thus, near-term, iterative ecological forecasting creates a model-data feedback loop that evaluates how effectively a model predicts future ecosystem states, with the forecasts evolving as the ecosystem experiences different environmental conditions.

Because of the need to constantly update forecasts with new data as they become available, most sub-daily to weekly iterative ecological forecasts have been generated for variables that can be measured with high-frequency sensors providing data in real-time or near real-time, thereby enabling rapid forecast updating on minute to daily time scales (e.g., Dietze 2017*b*, Barachinni et al. 2020, Nickl et al. 2020). For example, Dietze (2017*b*) used automated, high-frequency eddy-flux covariance sensor estimates of carbon fluxes in a deciduous forest to iteratively forecast net ecosystem exchange. Data were collected by sensors on 30-minute intervals and transmitted wirelessly through automated cyberinfrastructure to update future forecasts. Similar automated workflows have been implemented to forecast other variables that can be monitored with sensors, such as chlorophyll a (Page et al. 2018), water temperature (Baracchini et al. 2020, Thomas et al. 2020), and discharge (Ouellet-Proulx et al. 2017). However, many ecosystem-level biogeochemical states and processes (e.g., nutrient and carbon concentrations and fluxes) are still quantified using manual data collection, not sensors (Marcé et al. 2016). Additionally, because technology for automated high-frequency monitoring of many ecosystem processes is not affordable or fully developed (e.g., Porter et al. 2009, Marcé et al. 2016), there is a major need to develop new forecasting workflows which integrate manually collected data streams for iterative forecasting and thus advance improve our understanding of the predictability of ecosystem dynamics.

One biogeochemical process that near-term, iterative ecological forecasts could potentially improve estimates of **–** yet is challenging to both monitor and predict **–** is freshwater methane (CH4) ebullition, or bubble fluxes of CH4 from organic-rich sediments to the waterbody's surface. Freshwater ecosystems emit large quantities of CH4 to the atmosphere (currently estimated between 117 and 212 Tg CH4 yr-1; Saunois et al. 2020). Among the different types of freshwater CH4 emissions, estimates of ebullition are considered one of the most uncertain in regional CH4 budgets because of ebullition's high spatial and temporal variability within and among freshwater ecosystems (Wik et al. 2016, Saunois et al. 2020). Spatially, ebullition rates can vary substantially at sites located just a few meters apart (DelSontro et al. 2011, Beaulieu et al. 2016, Wik et al. 2016, McClure et al. 2020, Linkhorst et al. 2020). Temporally, CH4 ebullition rates within the same waterbody can vary substantially over monthly to interannual time scales (Burke et al., 2019, Männistö et al. 2019, Linkhorst et al. 2020). Because of the high spatiotemporal variation of ebullition, it is extremely challenging to monitor and predict ebullition, which is important for upscaling estimates (Wik et al. 2016). As a result, freshwater ebullition remains unaccounted for in the current global CH4 budget (Saunois et al. 2020).

Because freshwater CH4 ebullition is a highly variable flux for global emission estimates (Maher et al. 2019, Saunois et al. 2020), new methods are needed to evaluate how predictable CH4 ebullition emission rates are over shorter time scales (Butman et al. 2018, Saunois et al. 2020). Near-term iterative forecasting with data assimilation may be a particularly valuable method for estimating and evaluating CH4 ebullition predictability because this approach has been shown to successfully predict other CH4 emission pathways. For example, Nickl et al. (2020) developed iterative forecasts of CH4 emissions to the atmosphere from coal mining regions in Poland that predicted CH4 plumes up to four days into the future at the 7 km2 scale. Their CH4 emission forecast workflow used Earth system models that were updated with remote sensing data and airborne sampling quantifying atmospheric concentrations of CH4 (Nickl et al. 2020).

In addition to improving CH4 ebullition predictions, near-term iterative forecasting using data assimilation also has the potential to improve our basic understanding of the drivers of CH4 ebullition dynamics. There are multiple variables which drive freshwater CH4 ebullition rates at different time scales, including water temperature (DelSontro et al. 2016, Davidson et al. 2018, McClure et al. 2020), hydrostatic pressure and water table depth (Harrison et al. 2016, Liu et al. 2019), primary productivity (West et al. 2016), water turbulence (Joyce and Jewel 2002), sediment organic matter and organic matter quality (Wik et al. 2018, Zhou et al. 2019), plant abundance (Davidson et al. 2019), and zooplankton bioturbation (Bezerra et al. 2020). The relative importance of these predictor variables can vary substantially over days (Maher et al. 2019, Bezerra et al. 2020), seasons (DelSontro et al. 2016, McClure et al. 2020), and years (Männistö et al. 2019, Linkhorst et al. 2020). Developing iterative, near-term forecasts in which new observations regularly update a forecasting model could provide useful information on how the importance of different predictors changes under different environmental conditions, as indicated by the evolution of the model parameters over time.

In contrast to atmospheric CH4 plume estimates spanning large regions (e.g., Nickl et al. 2020), freshwater CH4 ebullition rates are usually estimated individually for single ecosystems and rarely are monitored using continuous remote sensing or live-streaming sensor data, necessitating a simpler iterative forecasting workflow using manually collected data and empirical models. While there have been recent advances in automated CH4 ebullition rate sensors (Delwiche et al. 2017, Maher et al. 2019) and methods to estimate CH4 ebullition rates via remote sensing in freshwater ecosystems (Engram et al. 2020), the vast majority of lake and reservoir CH4 ebullition emissions are still measured with passive traps that are deployed in one waterbody, retrieved manually, and later analyzed on a gas chromatograph (GC) in the laboratory (e.g., Klapstein et al. 2014, Beaulieu et al. 2016, Männistö et al. 2019, McClure et al. 2020). Consequently, to advance iterative forecasting of manually sampled CH4 ebullition, new iterative forecasting and data assimilation workflows must be developed.

We developed the first (to our knowledge) iterative forecast workflow for near-term (1 to 16-day horizon) CH4 ebullition rates in a freshwater ecosystem, using manually collected ebullition data. We used sequential data assimilation (Luo et al. 2011, Dietze 2017*a*) to update an empirical ebullition model's states and parameters on each model time step before generating the next CH4 ebullition rate forecast. Here, sequential data assimilation is a method in which newly acquired observations are used to refit a model before generating a new forecast (Dietze 2017*a*).

We tested our forecast workflow in a small, eutrophic freshwater reservoir in southwestern Virginia, USA. Our objectives were: 1) develop an iterative, near-term forecasting workflow for manually collected CH4 ebullition rates; 2) assess how well near-term forecasts with sequential data assimilation could predict future CH4 ebullition in comparison to forecasts without data assimilation as well as a persistence null model; and 3)quantify how forecast uncertainty estimates changed over a summer season as a metric of CH4 ebullition rate predictability. Our goal was to use near-term forecasting to see if we can advance our understanding of CH4 ebullition, specifically to as to both quantify the predictability of ebullition as well as examine how its drivers change over time.

**Methods**

*Site description*

We developed near-term, iterative CH4 ebullition rate forecasts in Falling Creek Reservoir (FCR, Figure 1) in real-time. FCR is a small (0.119 km2), shallow (Zmax = 9.3 m), eutrophic, drinking water reservoir located in southwestern Virginia, USA (37.30°N, 79.84°W). FCR is owned and operated by the Western Virginia Water Authority as a drinking water supply and is located in a completely forested watershed (Gerling et al. 2016). During our forecasting period of 18 April to 1 December 2019, FCR did not experience large fluctuations in water level (Carey et al. 2020*a*).

*Forecast model development*

A previous study by McClure et al. (2020) quantified the rates and drivers of CH4 ebullition at multiple sites in FCR from May to October 2017. That study demonstrated that ≥60% of the total reservoir-wide CH4 ebullition was emitted from a shallow upstream transect in FCR during the ice-free period, which was positively related to sediment-water interface (SWI) temperatures (Figure 1; McClure et al. 2020). That analysis developed an auto-regressive (AR) time series model (Eqn. 1) that predicted weekly CH4 ebullition at the upstream transect (Supporting figures, Figure S1; McClure et al. 2020):

Eqn. 1

where (mg CH4 m-2 d-1) is the predicted CH4 ebullition rate at the upstream transect, (mg CH4 m-2 d-1) is the observed CH4 ebullition rate at the previous time step (the AR term), (°C) is the water temperature averaged from measurements at the SWI below the upstream transect between each sampling date, and is the stochastic error term; see McClure et al. 2020 and Supporting information A. Bo is the intercept term, B1 is the parameter governing the effect of autoregressive term, and B2 is the parameter governing the effect of SWI temperature. In the McClure et al. (2020) analysis, these three parameters were fixed throughout the May to October 2017 period, at Bo=XXX, B1=XXX, and B2=XX. As our focus in this 2017 analysis was on the development of an iterative forecast workflow, not model selection, we used this same AR model structure (but with evolving parameters) for forecasting CH4 ebullition in FCR during 1 July to 7 November 2019.

Because we were generating forecasts in real-time, we needed forecasted, not previously observed, driver data (i.e., future SWI temperature) to run our AR model to generate forecasts of future CH4 ebullition rates. We leveraged the existing Forecasting Lake And Reservoir Ecosystems (FLARE; Thomas et al. 2020) water temperature forecasting framework and infrastructure deployed in FCR to generate predictions of future water temperatures that became the driver data for our CH4 ebullition rate forecasts. FLARE is composed of three components: water quality and meteorological sensors that wirelessly stream data, a data assimilation algorithm that uses sensor observations to update the predictions and re-calibrate the model’s parameters daily, and an ensemble forecast algorithm that generates an ensemble of daily water temperature forecasts at 29 depths on 0.33 m increments from the surface to sediments at the deepest site of the reservoir (the dam site) over a 16-day time horizon (Thomas et al. 2020). The hydrodynamic-water temperature model in FLARE (General Lake Model or GLM; Hipsey et al. 2019) uses NOAA GEFS 16-day meteorological forecasts as driver data. For our study, we generated 210 water temperature forecast ensembles in each FLARE forecast run. We refer interested readers to Thomas et al. (2020) and <https://smartreservoir.org/user_guide/> for a detailed description of FLARE and tutorial on the FLARE forecasting system.

