

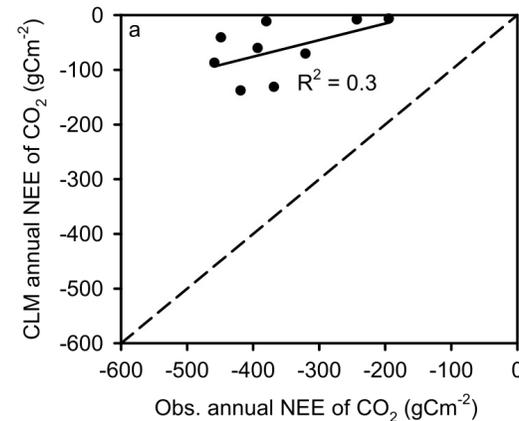
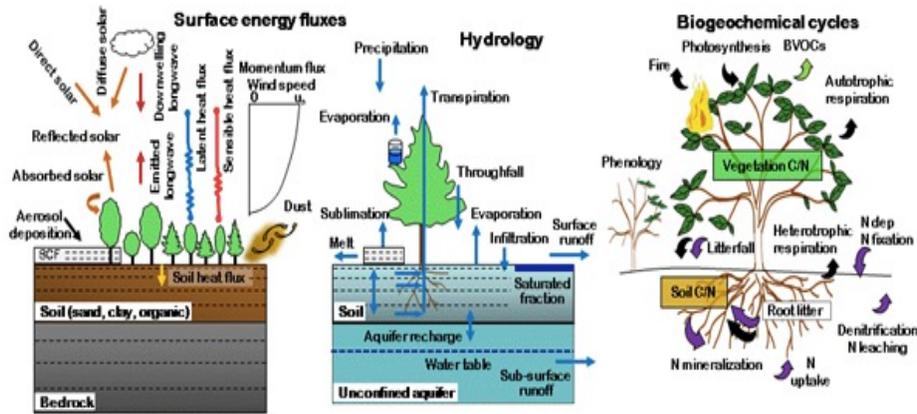
# Advancing Earth System Model Calibration: A diffusion-based method

Dan Lu (lud1@ornl.gov)

Yanfang Liu; Zezhong Zhang; Feng Bao; Guannan Zhang

# DBUQ method for efficient ELM calibration

- Energy Exascale Earth System Model (E3SM), land model (ELM) simulates terrestrial water, energy, and biogeochemical processes in terrestrial surfaces and is an important tool for improving our understanding of ecosystem responses to climate change.



There are significant discrepancies between ELM simulation with default parameter values and observed NEE at MOFLUX forest site. Model calibration is needed for improving prediction.

Gu et al., JGR, 2016

- ELM involves 65+ unknown parameters, and the use of default parameter values has shown large model discrepancy from site observations.
- Thus, ELM calibration is required at every site for improving prediction accuracy.
- However, calibrating ELM is challenging because of its strong nonlinearity, poorly constrained model parameters, and significant computational costs.
- We introduce a novel diffusion-based uncertainty quantification (DBUQ) method for efficient parameter estimation and model calibration.

# Diffusion-Based Uncertainty Quantification (DBUQ) method

- DBUQ is a score-based diffusion model.
- It draws samples to approximate posterior distribution

$$p(X|Y = y) \propto p(Y = y|X)p(X)$$

- DBUQ draws the parameter posterior samples by first formulating a parameterized generative model  $F$ ,

$$X|Y \approx F(Y, Z; \theta)$$

- DBUQ uses a NN to estimate  $F$  using supervised learning and then evaluates the trained  $F$  to generate the target parameter posterior samples,  $X|Y$  at  $Y = y$
- Procedure to estimate  $F$ :

- Generate  $\mathcal{D}_{\text{prior}} = \{(x_j, y_j)\}_{j=1}^J$  from prior distribution;
- Estimate the score function using Monte Carlo (MC) method based on  $\mathcal{D}_{\text{prior}}$
- Solve a reverse ODE in the diffusion model based on the estimated score function to generate labeled data
- Train a NN on the labeled data to learn the  $F$

