Embedded Framework for Model Error Quantification and Propagation

Khachik Sargsyan

Sandia National Laboratories
Livermore, CA

Second USACM Thematic Conference
on Uncertainty Quantification
for Machine Learning Integrated Physics Modeling

Arlington, VA
Aug 14, 2024

Acknowledgements

H. Najm – SNL
T. Casey, J. Oreluk, B. Debusschere, M. Eldred, C. Safta, C. Lacey, N. Iyengar – SNL

Y. Marzouk, C. Feng, K. Fisher — MIT
X. Huan – UMich
R. Ghanem – USC
D. Ricciuto – ORNL
J. Bender – LLNL
F. Ghahari – USGS
O. Cekmer – CSIRO

This work was supported by:

- DOE, Advanced Scientific Computing Research (ASCR), SciDAC
- DOE, Biological and Environmental Research (BER)
- DOE, Basic Energy Sciences (BES)
- DOE, Fusion Energy Sciences (FES)
- DOD, DARPA EQUiPS

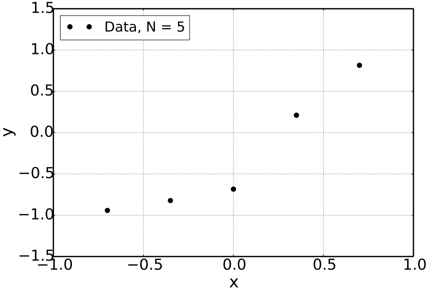
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

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

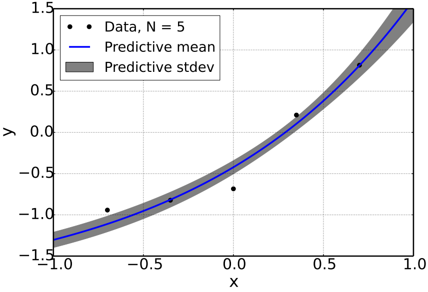
Ignoring model error leads to overconfident and biased predictions



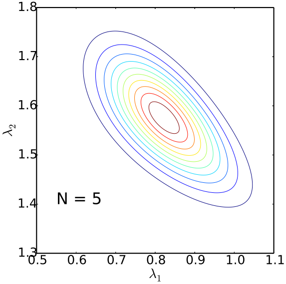
Model-data fit

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

Ignoring model error leads to overconfident and biased predictions



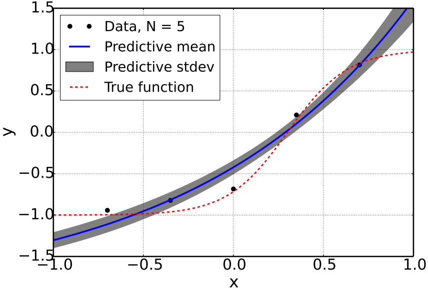
Model-data fit



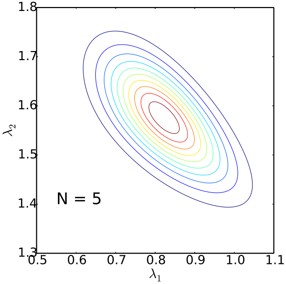
Posterior on parameters

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

Ignoring model error leads to overconfident and biased predictions



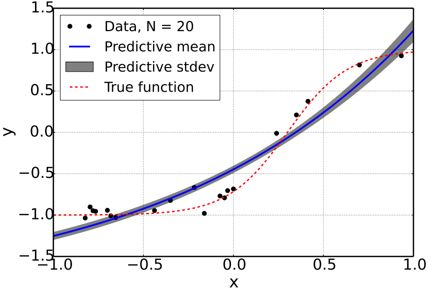
Model-data fit



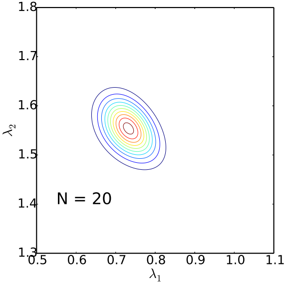
Posterior on parameters

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

Ignoring model error leads to overconfident and biased predictions



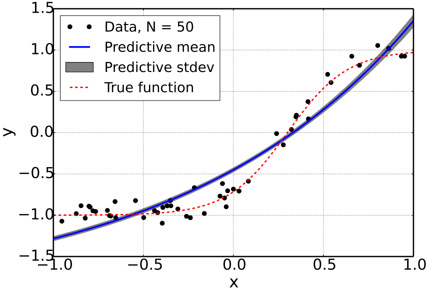
Model-data fit



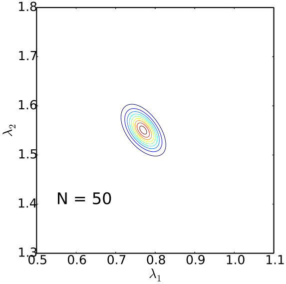
Posterior on parameters

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



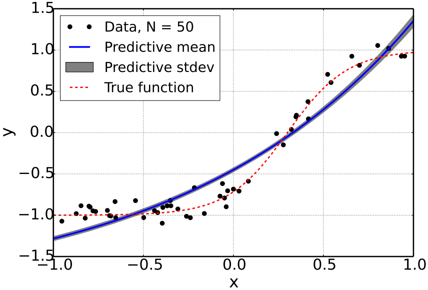
Model-data fit



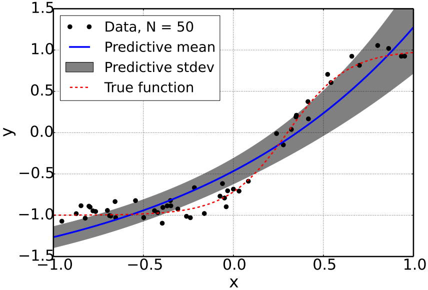
Posterior on parameters

- Given noisy data, calibrate an exponential model: $g(x) \approx f(x;\lambda)$
- Employ Bayesian inference to obtain posterior PDFs on λ
- 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 λ
- 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:
  - 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], [Rizzi et. al, 2018]...
  - 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]
  - Dynamic discrepancy [Bhat 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 ϵ_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 QoIs
- 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 λ and model-error parameters α.
- Bayes' theorem

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

- 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 δ_α.

