

# Embedded Model Error Representation for Bayesian Model Calibration

*Khachik Sargsyan, Xun Huan, Habib Najm*



Sandia National Laboratories

Livermore, CA

*Statistics When the Model is Wrong*

Radcliffe Institute

May 31 - June 1, 2018

# Acknowledgements

B. Debusschere, C. Safta, — SNL, CA

M. Eldred, G. Geraci — SNL, NM

R. Ghanem — USC

Y. Marzouk, C. Feng — MIT

D. Ricciuto, P. Thornton – ORNL

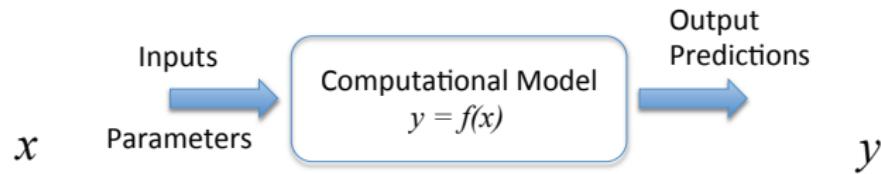
J. Bender – LLNL

This work was supported by:

- DOE, Advanced Scientific Computing Research (ASCR), SciDAC
- DOE, Basic Energy Sciences (BES)
- DOE, Biological and Environmental Research (BER)
- DOD, DARPA Enabling Quantification of Uncertainty in Physical Systems (EQUIPS) program

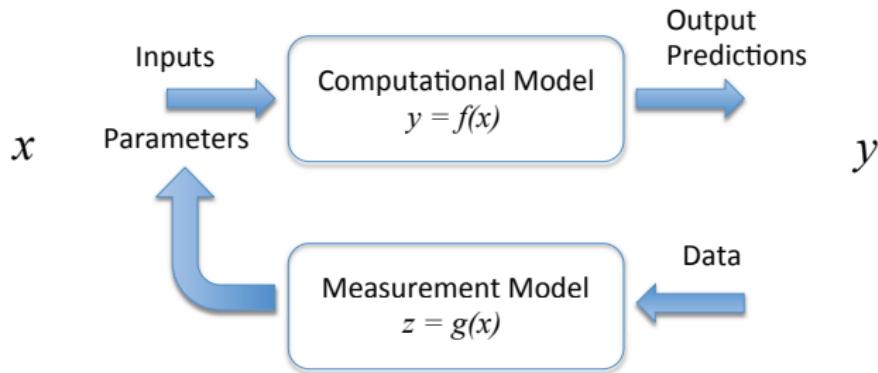
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

# Uncertainty Quantification and Computational Science



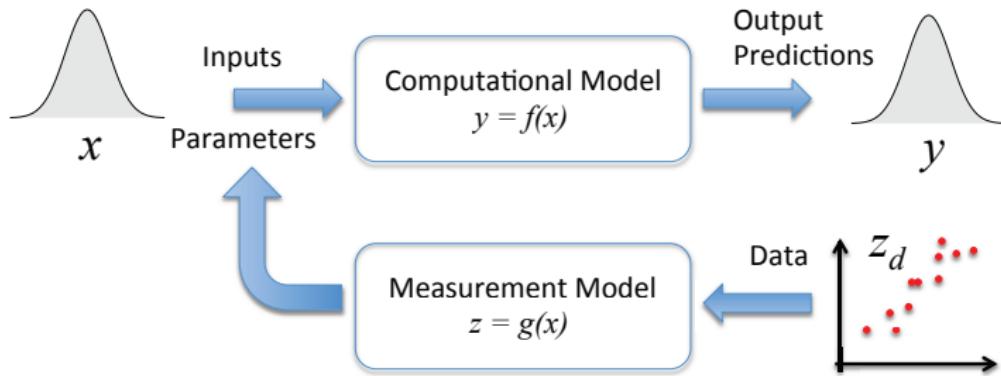
Forward problem

# Uncertainty Quantification and Computational Science



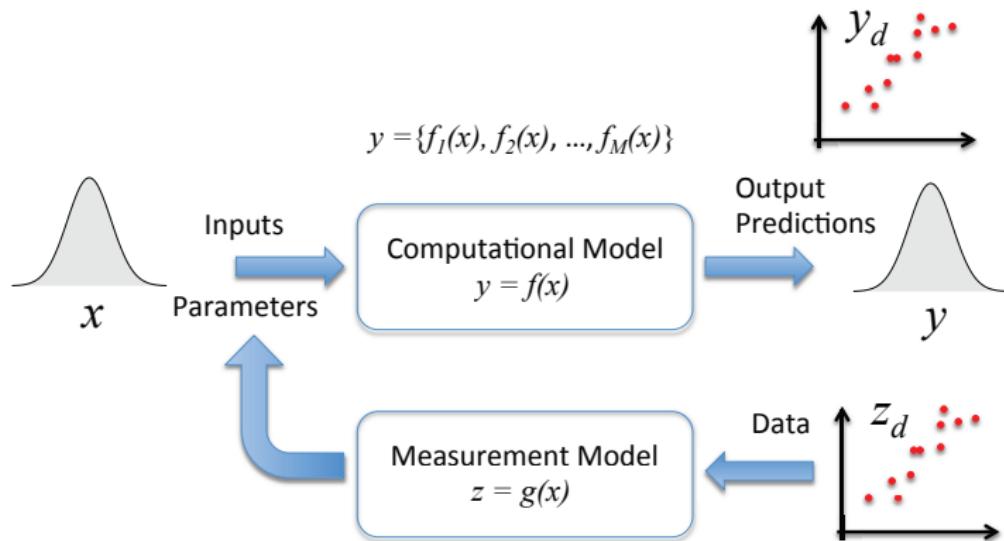
Inverse & Forward problems

# Uncertainty Quantification and Computational Science



Inverse & Forward UQ

# Uncertainty Quantification and Computational Science



Inverse & Forward UQ  
Model validation & comparison, Hypothesis testing

# The Case for Uncertainty Quantification

## UQ needed for...

- Model predictions
- Model validation and comparison
- Confidence assessment
- Reliability analysis
- Dimensionality reduction
- Optimal design
- Decision support
- (Noisy) data assimilation

## Uncertainty Sources

- Model parameters
- Initial/boundary conditions
- Model geometry/structure
- Lack of knowledge
- Data noise
- Intrinsic stochasticity
- Numerical errors, too

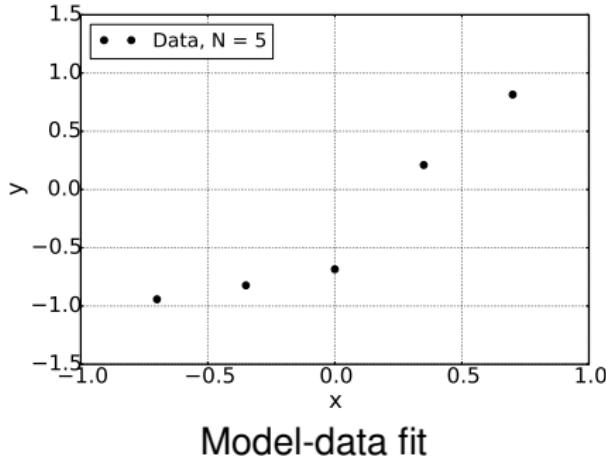
# Main target: model error

$$g(x) \approx f(x; \lambda)$$

deviation from ‘truth’ or from a higher-fidelity model

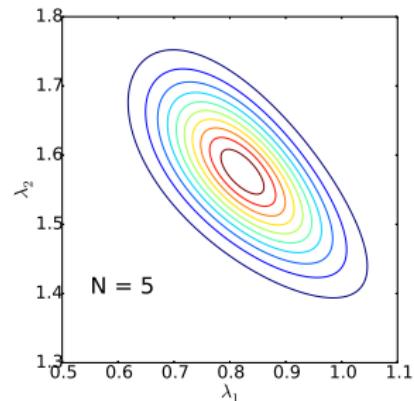
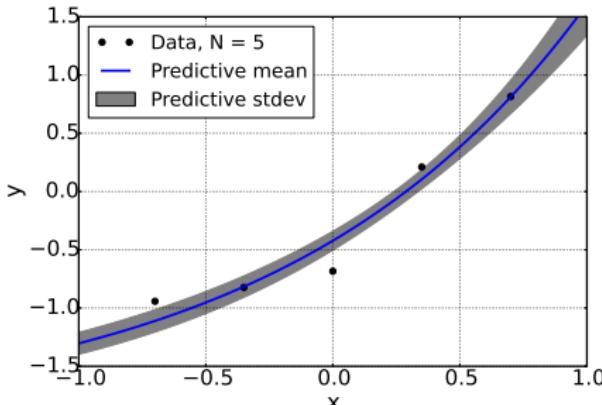
- ... otherwise called (with slightly altered meanings):  
model discrepancy, model structural error,  
model inadequacy, **model misspecification**,  
model form error, model uncertainty
- Inverse modeling context
  - Given experimental or higher-fidelity model data,  
estimate the model error
- Represent and estimate the error associated with
  - Simplifying assumptions, parameterizations
  - Mathematical formulation, theoretical framework
- ...will be useful for
  - Model validation
  - Model comparison
  - Scientific discovery and model improvement
  - Reliable computational predictions

# Ignoring model error leads to overconfident and biased predictions



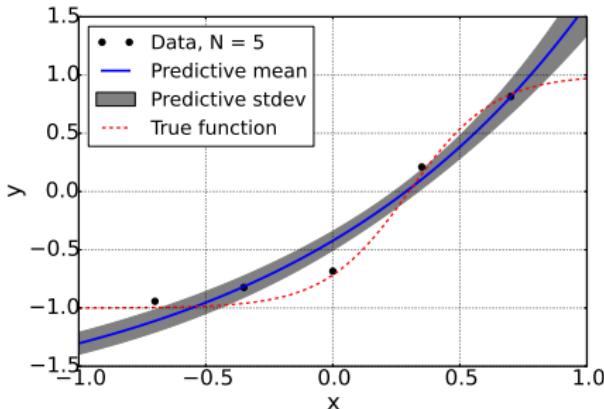
- Given noisy data, calibrate an exponential model:  $g(x) \approx f(x; \lambda)$

