USE OF EMULATORS IN UNCERTAINTY QUANTIFICATION

George Pau, William Riley, Chaopeng Shen, Daniel Walton, Yaning Liu
Lawrence Berkeley National Laboratory

2015 AGU Fall Meeting, San Francisco, CA.

Acknowledgment

- Funding for this study was provided by the US Department of Energy, BER Program, contract # DE-AC02-05CH11231.

Two aspects to consider

- From iLAMB paper (Luo et al., 2012):
  - Effective benchmarks should draw upon a broad set of independent observations spanning multiple temporal and spatial scales.
  - Calibration is an integral part of benchmarking.
- Difficulties:
  - Computationally demanding to simulate at observed temporal and spatial scales.

    Proper orthogonal decomposition (PODM) method
  - Repeated evaluations during calibration and UQ analysis compound the computational cost.

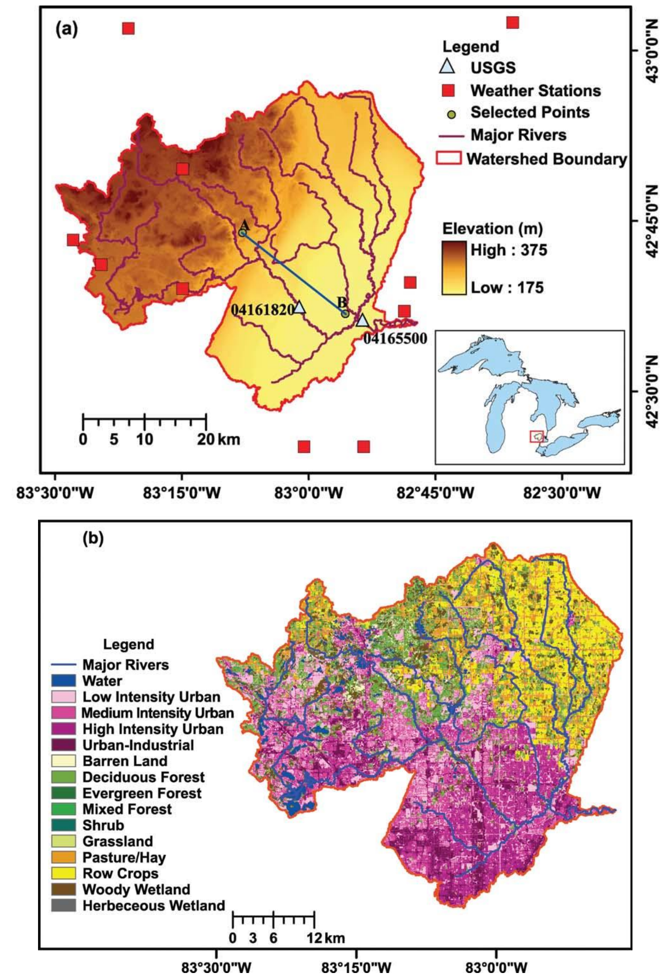
    Implicit sampling and surrogate models.

Surrogate models

- Reduced order models:
  - Efficient surrogates to expensive numerical models.
  - Numerical/statistical models.
  - Sampling-based construction: typically based on results from simulation.
- Proper orthogonal decomposition mapping method
  - Utilizes solutions from coarse-resolution models to reconstruct solutions from fine-resolution models.
  - Minimal changes in existing codes.
- Combination of POD and responses surface methods.
  - E.g. GPR, gPCE, HDMR.