The CH4 ebullition forecasts used FLARE’s 210 ensemble member water temperature forecasts for 2.3, 2.6, and 3.0 m, which encompassed the range in the SWI depths below four upstream ebullition traps deployed at the upstream transect in 2019 (Figure 1). Because FLARE’s water temperature forecasts were generated for the deepest site of FCR at its dam and the CH4 ebullition forecasts needed to be generated for the upstream transect, we developed a linear scaling model to convert the dam water temperatures to the upstream transect SWI temperature, generating 210 forecasted SWI temperatures for the upstream transect in FCR (Supporting information A).

The forecasted mean SWI temperature derived from FLARE and the SWI scaling model (Eqn. S1) was then used as driver data for a CH4 ebullition forecasting model (Eqn. 2) that was based on Eqn. 1. This model used observed CH4 ebullition rates as the AR term and the forecasted mean SWI temperature as a driver:

Eqn. 2

where (mg CH4 m-2 d-1) is the forecasted CH4 ebullition rates at the upstream transect on the next CH4 ebullition sampling date, (mg CH4 m-2 d-1) is the observed CH4 ebullition rates at the time of data collection, (°C) is the forecasted mean SWI temperature, and is random process noise (i.e., , where is the standard deviation on the priors). As we had 210 separate forecasted SWI temperatures, we ran the model 210 times for each ensemble member and used the ensemble member output from this model to quantify and partition the relative contributions of parameter, model process, and driver data uncertainty to the total CH4 ebullition rate forecast uncertainty (see below for more information on uncertainty partitioning).

We used a weekly time-step in our forecasts. While the maximum possible CH4 ebullition forecast horizon was 16 days (because the SWI temperature forecasts from FLARE had this maximum horizon), we sampled the ebullition traps approximately every 7 days. We did this to accommodate manual data collection, which was ultimately governed by weather and logistical constraints that determined when field personnel were able to sample. Thus, the time horizon of each forecast cycle ranged between 5 and 9 days between 1 July and 7 November, with a median forecast horizon of 7 days. Following McClure et al. (2020), we used the mean SWI temperature forecasted for the upstream transect traps on the next sampling date (usually 7 days in the future) as driver data for the AR model. Thus, it is important to note that our forecast model time step was the same as our forecast horizon. Finally, we separated the study into a training period (27 May – 24 June) and then a forecasting period (1 July – 7 November, n=19 forecast cycles) that was evaluated against forecasts without data assimilation and a persistence null model to the observations (see forecast evaluation section below).

*Detailed iterative forecasting workflow description*

Our iterative forecast workflow had three stages: data collection, data assimilation, and forecast generation (Figure 2). As described above, this workflow was repeated weekly in real-time. New CH4 ebullition forecasts for the next sampling day were initiated upon the collection and analysis of current CH4 ebullition rates. For example, the first CH4 ebullition forecast in our study used the mean transect CH4 ebullition rate measured on 24 June as initial conditions (the term in Eqn. 2) and the mean forecasted SWI temperatures (the term in Eqn. 2) to predict the CH4 ebullition rates for 1 July, when we manually collected the next ebullition data. With the new observations, we updated both SWI temperature and CH4 ebullition model states and parameters via sequential data assimilation and then used the newest water temperature forecasts from FLARE to generate the next CH4 ebullition rate forecast for the next ebullition collection date (8 July).

The first stage of our forecast workflow encompassed manual data collection and generation of forecast driver data from FLARE, which defined the first day of each forecast cycle. We started the first stage on the same day we manually collected CH4 ebullition and generated daily water temperature forecasts from FLARE for a 16-day forecast horizon (Thomas et al. 2020). Additionally, we accessed the observed water temperature observations measured at the dam site every 10 minutes by thermistors at 2 and 3 m depth (NexSens Technology, Inc.; Fairborn, Ohio, USA; Carey et al. 2020*b*; Figure 1). When we collected the CH4 ebullition samples at the upstream transect, we also manually downloaded 10-minute SWI temperature data from four HOBO temperature loggers (HOBO Pendant Temperature/Light Data Logger; Bourne, MA, USA) deployed just above the SWI at the upstream transect. The SWI temperatures were aggregated and scaled to the transect level (Supporting information A).

The second part of our data collection entailed manually collecting ebullition gas bubbles using four passive ebullition traps (described by McClure et al. 2020) to generate the AR term in our forecast model (Eqn. 2). The collected ebullition gas from the traps was analyzed for its CH4 concentration using a Shimadzu Nexus-2030 Gas Chromatography – Flame Ionization Detector (GC-FID; Shimadzu Corporation; Kyoto, Japan) within 24 hours of collection. We determined the CH4 ebullition rate from each trap by multiplying the total volume of ebullition collected in each trap by the concentration of CH4 in the ebullition gas and then dividing by the cross-sectional area of each ebullition trap (0.26 m2) and the duration of time between sampling days (Supporting information B).

The data assimilation stage consisted of using the data collected to update our forecasting model parameters weekly via data assimilation that refit Eqn. 2 using all existing data. When new observations of CH4 ebullition rates, deep site water temperatures, and upstream transect SWI temperatures became available, they were used to update the SWI temperature and CH4 ebullition models in a Bayesian framework (following Dietze 2017*a*), which generated new parameters for the SWI temperature scaling model (Eqn. S1) and the CH4 ebullition model (, , , and ), respectively. The Bayesian framework used three Monte Carlo Markov Chain simulations with 50,000 samples to generate the new parameter estimates that would be used for the next CH4 ebullition rate forecasts. We assessed for convergence of the chains using Gelman-Rubin diagnostics (Gelman and Rubin 1992). This framework ensured both forecast models (Eqn. S1 and 2) were probabilistic and accounted for initial conditions, parameter, and model process uncertainty (Dietze 2017*b*). We used the ‘rjags’ package (Plummer 2019) within the R statistical environment for the Bayesian analyses (R Core Development Team 2020).

The final stage generated forecasts using the newest parameter estimates (, , and in Eqn. 2), a normal distribution with mean equal to the observed ebullition rates from all four traps (the term in Eqn. 2), and the forecasted mean upstream SWI temperatures scaled from FLARE’s dam site forecasts (the term in Eqn. 2) to generate the weekly CH4 ebullition rate forecasts with partitioned uncertainty (Figure 2). First, the SWI temperature scaling model (Eqn. S1) generated 210 ensembles of mean upstream SWI temperatures from FLARE’s 2.3, 2.6, and 3.0 m water temperature forecasts at the dam site. These forecasted SWI temperatures () were then used as driver data for the forecasting CH4 model???.

Finally, the forecast analyses and results reported herein are from historical FLARE forecasts (i.e., hindcasts). The entire forecast workflow executing the CH4 ebullition rate forecasts during the 2019 forecasting period in FCR can be downloaded from GitHub (<https://github.com/ryanmclake/CH4cast>) and run in the R statistical environment (R Core Development Team 2020), following (McClure et al. 2020,

<https://zenodo.org/badge/latestdoi/275238644>).

*Forecast evaluation*

We tested how well our forecasts performed against out-of-sample observations, the same model as Eqn. 2 without any data assimilation (i.e., using fixed parameters), and a persistence null model (Eqn. S3). More information on how the forecasts without data assimilation and the persistence null model (i.e., forecasts generated with just the AR term () and random process noise ()) were developed are in Supporting information C.

First, we used Nash-Sutcliffe efficiency (NSE) to evaluate the performance of the CH4 ebullition rate forecasts. We calculated the NSE coefficient (Nash and Sutcliffe 1970) as follows:

Eqn. 3

where was the mean of the 210 CH4 ebullition rate forecast ensembles, was the observed mean daily CH4 ebullition rates, and is the average of all the observed mean daily CH4 ebullition rates. NSE values range between -∞ to 1, where 1 indicates a perfect score (i.e., the model perfectly recreates observations), a value of 0 indicates that the model predictions are as accurate as the mean of the observations, and a value <0 indicates that the model performs worse than the observed mean (Nash and Sutcliffe 1970, Moriasi et al. 2007).

Second, we used posterior predictive loss (Dpl, Gelfand and Ghosh 1998) as an additional forecast evaluation metric. Dpl is a decision theory approach based on prediction rather than parameter estimations and can be applied to multiple model dimensions (Gelfand and Ghosh 1998, Hobbs and Hooten 2015). We calculated the Dpl as follows:

Eqn. 4

where was the mean of the 210 CH4 ebullition rate forecast ensembles and was the observed mean daily CH4 ebullition rates.

Third, we compared the performance of the forecast cycles with data assimilation to forecasts without data assimilation and the persistence null model to predict the total amount of forecasted CH4 ebullition summed across the entire forecasting period (1 July – 7 November) to determine the utility of the iterative forecasting workflow for upscaling to seasonal CH4 budgets. We determined the total CH4 upstream ebullition emissions in mg CH4 by summing the forecasted and observed mean daily CH4 ebullition rates over the forecast evaluation period and then multiplied the summed rates by the area of all four ebullition traps on the transect (1.04 m2) and the total number of days of the forecast evaluation period (n=143). We also calculated the percentage difference of what??? CH4 ? and if so, aggregated on what time scale? between the forecasts with data assimilation, without data assimilation, and the persistence null model (Eqn. S3) from the observations to evaluate which approach better predicted the total CH4 ebullition emissions from the upstream transect.

*Forecast variance and partitioning uncertainty*

We quantified the total weekly uncertainty in the CH4 ebullition forecasts with data assimilation, without data assimilation, and the persistence null model. Total forecast uncertainty is a valuable metric of predictability, and provides insight to the overall predictability of CH4 ebullition (Petchey et al. 2015, Dietze 2017*a*, Dietze 2017*b*). We also partitioned the relative contributions of different sources of uncertainty (parameter, model process, driver data, and initial conditions uncertainty; (Dietze 2017*b*) weekly throughout the forecasting period. Partitioning uncertainty sources can help inform forecasting systems where improvements can be made (Dietze 2017*b*, Petchey et al. 2015), and may be particularly useful for developing a forecasting workflow for biogeochemical processes that still rely on manual data collection, like CH4 ebullition (Carey et al. Inland Waters paper).

We used a One-At-a-Time Sensitivity (OATS) analysis to determine the relative contribution of the four uncertainty sources to total uncertainty during each forecast cycle (Figure 2). An OATS analysis holds all sources of uncertainty at their mean except for one, and then numerically evaluates the sensitivity of the forecast to that specific source of uncertainty (Dietze 2017*a*). First, we quantified the total CH4 ebullition rate forecast variance for all 210 forecast ensembles using the var() function in the R statistical environment R (R Core Development Team, 2020).