# Ignoring model error leads to overconfident and biased predictions

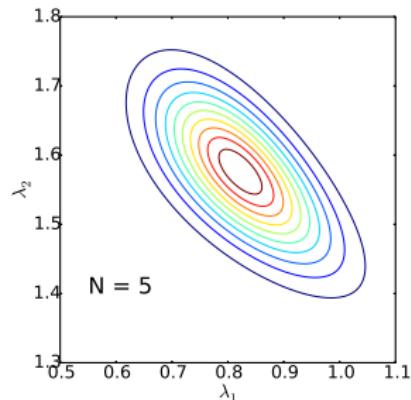


- Given noisy data, calibrate an exponential model:  $g(x) \approx f(x; \lambda)$
- Employ Bayesian inference to obtain posterior PDFs on  $\lambda$

# Ignoring model error leads to overconfident and biased predictions



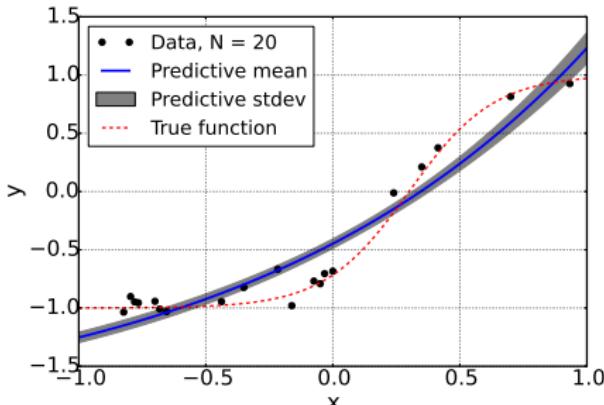
Model-data fit



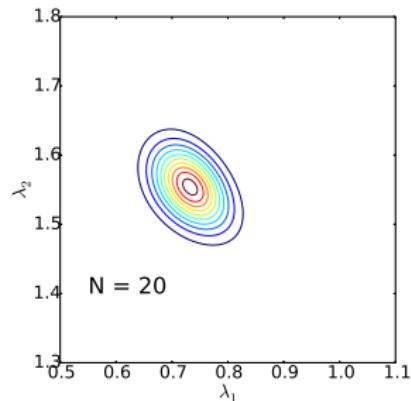
Posterior on parameters

- Given noisy data, calibrate an exponential model:  $g(x) \approx f(x; \lambda)$
- Employ Bayesian inference to obtain posterior PDFs on  $\lambda$
- True model – dashed-red – is *structurally* different from fit model  $f(x, \lambda)$

# Ignoring model error leads to overconfident and biased predictions



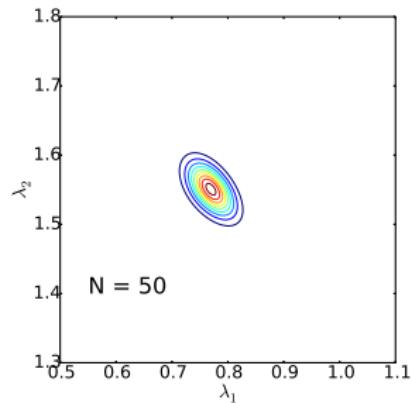
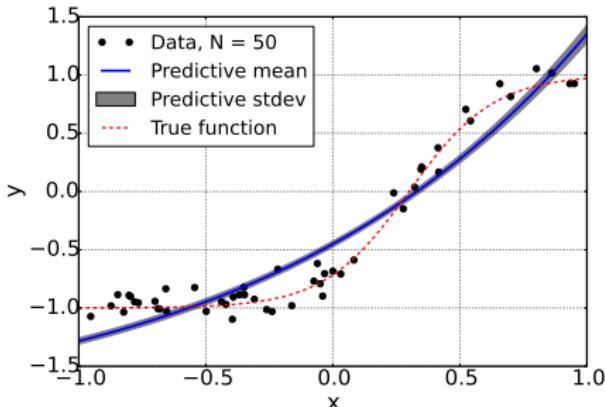
Model-data fit



Posterior on parameters

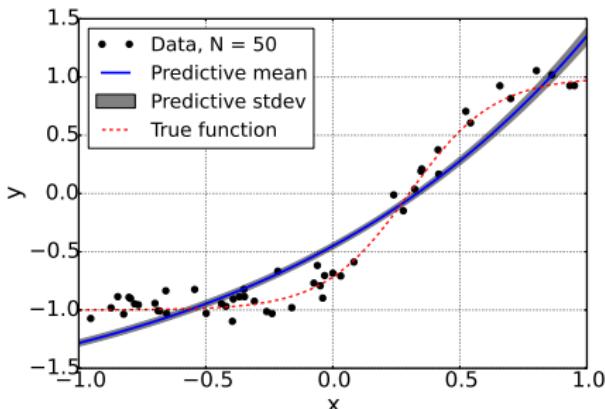
- Given noisy data, calibrate an exponential model:  $g(x) \approx f(x; \lambda)$
- Employ Bayesian inference to obtain posterior PDFs on  $\lambda$
- True model – dashed-red – is *structurally* different from fit model  $f(x, \lambda)$
- Higher data amount reduces posterior and predictive uncertainty
  - Increasingly sure about predictions based on the *wrong* model

# Ignoring model error leads to overconfident and biased predictions

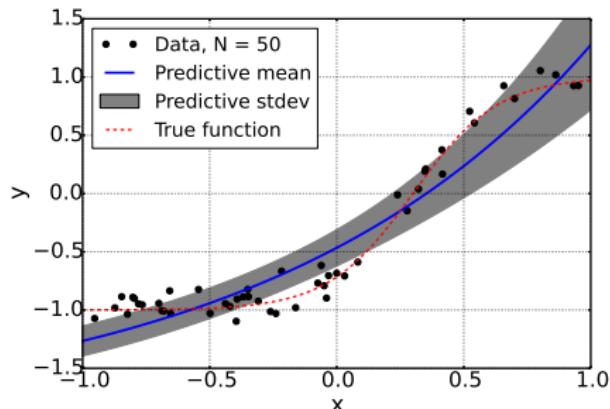


- Given noisy data, calibrate an exponential model:  $g(x) \approx f(x; \lambda)$
- Employ Bayesian inference to obtain posterior PDFs on  $\lambda$
- True model – dashed-red – is *structurally* different from fit model  $f(x, \lambda)$
- Higher data amount reduces posterior and predictive uncertainty
  - Increasingly sure about predictions based on the *wrong* model

# Ignoring model error leads to overconfident and biased predictions



No model error treatment



Model error accounted for

- Given noisy data, calibrate an exponential model:  $g(x) \approx f(x; \lambda)$
- Employ Bayesian inference to obtain posterior PDFs on  $\lambda$
- True model – dashed-red – is *structurally* different from fit model  $f(x, \lambda)$
- Accounting for model error allows extra uncertainty component to propagate through predictions

# Where to put model error?

$$y_i = f(x_i; \lambda) + \delta(x_i) + \epsilon_i$$

- Outside:

- Explicit GP representation [Kennedy-O'Hagan, 2001]
- See also [Higdon et. al, 2004], [Bayarri et. al, 2007]
- Usage: too many to cite
- Issues: see next slide
- Variants exist: multiplicative noise, non-linear maps etc.

$$y_i = \tilde{f}(x_i; \lambda, \delta(x_i)) + \epsilon_i$$

- Inside:

- Engineering/statistical adjustment [Joseph and Melkote, 2009]
- Additive corrections to submodels [Strong et. al, 2011]
- Validation of extrapolative predictions [Oliver et. al, 2014]
- Field inversion and machine learning [Duraisamy et. al, 2015-]
- Hybrid correction [He and Xiu, 2016]
- Random field correction [Brown and Atamturktur, 2016]
- Hierarchical mixture model [Feng, 2017]
- Parameter inflation [Pernot et. al, 2017]
- Hierarchical stochastic model [Wu et. al, 2017]
- Increased use, especially in physical models: [Emory et. al, 2011] [Oliver and Moser, 2011], [Morrison et. al, 2016], [Sondak et. al, 2017], [Huan et. al, 2017]...

# External correction often not satisfactory for physical models

$$y_i = \underbrace{f(x_i; \lambda) + \delta(x_i)}_{\text{truth } g(x_i)} + \epsilon_i$$

- Explicit additive statistical model for model error [KOH, 2001]
- Potential violation of physical constraints
- Disambiguation of model error  $\delta(x_i)$  and data error  $\epsilon_i$
- Yes, priors help: [Brynjarsdottir and O'Hagan, 2014], [Plumlee, 2017]
- Calibration of model error on measured observable does not impact the quality of model predictions on other Qols
- Physical scientists are unlikely to augment their model with a statistical model error term on select outputs
  - Calibrated predictive model:  $f(x; \lambda) + \delta(x)$  or  $f(x; \lambda)$  ?
- Problem is highlighted in model-to-model calibration ( $\epsilon_i = 0$ )
  - no a priori knowledge of the statistical structure of  $\delta(x)$

# Case for Model Error Embedding

Ideally, modelers want predictive *errorbars*:  
inserting randomness on the outputs has issues, so...

$$y_i = \tilde{f}(x_i; \lambda, \delta_\alpha) + \epsilon_i$$

- Embed model error in specific submodel phenomenology
  - a modified transport or constitutive law
  - a modified formulation for a material property
  - turbulent model constants
- Allows placement of model error term in locations where key modeling assumptions and approximations are made
  - as a correction or high-order term
  - as a possible alternate phenomenology
- Naturally preserves model structure and physical constraints
- Disambiguates model/data errors

# Embedded Model Error Options

- Explore different model forms,

*Intrusive*

$$y_i = \tilde{f}(x_i; \lambda, \delta_\alpha(x_i)) + \epsilon_i$$

- 
- Additive stochastic corrections to existing inputs

*Non-intrusive*

$$y_i = f(x_i; \lambda + \delta_\alpha(x_i)) + \epsilon_i$$

- ... even simpler,  $x$ -independent

$$y_i = f(x_i; \lambda + \delta_\alpha) + \epsilon_i$$

# Bayesian Framework for Model Error Estimation

$$y_i = f(x_i; \lambda + \delta_\alpha) + \epsilon_i$$

- Given data  $y_i$ , perform *simultaneous* estimation of  $\tilde{\alpha} = (\lambda, \alpha)$ , i.e. model parameters  $\lambda$  and model-error parameters  $\alpha$ .
- Bayes' theorem

$$\widehat{p(\tilde{\alpha}|y)} = \frac{\underbrace{p(y|\tilde{\alpha})}_{\text{Likelihood}} \underbrace{p(\tilde{\alpha})}_{\text{Prior}}}{\underbrace{p(y)}_{\text{Evidence}}}$$

- In order to estimate the likelihood  $L_y(\tilde{\alpha}) = p(y|\tilde{\alpha}) = p(y|\lambda, \alpha)$ , one needs uncertainty propagation through  $f(x_i; \underbrace{\lambda + \delta_\alpha}_{\text{stochastic}})$ ,
- ... hence, we employ Polynomial Chaos (PC) representation for  $\delta_\alpha$ .