Predict high-resolution NPP of a watershed model

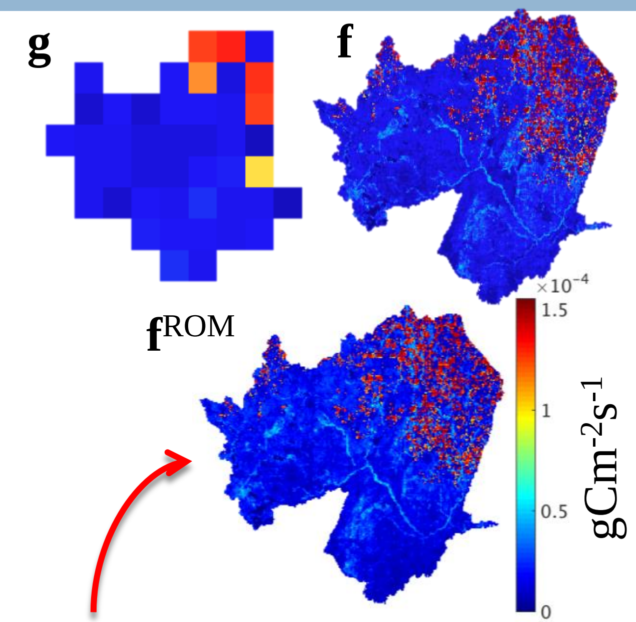
- Clinton River Watershed:
  - Large seasonal temperature change (22°C/summer and -3°C/winter).
  - Heterogeneous land covers.
  - Simulated by CLM+PAWS.
- No fitting of coarse model's parameters to fine model solution.
- PODM model's training data:
  - coarse (7km) and fine (0.2km) resolution.
  - daily solutions from 2001-2005.

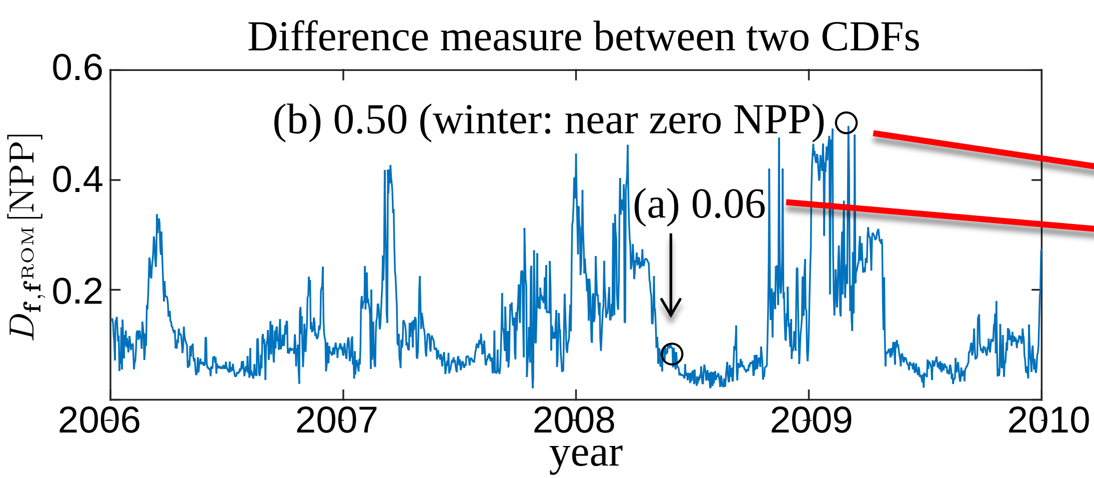


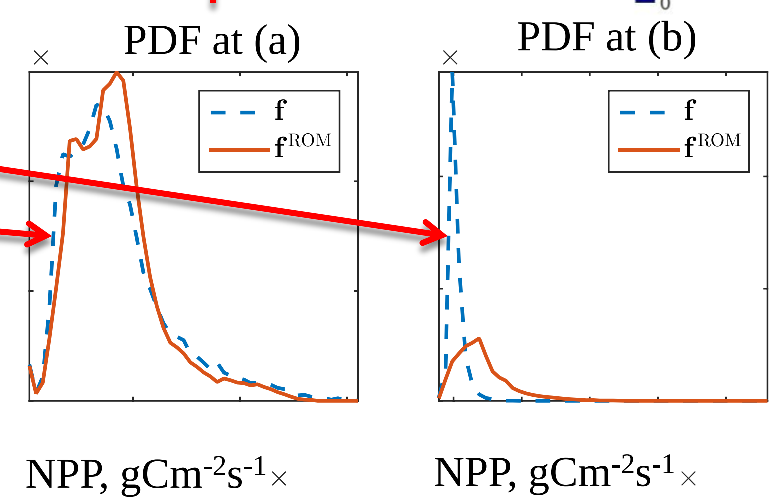
Pau et al., WRR, 52, 791-812, 2016.
Pau et al., Geosci. Model Dev., 7, 2091-2105, 2014.

