Machine Learning Assisted Bayesian Calibration of Model Physics Parameters for Wetland Methane Emissions: A Case Study at a FLUXNET-CH_4 Site

Sandeep Chinta[1], Xiang Gao[1], Qing Zhu[2]
[1]Center for Global Change Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
[2]Climate and Ecosystem Sciences Division, Climate Sciences Department, Lawrence Berkeley National Laboratory, Berkeley, California, USA.
sandeepc@mit.edu, xgao304@mit.edu, qzhu@lbl.gov

Abstract

Methane (CH_4) possesses a notably higher warming potential than carbon dioxide despite its lower atmospheric concentration, making it integral to global climate dynamics. Wetlands stand out as the predominant natural contributor to global methane emissions. Accurate modeling of methane emissions from wetlands is crucial for understanding and predicting climate change dynamics. However, such modeling efforts are often constrained by the inherent uncertainties in model parameters. Our work leverages machine learning (ML) to calibrate five physical parameters of the Energy Exascale Earth System Model (E3SM) land model (ELM) to improve the model's accuracy in simulating wetland methane emissions. Unlike traditional deterministic calibration methods that target a single set of optimal values for each parameter, Bayesian calibration takes a probabilistic approach and enables capturing the inherent uncertainties in complex systems and providing robust parameter distributions for reliable predictions. However, Bayesian calibration requires numerous model runs and makes it computationally expensive. We employed an ML algorithm, Gaussian process regression (GPR), to emulate the ELM's methane model, which dramatically reduced the computational time from 6 CPU hours to just 0.72 milliseconds per simulation. We exemplified the procedure at a representative FLUXNET-CH_4 site (US-PFa) with the longest continuous methane emission data. Results showed that the default values for two of the five parameters examined were not aligned well with their respective posterior distributions, suggesting that the model's default parameter values might not always be optimal for all sites, and that site-specific analysis is warranted. In particular, analyses at sites with different vegetation types and wetland characteristics could reveal more useful insights for understanding methane emissions modeling.

1 Introduction and Motivation

Greenhouse gas (GHG)-induced climate change poses unprecedented challenges and serious risks for human society and natural environment. Methane is the second most important GHG, with ~25 times stronger 100-year global warming potential than carbon dioxide [1, 2]. Since the Industrial Revolution, the atmospheric concentrations of CH_4 have sharply risen and doubled since pre-industrial times. Alarmingly, its growth rate in 2021 marked the highest record since 1984 [3–5]. Such increases have profound implications for global warming and highlight the urgency to manage its emissions effectively. Wetlands account for approximately 30% of global methane emissions and are the primary natural source [6]. However, wetland methane emission estimates based on biogeochemistry models remain highly uncertain [7, 8], mainly because CH_4 dynamics rely on a large number of poorly constrained model parameters to characterize a diverse array of physical,

Submitted to Tackling Climate Change with Machine Learning: workshop at NeurIPS 2023.

biological, and chemical processes [9, 10]. While these parameters typically have fixed (default) values, their exact values are often ambiguous and present large uncertainty. The parameter values are usually determined within their theoretically plausible uncertainty ranges based on the most reliable knowledge available. One of the essential paths to achieve reduced biases in methane emission estimates is to identify the most critical model parameters via sensitivity analysis and determine their optimal distribution through calibration.