# Polynomial Chaos Representation of Augmented Input

$$y_i = f(x_i; \lambda + \delta_\alpha) + \epsilon_i$$

- Zero-mean PC form  $\delta_\alpha = \sum_{k=1}^K \alpha_k \Psi_k(\xi)$
- Functional representation of a large class of random variables
- The PC *germ*  $\xi$  is a standard random variable
  - e.g. Uniform( $-1, 1$ ) or Normal( $0, 1$ )
- The PC bases (e.g. Legendre or Hermite polynomials) are orthogonal w.r.t. PDF of  $\xi$

$$\int \Psi_m(\xi) \Psi_k(\xi) \pi_\xi(\xi) d\xi = 0 \quad \text{for } m \neq k.$$

- PC representation allows efficient
  - Sampling
  - Moment estimation
  - Variance-based decomposition
  - Uncertainty propagation (via NISP)

# Model Error – Likelihood construction

$$y_i = f(x_i; \lambda + \delta_\alpha(\zeta)) + \epsilon_i = f_i(\tilde{\alpha}, \zeta) + \epsilon_i$$

- Likelihood  $\mathcal{L}_g(\tilde{\alpha}) = p(y|\tilde{\alpha})$  challenging, but can compute moments

$$\mu_i(\tilde{\alpha}) = \mathbb{E}_\zeta[f_i(\tilde{\alpha}, \zeta)] \quad \text{and} \quad \sigma_i^2(\tilde{\alpha}) = \mathbb{V}_\zeta[f_i(\tilde{\alpha}, \zeta)] + s_i^2$$

- Gauss-Marginal Approximate Likelihood compares data  $y_i$  and model predictions:

$$\mathcal{L}_g(\tilde{\alpha}) \approx \frac{1}{(2\pi)^{N/2}} \prod_{i=1}^N \frac{1}{\sigma_i(\tilde{\alpha})} \exp\left(-\frac{1}{2} \left(\frac{y_i - \mu_i(\tilde{\alpha})}{\sigma_i(\tilde{\alpha})}\right)^2\right)$$

- Non-intrusive spectral projection (NISP) with Polynomial Chaos

$$f_i(\tilde{\alpha}, \zeta) \stackrel{\text{NISP}}{\simeq} \sum_k f_{ik}(\tilde{\alpha}) \Psi_k(\zeta)$$

- ... provides easy access to mean and variance

$$\mu_i(\tilde{\alpha}) = f_{i0}(\tilde{\alpha}) \quad \text{and} \quad \sigma_i^2(\tilde{\alpha}) = \sum_{k \neq 0} f_{ik}^2(\tilde{\alpha}) ||\Psi_k||^2 + s_i^2$$

# Model Error – Surrogate and Prediction

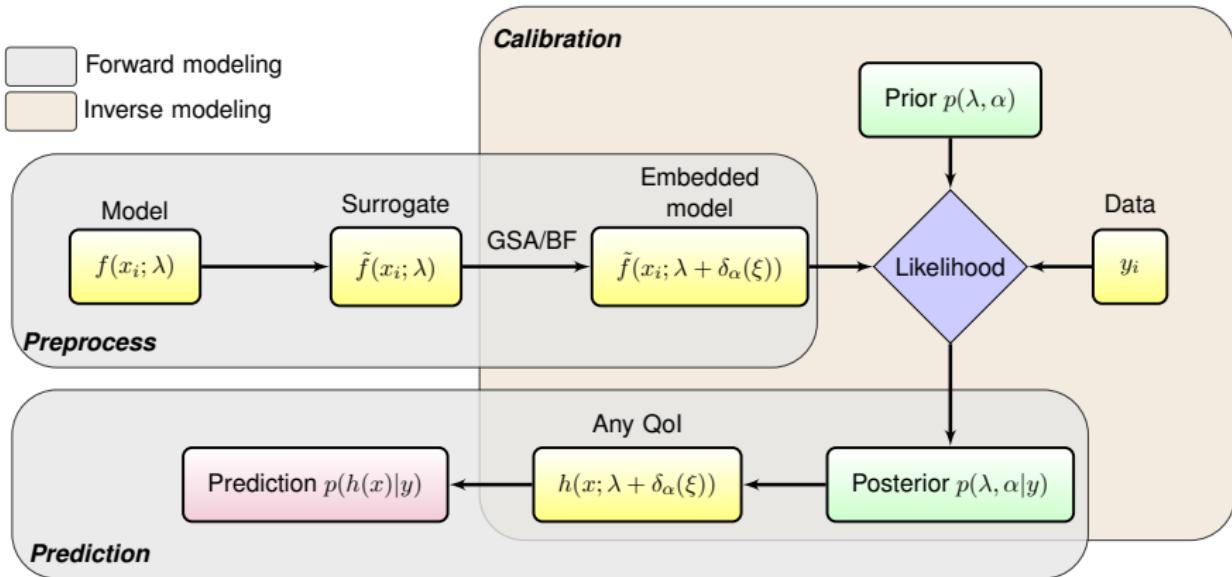
$$f_i(\lambda + \delta_\alpha(\zeta)) = f_i(\tilde{\alpha}, \zeta) \stackrel{\text{NISP}}{\simeq} \sum_k f_{ik}(\tilde{\alpha}) \Psi_k(\zeta)$$

- NISP is employed both for likelihood computation and for posterior/pushed-forward predictions in general
- In practice,  $f_i(\cdot)$  is replaced by a pre-constructed polynomial surrogate
- Note: NISP with finite truncation is exact,  
if one truncates NISP at the same order as the surrogate of  $f_i(\cdot)$
- Posterior predictive moments

$$\mu_i = \mathbb{E}_{\tilde{\alpha}} [\mu_i(\tilde{\alpha})]$$

$$\sigma_i^2 = \underbrace{\mathbb{E}_{\tilde{\alpha}} [\sigma_i^2(\tilde{\alpha})]}_{\text{Model error}} + \underbrace{\mathbb{V}_{\tilde{\alpha}} [\mu_i(\tilde{\alpha})]}_{\text{Posterior uncertainty}} + \underbrace{(\sigma_i^{LOO})^2}_{\text{Surrogate error}} + \underbrace{s_i^2}_{\text{Data noise}}$$

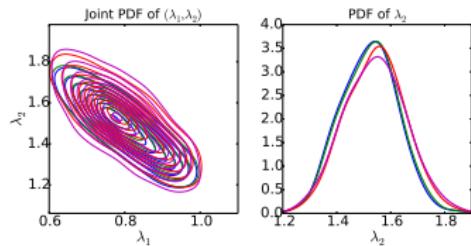
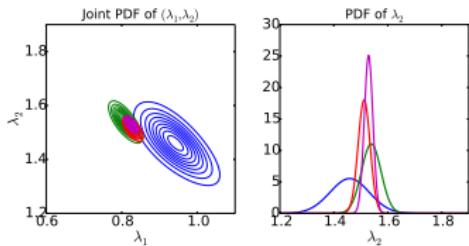
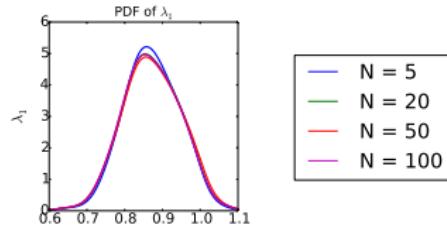
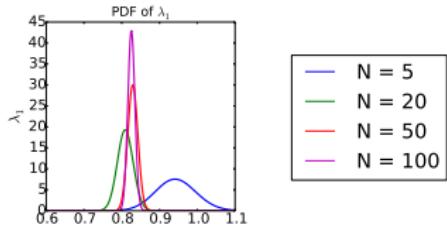
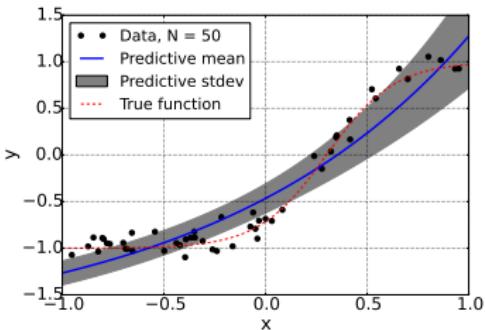
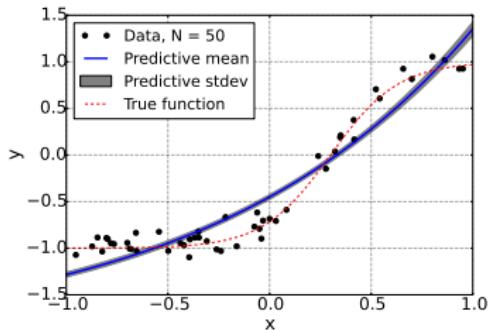
# Model error embedding – workflow