Next, we quantified the relative contribution of parameters (, , and in Eqn. 2), model process ( in Eqn. 2), driver data (), and initial conditions () uncertainty to the total CH4 ebullition rate forecast uncertainty. To account for initial conditions uncertainty, the standard error of the mean CH4 ebullition rates among all four ebullition traps () was used as initial conditions for the forecast, with uncertainty represented by a set of 210 random draws from a normal distribution which had a mean equal to the most recent observations and a standard error from the most recent observations. To account for parameter uncertainty, the ebullition forecast model sampled 210 different parameter values each from the three posterior distributions of , , and (Eqn. 2). To account for process uncertainty, random noise ( from Eqn. 2) was added to each ensemble member, drawn from a normal distribution with a mean of zero and a standard deviation equal to . DRIVER DATA UC??? Overall, this generated 210 ensembles of future CH4 ebullition rates for the next date we manually collected ebullition from FCR. The newest CH4 ebullition rate forecast ensembles were then archived on GitHub until the next manually observed CH4 ebullition rate was compared to the previously forecasted CH4 ebullition rate.

We applied the OATS analysis separately to all 19 forecast cycles that occurred after the forecast training period. Finally, the relative proportion of each source of uncertainty was determined by dividing the total variance of each isolated source of uncertainty by the sum of the variance of all uncertainty sources.

**Results**

*Near-term iterative CH4 ebullition forecasts*

We successfully developed near-term, iterative CH4 ebullition rate forecasts that predicted future CH4 ebullition rates at FCR’s upstream transect (Figure 3). Among all 19 forecast cycles that occurred between 1 July and 7 November 2019, the manually collected CH4 ebullition observations were always within the 95% quantile of the 210 forecast ensemble members (Figure 3a). Conversely, the forecasts without data assimilation and the persistence null model encompassed the observations over most of the forecasting period but had substantially higher maximum and minimum ranges among all 210 ensemble members (Figure 3b,c). The temperature scaling model (Eqn. S1) successfully predicted the SWI temperatures within 1.1°C of observations at the upstream transect throughout the forecasting period (Supporting figures; Figure S3 and S4).

*Observed upstream transect CH4 ebullition rates*

We observed high temporal variation in mean daily CH4 ebullition rates at FCR’s upstream transect during both the training and forecasting periods (Figure S2). Daily ebullition rates at the upstream increased from 0.8 mg CH4 m-2 d-1 to 14 mg CH4 m-2 d-1 from 27 May though 24 June (Figure S2). Between 24 June and 15 July, daily CH4 ebullition rates from the transect continued to increase until the peak maximum CH4 ebullition rate for the forecast period was observed on 15 July (49 mg CH4 m-2 d-1). After 15 July and until 16 October, the observed CH4 ebullition rates ranged between 13 and 42 mg CH4 m-2 d-1. After 16 October, CH4 ebullition rates dropped to ≤4.6 mg CH4 m-2 d-1 for the remainder of the forecasting period.

*CH4 forecasting model parameter evolution*

TOPIC SENTENCE ON HOW THE PARAMETERS EVOLVED OVER TIME For the first forecast on 1 July, the 210 ensemble estimates for , the parameter associated with the intercept term in Eqn. 2, had a range of -94 to 66 and mean of -1.9; by 15 July, the parameter range had decreased to -14 to 18 with a mean of -2.7. Similarly, the range for , the parameter associated with the autoregressive ebullition term in Eqn. 2, decreased from -11 to 20 with a mean of 0.15 in the 1 July forecast to a range of -1.0 to 2.0 with a mean of 0.14 on 15 July. The range for , the parameter associated with the SWI temperature term in Eqn. 2, decreased from -3.5 to 5.3 with a mean of 0.22 in the first forecast to -0.8 to 0.9 with a mean of 0.26 on 15 July. Finally, the range of associated with model process noise in the Eqn. 2, decreased from 0.1 to 60 with a mean of 0.57 in the first forecast to 0.1 to 4.1 with a mean of 0.31 on 15 July.

While the range around each of the parameters in Eqn. 2 did not substantially change after 15 July (Figure 6), the mean of each estimate continued to vary throughout the remainder of the forecasting period. The mean estimate of the intercept parameter (Eqn. 2) increased from -2.7 to -1.1 during July and early August and then remained between -1.0 and -1.5 between August and October before increasing to a maximum of -6.7 on 7 November. From 15 July through 16 October, the mean estimate of the autoregressive parameter was between 0.1 and 0.2 and then decreased to 0.006 and 0.015 on 23 and 30 Oct., respectively, before increasing to 0.15 on 7 November. The mean estimate for the SWI temperature parameter ranged between 0.26 and 0.17 throughout the summer before increasing to a maximum of 0.43 on 7 November. Finally, the mean estimate for ranged between 0.41 and 0.27 throughout the summer before increasing to 0.76 on 7 November.

*CH4 ebullition forecast performance evaluation*

The CH4 ebullition forecasts with data assimilation performed better than the forecasts without data assimilation and the persistence null model. During the forecasting period (1 July – 7 November), the forecasts with data assimilation better estimated the observed CH4 ebullition rates than the forecasts with data assimilation and the persistence null model (Table 1). The NSE and Dpl evaluation verified higher overall skill of the forecasts with data assimilation than the forecasts without data assimilation and the persistence null model (Table 1). The NSE of the CH4 ebullition forecasts with data assimilation was 0.83, much higher than the NSE of the forecasts without data assimilation (-0.02) and the persistence null model (-0.21, Table 1), suggesting that data assimilation substantially improved the forecasted mean CH4 ebullition rates. Additionally, the Dpl similarly highlighted the higher skill of the CH4 ebullition forecasts with data assimilation (Table 1). The Dpl of the forecasts with data assimilation was 1230 while the forecasts without data assimilation and the persistence null model was 6630 and 3660, respectively, also indicating that data assimilation substantially improved the forecast skill during the forecasting period.

In addition to the successful NSE and Dpl of the forecasts with data assimilation, the seasonally summed CH4 from the forecasts with data assimilation were closer to the observations than the seasonally summed CH4 from forecasts without data assimilation and the persistence null model (Table 1). The mean seasonally summed observed CH4 ebullition over the forecasting period (143 days) was 60000 g CH4, while the forecasts with data assimilation predicted a seasonal sum of 65000 g CH4, an 8% overestimate. In comparison, the forecasts without data assimilation estimated a total of 52000 g CH4, and the null persistence model estimated 89000 g CH4, representing a 14% underestimate and 39% overestimate, respectively.

*Forecast uncertainty*

With each iterative forecast cycle, we observed a substantial reduction in the uncertainty of the CH4 ebullition rate forecasts with data assimilation, indicating relatively high predictability in CH4 ebullition throughout most of the study period? (Figure 4). On 1 July, the total forecast variance was 4.1 mg CH4 m-2 d-1 and within one forecasting cycle, the total variance in the forecast had decreased to 0.5 mg CH4 m-2 d-1 on 15 July (Figure 5). Between 15 July and 30 October, the total forecast variance decreased to ≤0.4 mg CH4 m-2 d-1 for each cycle and then increased to 0.6 mg CH4 m-2 d-1 on the final forecast that was generated for 7 November. Conversely, the forecasts without data assimilation did not reduce in uncertainty during the 19 forecast cycles and ranged from a minimum variance of 1.5 mg CH4 m-2 d-1 and a maximum variance of 20 mg CH4 m-2 d-1 (Figure 4). Finally, the persistence null model exhibited a consistent trend during the forecast period, starting at 1.2 mg CH4 m-2 d-1 on the 1 July forecast and steadily increasing to a maximum of 2.6 mg CH4 m-2 d-1 on 16 October.

As the forecast without data assimilation variance declined within the first three forecast cycles (Figure 4), there were also substantial shifts in the relative contributions of initial conditions, parameter, model process, and driver data uncertainty to the forecast variance (Figure 5). Between 1 July and 15 July, the contribution of parameter and model process uncertainty were the two largest contributors to total forecast variance, contributing between 24-60% and 33-51%, respectively. After 15 July, parameter uncertainty decreased and was ≤15% until 23 October. During this same time period (15 July – 23 October), model process uncertainty remained elevated (44-60%), driver data uncertainty increased and ranged between 22-42%, and initial conditions uncertainty increased and ranged between 5-20%. On 23 October, parameter and driver data uncertainty increased to 25% and 23%, respectively, while initial conditions decreased to 6% and model process was 40%. Finally, for the last two forecasts on 30 October and 7 November, the contribution from parameter uncertainty decreased to ≤18%, model process and driver data uncertainty were ≤54 and ≤31%, respectively, and initial conditions uncertainty was negligible (<1%, Figure 5).

<this is good text that should be reworked into the paragraph above> Finally, the decrease in the total variance of the CH4 ebullition rate forecasts with data assimilation (Figure 4) and the substantial shift in the contribution of parameter uncertainty (Figure 5) between 1 and 15 July corresponded with a substantial decrease in the range of the parameters in the CH4 ebullition rate forecasts with data assimilation throughout the forecasting period (Eqn. 2; Figure 6).

**Discussion**

Our CH4 ebullition forecasting workflow demonstrates that near-term, iterative ecological forecasts with data assimilation are manageable without automated sensor systems and provides evidence that CH4 ebullition rates can be highly predictable on the weekly scale using this approach. Because many biogeochemical states and fluxes are still quantified using manual data collection in both aquatic and terrestrial ecosystems, creating near-term, iterative ecological forecasting workflows that use manually collected observations for data assimilation may improve our understanding of their predictability. Our case study highlights how data assimilation at each forecast cycle substantially improved the performance ofCH4 ebullition rate forecasts over time (Figures 3-6). The forecasts with data assimilation performed better than forecasts without data assimilation and a persistence null model (Table 1, Figure 3 and 4). This suggests that iteratively forecasting CH4 ebullition with data assimilation in freshwater ecosystems has the potential to rapidly improve model performance with time (Figure 5) and respond to changing environmental conditions (Figure 6). The improvement in our CH4 ebullition rate forecasts with data assimilation over those without data assimilation and the persistence null model (Table 1, Figures 3 and 4), the partitioned uncertainty (Figures 5), and the changes in parameters over the study period altogether indicate that near-term, iterative forecasting can improve our quantification and understanding of this biogeochemical process.