- For an observation  $y$ , we evaluate  $F$  at standard Gaussian samples  $Z$  to generate target samples of  $X|Y$  at  $Y = y$

---

## Our DBUQ method for parameter uncertainty quantification

---

**Input:** Prior sample set  $\mathcal{D}_{\text{prior}} = \{(x_j, y_j)\}_{j=1}^J$ ;

**Output:** Trained generative model  $F(Y, Z; \theta)$ ;

**Procedure:**

1. for  $m = 1, \dots, M$ 
    - Estimate score function using Monte Carlo estimation through Eq. (13)–(15);
    - Solve the ODE in Eq. (16) with the estimated score function;
    - Obtain one sample  $(x_m, y_m, z_m)$  in the dataset  $\mathcal{D}_{\text{label}}$ ;
- end
2. Train a NN to approximate the generative model  $X = F(Y, Z; \theta)$  using the training data  $\mathcal{D}_{\text{label}} = \{(x_m, y_m, z_m)\}_{m=1}^M$ .

**Generate parameter posterior samples:** for a given observation  $y$ , evaluate the trained  $F$  at standard Gaussian samples  $Z$  to generate parameter posterior samples to approximate the target distribution  $p(X|Y = y)$

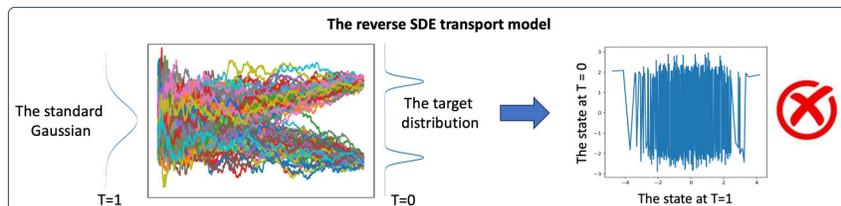
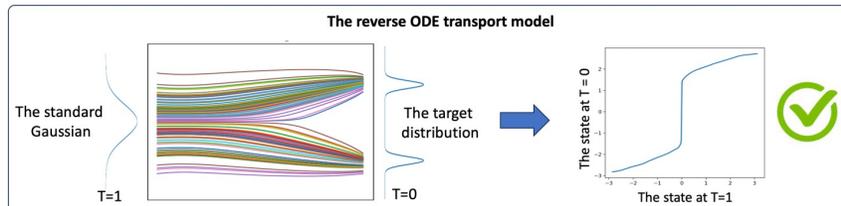
---

# Advantages of DBUQ method

## Traditional score-based diffusion models

### Computationally expensive, memory-intensive

- ❖ Use a NN to learn the score function under unsupervised learning due to lack of labeled data;
  - The unsupervised learning requires storing a large number of stochastic paths of the forward SDE, which is computationally expensive and memory intensive;
- ❖ Solve a reverse SDE repeatedly using the learned score function to generate target samples;
  - For each sample generation, it requires to solve the SDE for a large number of time steps



## Our DBUQ method

### Computationally and memory efficient Amortized Bayesian inference

- ❖ Formulate a supervised learning problem to estimate the sample generator  $F$ ;
  - After the generator is trained, it can quickly generate numerous parameter posterior samples for any given observations;
- ❖ Use MC method to estimate the score function;
  - Computationally efficient;
- ❖ Solve a reverse ODE based on the estimated score function to generate the labeled data to train  $F$ ;
  - Computationally and memory efficient as solving the ODE only needs to store the initial and terminal states of the transport path;

# An illustrative example of DBUQ

## Problem

We use a simple 1D problem to illustrate our DBUQ method. The forward model  $g(X)$  in the likelihood function of Eq. (2) is defined by