- Predictive uncertainty decomposition: Total Variance =

Posterior uncertainty + Data noise + Model error + Surrogate error

# .. back to toy example



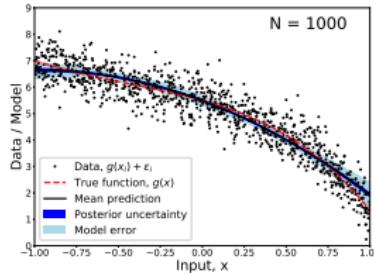
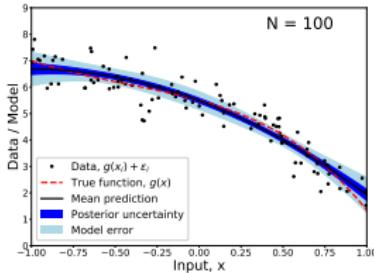
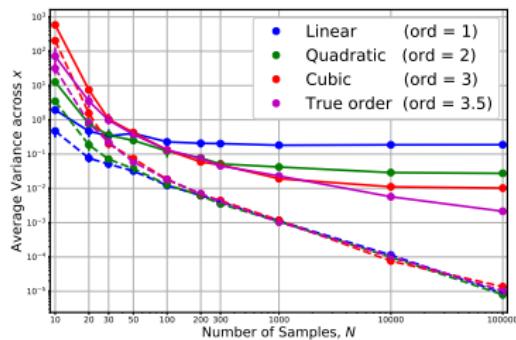
# More data leads to ‘leftover’ model error

Calibrating a quadratic  $f(x) = \lambda_0 + \lambda_1 x + \lambda_2 x^2$

w.r.t. ‘truth’  $g(x) = 6 + x^2 - 0.5(x + 1)^{3.5}$  measured with noise  $\sigma = 0.1$ .

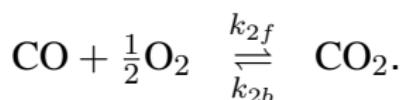
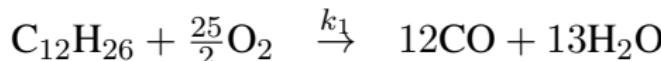
## Summary of features:

- Well-defined model-to-model calibration
- Model-driven discrepancy correlations
- Respects physical constraints
- Disambiguates model and data errors
- Calibrated predictions of multiple QoIs



# Ignition time in chemical kinetics

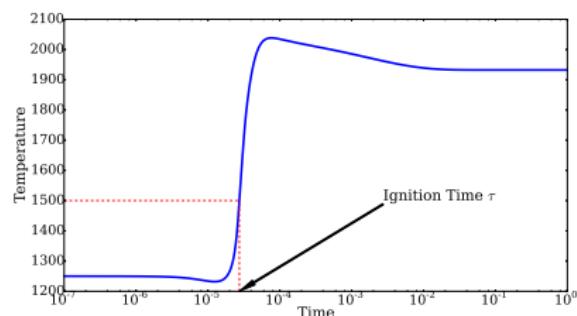
- Two-step global reaction model calibrated against shock tube experimental data
- Operating conditions: pressure  $P$ , initial temperature  $T_0$  & equivalence ratio  $\phi$



$$k_1 = A e^{-\left(\frac{E}{RT}\right)} [\text{C}_{12}\text{H}_{26}]^{0.25} [\text{O}_2]^{1.25}$$

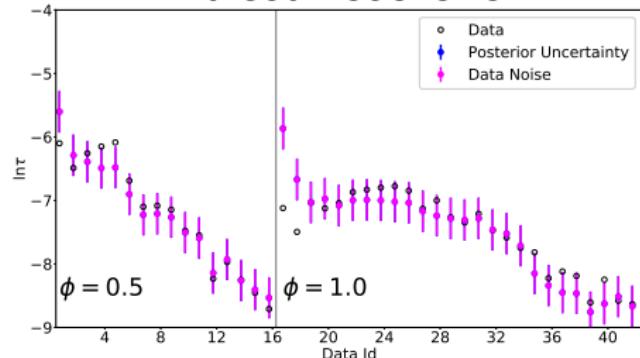
- Data:  $\log(\text{ignition time})$
- Embedding

$$(\ln A, E) = \sum_k \alpha_k \Psi_k(\xi)$$

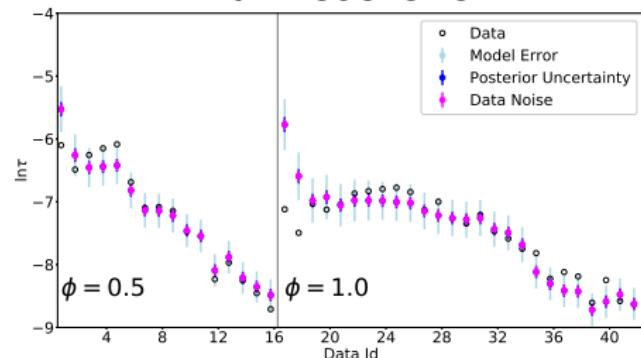


# Ignition time in chemical kinetics

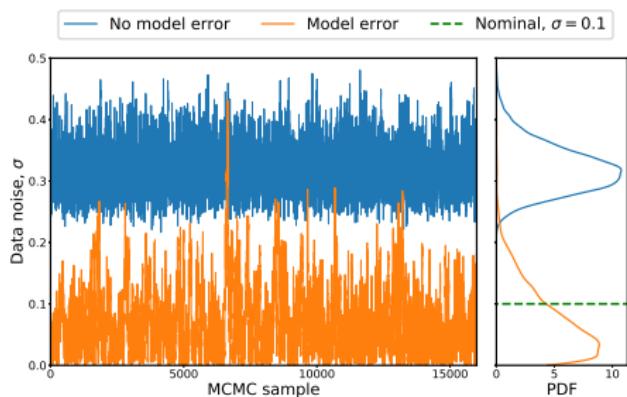
Without model error



With model error

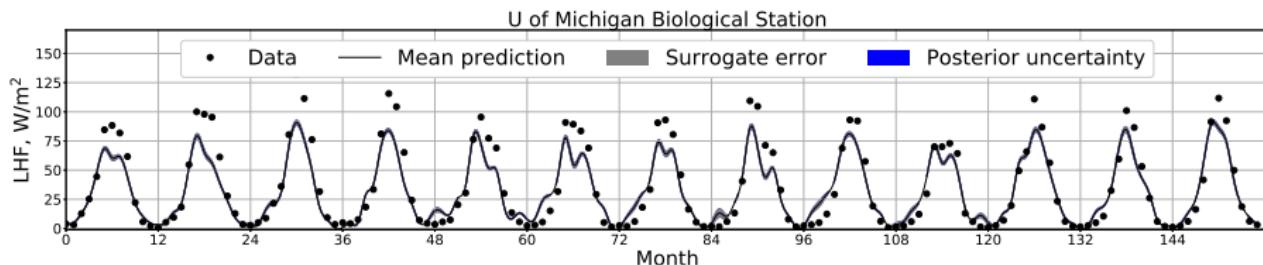


- Model error disambiguated from data error
- Data error correctly captured
- Meaningful extrapolative predictions



# E3SM Land Model (ELM)

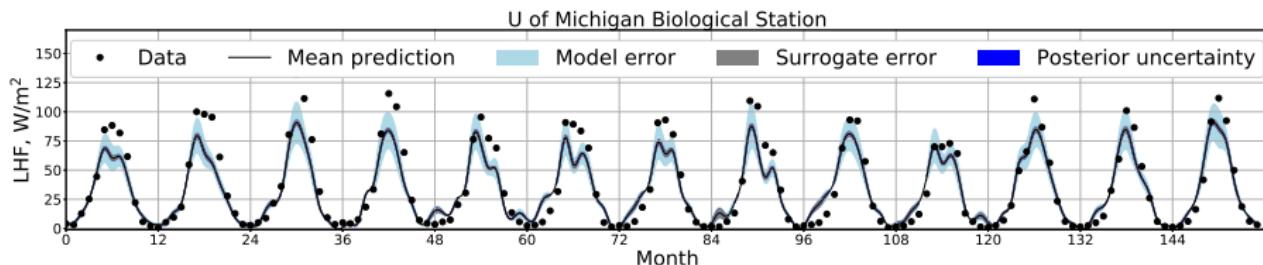
- US Department of Energy (DOE) sponsored Earth system model
- Land, atmosphere, ocean, ice, human system components
- High-resolution, employ DOE leadership-class computing facilities



- Conventional calibration without model error

# E3SM Land Model (ELM)

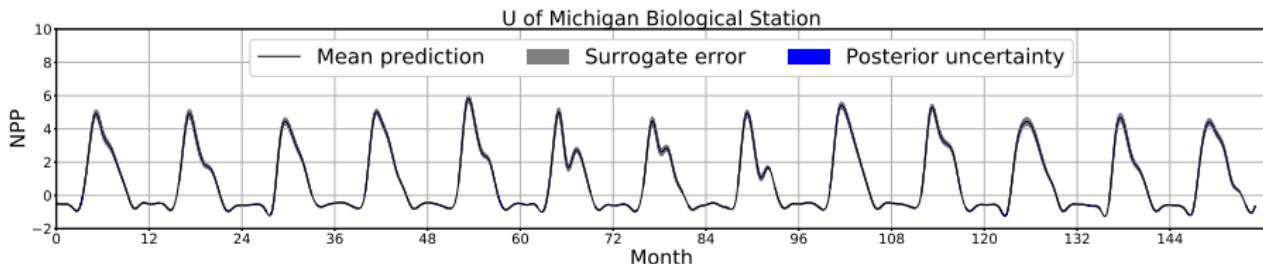
- US Department of Energy (DOE) sponsored Earth system model
- Land, atmosphere, ocean, ice, human system components
- High-resolution, employ DOE leadership-class computing facilities



- Predictive variance decomposition with model-error component
- ... with predictive uncertainty that captures model error