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* ξ 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 ξ

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

- 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

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

GM Gauss-Marginal Approximate Likelihood:

$$\log \mathcal{L}_g(\tilde{\alpha}) \approx \sum_{i=1}^N \left[- \log \sigma_i(\tilde{\alpha}) - \frac{1}{2} \left(\frac{y_i - \mu_i(\tilde{\alpha})}{\sigma_i(\tilde{\alpha})} \right)^2 \right]$$

ABC Moment-matching / Approximate Bayesian Computation:

$$\log \mathcal{L}_g(\tilde{\alpha}) \approx \sum_{i=1}^N \left[-\frac{1}{2} \left(\frac{y_i - \mu_i(\tilde{\alpha})}{\epsilon_\mu} \right)^2 - \frac{1}{2} \left(\frac{\sigma_i(\tilde{\alpha}) - \gamma |y_i - \mu_i(\tilde{\alpha})|}{\epsilon_\sigma} \right)^2 \right]$$

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

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

- ... provides easy access to mean and variance

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

Model Error – Surrogate and Prediction

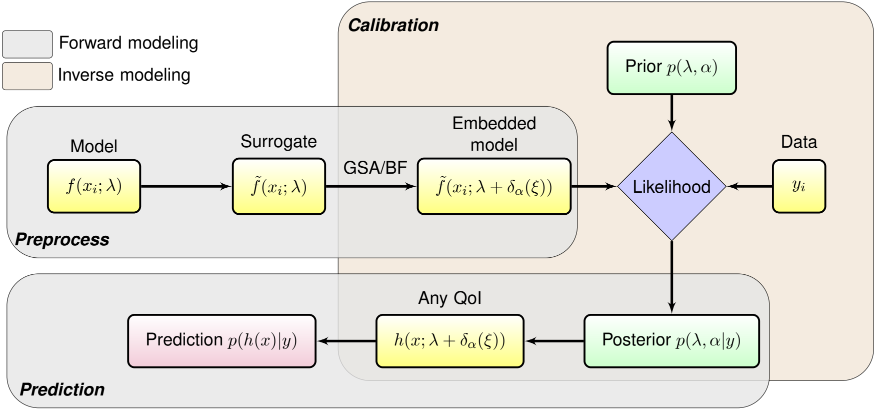
$$f_i(\lambda + \delta_\alpha(\zeta)) = f_i(\tilde{\alpha}, \zeta) \quad \overset{\text{NISP}}{\simeq} \quad \sum_k f_{ik}(\tilde{\alpha}) \Psi_k(\zeta)$$

- 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

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

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

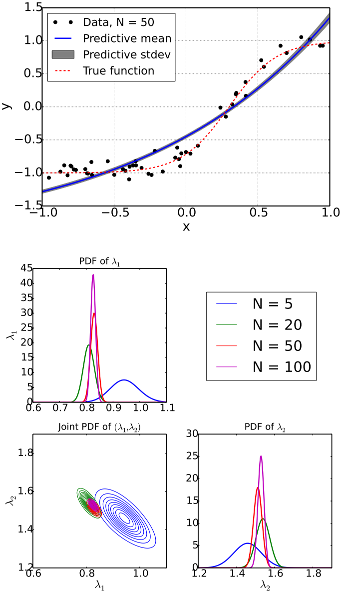
Model error embedding – workflow

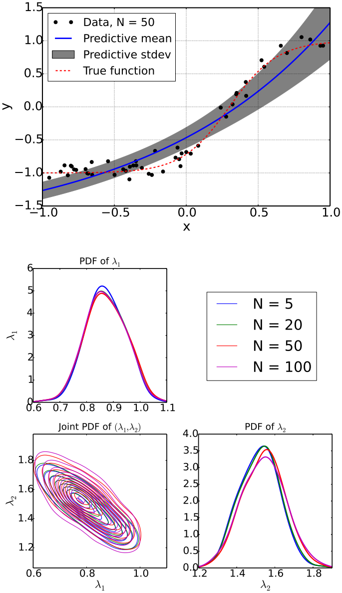


- Predictive uncertainty decomposition: Total Variance =

  Posterior uncertainty + Data noise + Model error + Surrogate error

.. back to toy example



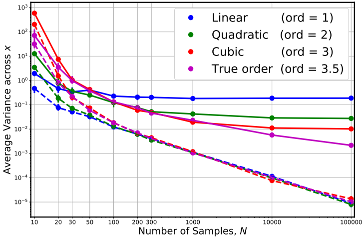


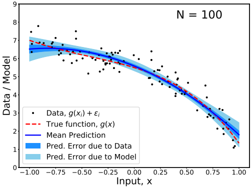
More data leads to ‘leftover’ model error

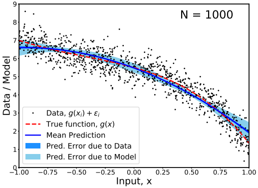
Calibrating a quadratic $f(x) = \lambda_0 + \lambda_1 x + \lambda_2 x^2$
w.r.t. ‘truth’ $g(x) = 6 + x^2 - 0.5(x+1)^{3.5}$ measured with noise $\sigma = 0.1$.