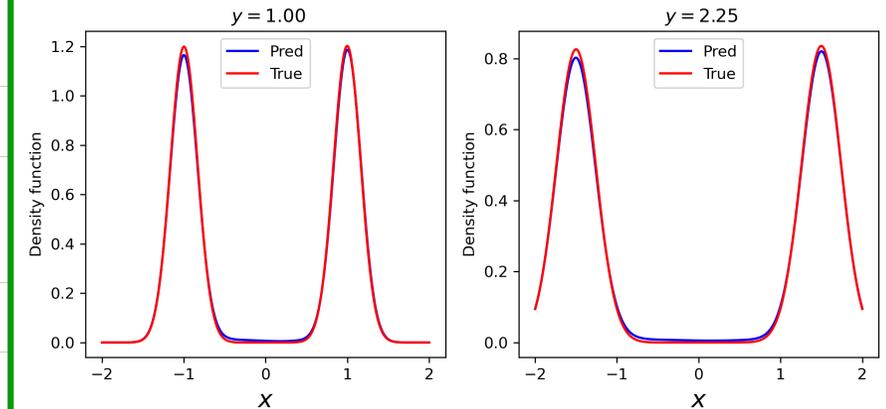
$$g(X) = X^2,$$

where the prior distribution of  $X$  is defined by a uniform distribution  $\mathcal{U}([-2, 2])$  in the domain  $[-2, 2]$ . The observation variable  $Y$  is defined by

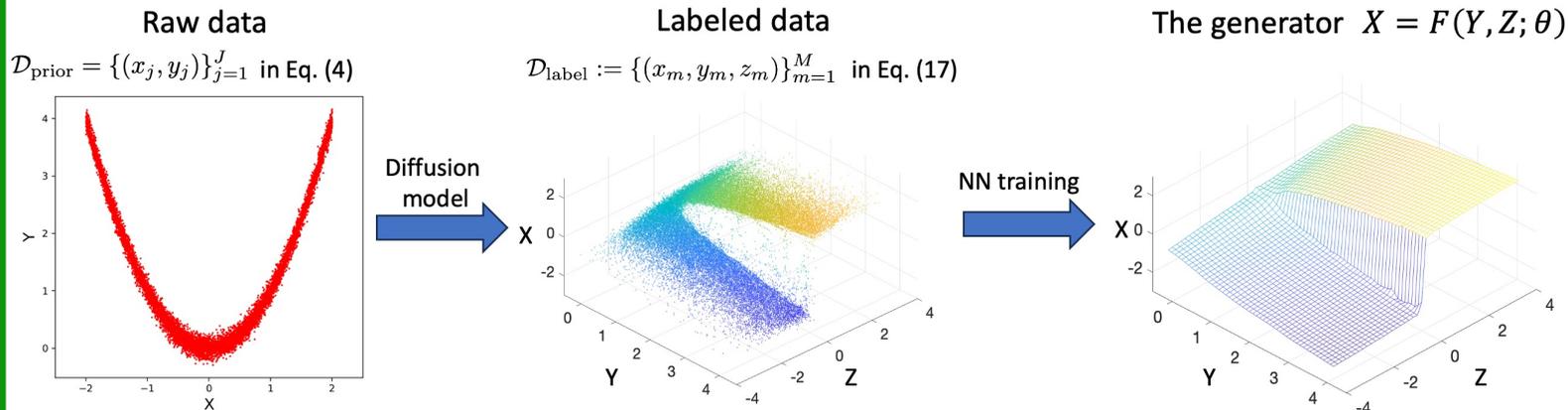
$$Y = g(X) + \varepsilon,$$

where  $\varepsilon$  follows the Gaussian distribution  $\mathcal{N}(0, \sigma^2)$  with  $\sigma = 0.1$ .

## Result



## Procedure



- Multimodal distributions can be common for earth system model parameter estimation;
- It is challenging for UQ methods to capture all the possible modes;
- DBUQ accurately approximates the target bi-modal distributions;
- DBUQ is also computationally efficient, taking  $< 2$ min.