*Iterative forecasting suggests CH4 ebullition may be highly predictable for some time scales*

Contrary to the expectation that CH4 ebullition is difficult to predict at varying spatial and temporal scales (Deemer et al. 2016, Wik et al. 2016, Maher et al. 2019, Saunois et al. 2020), our near-term, iterative ecological forecasts with data assimilation showed that CH4 ebullition at this specific transect in FCR waspredictable on the weekly time scale. While we note that our forecasts were for just one site of a small reservoir, other studies have also recently developed models that were able to successfully predict ebullition on different temporal and spatial scales, from the mesocosm (Davidson et al. 2018) to whole-ecosystem (Nachimuthu et al. 2016, Walter-Anthony 2013), and on temporal scales varying from hours (Bezerra et al. 2020) to years (Männistö et al. 2019). While these studies generated prediction models after data were collected, not updating their models and predictions in real-time, this previous work in combination with ours suggests that freshwater CH4 ebullition may be more predictable as a biogeochemical process than previously thought. As a start, developing models that include additional predictor variables (e.g., chlorophyll *a*, nutrients, pressure; DelSontro et al. 2016, Harrison et al., 2016, West et al. 2016, Davidson et al. 2018) and increase in complexity over our autoregressive forecast model (Eqn. 2) are likely needed to predict CH4 ebullition at other spatial and temporal scales. However, our study demonstrates high weekly predictability of CH4 ebullition at least at this one site and the utility of near-term iterative ecological forecasting as an approach for quantifying ebullition predictability.

The substantial reduction in total forecast variance over the forecasting period further demonstrates the value of the iterative forecasting for improving our understanding of CH4 ebullition predictability. Following Petchey et al. (2015), we assessed predictability by using the metric of total forecast variance over the forecasting period (Figure 4). Other studies that followed this approach found that total forecast variance can decrease after a relatively short training period if the chosen prediction model is appropriate (Luo et al. 2011, Petchey et al. 2015, Dietze 2017*b*). The rapid decrease in total forecast variance between 1 and 15 July (Figure 5) after a one-month training period suggests that our AR model with forecasted SWI temperature as a covariate was appropriate for forecasting CH4 ebullition weekly at FCR. For other near-term, iterative forecasting applications of CH4 ebullition, rapid decreases in total forecast variance can be used to check if a chosen predictive model is appropriate (following Dietze 2017b).

*Data assimilation improves forecast skill*

The use of data assimilation to update the forecast model (Eqn. 2) was critically important for improving the skill of the forecasted CH4 ebullition rates in comparison to forecasts without data assimilation and the persistence null model (Eqn. S3). Sequential data assimilation improved the mean forecast of the ensemble members relative to observations over time; in contrast, the forecasts without data assimilation and the persistence null model did not predict CH4 ebullition rates as well throughout the summer (Table 1, Figure 3). The poor performance by the forecasts without data assimilation and the persistence null model underscores the value of iterative forecasting with data assimilation, which can update model parameters at each forecast cycle to account for substantial within-season variability in CH4 ebullition rates (Linkhorst et al. 2020).

SECOND PARAGAPH HERE ON IMPORTANCE OF DA

*NEW HEADER HERE ON WHAT WE LEARNED ABOUT CH4*

Importantly, the evolution of model parameters throughout the forecasting period (Figure 6) highlights the ability of our forecasting framework to improve our understanding of how CH4 ebullition rates can respond to seasonal environmental conditions. For example, an increase in the observed ebullition rates (Figure S2) between 1 July and 8 July was concomitant with an increase in the SWI temperature parameter and decrease in the AR parameter (Figure 6) as the reservoir warmed, indicating that SWI temperature is likely an important predictor of CH4 ebullition rates during periods of sediment warming periods while previously observed ebullition rates are less so. However, between 15 and 22 July, the SWI temperature parameter estimate decreased and AR term parameter estimate increased and then remained stationary until 11 October. This shift between the parameter estimates followed by the stationary trend were concomitant with the seasonal increase and decrease in the observed CH4 ebullition rates (Figure S2), indicating that the AR term in our model was the important forecast variable for CH4 ebullition rates throughout the middle of the forecast period when seasonal trends are observed. Finally, after 11 October, the SWI temperature parameter substantially increased again while the AR term parameter decreased and then increased to a similar value that was observed on the first 1 July forecast. This dynamic change in both parameter estimates highlight how SWI temperature is likely most important for predicting CH4 ebullition in spring and late autumn, when there are large changes occurring in SWI temperature week to week as water at the sediment water interface cools during fall mixing. Conversely, the AR term was most important during mid-summer when there are consistent weekly autocorrelated changes in ebullition.

ADD PARAGRAPH HERE ON HOW THE RELATIONSHIP OF CH4 AND TEMP THAT THE MODEL SHOWS VIA PARAMETER EVOLUTION CHANGES OUR UNDERSTANDING OF THE DRIVERS OF CH4 EBULLITION

*Uncertainty partitioning provides insight on forecasting system performance*

Partitioning different sources of uncertainty (initial conditions, parameter, model process, and driver data; Dietze 2017*b*) in our CH4 ebullition rate forecasts provides valuable information on how best to improve future forecasts that operate on a similar time scale (Figure 5). Process uncertainty was the largest contributor to total forecast uncertainty during the entire forecasting period, followed by driver data uncertainty, then initial conditions uncertainty, and then parameter uncertainty (Figure 5). Process uncertainty encompasses the ability of the model to recreate ecosystem dynamics (Dietze 2017*b*). The dominance of model process uncertainty during the middle of the forecasting period (when the SWI temperature parameter estimate was lowest and the AR term parameter estimate was highest) suggests that additional environmental variables that have been shown to be important predictors of CH4 ebullition in other studies (e.g., chlorophyll *a*, sediment organic quality, water-level/hydrostatic pressure; West et al. 2016, Beaulieu et al. 2018, Wik et al. 2018) into our model would potentially reduce the model process uncertainty and improve the CH4 ebullition forecasts. However, it is important to also recognize that increasing the number of forecasted predictors into our model may reduce model process uncertainty but also increase uncertainty in other areas like driver data and parameter uncertainty, especially if there is a lot of uncertainty in forecasted the driver variables which can be challenging to generate and likely increase total uncertainty for the CH4 ebullition forecast.

We also observed an increase in the relative contribution of driver data uncertainty in the forecasting period (Figure 5). The driver data uncertainty in the CH4 ebullition forecasts was due to both uncertainty in the forecasted water temperatures from FLARE and the SWI temperature scaling model (Figure S5), with most of the uncertainty likely from FLARE’s forecasted water temperatures. Our results are similar to other work who found that driver data uncertainty can contribute a large part of total forecast uncertainty. For example, Thomas et al. (2020) found that meteorological forecasts greater than 2 days into the future contributed the largest proportion to total uncertainty in the water temperature forecasts in the reservoir’s surface waters in the FLARE forecasting system. Similarly, Dietze (2017*b*) found that meteorological driver data uncertainty dominated forest net ecosystem exchange forecasts at longer forecast horizons. The larger driver data uncertainty in our CH4 ebullition rate forecasts is likely a result of propagated meteorological uncertainty from the FLARE forecasting system into the SWI temperature forecasts. Additionally, developing CH4 ebullition forecasts over shorter forecast horizons (e.g. < 2-days) may decrease driver data uncertainty overall but likely increase uncertainty in other areas.

*Forecasting system limitations and improvements*

There are limitations of our near-term, iterative CH4 ebullition rate forecasts. As noted above, there are many other potential predictors of CH4 ebullition and other model structures that could be integrated into our iterative workflow for forecasting CH4 ebullition in other freshwater ecosystems at different temporal and spatial scales. Because our forecasts used only one statistical model developed for FCR in earlier work (McClure et al. 2020), we were unable to quantify model selection uncertainty by using different models in our forecasting workflow (e.g., process-based ebullition models; Peltola et al. 2018), which is an important step for future work. Testing different CH4 ebullition models at the weekly time scale in FCR and varying temporal and spatial scales may further improve our understanding of this biogeochemical process.

Additional improvements to our forecasting system include the addition of automated sensors and other statistical methods into the forecasting workflow. Recent technological improvements in automated ebullition sensors hold great potential for advancing the future of CH4 ebullition forecasting at daily or subdaily scales (e.g., Varadharajan et al. 2010, Delwiche et al. 2015, Maher at al. 2019). Similarly, alternate data assimilation methods (e.g., ensemble Kalman filters (e.g., Page et al. 2018, Thomas et al. 2020) or particle filters (Rowe et al. 2016)), could be useful for future applications of this work that uses more complex models.

*Scaling CH4 ebullition forecasting to other sites for other manual-collection field programs*

Although we designed near-term, iterative CH4 ebullition forecasts specifically for FCR at weekly forecast horizons, our forecasting iterative workflow could be adapted for predicting CH4 ebullition in other freshwater ecosystems. Here, we provide a list of minimum requirements to begin forecasting CH4 ebullition rates from other lakes and reservoirs:

**1) A CH4 ebullition field monitoring program.** The foundation of near-term, iterative forecasting is a coupled model-data feedback loop (Dietze 2017*a*). Thus, developing a field monitoring program that routinely samples ebullition from the waterbody of interest is necessary to forecast ebullition. Initial observations are needed to develop and train a forecast model, as we did in this study using observations collected in summer 2017 by McClure et al. (2020) to develop a model that was trained for summer 2019 forecasts using 27 May to 24 June 2019 data. Additionally, data from a monitoring program can help prioritize CH4 ebullition spatial hotspots for modeling and forecasting, as we did in this study by focusing on the upstream transect site that emitted most of the reservoir-wide CH4 ebullition in 2017. Finally, it is ideal if the monitoring data collection can be collected *within* the maximum time horizon of the forecast driver data, to enable forecast evaluation. In our case, the NOAA GEFS maximum horizon is 16 days, so we had a 16-day window from when the forecast was generated to collect new observations to evaluate forecast skill.

**2) A model to generate future predictions of CH4 ebullition that can be updated as observations become available.** Models are the second essential component of the model-data feedback loop in an iterative forecast workflow. While we used a relatively simple statistical model for this study (McClure et al. 2020), any model type can be applied as long as its drivers can also be forecasted, and the model can iteratively update as new observations become available. Multiple model types can be used: e.g., time series models, process-based models, and even neural networks (Abbasi et al. 2020) can all be used to develop CH4 ebullition forecasts.

