

Scaling Genes to Global Methane Modeling Through Artificial Intelligence

Xiaofeng Xu (xxu@sdsu.edu), San Diego State University
Jorge Rodrigues (jmrodrigues@ucdavis.edu), University of California Davis

Focal areas:

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 combining data across spatial or temporal scales or scientific domains may lead to new scientific insights, either within or across fields.

Science challenges:

Methane (CH₄) is produced at the micro-scale, while policymaking relies on macro-scale CH₄ information. The ability to understand and predict CH₄ cycling across distinct scales is essential but remains a grand challenge. Microbiologists have produced a huge amount of metagenomic data on CH₄-relevant functional genes (Freitag and Prosser 2009, Kroeger et al. 2020), which can inform the individual CH₄ processes at the microscale. Correspondingly, land surface CH₄ flux data can infer the CH₄ cycling at an ecosystem level, and atmospheric CH₄ concentration fluctuation implies CH₄ cycling at a regional scale. Building a multiscale modeling capability can benefit from integrating data obtained at various scales but is particularly limited by the modeling capability for assimilating metagenomic data on methanogenesis and methanotroph (Xu et al. 2015, Sihi et al. 2021b).

Rationale:

Physical and chemical processes are predictable with high confidence, while biological processes remain challenging for accurate prediction which heavily relies on massive datasets and robust models. Methane modeling techniques have been developed and applied for more than 40 years, and more than 40 CH₄ models have been developed from 1987 to 2016 (Xu et al. 2016); considering the CH₄ models developed in the past 6 years (Song et al. 2020, Ricciuto et al. 2021, Sihi et al. 2021a, Yuan et al. 2021), substantial progress has been made in modeling CH₄ cycling processes. Meanwhile, progress has been made in gathering large datasets of functional genes encoding proteins responsible for CH₄ production and oxidation in various biomes. Knowledge gaps still exist in three aspects: (1) identifying the driving factors for microbial mechanisms associated with CH₄ production and oxidation; (2) connecting the microscale processes with large-scale CH₄ fluctuation with high predictability; and (3) parameterizing multiscale CH₄ models for simulating CH₄ cycling within an Earth System modeling framework. Because increased data cumulation did not bring a significant improvement in our confidence in predicting CH₄ fluxes in terrestrial ecosystems (and in aquatic ecosystems as well), we propose that integrating microbial genomic data with ecosystem-level measurements through advanced artificial intelligence would significantly improve our predictability of CH₄ flux. This should be an achievable key task for the next 10 years. The mechanistic modeling approach carries the advantage of representing each CH₄ process individually while allowing for the integration of multiple sources of data (Xu et al. 2016). Advanced Artificial Intelligence (AI) algorithms can also be used for identifying genetic makers with direct association with CH₄ emissions within large metagenomic datasets (Khan et al. 2023), but not previously linked to the processes of methanogenesis and methanotrophy. These agnostic approaches being performed through AI are anticipated to better support the model parameterization and application in predicting CH₄ cycling.

Narrative:

In order to develop robust predictability, the research community needs to enhance collaborative research for CH₄ modeling on and across three scales – (1) at micro-scale, where different microbial processes are occurring to understand hot moments of emissions; (2) at the ecosystem scale, where CH₄ emissions are being measured to capture ecosystem drivers of methanogenesis and methanotrophic processes and validate

models, and (3) at regional/global scale, to upscale and predict changes over time. AI serves as a powerful tool to expedite the development of modeling capability by assisting in distilling information from massive data and further supporting model development and application. Specifically, we envision in-depth AI approaches to be used in three areas.

- 1) AI assistance in processing and integrating micro-scale meta'omics (metagenomics, metatranscriptomics, metaproteomics, and metabolomics) data with CH₄ models. Massive metagenomic data have been produced, but specific drivers of biological processes are challenging to retrieve. Artificial intelligence can be a powerful tool for understanding microbial physiology that is fundamental for methane production and oxidation processes occurring at the microscale. A study has applied a multifactorial strategy of deep sequencing and a machine learning approach to compare taxonomic differences and generated metabolic maps with differential representations of genes involved in the cycling of nutrients and CH₄ in forest and pasture soils in the Amazon forest (Khan et al. 2023). This is an area that deserves further exploration as datasets have already been collected.
- 2) AI can assist in building ecosystem-level predictability based on plot-level observational data and micro-scale meta'omic datasets. Our group is working on a project to integrate metagenomic data with an ecosystem model to better parameterize the model for simulating individual CH₄ production processes rather than solely focusing on land surface CH₄ flux (Zuo et al. 2023). Our AI approach assists with model parameterization on meta'omic data and ecosystem-level CH₄ flux.
- 3) Enhance the earth system model by including a microbial functional group based CH₄ module with the capability of assimilating data of functional genes, ecosystem level CH₄ flux, and atmospheric CH₄ concentration. Artificial Intelligence algorithms can be used to improve model efficiency and data assimilation.