# E3SM Land Model (ELM)

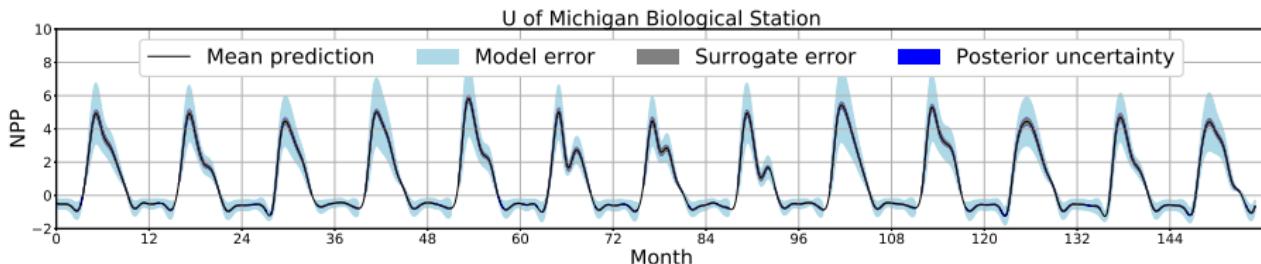
- US Department of Energy (DOE) sponsored Earth system model
- Land, atmosphere, ocean, ice, human system components
- High-resolution, employ DOE leadership-class computing facilities



- Predictive variance decomposition with model-error component
- Allows meaningful prediction of other QoIs  
(e.g. no data/observable)

# E3SM Land Model (ELM)

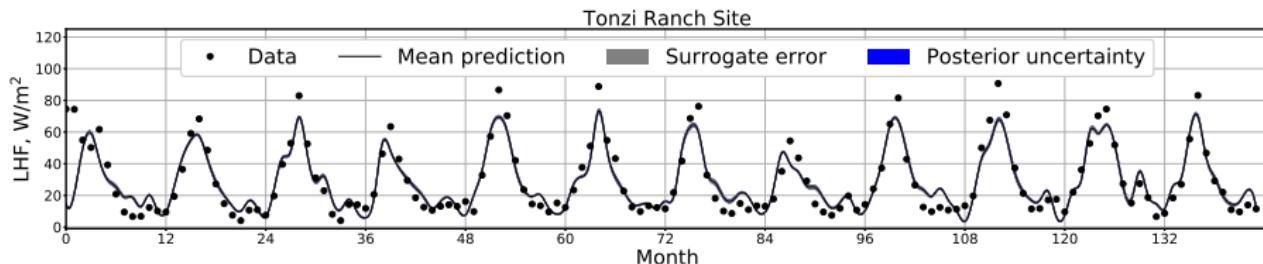
- US Department of Energy (DOE) sponsored Earth system model
- Land, atmosphere, ocean, ice, human system components
- High-resolution, employ DOE leadership-class computing facilities



- Predictive variance decomposition with model-error component
- Allows meaningful prediction of other QoIs  
(e.g. no data/observable)
- ... with predictive uncertainty that captures model error

# E3SM Land Model (ELM)

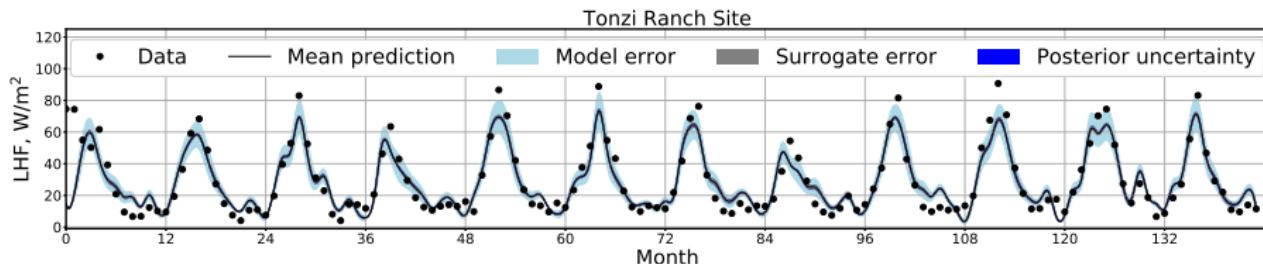
- US Department of Energy (DOE) sponsored Earth system model
- Land, atmosphere, ocean, ice, human system components
- High-resolution, employ DOE leadership-class computing facilities



- Predictive variance decomposition with model-error component
- Allows (a more dangerous) extrapolation to other sites

# E3SM Land Model (ELM)

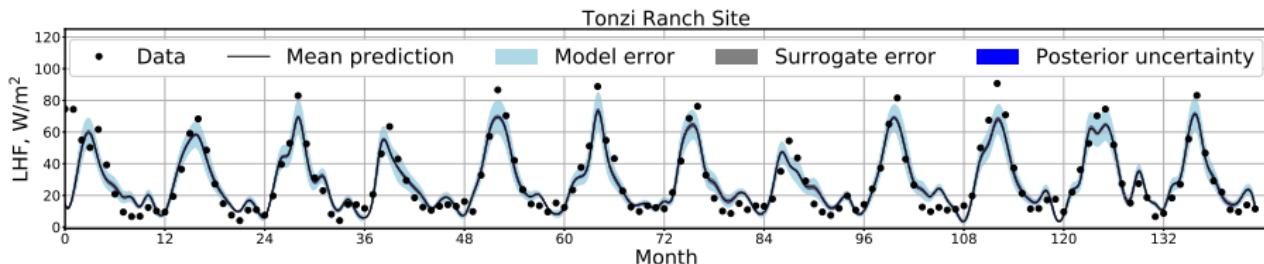
- US Department of Energy (DOE) sponsored Earth system model
- Land, atmosphere, ocean, ice, human system components
- High-resolution, employ DOE leadership-class computing facilities



- Predictive variance decomposition with model-error component
- Allows (a more dangerous) extrapolation to other sites
- ... with predictive uncertainty that captures model error

# E3SM Land Model (ELM)

- US Department of Energy (DOE) sponsored Earth system model
- Land, atmosphere, ocean, ice, human system components
- High-resolution, employ DOE leadership-class computing facilities



# Summary

- Embedded, non-intrusive model error quantification
- PC-based representation and propagation
- Bayesian framework for simultaneous estimation of model inputs and model error parameters
- All developments done within UQTk, lightweight C++/Python library out of SNL-CA ([www.sandia.gov/uqtoolkit](http://www.sandia.gov/uqtoolkit))
- Challenges:
  - Identifiability
  - Respecting the physics
  - Extrapolation/generalization
  - Where/how to embed
  - Likelihood degeneracy
  - Priors
- Opportunities:
  - Handling discrete inputs, relation to BMA
  - Hierarchical Bayesian viewpoint: latent variable models

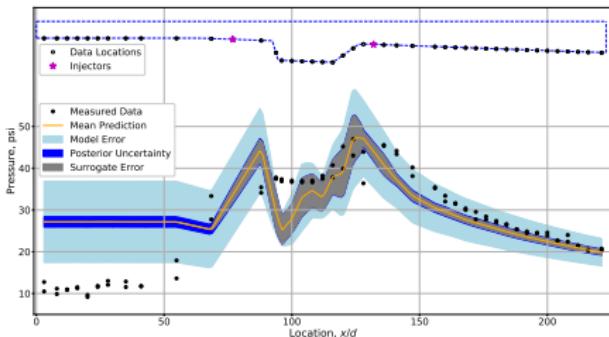
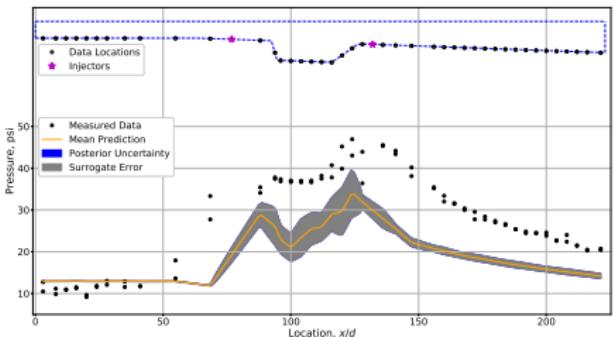
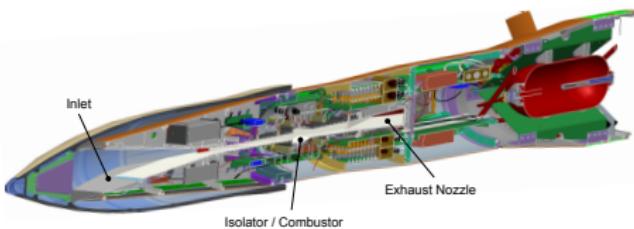
- M. Kennedy and A. O'Hagan, "Bayesian Calibration of Computer Models", *Journal of the Royal Statistical Society, Series B*. 63, 425-464, 2001.
- D. Higdon, M. Kennedy, J. C. Cavendish, J. A. Cafeo, and R. D. Ryne. "Combining Field Data and Computer Simulations for Calibration and Prediction", *SIAM Journal on Scientific Computing*, 26(2):448-466, 2004.
- V. R. Joseph and S. N. Melkote. "Statistical Adjustments to Engineering Models", *Journal of Quality Technology*, 41(4):362, 2009.
- T. A. Oliver, G. Terejanu, C. S. Simmons, and R. D. Moser, "Validating Predictions of Unobserved Quantities", *Computer Methods in Applied Mechanics and Engineering*, 283:1310-1335, 2015.
- J. Brynjarsdottir and A. O'Hagan. Learning about Physical Parameters: The Importance of Model Discrepancy". *Inverse Problems*, 30, 2014.
- K. Sargsyan, H. Najm, R. Ghanem, "On the Statistical Calibration of Physical Models", *Int. J. Chem. Kinetics*, 47(4), 246-276, 2015.
- K. Sargsyan, X. Huan, H. Najm. "Embedded Model Error Representation for Bayesian Model Calibration", arXiv:1801.06768, 2018.