**3) Forecasted driver data.** Depending on model structure, real-time ebullition forecasts may need forecasted driver data or an autoregressive term to predict CH4 ebullition at the next time step. This, indeed, may be a hurdle to begin forecasting CH4 ebullition. In our case study, we had access to FLARE’s existing infrastructure, which generated daily forecasts of water temperature that were scaled to become driver data for our model (Eqn. 2). Although FLARE can be deployed to other lakes and reservoirs (Thomas et al. 2020), other waterbodies may not have the cyberinfrastructure or temperature sensors needed to run that forecasting system. However, in lieu of having a forecasting system like FLARE, weather forecasts by themselves can be integrated into a CH4 ebullition forecast workflow if the ebullition prediction model is driven by meteorological variables like barometric pressure or shortwave radiation (Tokida et al. 2007, Wik et al. 2014, Peltola et al. 2018). For example, Tokida et al. (2007) showed how falling atmospheric pressure triggered substantial increases in CH4 ebullition from natural wetlands and Wik et al. (2014) associated seasonal CH4 ebullitive fluxes from thermokarst lakes with shortwave radiation. Thus, a forecast workflow for CH4 ebullition is possible, even if the site has no pre-existing forecasting infrastructure.

*Conclusions*

We were able generate successful forecasts of CH4 ebullition rates, a highly variable biogeochemical process in freshwater ecosystems, using using forecasted SWI temperature data and manually collected CH4 ebullition rates (Figure 3a). Sequential data assimilation in each forecast cycle substantially improved model performance and the forecasts performed better than forecasts without data assimilation and a persistence null model (Table 1, Figure 3 and 4). Finally, data assimilation decreased total forecast variance and uncertainty partitioning showed how the drivers of total forecast uncertainty changed during the forecast period, which improved and changed the parameter estimates through time (Figures 4 – 6). 2 SENTENCES ON WHAT YOU LEARNED The success of our case study suggests that iterative ecological forecasting with data assimilation may be a valuable approach for better quantifying biogeochemical processes that are difficult to monitor with automated sensors simultaneously while providing insight to their predictability.

**Acknowledgements**

*Data and Code Availability*

All data used in this manuscript will be available in the Environmental Data Initiative repository. The code, directions, and data to run the CH4 ebullition rate forecasts are available on GitHub (McClure et al. 2020): <https://github.com/ryanmclake/CH4cast>.

*Author Contributions*

RPM and CCC developed the original research idea.RPM developed the forecasting system framework with substantial input on structure and code development from RQT. WMW and MEL contributed to code development, field data collection, and model testing. RPM wrote the manuscript with CCC and RQT. All authors provided feedback and approved the final version. The authors declare no conflicts of interest.

*Funding and Support*

This work was supported by the U.S. National Science Foundation (CNS-1737424, DEB-1753639, DBI-1933016, DEB-1926050); the Virginia Tech Global Change Center; and the Fralin Life Sciences Institute. We thank the Smart Reservoir and FCR Carbon teams for their helpful feedback during this project and the Western Virginia Water Authority for their long-term support and access to the reservoir. We thank Ashley Mickens for her critical help with data collection and GC-FID analyses. Additionally, we thank Bethany Bookout, Alexandria Hounshell, Abby Lewis, Heather Wander, Rachael Corrigan, James Maze, Dexter Howard, Jacob Wynne, and Nick Hammond for their help with field data collection and Vahid Daneshmand and Renato Figueiredo for their development and maintenance of the FLARE forecast system. We also thank Paul Hanson, Erin Hotchkiss, Madeline Schreiber, and GLEON colleagues for their feedback on manuscript development. Finally, as always, we thank Bobbie Niederlehner for helping with the GC-FID and laboratory analysis methods development.

**References**

Abbasi, T., Luithui, C., & Abbasi, S. A. (2020). A Model to Forecast Methane Emissions from

Topical and Subtropical Reservoirs on the Basis of Artificial Neural Networks. Water, 12(1), 145.

Baracchini, T., Wüest, A., & Bouffard, D. Meteolakes: An operational online three-dimensional

forecasting platform for lake hydrodynamics. (2020). Water Research, 172, 115529.

Beaulieu, J. J., McManus, M. G., & Nietch, C. T. (2016). Estimates of reservoir methane

emissions based on a spatially balanced probabilistic‐survey. Limnology and Oceanography, 61, 27-40.

Beaulieu, J. J., Balz, D. A., Birchfield, M. K., Harrison, J. A., Nietch, C. T., Platz, M. C., et al.

(2018). Effects of an experimental water-level drawdown on methane emissions from a eutrophic reservoir. Ecosystems, 21(4), 657-674.

Bezerra, M. P., McGinnis, D. F., Bezerra-Neto, J. F., & Barbosa, F. A. R. Is it stochastic?

Chaoborus larvae bioturbation likely affect the timing of daily methane (CH4) ebullitive flux in a tropical reservoir. (2020). Hydrobiologia, 1-18.

Butman, D. E., Streigel, R. G., Stackpoole, S. M., Del Giorgio, P., Prairie, Y., Pilcher, D.,

Raymond, P., Paz Pellet, F., Alcocer, J. (2018). Inland Waters: Second State of the Carbon Cycle Report (SOCCR2): A Sustained Assessment Report, 28 pp.

Carey, C.C., A.G. Hounshell, M.E. Lofton, B.J. Bookout, R.S. Corrigan, A.B. Gerling, R.P.

McClure, and W.M. Woelmer. (2020a). Discharge time series for the primary inflow tributary entering Falling Creek Reservoir, Vinton, Virginia, USA 2013-2020 Environmental Data Initiative.

Carey, C.C., B.J. Bookout, W.M. Woelmer, and A.S. Lewis. (2020b). Time series of high-

frequency sensor data measuring water temperature, dissolved oxygen, conductivity, specific conductivity, total dissolved solids, chlorophyll a, phycocyanin, and fluorescent dissolved organic matter at discrete depths in Falling Creek Reservoir, Virginia, USA in 2018-2019 ver 4. Environmental Data Initiative.

Clark, J. S., Carpenter, S. R., Barber, M., Collins, S., Dobson, A., Foley, J. A., Lodge, D. M.,

Pascual, M., Pielke, R., Pizer, M., Pringle, C., Reid, W. V., Rose, K. A., Sala, O., Schlesinger, W. H., Wall, D. H., Wear, D. (2001). Ecological forecasts: an emerging imperative. Science, 293, 657-660.

Davidson, T. A., Audet, J., Jeppesen, E., Landkildehus, F., Lauridsen, T. L., Søndergaard, M., &

Syväranta, J. (2018). Synergy between nutrients and warming enhances methane ebullition from experimental lakes. Nat. Clim. Change, 8(2), 156-160.

Deemer, B. R., Harrison, J. A., Li, S., Beaulieu, J. J., DelSontro, T., Barros, N., et al. (2016).

Greenhouse gas emissions from reservoir water surfaces: a new global synthesis. BioScience, 66, 949-964.

DelSontro, T., Boutet, L., St-Pierre, A., del Giorgio, P. A., & Prairie, Y. T. (2016). Methane

ebullition and diffusion from northern ponds and lakes regulated by the interaction between temperature and system productivity. Limnol. Oceanogr., 61, 62-77.

Delwiche, K., & Hemond, H. F. An enhanced bubble size sensor for long‐term ebullition studies.

(2017). Limnology and Oceanography: Methods, 15, 821-835.

Dietze M. C. (2017a). Ecological Forecasting. Princeton University Press. 288 pp. ISBN

9781400885459

Dietze, M. C. Prediction in ecology: a first-principles framework. (2017b). Ecological

Applications, 27: 2048-2060.

Dietze, M. C., Fox, A., Beck-Johnson, L. M., Betancourt, J. L., Hooten, M. B., Jarnevich, C. S.,

Keitt, T. H., Kenney, M. A., Laney, C. M., Larsen, L. G., et al. (2018). Iterative near-term ecological forecasting: Needs, opportunities, and challenges. PNAS, 115, 1424-1432.

Engram, M., Walter Anthony, K. M., Sachs, T. et al. (2020). Remote sensing northern lake

methane ebullition. Nat. Clim. Chang. 10, 511–517, doi:10.1038/s41558-020-0762-8.

Gerling, A. B., Munger, Z. W., Doubek, J. P., Hamre, K. D., Gantzer, P. A., Little, J. C., and

Carey, C. C. (2020). Whole-catchment manipulations of internal and external loading reveal the sensitivity of a century-old reservoir to hypoxia. Ecosystems, 19, 555-571.

Harrison, J. A., Deemer, B. R., Birchfield, M. K., & O’Malley, M. T. (2017). Reservoir water-

level drawdowns accelerate and amplify methane emission. Environmental science & technology, 51, 1267-1277.

Hipsey, M. R., Bruce, L. C., Boon, C., Busch, B., Carey, C. C., Hamilton, D. P., et al. (2019). A

General Lake Model (GLM 3.0) for linking with high-frequency sensor data from the Global Lake Ecological Observatory Network (GLEON). Geoscientific Model Development, 12(1), 473-523.

Joyce, J., & Jewell, P. W. (2003). Physical controls on methane ebullition from reservoirs and

lakes. Environ. Eng. Geosci., 9(2), 167-178.

Klapstein, S. J., Turetsky, M. R., McGuire, A. D., Harden, J. W., Czimczik, C. I., Xu, X., et al.

(2014). Controls on methane released through ebullition in peatlands affected by

permafrost degradation. Journal of Geophysical Research: Biogeosciences, 119, 418-431.

Linkhorst, A., Hiller, C., DelSontro, T., M. Azevedo, G., Barros, N., Mendonça, R., & Sobek, S.

(2020). Comparing methane ebullition variability across space and time in a Brazilian reservoir. Limnology and Oceanography, 65(7), 1623-1634.

Liu, L., Sotiri, K., Dück, Y. et al. (2019). The control of sediment gas accumulation on spatial

distribution of ebullition in Lake Kinneret. Geo-Mar Lett 40, 453-466.

Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., et al. (2011). Ecological forecasting

and data assimilation in a data‐rich era. Ecological Applications, 21, 1429-1442.

Maavara, T., Chen, Q., Van Meter, K., Brown, L. E., Zhang, J., Ni, J., & Zarfl, C. (2020). River

dam impacts on biogeochemical cycling. Nature Reviews Earth & Environment, 1-14.

Maher, D. T., Drexl, M., Tait, D. R., Johnston, S. G., & Jeffrey, L. C. (2019). iAMES: An

inexpensive, Automated Methane Ebullition Sensor. Environ. Sci. Tech., 53, 6420-6426.

Männistö, E., Korrensalo, A., Alekseychik, P., Mammarella, I., Peltola, O., Vesala, T., &

Tuittila, E. S. (2019). Multi-year methane ebullition measurements from water and bare peat surfaces of a patterned boreal bog. Biogeosciences, 16, 2409-2421.