Traditional deterministic calibration methods focus on identifying an optimal parameter value by comparing model output to observations, frequently neglecting the inherent predictive uncertainties and incorrectly assuming that there is always a single set of optimal values for all parameters [11]. Nevertheless, a single optimal set may not always exist for a problem. Even if it exists, the associated uncertainties could be large. Moreover, several alternative parameter configurations may yield similarly accurate results, challenging the notion of a single optimal parameter set [12, 13]. Considering these challenges posed by the deterministic calibration methods and the inherent complexities of methane models, there is a need for a more nuanced and statistically rigorous approach to parameter calibration. Probabilistic approaches like Bayesian calibration handle these issues by statistically representing parameter uncertainties [14]. The input parameter space is represented as probability distributions of parameters. Multiple parameter samples are drawn from this distribution, and simulations are conducted for these samples, inherently creating an ensemble of model predictions. An objective function is used to evaluate the prediction range by comparing these simulation outputs with observations. After calibration, final posterior distributions of the parameters are obtained. When samples are drawn from these distributions, the resulting simulations, forming an ensemble, more closely match the observations. However, it is important to note that this approach generally requires massive computational power to simulate the methane model multiple times. Hence, there is a compelling need for strategies like machine learning that can facilitate this process but also retain accuracy. ML-based emulators have emerged as a pivotal tool to emulate the behavior of complex earth system models for achieving this goal. These emulators are first trained on a subset of model simulations to learn the intricate relationships between inputs and outputs. The trained emulators can then be used to predict the model's response for a new set of parameters, effectively eliminating the need for exhaustive simulations every time there is a change in parameter values. This approach is particularly advantageous for Bayesian calibration, where thousands of simulations are needed to explore the parameter uncertain space adequately. Until now the attempt is still very few for applying this probabilistic approach to predict wetland methane emissions from land models.

This study aims to bridge the gap between the intensive computational demands of Bayesian calibration and the desired accuracy in wetland methane emission modeling. This is achieved by emulating the ELM with the Gaussian Process Regression (GPR) ML algorithm. We leverage observed CH_4 emission data from a specific FLUXNET-CH_4 site (US-PFa) as a case study. We select five model parameters and train GPR for estimating the error associated with methane emission outputs. Our primary objective is to minimize the error between the emulator-estimate emissions and corresponding observations. While we only focus on one site, the demonstrated methodology is expected to be easily extended to other FLUXNET-CH_4 sites with much broader applicability.

2 Model parameters and FLUXNET-CH_4 site (US-PFa) data

We configured the most recent version of the Energy Exascale Earth System Model (E3SM) land model (ELM) (https://github.com/E3SM-Project/E3SM), which contains many added new features for the CH_4 dynamics modeling [15]. In previous studies, a comprehensive sensitivity analysis was conducted for 19 ELM parameters associated with different CH_4 dynamic processes and five parameters are identified as sensitive for methane emission [16]. These five sensitive parameters are presented in Table 1 along with their default values, theoretical ranges, and brief description. A uniform prior distribution is assumed for these parameters. Our objective is to identify a posterior distribution of these five parameters to minimize the emission prediction error.

FLUXNET-CH_4 is a global network of sites that provides continuous, high-frequency, and quality-checked CH_4 emission flux measurements. US-PFa [17](https://ameriflux.lbl.gov/sites/siteinfo/US-PFa) with longest continuous available data is chosen for our study. The monthly averaged data

from this site is used to evaluate the monthly averaged site-specific simulated emissions from ELM model for various parameter samples.

Table 1: List of 5 ELM parameters used and their default values, ranges, and brief description.

Mechanism	Parameter	Default	Range	Units	Description
Production	Q_{10}	2	[1.5 4]	-	CH_4 production Q_{10}
	f_{CH_4}	0.2	[0.1 0.3]	-	Ratio between CH_4 and CO_2 production below the water table
Substrate availability	z_τ	0.5	[0.1 0.8]	m	e-folding depth for decomposition
Diffusion	f_{D_o}	1	[1 10]	$m^2\ s^{-1}$	Diffusion coefficient multiplier
Oxidation	K_{O_2}	2×10^{-2}	$[2\times10^{-3}\ 2\times10^{-1}]$	$mol\ m^{-3}$	O_2 half-saturation oxidation coefficient