Predict high-resolution NPP of a watershed model

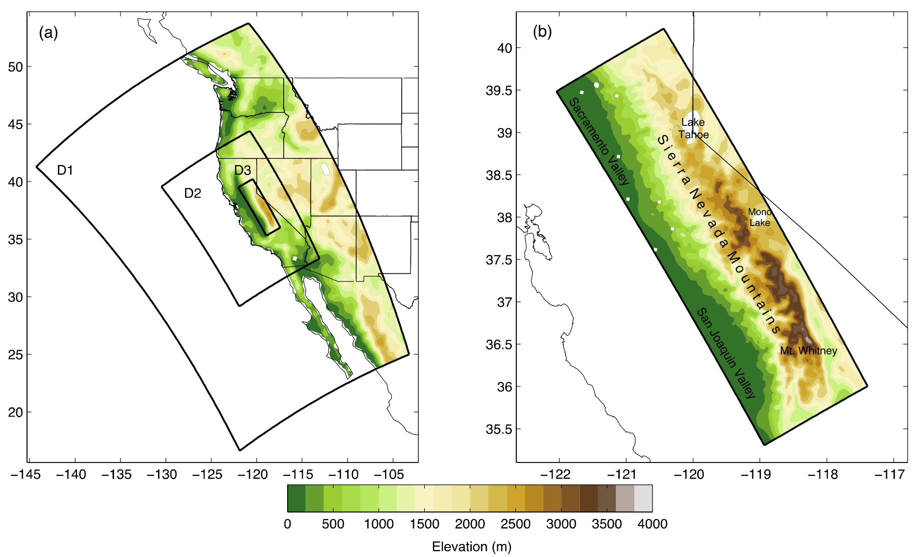
- Good prediction of daily fine scale solution for 2006-2009 using coarse model.
- But what is good?
  - We used the Kolmogorov-Smirnov measure.
  - Potential use of diagnostics in CODA R package.
  - Relevant for statistical benchmarking against fine-scale observations.