# Additional Material

# LES: Turbulent Combustion in Scramjet Engine



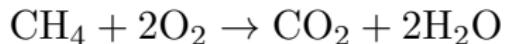
- HIFiRE (Hypersonic International Flight Research and Experimentation) scramjet
- Pressure data from NASA Langley Research Center
- Highly complex LES model



- Augmenting model error leads to more ‘physical’ likelihood

# Chemistry problem – ABC

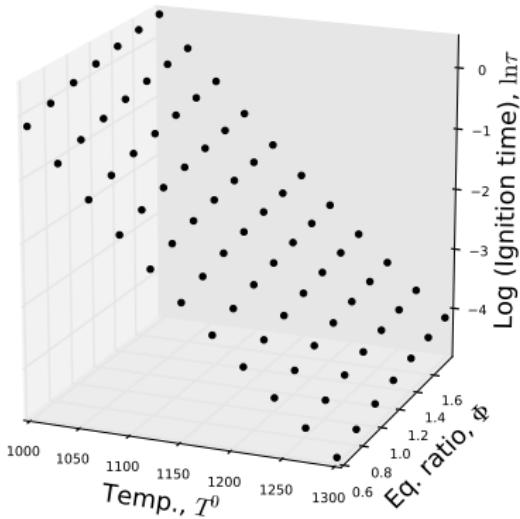
- Homogeneous ignition, methane-air mixture
- Single-step global reaction model calibrated against a detailed chemical kinetic model
- Data: ignition time; range of initial  $T$  & equivalence ratio
- Single-step model:



$$\mathfrak{R} = [\text{CH}_4][\text{O}_2]k_f$$

$$k_f = A \exp(-E/R^o T)$$

- $(\ln A, E) = \sum_k \alpha_k \Psi_k(\xi)$

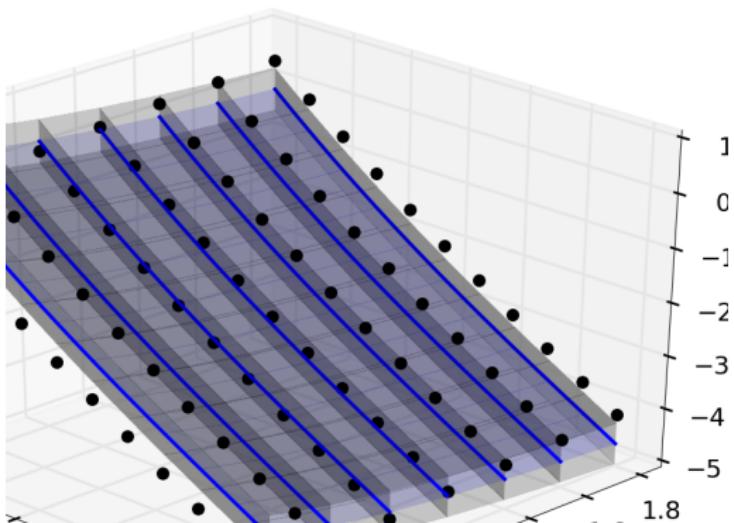


# Quality of Uncertain Calibrated Model Predictions

Calibrated uncertain fit model  
is consistent with the  
detailed-model data.

Over the range of  $(T^0, \Phi)$ :

- MAP predictive mean  
ignition-time is centered  
on the data



# TransCom3 Experiment of $CO_2$ Flux Inversion

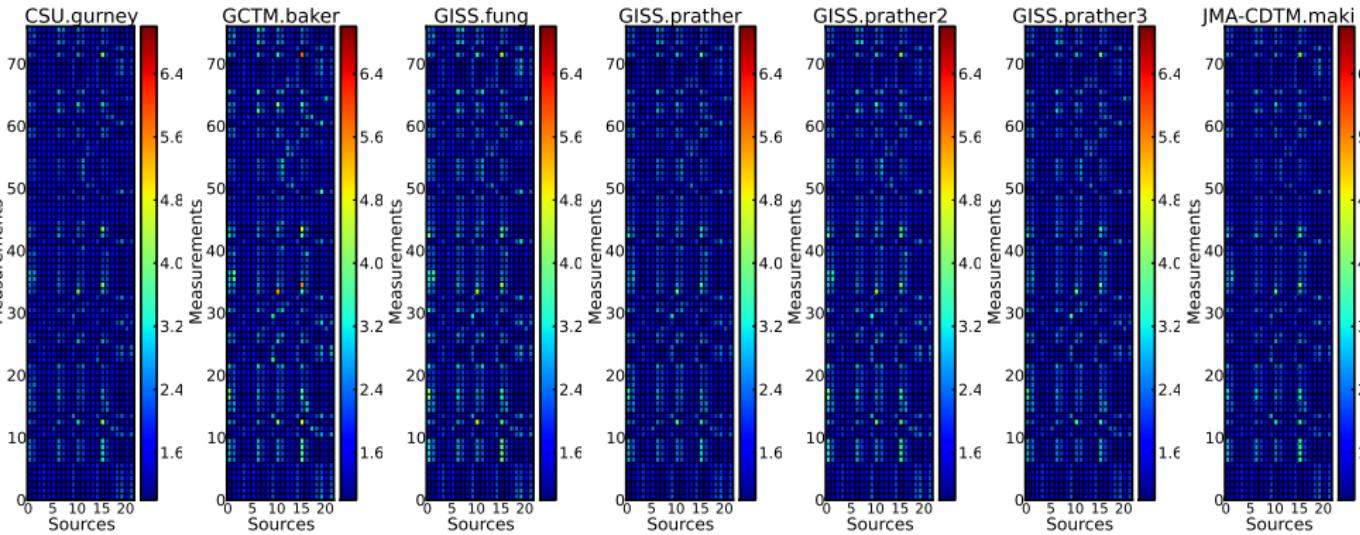
[Gurney *et al.*, Tellus B, 2003]

- Observations  $\mathbf{d}$  at  $N = 77$  sites around the world
- Inverse problem: find fluxes  $\mathbf{s}$  at  $M = 22$  locations
- Linearized ‘response’ model  $\mathbf{R}$ , such that  $\mathbf{d} \approx \mathbf{Rs}$

$$\mathbf{d} = \mathbf{Rs} + \epsilon_{\mathbf{d}}$$

- Model  $\mathbf{R}$  is never perfect thus contaminating the inversion
- The inferred values of  $\mathbf{s}$  compensate for model deficiencies
- $\epsilon_{\mathbf{d}}$  is meant to capture data errors, but is ‘entangled’ with model errors

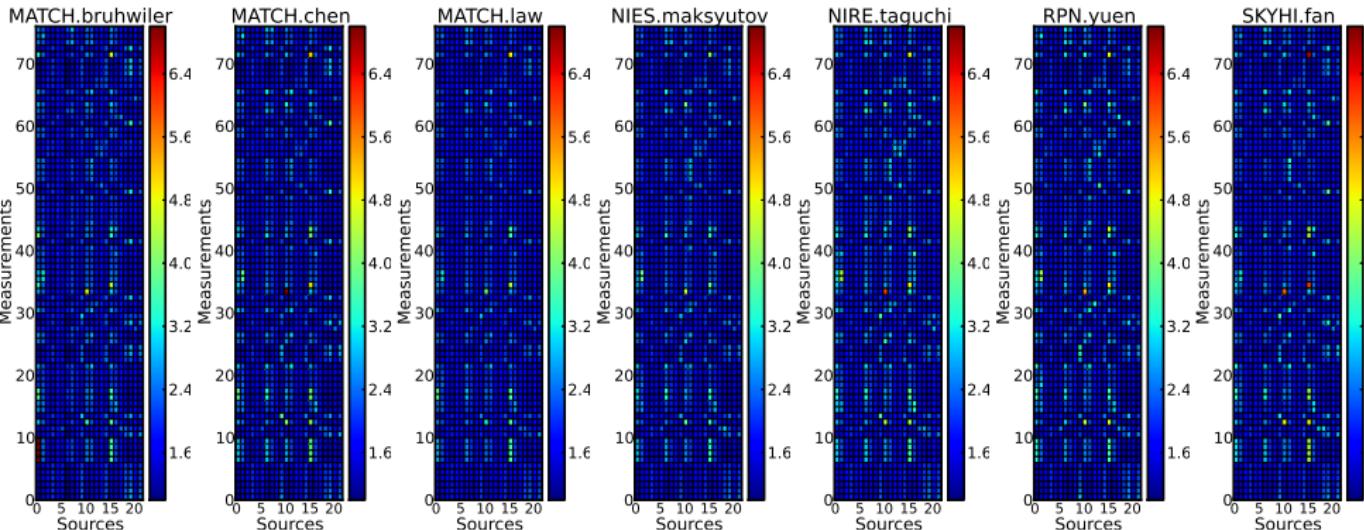
# Consider 14 different response models R



Infer fluxes  $s$ , given measurements  $d$  to satisfy  $d \approx Rs$

- Conventional additive Gaussian error (least-squares):  $d = Rs + \xi$
- Embed probabilistic model for fluxes  $s$ :  $d = R(\mu_s + C_s \xi)$