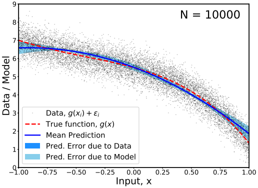
Summary of features:

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

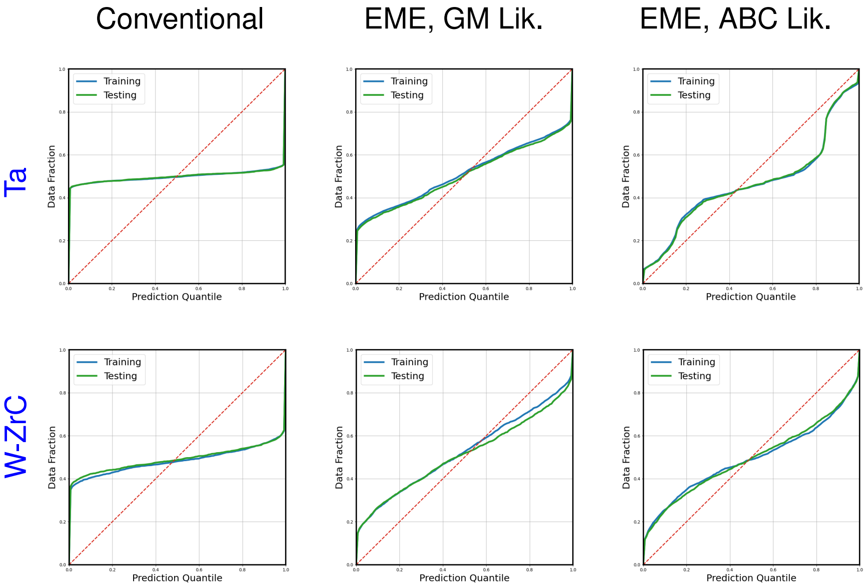








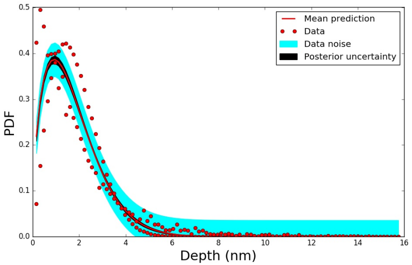
App.: Interatomic Potential Modeling



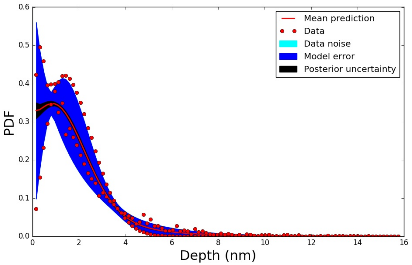
App.: Fusion Science

- Evolution of implanted gas in tungsten plasma-facing component
- Assess the uncertainties in helium flux
- Account for variability due to the two different data sources

Without EME



With EME



[O. Cekmer, K. Sargsyan, S. Blondel, H. Najm, D. Bernholdt,, B.D. Wirth, “Uncertainty quantification for incident helium flux in plasma-exposed tungsten”, *Int. J. Uncertainty Quantification*, 8(5), 2018]

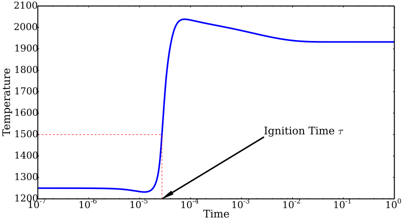
App.: Chemical Kinetics

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

$$C_{12}H_{26} + \frac{25}{2}O_2 \xrightarrow{k_1} 12CO + 13H_2O$$

$$CO + \frac{1}{2}O_2 \underset{k_{2b}}{\overset{k_{2f}}{\rightleftharpoons}} CO_2.$$

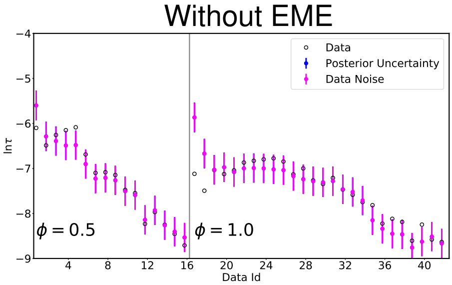
$$k_1 = Ae^{\left(-\frac{E}{RT}\right)}[C_{12}H_{26}]^{0.25}[O_2]^{1.25}$$

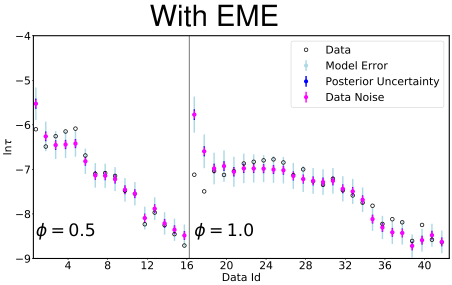


- Data: log(ignition time)
- Embedding
  $(\ln A, E) = \sum_k \boldsymbol{\alpha}_k \Psi_k(\boldsymbol{\xi})$