# Apply DBUQ to improve ELM calibration

- Problem: Use DBUQ to estimate 8 ELM parameters;
- Observation: Annual averaged latent heat flux (LH) for 5 years at the Missouri Ozark AmeriFlux site in 2006-2010;
- Prior sample: 1000 paired samples  $\mathcal{D}_{\text{prior}} = \{(x_j, y_j)\}_{j=1}^J$
- Two case studies:
  - Synthetic case for method verification
  - Real observations application
- Compare DBUQ with MCMC for performance evaluation

Parameter name	Parameter range
<i>rootb_par</i>	[0.5, 4]
<i>slatop</i>	[0.01, 0.05]
<i>flnr</i>	[0.1, 0.4]
<i>frootcn</i>	[25, 60]
<i>froot_leaf</i>	[0.3, 1.5]
<i>br_mr</i>	[1.5e-6, 4e-6]
<i>crit_dayl</i>	[35000, 45000]
<i>crit_onset_gdd</i>	[600, 1000]

## DBUQ

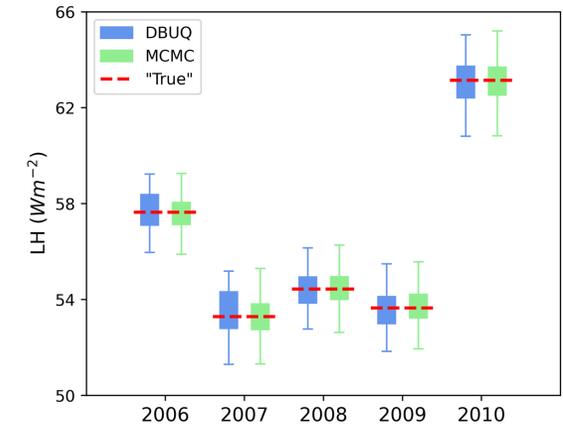
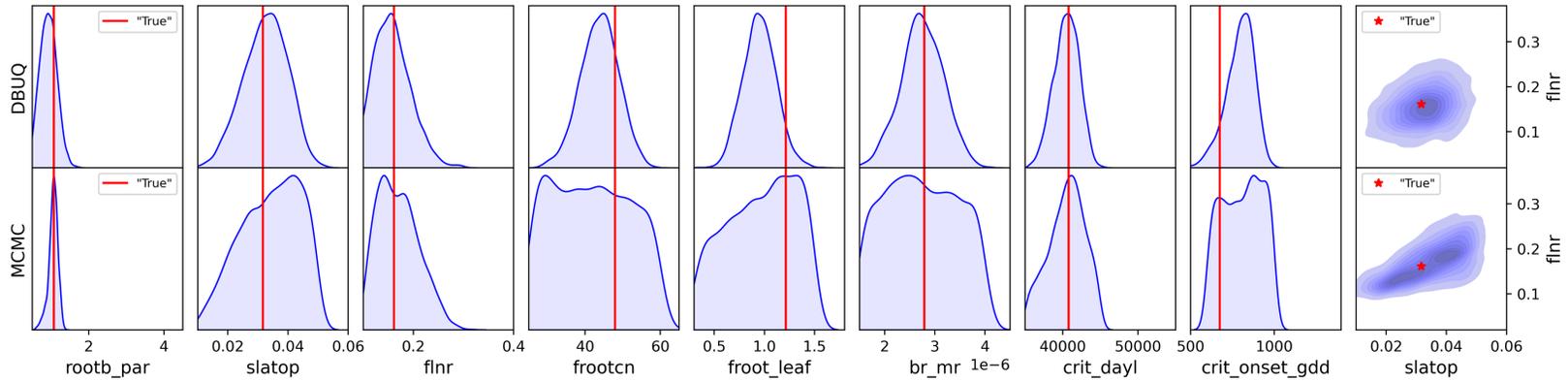
- Input: 1000 prior samples  $\mathcal{D}_{\text{prior}} = \{(x_j, y_j)\}_{j=1}^J$
- Output: a **trained generator** which can be quickly evaluated to generate target samples for any given observations;
- Computing time: < 10 min for solving both cases
- **Particularly suitable for site-specific earth system model calibration at a global scale** due to its computational efficiency and amortized inference.