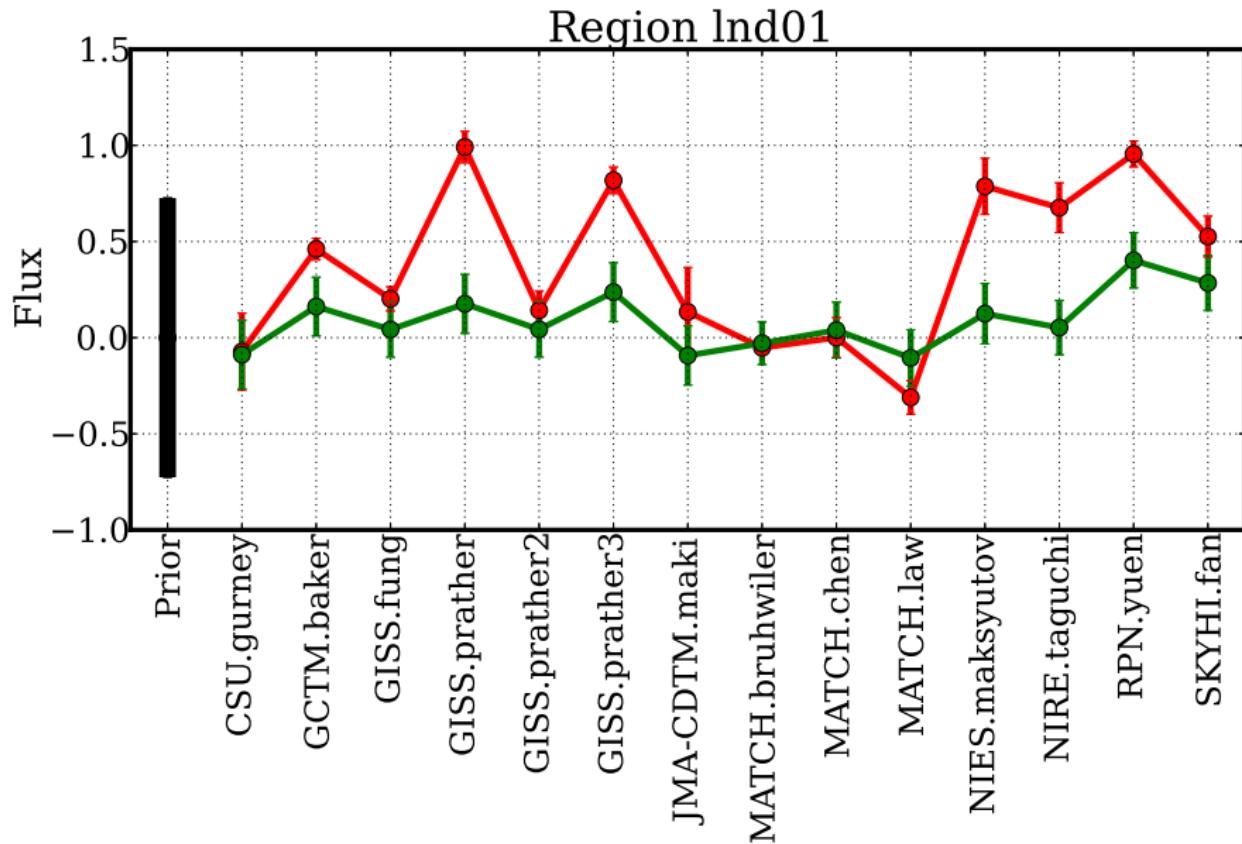
# Consider 14 different response models R



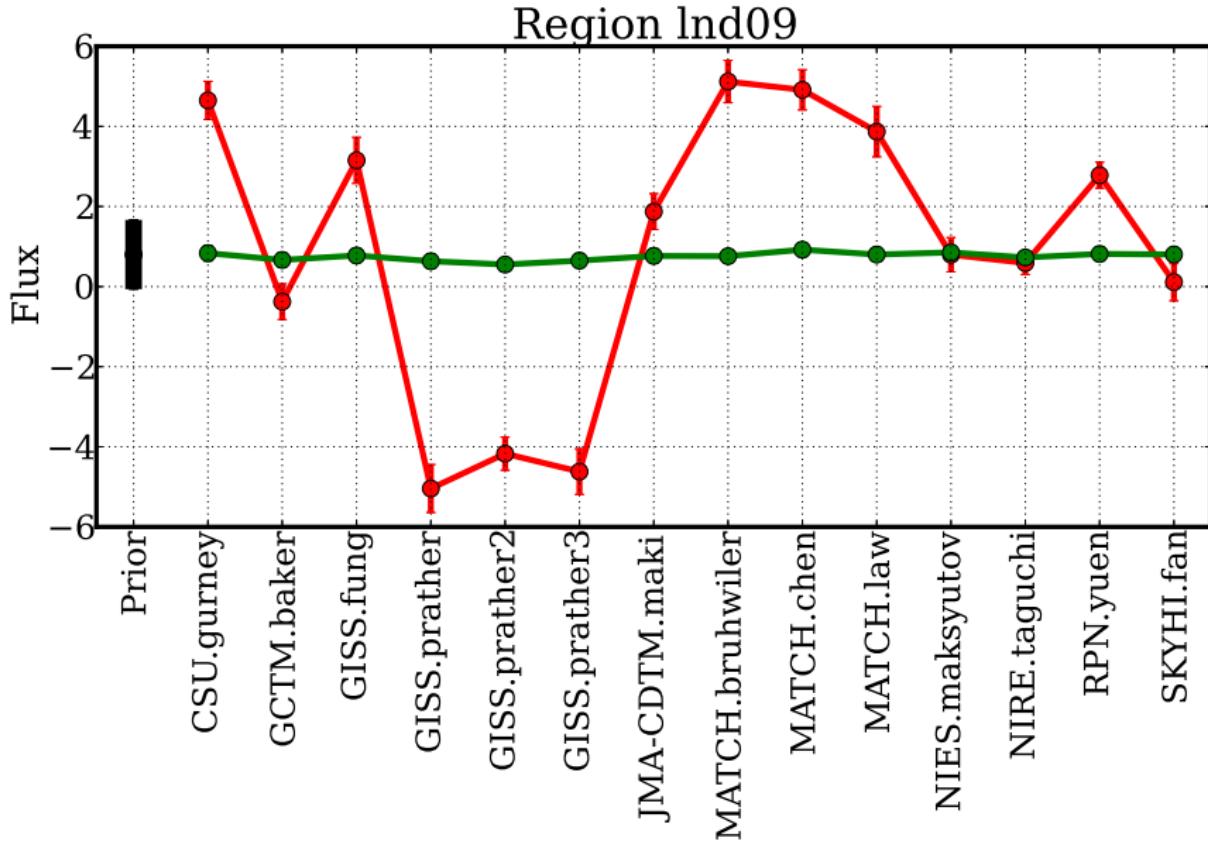
Infer fluxes  $s$ , given measurements  $d$  to satisfy  $d \approx Rs$

- Conventional additive Gaussian error (least-squares):  $d = Rs + \xi$
- Embed probabilistic model for fluxes  $s$ :  $d = R(\mu_s + C_s \xi)$

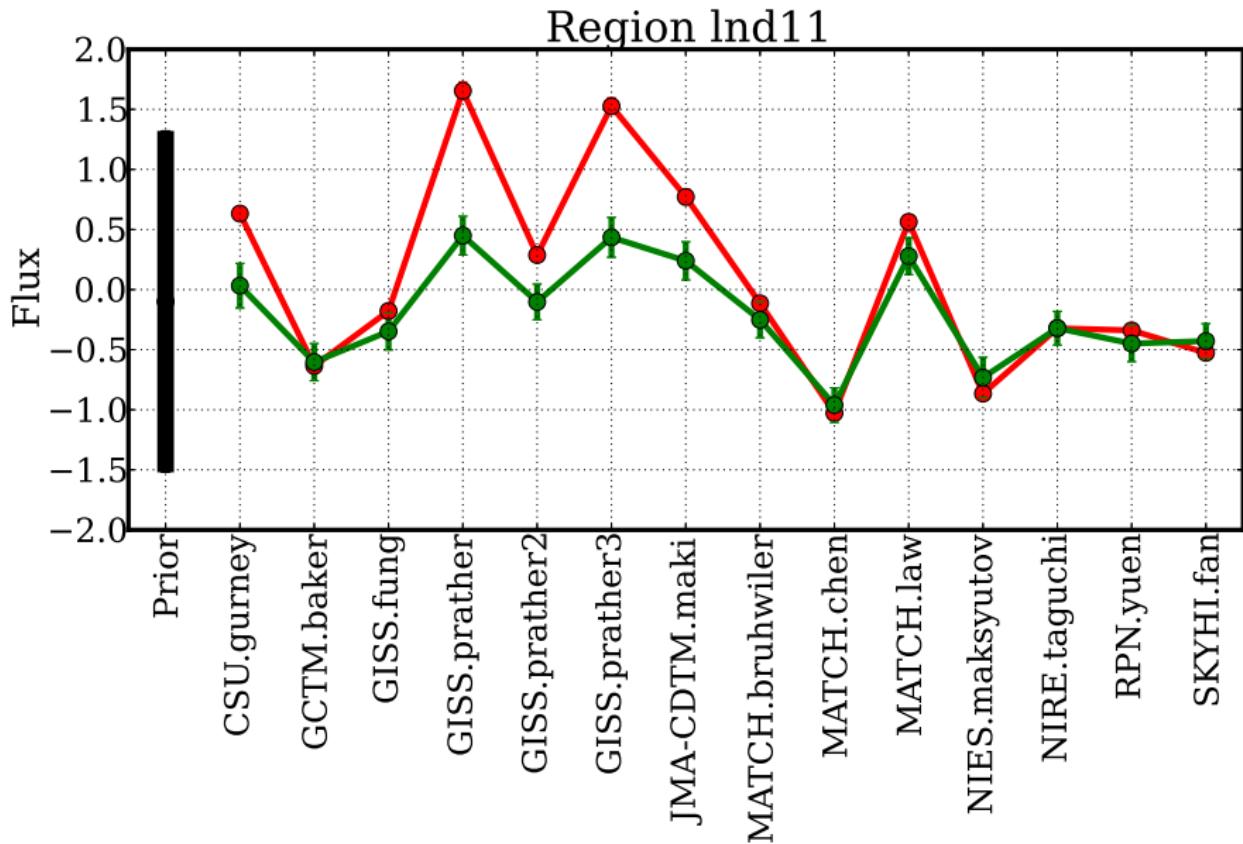
# Inferred fluxes show less variability across models