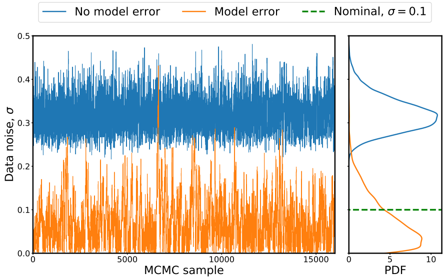
[K. Sargsyan, X. Huan, H. Najm. "Embedded Model Error Representation for Bayesian Model Calibration", arXiv:1801.06768, in press, *Int. J. Uncert. Quant.*, 9(4), 2019]

App.: Chemical Kinetics



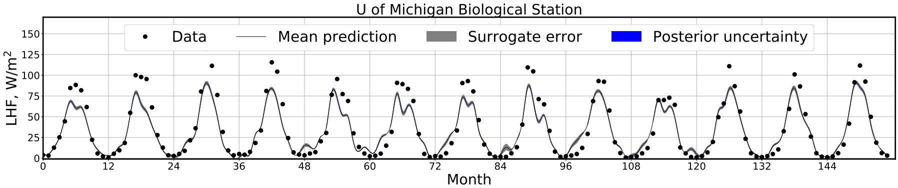


- Model error disambiguated from data error
- Data error correctly captured
- Meaningful extrapolative predictions



App.: E3SM Land Model (ELM)

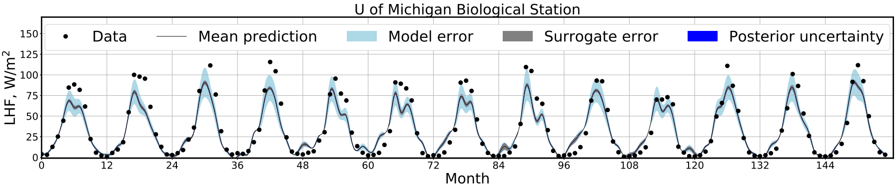
- US Department of Energy (DOE) Earth system model, *e3sm.org*
- Land, atmosphere, ocean, ice, human system components
- High-resolution, employ DOE leadership-class computing facilities



- Conventional calibration without model error

App.: E3SM Land Model (ELM)

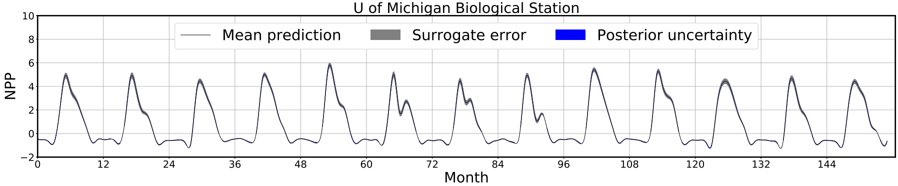
- US Department of Energy (DOE) Earth system model, *e3sm.org*
- Land, atmosphere, ocean, ice, human system components
- High-resolution, employ DOE leadership-class computing facilities



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

App.: E3SM Land Model (ELM)

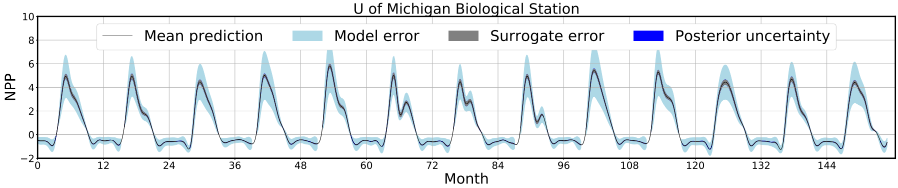
- US Department of Energy (DOE) Earth system model, *e3sm.org*
- Land, atmosphere, ocean, ice, human system components
- High-resolution, employ DOE leadership-class computing facilities



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

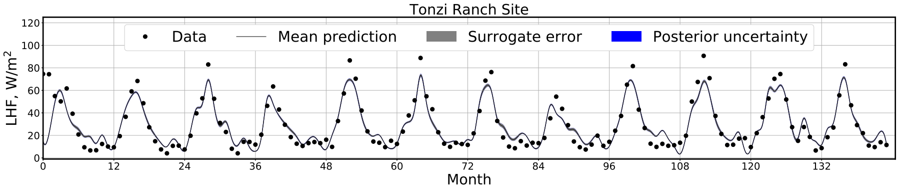
App.: E3SM Land Model (ELM)

- US Department of Energy (DOE) Earth system model, *e3sm.org*
- 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

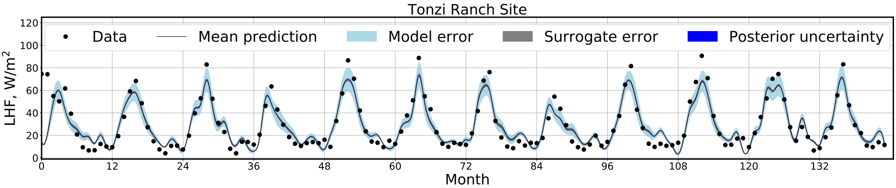
- US Department of Energy (DOE) Earth system model, *e3sm.org*
- Land, atmosphere, ocean, ice, human system components
- High-resolution, employ DOE leadership-class computing facilities



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

App.: E3SM Land Model (ELM)

- US Department of Energy (DOE) Earth system model, *e3sm.org*
- Land, atmosphere, ocean, ice, human system components
- High-resolution, employ DOE leadership-class computing facilities



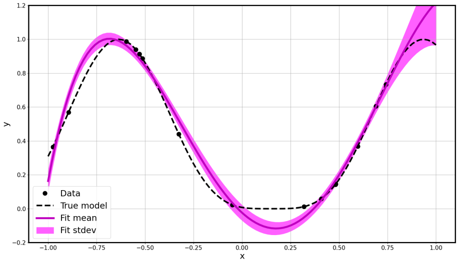
- Predictive variance decomposition with model-error component
- Allows (a more dangerous) extrapolation to other sites
- ... with predictive uncertainty that captures model error