## MCMC

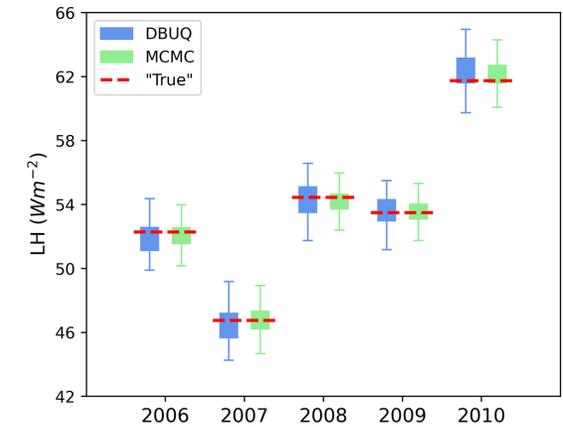
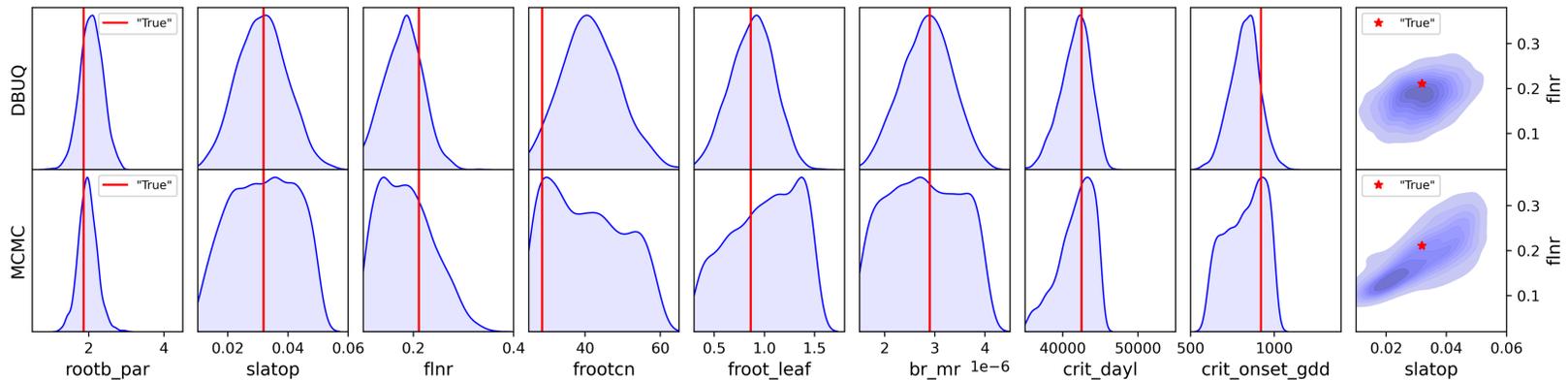
- Input: 1000 prior samples  $\mathcal{D}_{\text{prior}} = \{(x_j, y_j)\}_{j=1}^J$
- Procedure: build a surrogate model on the prior samples, and then perform MCMC simulations on the surrogate;
- Output: a **set of posterior samples**; For a different observation, we need to re-run MCMC;
- Computing time: ~ 5 hours for one case to generate the same number of posterior samples as DBUQ.