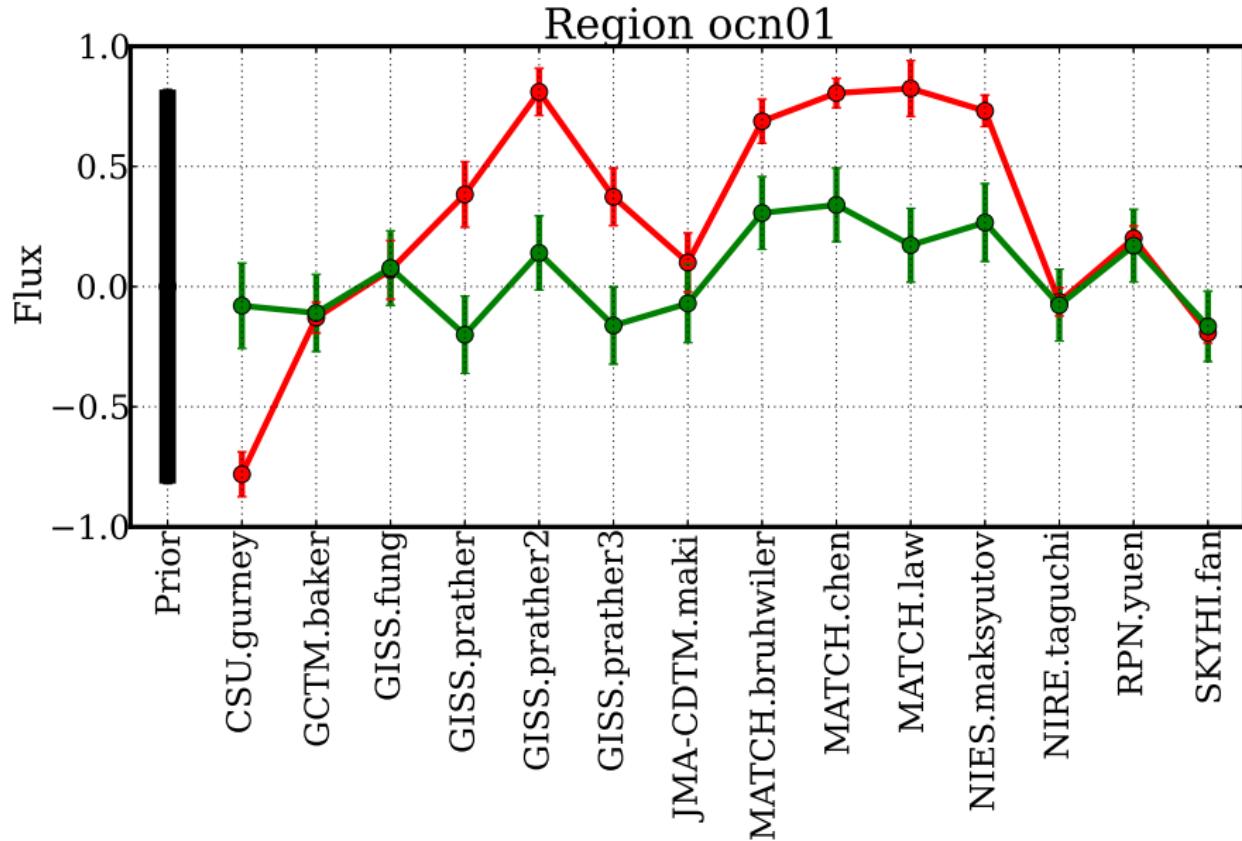
# Inferred fluxes show less variability across models



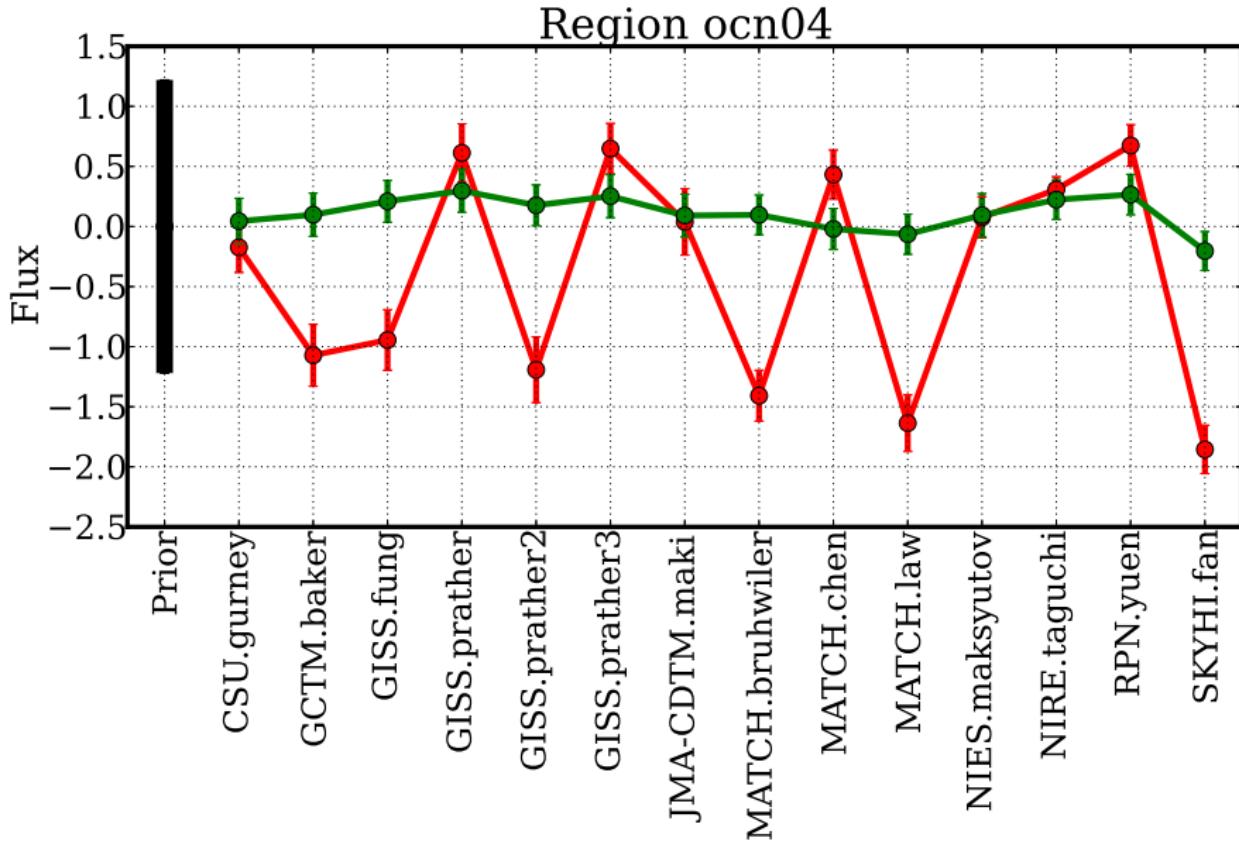
# Inferred fluxes show less variability across models



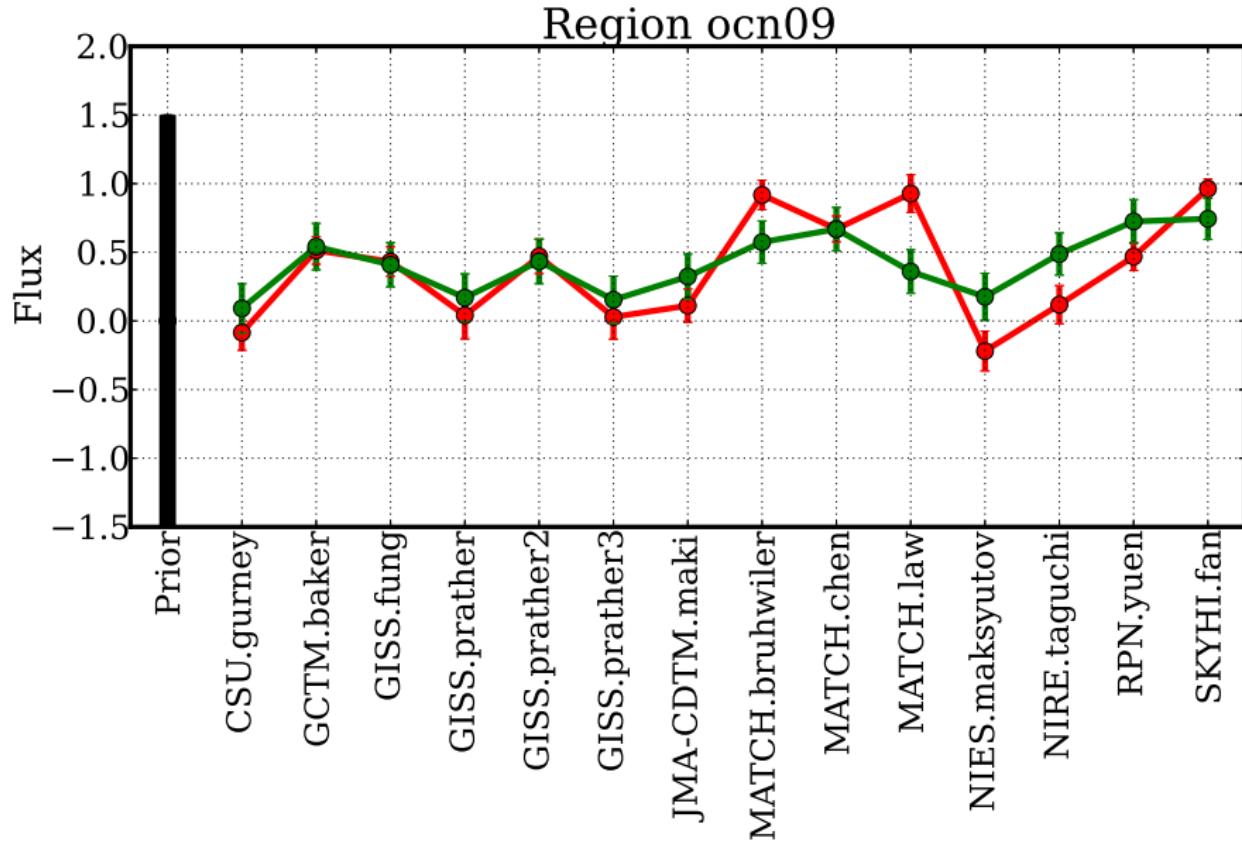
# Inferred fluxes show less variability across models



# Inferred fluxes show less variability across models



# Inferred fluxes show less variability across models



# Inferred fluxes show less variability across models