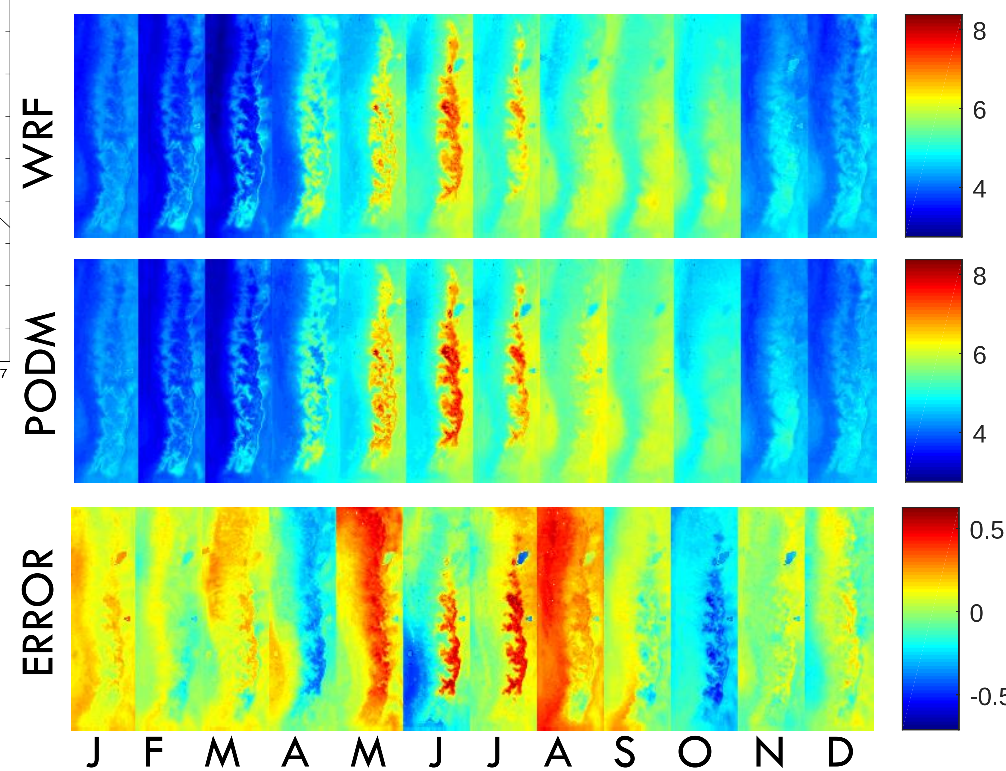




Quantifying impacts of climate change on Sierra Nevada



Average monthly temperature change between 2091-2100 and 1991-2000



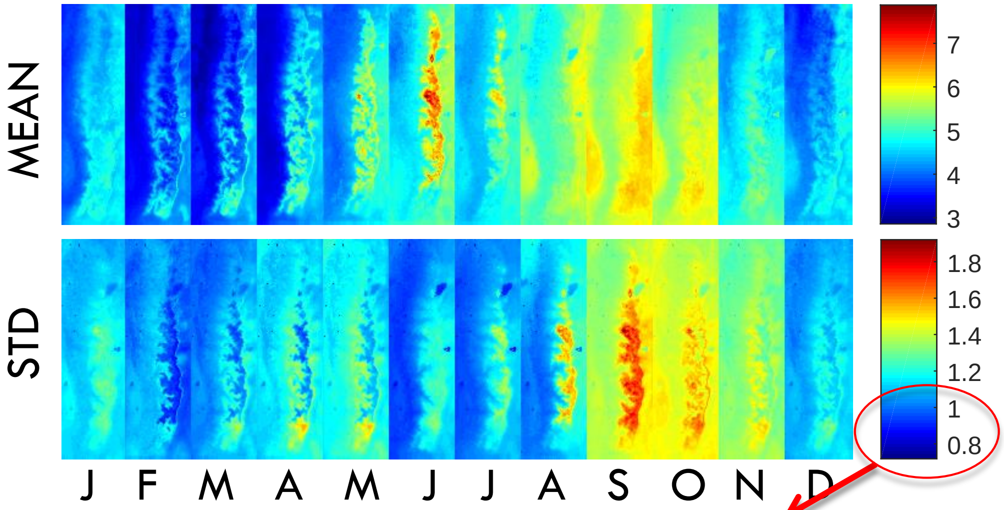
- Training data:

  Monthly air temperature from
  - climate change scenarios (5 different GCM models),
  - dynamically downscaled WRF solutions of Sierra Nevada.

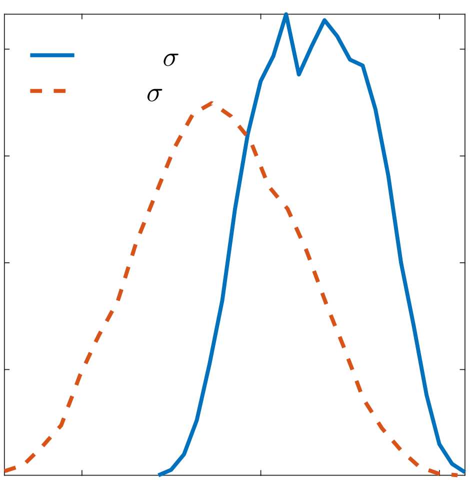
- Leave-one-out validation:
  - average MAE error = 0.4°C.