Marcé, R., George, G., Buscarinu, P., Deidda, M., Dunalska, J., de Eyto, E., et al. (2016).

Automatic high frequency monitoring for improved lake and reservoir management. Environmental Science & Technology, 50(20), 10780-10794.

Mendonça, R., Müller, R. A., Clow, D., Verpoorter, C., Raymond, P., Tranvik, L. J., & Sobek, S.

(2017). Organic carbon burial in global lakes and reservoirs. Nat. Comm., 8(1), 1-7.

McClure, R. P., Lofton, M. E., Chen, S., Krueger, K. M., Little, J. C., Carey, C. C. (2020). The

magnitude and drivers of methane ebullition and diffusion vary on a longitudinal gradient in a small freshwater reservoir. JGR: Biogeosciences, 125, e2019JG005205.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L.

(2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE, 50, 885-900.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I-

A discussion of principles. Journal of Hydrology, 10, 282-290.

Natchimuthu, S., Sundgren, I., Gålfalk, M., Klemedtsson, L., Crill, P., Danielsson, Å., &

Bastviken, D. (2016). Spatio-temporal variability of lake CH4 fluxes and its influence on annual whole lake emission estimates. Limnol. Oceanogr., 61, 13-26.

Nickl, A. L., Mertens, M., Roiger, A., Fix, A., Amediek, A., Fiehn, A., et al. (2020). Hindcasting

and forecasting of regional methane from coal mine emissions in the Upper Silesian Coal Basin using the online nested global regional chemistry–climate model MECO (n)(MESSy v2. 53). Geoscientific Model Development, 13(4), 1925-1943.

Ouellet-Proulx, S., Chiadjeu, O. C., Boucher, M. A., & St-Hilaire, A. (2017). Assimilation of

water temperature and discharge data for ensemble water temperature forecasting. Journal of Hydrology, 554, 342-359.

Page, T., Smith, P. J., Beven, K. J., Jones, I. D., Elliott, J. A., Maberly, S. C., Mackay, E. B., De

Ville, M., & Feuchtmayr, H. (2018). Adaptive forecasting of phytoplankton communities. Water Research, 134, 74-85.

Peltola, O., Raivonen, M., Li, X., and Vesala, T. (2018). Comparison of methane ebullition

modelling approaches used in terrestrial wetland models. Biogeosciences, 15, 937.

Petchey, O. L., Pontarp, M., Massie, T. M., Kéfi, S., Ozgul, A., Weilenmann, M., et al.

Z. (2015). The ecological forecast horizon, and examples of its uses and determinants. Ecology letters, 18(7), 597-611.

Porter, J. H., Nagy, E., Kratz, T. K., Hanson, P., Collins, S. L., & Arzberger, P. (2009). New

eyes on the world: advanced sensors for ecology. BioScience, 59(5), 385-397.

Rowe, M. D., Anderson, E. J., Wynne, T. T., Stumpf, R. P., Fanslow, D. L., Kijanka, K., et al.

(2016). Vertical distribution of buoyant Microcystis blooms in a Lagrangian particle tracking model for short‐term forecasts in Lake Erie. Journal of Geophysical Research: Oceans, 121, 5296-5314.

Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., Raymond,

P. A., Dlugokencky, E. J., Houweling, S., Patra, P. K., et al. (2020). The Global Methane

Budget 2000–2017. Earth Syst. Sci. Data Discuss.

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram.

Journal of Geophysical Research: Atmospheres, 106, 7183-7192.

Thomas, R. Q., Figueiredo, R. J., Daneshmand, V., Bookout, B. J., Puckett, L. K., and Carey, C.

C. (2020). A near-term iterative forecasting system successfully predicts reservoir hydrodynamics and partitions uncertainty in real time. bioRxiv.

Tokida, T., Miyazaki, T., Mizoguchi, M., Nagata, O., Takakai, F., Kagemoto, A., and Hatano, R.

(2007), Falling atmospheric pressure as a trigger for methane ebullition from peatland, Global Biogeochem. Cycles, 21, GB2003.

West, W. E., Creamer, K. P., & Jones, S. E. (2016). Productivity and depth regulate lake

contributions to atmospheric methane. Limnol. Oceanogr., 61 (S1), S51-S61.

Wik, M., Thornton, B. F., Bastviken, D., MacIntyre, S., Varner, R. K., & Crill, P. M. (2014).

Energy input is primary controller of methane bubbling in subarctic lakes. Geophysical Research Letters, 41, 555– 560.

Wik, M., Thornton, B. F., Bastviken, D., Uhlbäck, J., & Crill, P. M. (2016). Biased sampling of

methane release from northern lakes: A problem for extrapolation. Geophys. Res. Lett., 43(3), 1256-1262.

Wik, M., Johnson, J. E., Crill, P. M., DeStasio, J. P., Erickson, L., Halloran, M. J., et al. (2018).

Sediment characteristics and methane ebullition in three subarctic lakes. J. Geophys. Res. Biogeosci., 123, 2399-2411.

Zhou, Y., Zhou, L., Zhang, Y., de Souza, J. G., Podgorski, D. C., Spencer, R. G., et al. (2019).

Autochthonous dissolved organic matter potentially fuels methane ebullition from experimental lakes. Water Res., 115048.

Table 1. Nash-Sutcliffe efficiency (NSE), seasonally-summed CH4 ebullition emissions, and the percent difference of the forecasts with data assimilation, forecasts without data assimilation, and persistence null model to the observations during the forecasting period (1 July – 7 November 2019).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **NSE** | **Dpl** | **Total seasonal emissions (g CH4)** | **% difference from observations** |
| Observations | - | - | 60000 | - |
| Forecasts with data assimilation | 0.83 | 1230 | 65000 | 8 |
| Forecasts without data assimilation | -0.02 | 6629 | 52000 | 14 |
| Persistence null model | -0.21 | 3660 | 89000 | 39 |

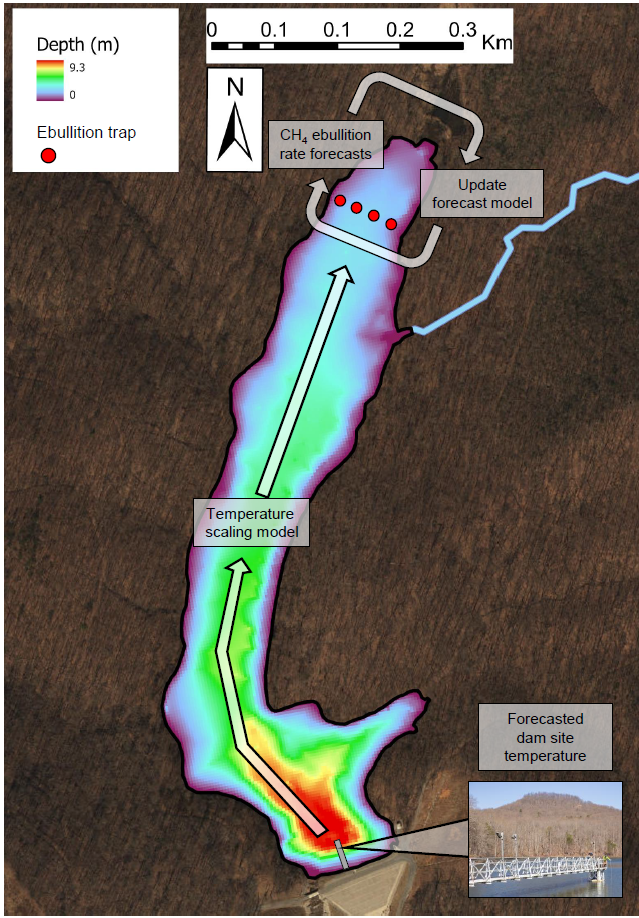


Figure 1. The bathymetry of Falling Creek Reservoir superimposed by a depiction of how our forecast workflow mapped spatially from the dam to the upstream site where CH4 ebullition rates were forecasted. Our forecast workflow used water temperature forecasts from the dam site in the reservoir (following Thomas et al. 2020). The water temperature forecasts were scaled to an upstream transect (denoted by red dots on map) to generate near-term iterative CH4 ebullition rate forecasts. Simultaneously, direct observations from ebullition traps and temperature loggers near the sediments were used to iteratively update the forecast model via sequential data assimilation.