3 Methodology and Results

A commonly adhered guideline [18] suggests using a sample size ten times the number of model parameters for training ML models. Following this, we generated sixty samples (exceeding the 10*5 criterion) of diverse parameter values using Latin Hypercube Sampling (LHS) [19] to train the GPR. Another set of 20 independent LHS samples are generated to test the GPR fit. Simulations are also conducted with the ELM methane model based on these 80 samples and the corresponding root-mean-square-error (RMSE) for each simulation is evaluated using observed monthly averaged emissions. The RMSE values from these 80 simulations is normalized by dividing with RMSE value from the simulation with default parameter values. This is referred to as normalized RMSE (nRMSE). Any value of nRMSE < 1 implies that the parameter sample is better than default. Ideally, we would want nRMSE closer to 0. GPR is trained with normalized parameter values in the ranges from Table 1 (normalized to [0 1] using MinMax scaling) as inputs and nRMSE values as outputs with the 60 simulations. Subsequent testing of the trained GPR on 20 test samples results in an R^2 value of 0.92, indicating a strong model fit. This GPR model is subsequently employed for Bayesian calibration.

Bayesian calibration is a process of updating our beliefs about model parameters based on observed data. We start with prior beliefs (priors), minimize nRMSE (likelihood), and then update our beliefs to obtain the posterior distribution. Markov Chain Monte Carlo (MCMC) [20] is employed (using 'emcee' package [21], which is particularly efficient for multi-dimensional problems) to construct a Markov chain where the stationary distribution (the distribution to which the chain converges over time) is the desired posterior distribution. By running the chain for a sufficient number of steps (12000) and discarding an initial set of "burn-in" samples (3000), we obtain samples that approximate the posterior distribution. Convergence of the MCMC chains is ensured by implementing the Gelman-Rubin diagnostic ($\hat{R}$) to check for convergence [22, 23]. A value of $\hat{R}$ close to 1 for all parameters indicate convergence. $\hat{R}_{1,2,3,4,5}$ = (1.022, 1.032, 1.042, 1.011, 1.019). This implies that the parameters converged to a posterior distribution minimizing the nRMSE and the distributions are presented in Figure 1. The mean nRMSE of the posterior distribution of parameters is 0.228, which is a remarkable 77.2% reduction in error compared to default. The 2σ range (95% confidence) of nRMSE is [0.199 0.254], which depicts a significant improvement from default.

The posterior distribution of parameters in Fig. 1 offers several insights. The default parameter values for Q_{10} and f_{CH4} are outside the 2σ interval, which indicates these parameters likely have true values different from the default values. Even though z_τ has its default value within the 2σ interval, it is less likely than the values near the bounds (a bi-modal distribution). The posterior distribution of K_{O2} is closer to uniform distribution, which is our prior distribution, implying that this parameter is not very sensitive to methane emission at US-PFa site.