Quantifying impacts of climate change on Sierra Nevada

- Impact of climate change on high-resolution temperature of the Sierra Nevada:
  - PODM model rapidly determines the temperature based on RCP8.5 results from 35 different GCMs.



The standard deviations are greater than the error of the PODM model: the accuracy of the PODM model is sufficient for statistical analysis.



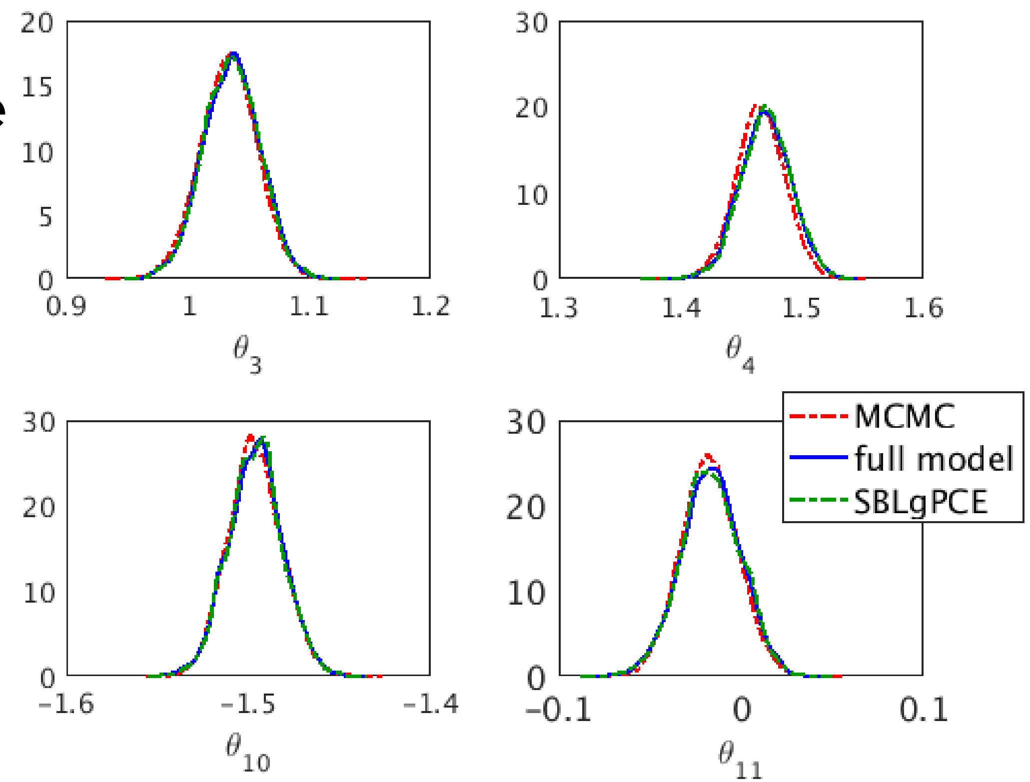
Resampling the results from the 35 GCMs allows rapid reconstruction of the PDF at point locations, potentially at resolution consistent with observations.