# DBUQ accurately and efficiently estimates parameter PDFs

## Synthetic case I



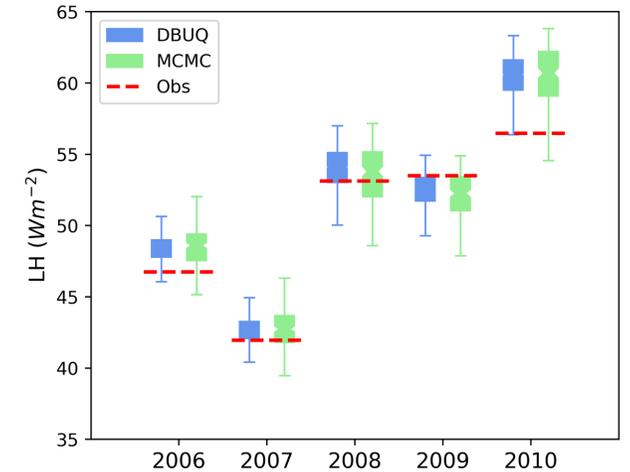
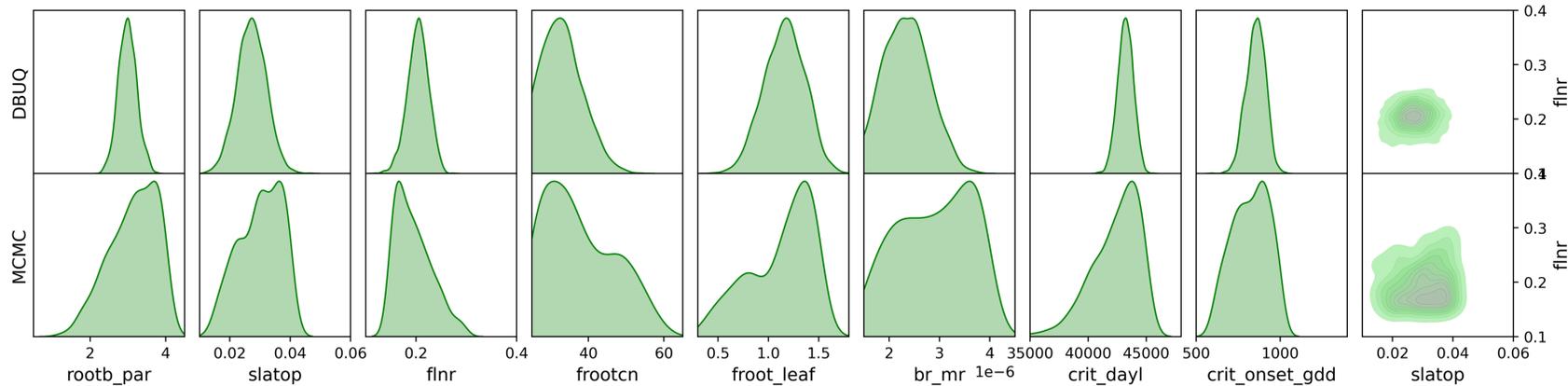
## Synthetic case II



- ❖ DBUQ shows high accuracy in approximating the parameter posterior distributions.
  - Similar to the MCMC results, both accurately estimate the “true” parameter values with high probability.
- ❖ DBUQ demonstrates an accurate model calibration, as the prediction samples simulated from the parameter posterior samples are closely around the “true” observation.

# DBUQ accurately and efficiently estimates parameter PDFs

## Real observation case



- ❖ DBUQ again shows high accuracy in approximating the parameter posterior distributions.
  - Similar to the MCMC results in estimating both the marginal and joint PDFs.
- ❖ DBUQ demonstrates an accurate model calibration, as the prediction samples simulated from the parameter posterior samples are closely around the observation.
- ❖ Note, DBUQ achieves the similar accurate results with MCMC but using much less computational time.
  - DBUQ: 10 min for all the three case studies;
  - MCMC: 5 hours for one case studies;

# Summary

- We developed a novel diffusion-based uncertainty quantification method for parameter estimation and model calibration.
- The method is accurate and computationally efficient.
- It can be generally applied to site-specific earth system model calibration on a global scale, which paves the way for more effective and timely climate impact analyses.