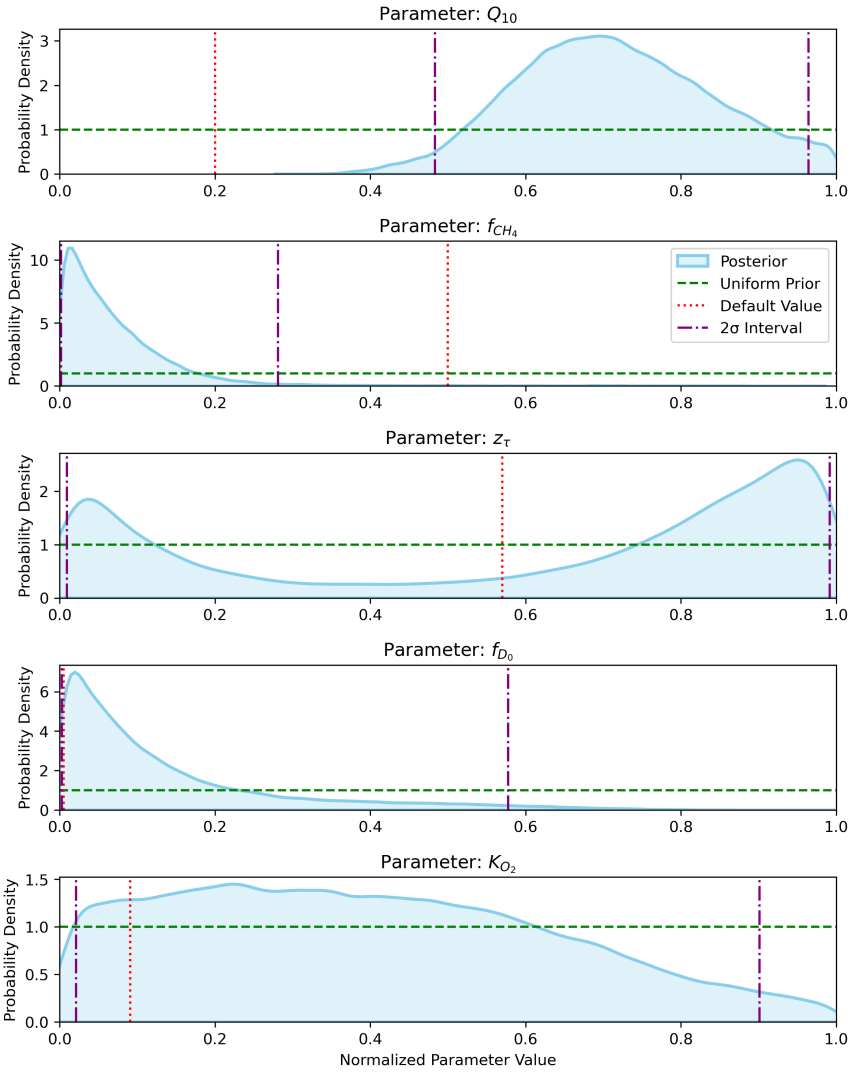

Fig. 1: Posterior distribution of five parameters following Bayesian calibration, plotted over a normalized parameter range. The red vertical line denotes the default value of the parameter. The green horizontal line denotes the prior distribution, which was the uniform distribution. The 2σ interval, encompassing 95% of the distribution's values, is marked by violet vertical lines.

4 Conclusions and Future Work

Our study delves into the crucial challenge of accurately predicting wetland methane emissions by adjusting ELM parameters using Bayesian calibration. Employing a machine learning algorithm (GPR) as an emulator for the ELM demonstrates significant potential in reducing computational demand. A remarkable improvement of 77% (normalized RMSE) was achieved compared to the default model parameters. This deviation of posterior distributions from the default values highlights that models may require fine-tuning to address site-specific nuances, particularly when addressing heterogeneous systems like wetlands. While our study focused on the US-PFa FLUXNET-CH_4 site, the presented methodology is universally applicable. This methodology can be easily extended to global wetland ecosystems with different vegetation types to get further insights into wetland methane emissions dynamics and relevant parameters. As methane emissions have seasonal fluctuations, it is our future interest to explore whether the posterior distributions change relative to seasons.