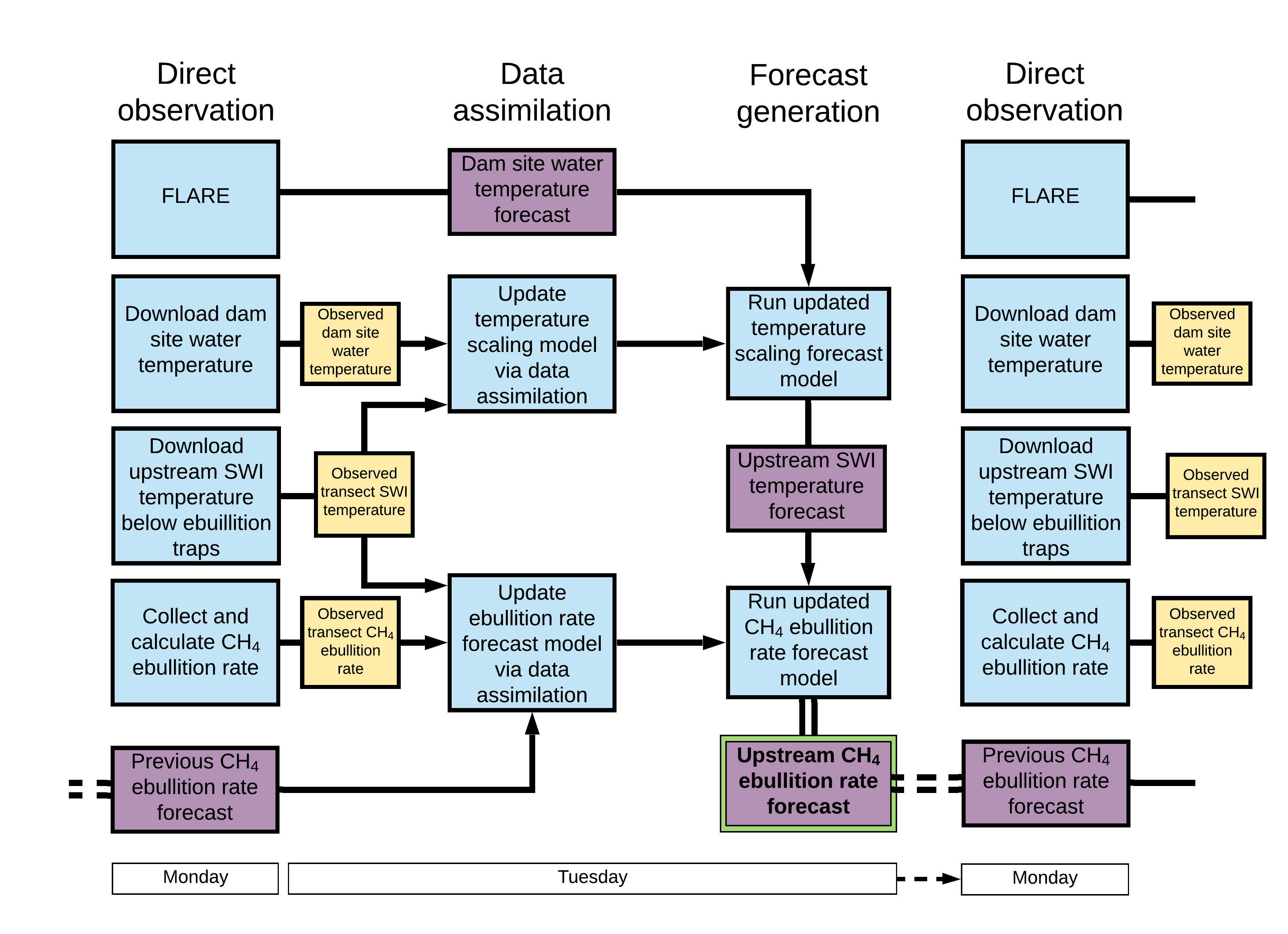


Figure 2. Conceptual illustration of the workflow generating near-term, iterative CH4 ebullition forecasts. The direct observation, data assimilation, and forecast generation stages were all completed within one business week. The blue boxes represent processes, and the yellow and purple boxes represent the data products that resulted from each process. The yellow boxes are manually collected data products that were used to update models and the purple boxes represent forecasted data products.

Graphical user interface, chart, histogram

Description automatically generated

Figure 3. A comparison of weekly forecasted CH4 ebullition rates in mg CH4 m-2 d-1 with data assimilation (a), without data assimilation (b), and with a persistence null model (c) from 1 July to 7 November 2019. The vertical blue (a) , magenta (b), and green (c) histograms represent the distribution of forecasted CH4 ebullition rates in Falling Creek Reservoir’s upstream transect from 210 ensemble members, and the small triangles matching the color on each histogram represents the mean of the forecast ensembles. The red diamonds represent the observed mean transect CH4 ebullition rate from four ebullition traps and the small red circles represent the directly observed ebullition rates of each trap. The vertical dotted lines denote the transition between the training period (27 May – 24 June) and forecasting period (1 July – 7 November).

Chart, histogram

Description automatically generated

Figure 4. A comparison of weekly CH4 ebullition rate forecast variance with data assimilation (blue), without data assimilation (magenta), and with a persistence null model (green) from 1 July to 7 November 2019 (left panel) and the total forecast variance over the forecasting period (right panel).

Chart, histogram

Description automatically generated

Figure 5. The proportion of total variance contributed among initial conditions, parameter, process, and driver data uncertainty from 1 July to 7 November 2019 of the CH4 ebullition rate forecasts with data assimilation.

A close up of a map

Description automatically generated

Figure 6. Parameter estimates of (a), (b), (c), and (d), of the CH4 ebullition rate forecasting model (Eqn 2: . CH4 ebullition rate forecasts with data assimilation are represented by the grey lines (±1 S.D.) and the static parameter estimates of the forecasts without data assimilation are represent by the horizontal blue lines (±1 S.D., (light blue dashed lines)). The forecasts with data assimilation updated iteratively as new data became available weekly from 1 July to 7 November, while the forecasts without data assimilation estimated future CH4 ebullition rates with the static 1 July model parameters over the whole forecasting period.

**Supporting information**

**Chapter V: Near-term, iterative forecasting suggests high predictability of reservoir methane ebullition at weekly time scales**

Ryan P. McClure1, R. Quinn Thomas2, Mary E. Lofton1, Whitney M. Woelmer1 and Cayelan C. Carey1

1Department of Biological Sciences, Virginia Tech, Blacksburg, VA, 24061, USA

2Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA, 24061, USA

*Contents of this file*

Supporting information A

Supporting information B

Supporting information C

Figures S1 to S4

Equations 1 to 3

In preparation for submission to *Biogeosciences*

Supporting information A: Details of the sediment-water interface (SWI) temperature scaling model and manual data collection procedures for the HOBO temperature loggers, SWI transect-level scaling, and relationship between the dam temperature sensors and the upstream SWI temperatures.

The temperature scaling model consisted of a first-order linear model:

Eqn. S1

where is the forecasted upstream transect SWI temperature until the next sampling date, represents the mean forecasted temperatures of 2.3, 2.6, and 3.0 m from FLARE, and represented random process noise added to the model (i.e., , where is the standard deviation on the priors). The original parameter estimates of and were fitted from temperature observations that were measured at the dam site and the SWI transect site between 11 July to 30 October 2018, 27 May to 24 June 2019, and then updated via sequential data assimilation during the 2019 forecasting period (1 July – 7 November; Figure S5).

We collected water temperature data at the SWI below each ebullition trap during the training period (11 July to 30 October 2018 and 27 May to 24 June 2019) and forecasting period (1 July – 7 November). We deployed one temperature logger (HOBO Pendant Temperature/Light Data Logger, Bourne, MA, USA) below each of the n=4 ebullition traps. All temperature loggers recorded data on 10-minute intervals. The loggers were attached to a stainless-steel weight and a nylon string to hold them at the sediments ~1.0 m horizontally away from each ebullition trap to prevent disturbance of sediments under the traps. On every CH4 ebullition sampling day (Supporting information B), we simultaneously pulled the temperature logger at the trap to the surface and downloaded the data since the previous sampling. The temperature loggers were downloaded using HOBOware version 3.7.13.

We calculated the daily observed SWI temperature separately for each trap and then averaged the rates from the SWI temperatures measured below all four traps within the transect to determine a mean daily transect SWI temperature. We aggregated the upstream transect data together to match FLARE’s dam site water temperature observations and forecasts, which were also just for one site. The dam site temperature data were downloaded the same day that the SWI temperature loggers were downloaded in FCR and available at Carey et al. (2020).

Like the data assimilation steps described in the main body of the text, the model parameter estimates were updated iteratively when new data were available. Before the next ebullition forecasts were generated, the previous week of daily water temperature data from the dam site platform and upstream were appended to the dataset of previous observations and used to refit the parameters (, , and in Eqn. S1). When the model parameter estimates were updated based on the newest dam site temperature data and the mean daily transect SWI temperature, the 210 FLARE water temperature ensemble member forecasts for 2.3, 2.6, and 3.0 m were averaged together and then applied to the model (Eqn. S1) to predict what the forecasted SWI temperatures were at FCR’s upstream transect for the next sampling day.

Supporting information B: Manual data collection procedures for CH4 ebullition rates and transect-level scaling

We used passive traps to measure CH4 ebullition at the upstream transect site following McClure et al. (2020). The ebullition samples from each trap were extracted across a septum stopper using a needle attached to a 10-mL syringe. We injected 10 mL of ebullition gas into a 20-mL crimped top glass vial that was pre-filled with saturated salt brine solution. A secondary exit syringe extracted the salt brine solution as the sample was injected to generate 10 mL of gas headspace in the vial. If enough gas sample was available, two replicates were collected from each trap on a sampling day. The vials were stored upside down until analysis, so the remaining 10 mL of salt brine solution acted as a barrier to prevent any gas from escaping. We extracted any remaining gas from each ebullition trap using a 30 mL syringe and summed the total volume of ebullition gas collected each week.

The gas samples were analysed using a gas chromatograph coupled with a flame ionization detector (GC-FID) within 24 hours of collection (following McClure et al. 2018, 2020). We determined the directly observed CH4 ebullition rate (Et) as follows:

Eqn. S2

where was the volume of ebullition collected in the trap (liters), is the CH4 concentration of the gas (mg CH4 L−1), is the duration of time the trap was deployed (in days), and is the cross-sectional area of the funnel (0.26 m2). As with the forecasted ebullition rates, the observed ebullition rate therefore represents an integration of the total gas captured per unit area during . Following McClure at al. (2020), we calculated the daily ebullition rate separately for each trap every week and then averaged the rates from the four traps within the transect to determine a mean daily transect ebullition rate.

Supporting information C: Generation of the forecasts without data assimilation and the persistence null model.

We developed a persistence null model (Eqn. S3) and forecasts without iterative data assimilation (using Eqn. 2 in main document) that forecasted ebullition to evaluate the performance of our iterative CH4 ebullition forecasts with data assimilation. The persistence null model forecasted CH4 ebullition rates () as follows:

Eqn S3:

where is a set of 210 random draws from a normal distribution which had a mean equal to the mean of the most recent observations and a standard error from the standard error of the observations and is random process noise that was added to the model (i.e., , where is the standard deviation on the priors). As a result, the persistence null model generated 210 CH4 ebullition rate forecast ensemble members for one week into the future using the observed ebullition rates and added process noise. The persistence null model forecasted ebullition during the same time period as the forecasts with data assimilation (1 July – 7 November).

Finally, we ran the CH4 ebullition rate forecasts without data assimilation. When new data were available, we did not update the model states and parameter estimates from the original forecast model (Eqn. 2 in the main document) developed after a training period. Thus, after the model trained from 27 May to 24 June, we sampled from 210 different parameter values each from the three posterior distributions of , , and (Eqn. 2) that were estimated for the 1 July forecast to predict future CH4 ebullition rates the remainder of the forecasting period (1 July – 7 November). Additionally, because the observed ebullition was included in the original forecast model as the AR term and updated with the forecast model at each forecast cycle, we instead sampled from 210 different values from the distribution of the previous week’s CH4 ebullition rate forecast as the term. Together, the posterior distributions of the parameters and use of the previous week’s forecast ensembles instead of observed ebullition ensured that no data assimilation occurred throughout the 1 July to 7 November forecasting period.