Perspectives

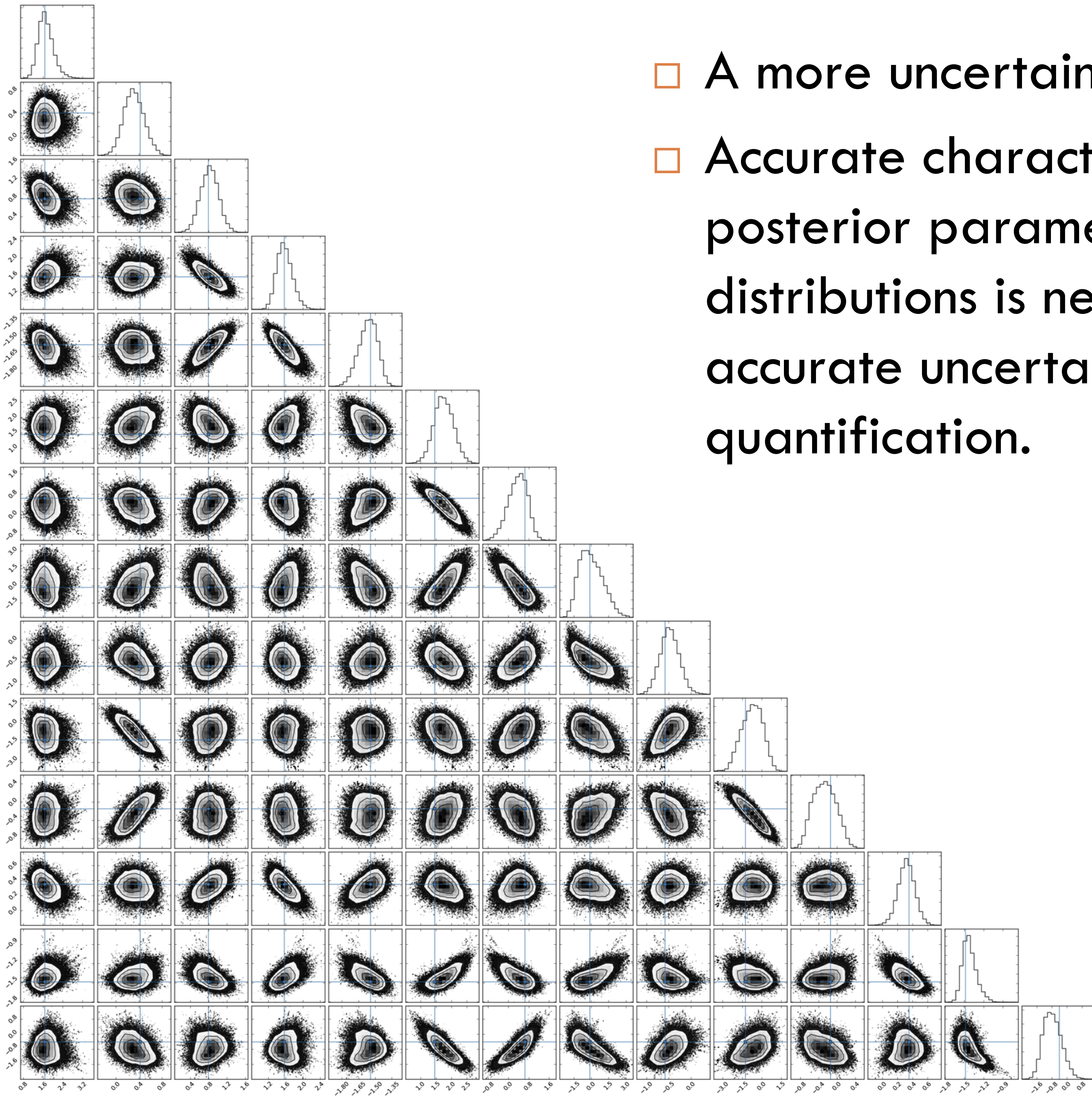
- Use of ROM in uncertainty analysis
  - Efficient high resolution reconstruction allows uncertainties to be quantified at scales consistent with the observations.
  - Error estimation and robust validation are needed, similar to a physics-based model.
- Construction of ROM
  - ROM must be paired with appropriate UQ methods to allow efficient sampling and thus reduce computational overhead.
  - High-resolution reconstruction is still a challenging problem: appropriate evaluation criteria consistent with the application ensure the ROM has the required accuracy.

Bayesian calibration

- Implicit sampling: a particle filtering technique
  - deterministic optimization + importance sampling techniques.
- Work well with ROM.
  - Constrained parameter space for training dataset
  - Highly parallelizable.
  - Construct ROM for outputs, not likelihood function.