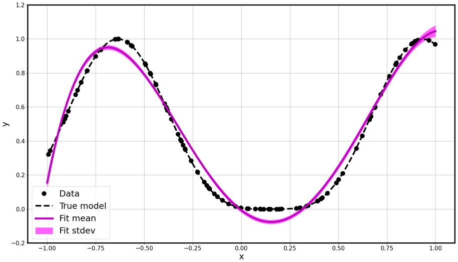
Connection to Variational Inference

Embedded Model Error	Variational Inference
Casting physical parameters λ as random variable	Casting (NN) weights w in a variational family
Infer (λ_0, λ_1)	Infer (μ, σ)
Bayesian/MCMC	Optimization/SGD
Uses ABC/Moment matching in the output	Uses Kullback-Leibler (KL) divergence in the input
PC-based propagation	Functional VI: KL in the output

EME uncertainty is well-calibrated compared to LSQ or VI

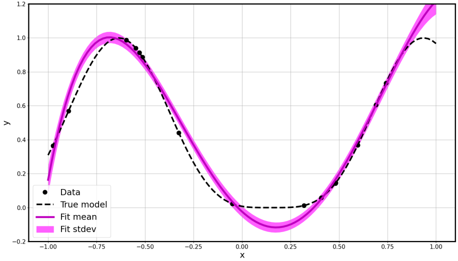
LSQ

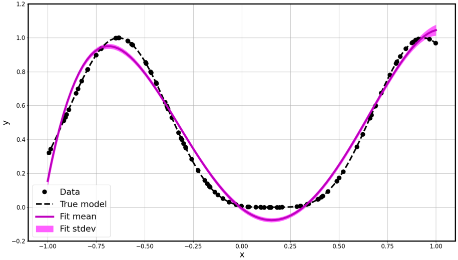




EME uncertainty is well-calibrated compared to LSQ or VI

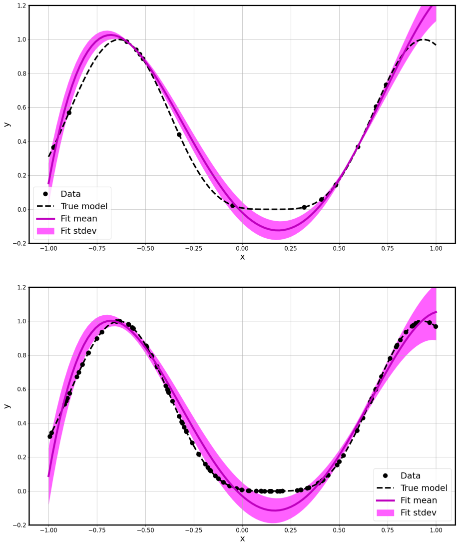
VI



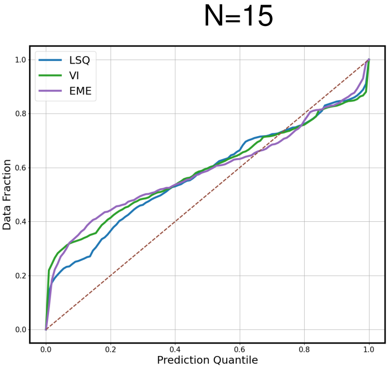


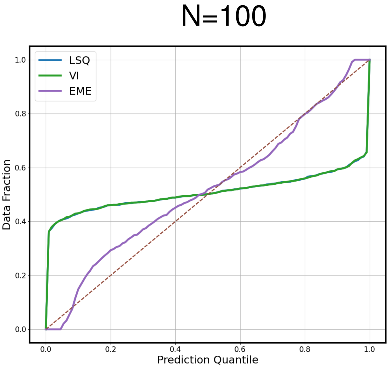
EME uncertainty is well-calibrated compared to LSQ or VI

EME



EME uncertainty is well-calibrated compared to LSQ or VI





Summary

- Embedded, but *non-intrusive* model error quantification
- PC-based representation and propagation
- Simultaneous estimation of model inputs and model error
- Predictive uncertainty attribution due to model error

- Challenges:
  - Higher-d inference problem
  - Identifiability
  - Prior selection
  - Where/how to embed
  - Likelihood degeneracy
  - Extrapolation/generalization

- Opportunities:
  - Field (x-dependent) correction
  - Intrusive, domain-knowledge corrections
  - Non-physical models: parallel to VI
  - Handling discrete/categorical inputs: parallel to BMA

Literature – Methodological

Thank you!

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.

M. Bayarri, J. Berger, R. Paulo, J. Sacks, J. Cafeo, J. Cavendish, C. Lin, and J. Tu. "A Framework for Validation of Computer Models", *Technometrics*, 49(2):138-154, 2007.

V. R. Joseph and S. N. Melkote. "Statistical Adjustments to Engineering Models", *Journal of Quality Technology, 41(4):362*, 2009.

J. Brynjarsdottir and A. O'Hagan. "Learning about Physical Parameters: The Importance of Model Discrepancy". *Inverse Problems*, 30, 2014.

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.

M. Strong, J. E. Oakley, "When is a model good enough? Deriving the expected value of model improvement via specifying internal model discrepancies", *SIAM/ASA Journal on Uncertainty Quantification*, 2, 2014.

Y. He, D. Xiu, "Numerical strategy for model correction using physical constraints", *Journal of Computational Physics*, 313, 2016.