Supporting Figures

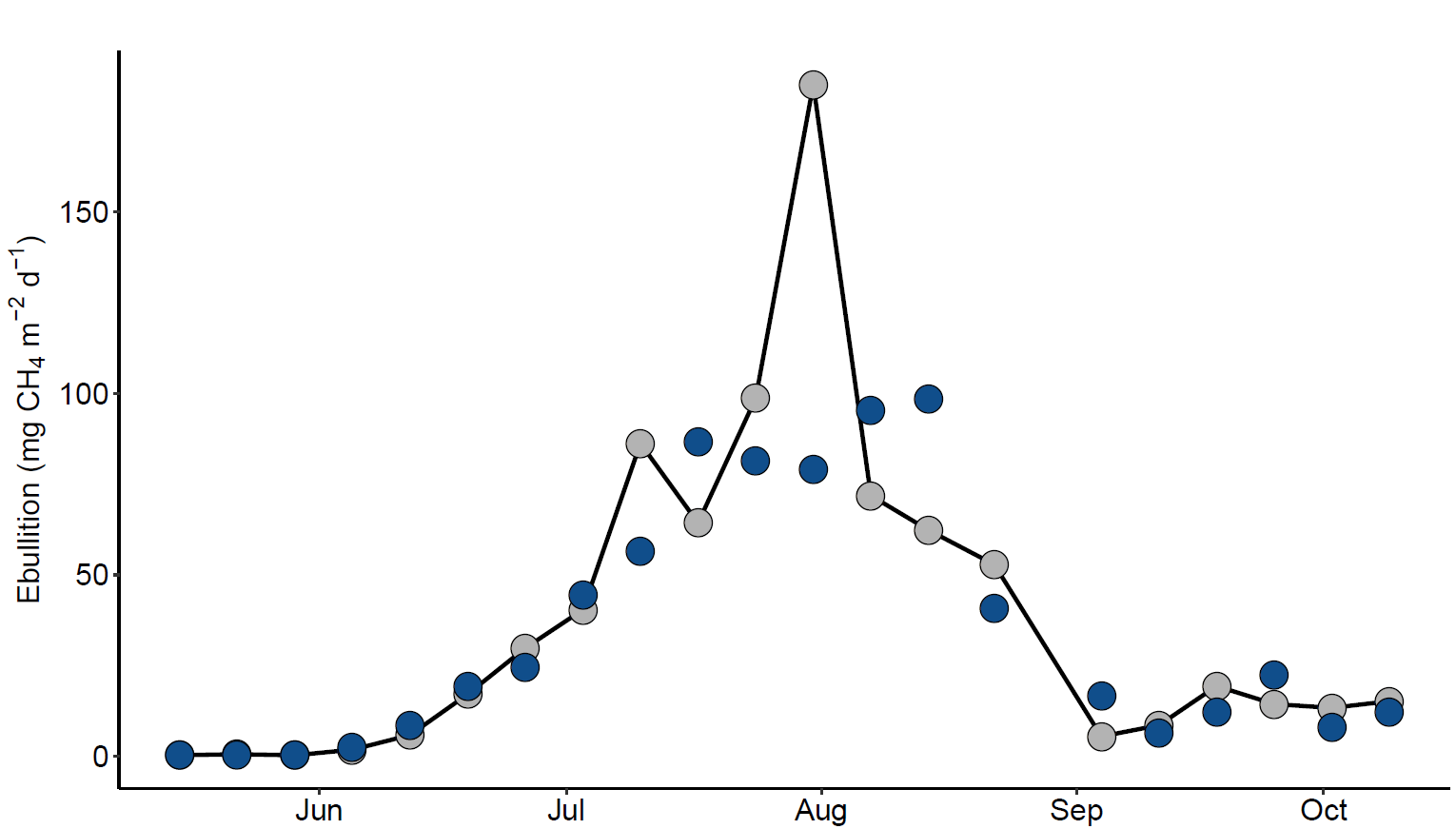
****

Figure S1. The weekly time series of the observed (grey) in comparison to predicted (blue) ebullition rates from the upmost transect in Falling Creek Reservoir in summer 2017. The predictions were derived from the autoregressive time series model in Eqn. 1 in the main text using static parameters. The parameters of the CH4 ebullition rate model in Eqn. 1 for the entire May to October 2017 study duration were = -6.58 ± 1.27 (1 S.D.), = 0.26 ± 0.05, and = 0.38 ± 0.06. The root-mean-square-error (RMSE) for the model in 2017 was 26.9 mg CH4 m-2 d-1. More information on the development of the model can be found in McClure et al. (2020).

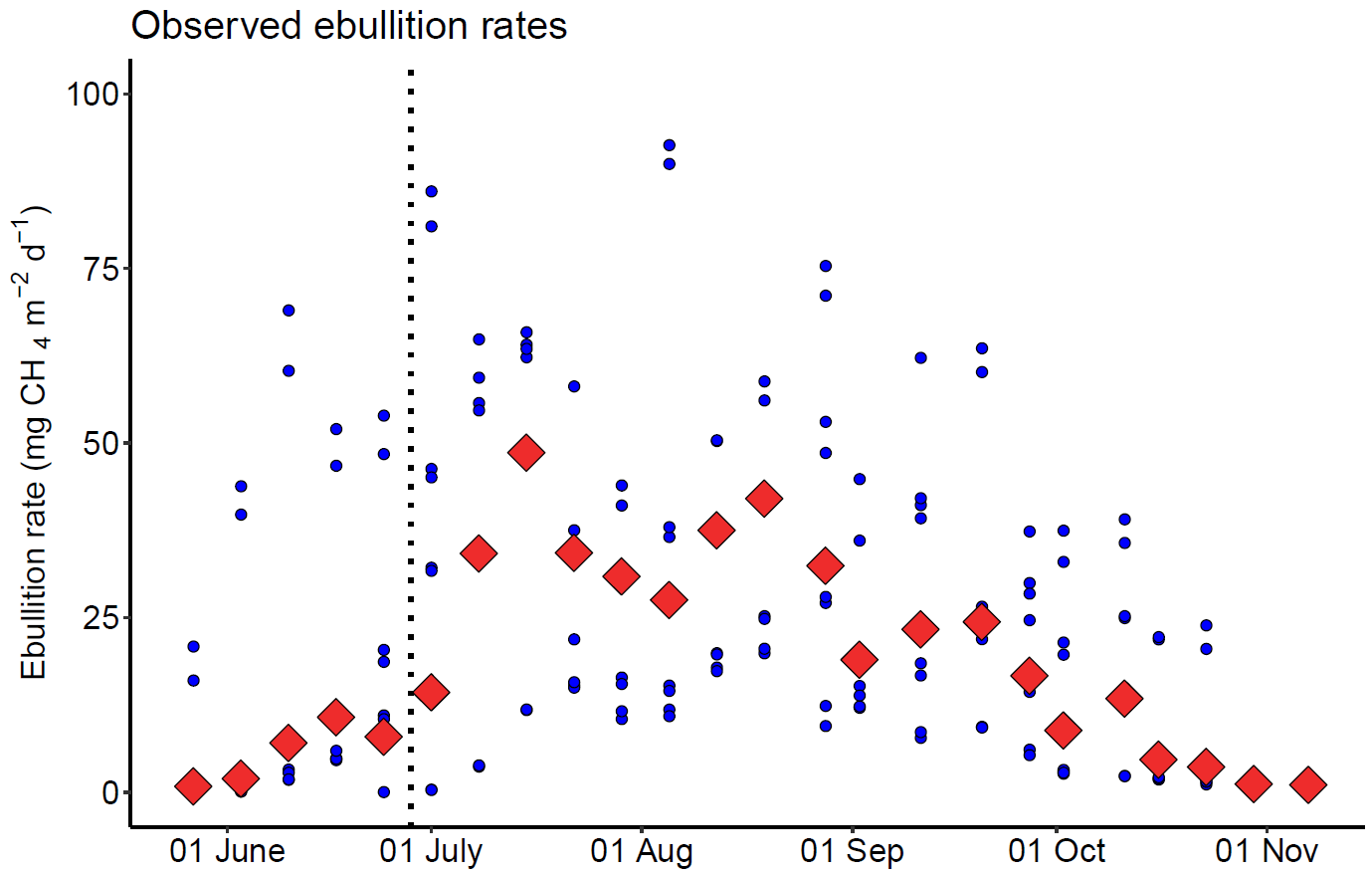


Figure S2. The observed CH4 ebullition rates (mg CH4 m-2 d-1) from all four ebullition traps (blue circles) and the mean CH4 ebullition rate among all four traps (red diamonds) during the 2019 training and forecasting periods. The vertical dotted line denotes the shift from the training (left) to forecasting period (right).

****

Figure S3. Weekly SWI temperature forecasts for the upstream transect in Falling Creek Reservoir in 2019 in both training and forecasting periods. The red diamond represents the weekly mean observed SWI temperature (averaged from the loggers below all four ebullition traps during the interval between sampling days) and the orange horizontal histograms represent the weekly SWI temperature forecast for the transect from 210 ensemble members.

****

Figure S4. Evaluation of the observed and forecasted SWI temperatures at the upstream transect for each forecasted sampling day from 1 July to 7 November 2019. The Nash-Sutcliffe efficiency (NSE) of the forecasts and observations was 0.9.



Figure S5. Linear relationship between the mean daily temperatures from 2.0 and 3.0 m water temperature sondes at the dam site and the upstream transect daily averaged SWI temperatures measured from the HOBO data loggers. The cyan points represent the trained data from 11 July to 10 October 2018 and 27 May to 24 June. The red points represent the daily observations from 1 July to 07 November 2019 forecasting period. Refer to Supporting information A on the SWI scaling model development.

**References**

Carey, C.C., B.J. Bookout, W.M. Woelmer, and A.S. Lewis. (2020). Time series of high-

frequency sensor data measuring water temperature, dissolved oxygen, conductivity, specific conductivity, total dissolved solids, chlorophyll a, phycocyanin, and fluorescent dissolved organic matter at discrete depths in Falling Creek Reservoir, Virginia, USA in 2018-2019 ver 4. Environmental Data Initiative.

McClure, R. P., Hamre, K. D., Niederlehner, B. R., Munger, Z. W., Chen, S., Lofton, M. E.,

Schreiber, M. E., Carey, C. C. (2018). Metalimnetic oxygen minima alter the vertical profiles of carbon dioxide and methane in a managed freshwater reservoir. Science of the Total Environment, 636, 610-620.

McClure, R. P., Lofton, M. E., Chen, S., Krueger, K. M., Little, J. C., & Carey, C. C. (2020). The

magnitude and drivers of methane ebullition and diffusion vary on a longitudinal gradient in a small freshwater reservoir. Journal of Geophysical Research: Biogeosciences, 125(3), e2019JG005205.