References

1. Solomon, S., others: Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. 4, (2007)
2. IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Presented at the (2013)
3. 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., others: The Global Methane Budget 2000–2017. Earth Syst Sci Data. 12, 1561–1623 (2020)
4. Lan, X., Thoning, K.W., Dlugokencky, E.J.: Trends in globally-averaged CH_4, N_2O, and SF_6 determined from NOAA Global Monitoring Laboratory measurements, (2023)
5. Dlugokencky, E.J., Bruhwiler, L., White, J.W.C., Emmons, L.K., Novelli, P.C., Montzka, S.A., Masarie, K.A., Lang, P.M., Crotwell, A.M., Miller, J.B., others: Observational constraints on recent increases in the atmospheric CH4 burden. Geophys Res Lett. 36, (2009)
6. Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J.G., Dlugokencky, E.J., Bergamaschi, P., Bergmann, D., Blake, D.R., Bruhwiler, L., others: Three decades of global methane sources and sinks. Nat Geosci. 6, 813–823 (2013)
7. Jackson, R.B., Saunois, M., Bousquet, P., Canadell, J.G., Poulter, B., Stavert, A.R., Bergamaschi, P., Niwa, Y., Segers, A., Tsuruta, A.: Increasing anthropogenic methane emissions arise equally from agricultural and fossil fuel sources. Environmental Research Letters. 15, 71002 (2020)
8. Rosentreter, J.A., Borges, A. V, Deemer, B.R., Holgerson, M.A., Liu, S., Song, C., Melack, J., Raymond, P.A., Duarte, C.M., Allen, G.H., others: Half of global methane emissions come from highly variable aquatic ecosystem sources. Nat Geosci. 14, 225–230 (2021)
9. Riley, W.J., Subin, Z.M., Lawrence, D.M., Swenson, S.C., Torn, M.S., Meng, L., Mahowald, N.M., Hess, P.: Barriers to predicting changes in global terrestrial methane fluxes: analyses using CLM4Me, a methane biogeochemistry model integrated in CESM. Biogeosciences. 8, 1925–1953 (2011)
10. Ricciuto, D.M., Xu, X., Shi, X., Wang, Y., Song, X., Schadt, C.W., Griffiths, N.A., Mao, J., Warren, J.M., Thornton, P.E., others: An integrative model for soil biogeochemistry and methane processes: I. Model structure and sensitivity analysis. J Geophys Res Biogeosci. 126, e2019JG005468 (2021)
11. Hoversten, G.M., Cassassuce, F., Gasperikova, E., Newman, G.A., Chen, J., Rubin, Y., Hou, Z., Vasco, D.: Direct reservoir parameter estimation using joint inversion of marine seismic AVA and CSEM data. Geophysics. 71, C1–C13 (2006)
12. Gupta, H.V., Sorooshian, S., Yapo, P.O.: Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. Water Resour Res. 34, 751–763 (1998)
13. Van Straten, G.T., Keesman, K.J.: Uncertainty propagation and speculation in projective forecasts of environmental change: A lake-eutrophication example. J Forecast. 10, 163–190 (1991)
14. Beven, K., Binley, A.: The future of distributed models: model calibration and uncertainty prediction. Hydrol Process. 6, 279–298 (1992)
15. Zhu, Q., Riley, W.J., Tang, J., Collier, N., Hoffman, F.M., Yang, X., Bisht, G.: Representing nitrogen, phosphorus, and carbon interactions in the E3SM land model: Development and global benchmarking. J Adv Model Earth Syst. 11, 2238–2258 (2019)
16. Chinta, S., Gao, X., Zhu, Q.: Sensitivity Analysis and Uncertainty Quantification of E3SM Methane model using Machine Learning. In: AGU Fall Meeting Abstracts. pp. B16D-02 (2022)
17. Desai, A.R., Xu, K., Tian, H., Weishampel, P., Thom, J., Baumann, D., Andrews, A.E., Cook, B.D., King, J.Y., Kolka, R.: Landscape-level terrestrial methane flux observed from a very tall tower. Agric For Meteorol. 201, 61–75 (2015)
18. Loeppky, J.L., Sacks, J., Welch, W.J.: Choosing the sample size of a computer experiment: A practical guide. Technometrics. 51, 366–376 (2009)
19. McKay, M.D., Conover, W.J., Beckman, R.J.: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics. 21, 239–245 (1979)
20. Brooks, S., Gelman, A., Jones, G., Meng, X.-L.: Handbook of markov chain monte carlo. CRC press (2011)
21. Foreman-Mackey, D., Hogg, D.W., Lang, D., Goodman, J.: emcee: the MCMC hammer. Publications of the Astronomical Society of the Pacific. 125, 306 (2013)
22. Gelman, A., Rubin, D.B.: Inference from iterative simulation using multiple sequences. Statistical science. 7, 457–472 (1992)
23. Vats, D., Knudson, C.: Revisiting the gelman–rubin diagnostic. Statistical Science. 36, 518–529 (2021)