P. Pernot, F. Cailliez, "A critical review of statistical calibration/prediction models handling data inconsistency and model inadequacy". *AIChE Journal*, 63, 2017.

M. Plumlee, "Bayesian calibration of inexact computer models". *Journal of the American Statistical Association*, 112, 2017.

J.-L. Wu, M. Levine, T. Schneider, A. Stuart, "Learning about structural errors in models of complex dynamical systems", *Journal of Computational Physics*, 513, 2024,

N. Leoni, O. Le Maitre, M.-G. Rodio, P. M. Congedo, "Bayesian calibration with adaptive model discrepancy", *International Journal for Uncertainty Quantification*, 14(1), 2024.

K. Sargsyan, H. Najm, R. Ghanem, "On the Statistical Calibration of Physical Models", *Int. J. Chem. Kinetics*, 47(4), 2015.

K. Sargsyan, X. Huan, H. Najm. "Embedded Model Error Representation for Bayesian Model Calibration", arXiv:1801.06768, in press, *Int. J. Uncert. Quant.*, 9(4), 2019.

 Thank you!

F. Ghahari, K. Sargsyan, E. Taciroglu, “Quantification of Modeling Uncertainty in the Rayleigh Damping Model”, *Earthquake Engineering & Structural Dynamics*, 53(9), 2024.

A. Hegde, E. Weiss, W. Windl, H. Najm, C. Safta, “A Bayesian Calibration Framework with Embedded Model Error for Model Diagnostics”, *International Journal for Uncertainty Quantification*, 14(6), 2024.

F. Ghahari, K. Sargsyan, G. Parker, D. Swensen, M. Celebi, H. Haddadi, E. Taciroglu, “Performance-Based Earthquake Early Warning for Tall Buildings”, *Earthquake Spectra*, 40(2), 2024.

M. Savarese, L. Giuntini, R. M. Galassi, S. Iavarone, C. Galletti, W. De Paepe, A. Parente, “Model-to-model Bayesian calibration of a Chemical Reactor Network for pollutant emission predictions of an ammonia-fuelled multistage combustor”, *International Journal of Hydrogen Energy*, 49(B), 2024.

F. Ghahari, K. Sargsyan, M. Celebi, E. Taciroglu, “Quantifying Modeling Uncertainty in Simplified Beam Models for Building Response Prediction”, *Structural Control and Health Monitoring*, 29(11), 2022.

A. Frankel, E. Wagman, R. Keedy, B. Houchens, S. N. Scott “Embedded-Error Bayesian Calibration of Thermal Decomposition of Organic Materials ”, *J. Verif. Valid. Uncert.*, 6(4), 2021.

R.S.M. Freitas, F. Rochinha, D. Mira, X. Jiang, “Parametric and model uncertainties induced by reduced order chemical mechanisms for biogas combustion”, *Chemical Engineering Science*, 227, 2020.

F. Rizzi, M. Khalil, R.E. Jones, J.A. Templeton, J.T. Ostien, B.L. Boyce, “Bayesian modeling of inconsistent plastic response due to material variability”, *Computer Methods in Applied Mechanics and Engineering*, 353, 2019.

O. Cekmer, K. Sargsyan, S. Blondel, H. Najm, D. Bernholdt, B. Wirth, “Uncertainty Quantification for Incident Helium Flux in Plasma-Exposed Tungsten”, *International Journal for Uncertainty Quantification*, 8(5), 2018.

L. Hakim, G. Lacaze, M. Khalil, K. Sargsyan, H. Najm, J. Oefelein, “Probabilistic Parameter Estimation in a 2-Step Chemical Kinetics Model for n-Dodecane Jet Autoignition”, *Combustion Theory and Modeling*, 22(3), 2018.

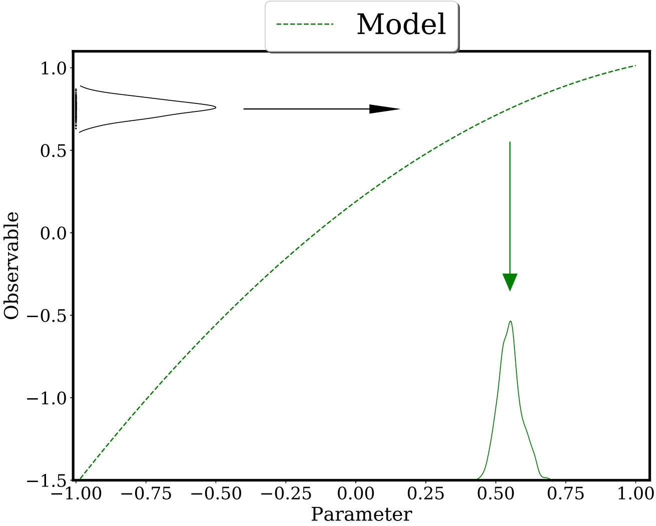
X. Huan, C. Safta, K. Sargsyan, G. Geraci, M. Eldred, Z. Vane, G. Lacaze, J. Oefelein, Habib N. Najm, “Global Sensitivity Analysis and Estimation of Model Error, toward Uncertainty Quantification in Scramjet Computations”, *AIAA Journal*, 56(3), 2018.

C. Safta, M. Blaylock, J. Templeton, S. Domino, K. Sargsyan, H. Najm, “Uncertainty Quantification in LES of Channel Flow”, *International Journal for Numerical Methods in Fluids*, 83, 2016.