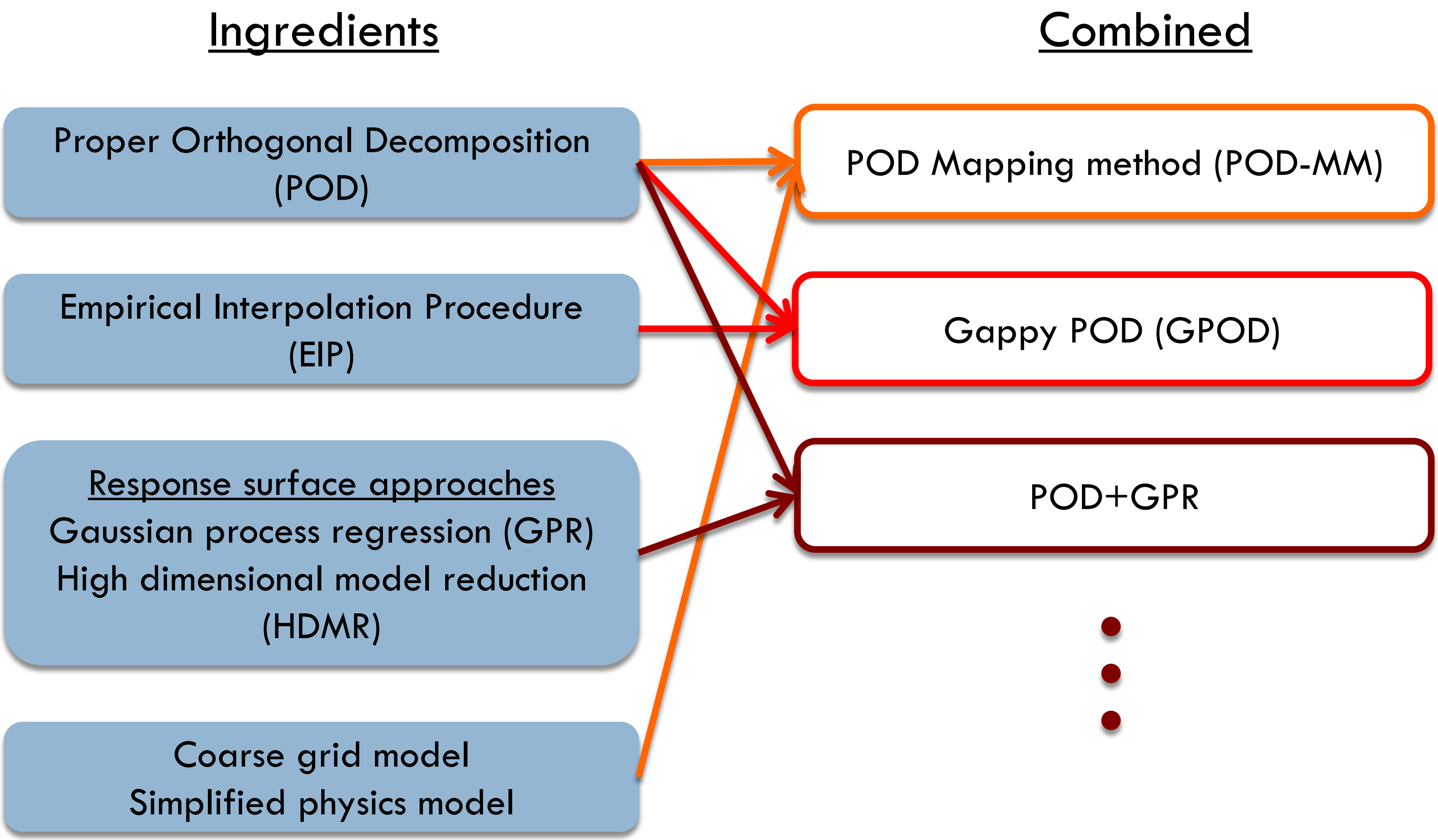
Implemented within Agni/ASCEM



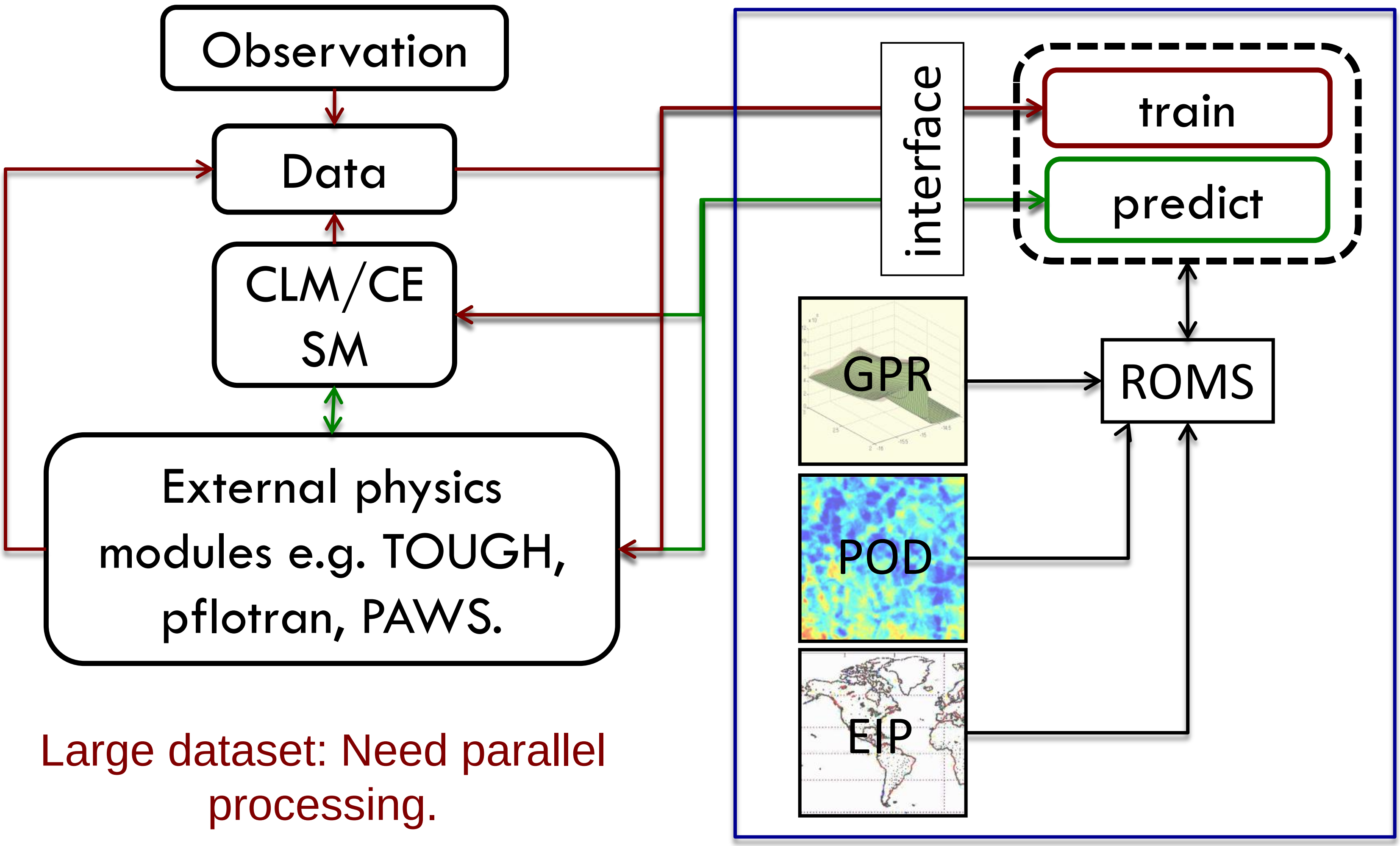
- A more uncertain problem.
- Accurate characterizations of posterior parameter distributions is needed for accurate uncertainty quantification.

Thank you.

ROM Approaches



pROME Parallel Reduced Order Models for Earth Systems



Large dataset: Need parallel processing.