Title: Partitioning net wetland CH_4 emissions into production and oxidation components using ecosystem scale flux measurements and physically guided machine learning

Authors/Affiliations:

Qing Zhu, William Riley, Jinyun Tang, Robinson Negron-Juarez, Housen Chu, Kunxiaojia Yuan
Lawrence Berkeley National Laboratory, Climate & Ecosystem Sciences Division

Gavin McNicol
University of Illinois at Chicago, Department of Earth and Environmental Sciences

Min Chen, Li Fa
University of Wisconsin-Madison Madison

Focal Area(s):

- Key uncertainties and knowledge gaps where new methodology, infrastructure, or technology can advance predictive understanding of the methane cycle.
- The importance of high potential datasets or how the combination of data across spatial or temporal scales or scientific domains may lead to new scientific insights

Science or Technological Challenge:

Wetland CH_4 emissions result from production source and oxidation sink that are controlled by different microbial groups (*i.e.*, methanogens and methanotrophs). These production and oxidation rates have different dynamics and exhibits a wide range of responses to environmental changes. The net emissions also depend on transport processes (aerencnyma, ebullition, diffusion). However, ecosystem scale long-term observations (e.g., FLUXNET-CH_4; Delwiche et al., 2021) only measure net CH_4 emissions, which hinders predictive understanding of wetland CH_4 emissions across space and time.

A robust flux partitioning algorithm is urgently needed to decompose observed net emissions into gross production and oxidation rates. Such new datasets can be used to improve predictions of future wetland CH_4 emissions as well as spatial upscaling across heterogeneous landscape.

Rationale:

Wetland CH_4 emission represent ~20-30% of global CH_4 emissions, and these emissions are increasing due to ongoing climate warming since the radiative power of CH_4 is ~30 times stronger than CO_2 over a 100-year time horizon. Classic approaches to estimate wetland CH_4 emissions used either process-based bottom-up (BU) models that directly simulated wetland biogeochemical processes or top-down (TD) transport models that indirectly inferred wetland CH_4 emissions based on atmospheric CH_4 concentrations. Existing BU modeling studies showed some progress in capturing the observed CH_4 emissions at a handful of FLUXENT-CH4 sites after careful calibration. However, BU models still suffer from large parametric uncertainty and use incomplete biogeochemical theories. Furthermore, TD models often use prior information derived from BU model estimates and other non-wetland surface CH_4 emissions that inevitably introduce uncertainties. The most recent Global Carbon Project methane budget revealed ~ 30 TgCH_4 discrepancies in the magnitude, inter-annual variability, and long-term trends of BU and TD model estimates of wetland CH_4 budgets.

The ongoing synthesis efforts at FLUXNET-CH4 sites provides useful data to parameterize BU models. However, the net CH_4 emissions from the FLUXNET-CH4 dataset do not provide sufficient constraints on the CH_4 biogeochemical cycle. Minimally, a BU model requires methane gross production and oxidation rates to constrain methanogenesis and methanotrophic processes, respectively. Thus, developing a robust partitioning algorithm for FLUXNET-CH4 CH_4 emissions becomes a critical research need to improve process understanding and model predictability of wetland CH_4 cycle.

Narrative:

Our overall objective is to robustly partition observed FLUXNET-CH4 net CH_4 emissions into production sources and oxidation sinks using physically guided machine learning (PGML). We will leverage the existing FLUXNET-CH4 dataset (Delwiche et al., 2021) and new wetlands sites in South America to generate global datasets of wetland CH_4 production and oxidation rates. Although the measurements of CH_4 emissions are far fewer than those of CO_2 fluxes, we expect to overcome this data limitation by developing and advanced PGML model. Dislike traditional ML that depends entirely on the information content of a big dataset, a PGML can combine physical principles and ecological theory to leverage the information contained in a more limited dataset to understand and predict the dynamics of target processes. Our previous work on PGML has demonstrated promising model performance at temperate and high latitude wetland sites (Yuan et al., 2022). In our previous version of PGML, we have successfully integrated causal knowledges of how CH_4 emissions interact with physical and biological factors. Here, we will further develop the PGML model to include 1) distinct temperature sensitivities of methanogens and methanotrophs in our PGML model; 2) pre-training with synthetic data from a more mechanistic microbial model; and 3) constraints on model structure based on the knowledge of methane process interactions.

References

Delwiche, K.B., Knox, S.H., Malhotra, A., Fluet-Chouinard, E., McNicol, G., Feron, S., Ouyang, Z., Papale, D., Trotta, C., Canfora, E. and Cheah, Y.W., 2021. FLUXNET-CH 4: a global, multi-ecosystem dataset and analysis of methane seasonality from freshwater wetlands. Earth system science data, 13(7), pp.3607-3689.

Yuan, K., Zhu, Q., Li, F., Riley, W.J., Torn, M., Chu, H., McNicol, G., Chen, M., Knox, S., Delwiche, K. and Wu, H., 2022. Causality guided machine learning model on wetland CH4 emissions across global wetlands. Agricultural and Forest Meteorology, 324, p.109115.