Additional Material

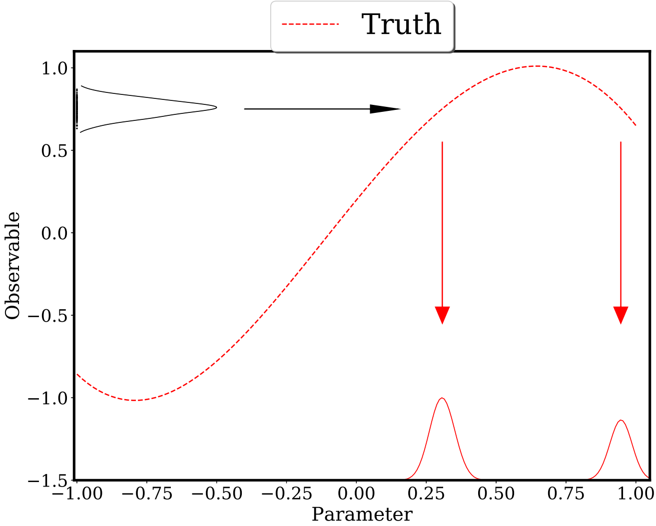
Data informs model parameters:

but what if the model is only an approximation?



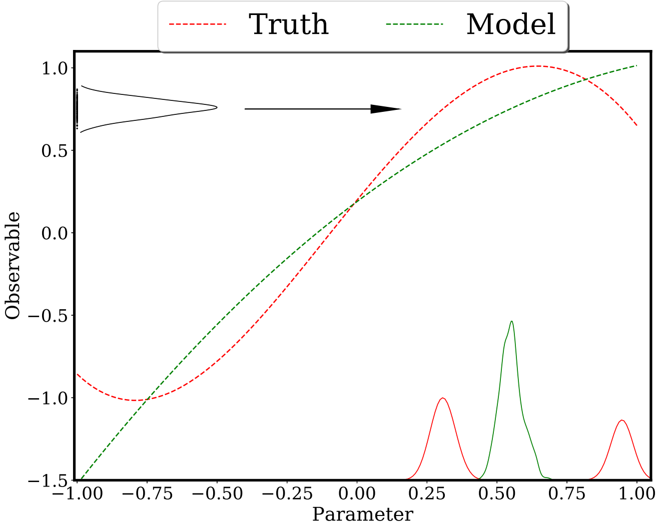
Data informs model parameters:

but what if the model is only an approximation?



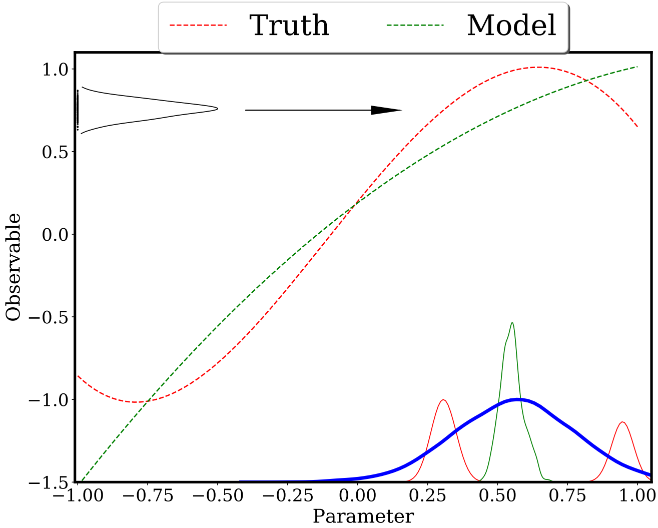
Data informs model parameters:

but what if the model is only an approximation?



Data informs model parameters:

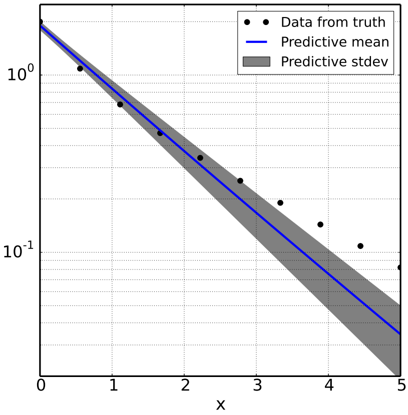
but what if the model is only an approximation?



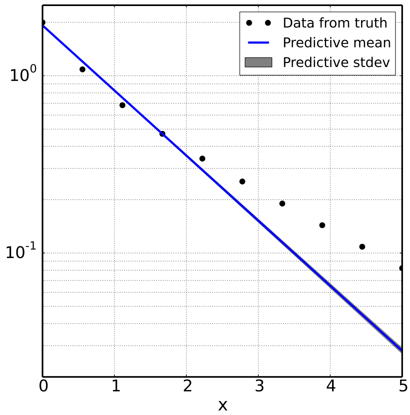
Predictions account for model error

Calibrating single-exponential models
with data from a double exponential model $g(x) = e^{-0.5x} + e^{-2x}$