References:

- Freitag, T. E., and J. I. Prosser. 2009. Correlation of methane production and functional gene transcriptional activity in a peat soil. *Applied and Environmental Microbiology* **75**:6679-6687.
- Khan, M. A. W., B. J. Bohannan, K. M. Meyer, A. M. Klein, K. Nusslein, J. P. Grover, and J. L. Rodrigues. 2023. Amazonian soil metagenomes indicate different physiological strategies of microbial communities in response to land use change. *Applied Environmental Microbiology* **Under revision**.
- Kroeger, M. E., L. K. Meredith, K. M. Meyer, K. D. Webster, P. B. De Camargo, L. F. De Souza, S. M. Tsai, J. Van Haren, S. Saleska, B. J. Bohannan, J. L. Rodrigues, E. Berenguer, J. Barlow, and K. Nusslein. 2020. Rainforest-to-pasture conversion stimulates soil methanogenesis across the Brazilian Amazon. *The ISME Journal* **15**:658-672.
- Ricciuto, D., X. Xu, X. Shi, Y. Wang, X. Song, C. W. Schadt, N. A. Griffiths, J. Mao, J. Warren, P. Thornton, J. Chanton, J. K. Keller, S. Bridgham, J. Gutknecht, S. Sebestyen, A. C. Finzi, R. K. Kolka, and P. J. Hanson. 2021. An Integrative Model for Soil Biogeochemistry and Methane Processes: I. Model Structure and Sensitivity Analysis. *Journal of Geophysical Research-Biogeosciences* **10.1029/2019JG005468**.
- Sihi, D., X. Xu, M. Salazar Ortiz, C. S. O'Connell, W. L. Silver, C. López-Lloreda, J. M. Brenner, R. K. Quinn, J. R. Phillips, and B. D. Newman. 2021a. Representing methane emissions from wet tropical forest soils using microbial functional groups constrained by soil diffusivity. *Biogeosciences* **18**:1769-1786.
- Sihi, D., X. Xu, M. Salazar Ortiz, C. S. O'Connell, W. L. Silver, C. López-Lloreda, J. M. Brenner, R. K. Quinn, J. R. Phillips, B. D. Newman, and M. A. Mayes. 2021b. Representing methane emissions from wet tropical forest soils using microbial functional groups constrained by soil diffusivity. *Biogeosciences* **18**:1769-1786.
- Song, C., J. Luan, X. Xu, M. Ma, M. Aurela, A. Lohila, I. Mammarella, P. Alekseychik, E. S. Tuittila, and W. Gong. 2020. A microbial functional group-based CH₄ model integrated into a terrestrial ecosystem model: Model structure, site-level evaluation, and sensitivity analysis. *Journal of Advances in Modeling Earth Systems* **12**:e2019MS001867.
- Xu, X., D. A. Elias, D. E. Graham, T. J. Phelps, S. L. Carrol, S. D. Wullschleger, and P. E. Thornton. 2015. A microbial functional group based module for simulating methane production and consumption: application to an incubation permafrost soil. *Journal of Geophysical Research-Biogeosciences* **120**:1315-1333.
- Xu, X., F. Yuan, P. J. Hanson, S. D. Wullschleger, P. E. Thornton, W. J. Riley, X. Song, D. E. Graham, C. Song, and H. Tian. 2016. Review and Synthesis: Four decades of modeling methane cycling within terrestrial ecosystems. *Biogeosciences* **13**:3735-3755.
- Yuan, F., Y. Wang, D. M. Ricciuto, X. Shi, F. Yuan, P. J. Hanson, S. Bridgham, J. K. Keller, P. E. Thornton, and X. Xu. 2021. An Integrative Model for Soil Biogeochemistry and Methane Processes: II. Warming and elevated CO₂ Impacts on Peatland CH₄ Emission. *Journal of Geophysical Research-Biogeosciences* **10.1029/2020JG005963**.
- Zuo, Y., X. Zhu, N. Wang, F. Yuan, J. Liu, K. Li, C. Song, L. Sun, and X. Xu. 2023. Metagenomic data for improving ecosystem model in simulating methane cycling in temperate wetlands. *Journal of Advances in Modeling Earth Systems* **In preparation**.