Linear-exponential $f(x, \lambda) = e^{\lambda_1 + \lambda_2 x}$



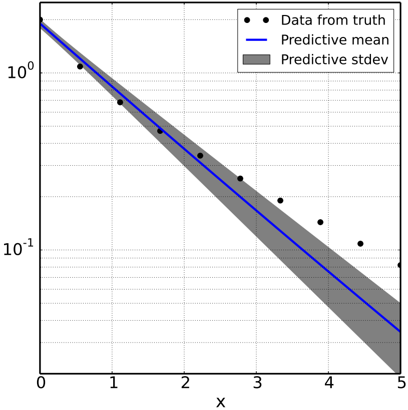
Additive Gaussian error



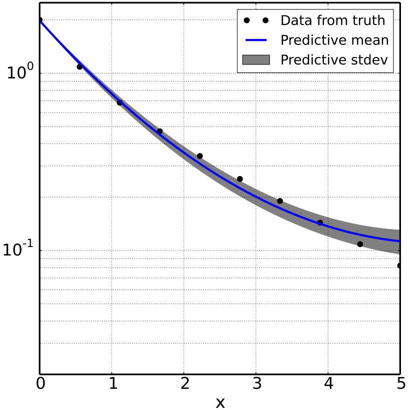
Predictions account for model error

Calibrating single-exponential models
with data from a double exponential model $g(x) = e^{-0.5x} + e^{-2x}$

Linear-exponential $f(x, \lambda) = e^{\lambda_1 + \lambda_2 x}$



Quadratic-exponential $f_2(x, \lambda) = e^{\lambda_1 + \lambda_2 x + \lambda_3 x^2}$



Chemistry problem – ABC

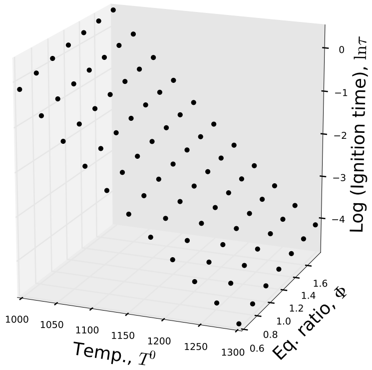
- Homogeneous ignition, methane-air mixture
- Single-step global reaction model calibrated against a detailed chemical kinetic model

- Data: ignition time; range of initial T & equivalence ratio
- Single-step model:

$$\mathrm{CH_4} + 2\mathrm{O_2} \rightarrow \mathrm{CO_2} + 2\mathrm{H_2O}$$

$$\begin{aligned} \mathfrak{R} &= [\mathrm{CH_4}][\mathrm{O_2}]k_f \\ k_f &= A\exp(-E/R^oT) \end{aligned}$$

- $(\ln A, E) = \sum_k \boldsymbol{\alpha}_k \Psi_k(\boldsymbol{\xi})$

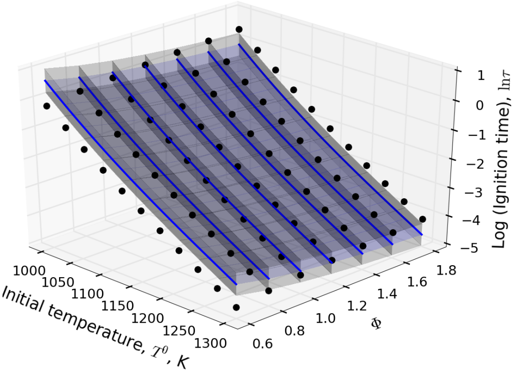


Quality of Uncertain Calibrated Model Predictions

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

Over the range of (T^0, Φ):

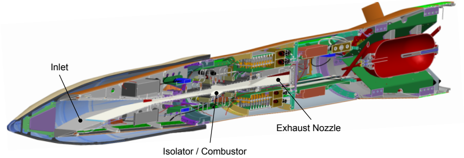
- MAP predictive mean ignition-time is centered on the data
- MAP predictive stdv is consistent with the scatter of the data

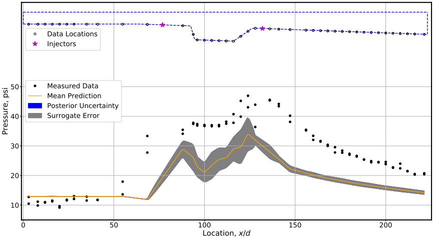


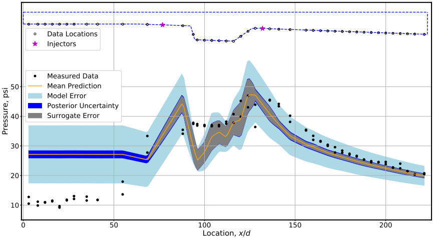
K. Sargsyan, H.N. Najm, and R. Ghanem
”On the Statistical Calibration of Physical Models”
Int. J. Chem. Kin., 47(4): 246-276, 2015

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

TransCom3 Experiment of CO_2 Flux Inversion

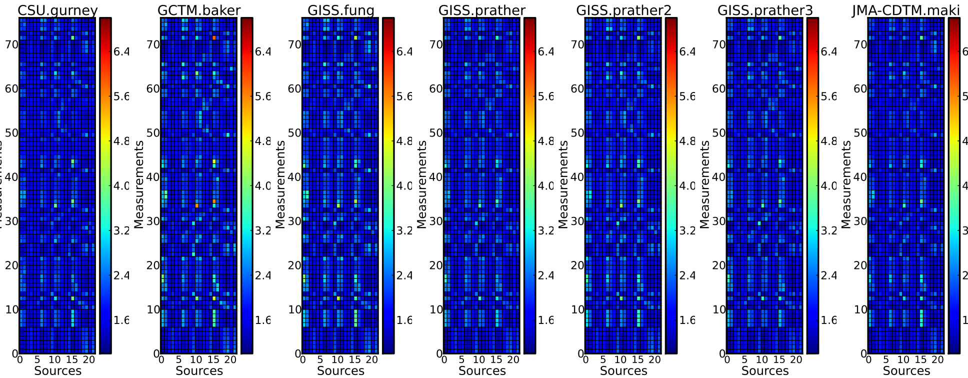
[Gurney *et al.*, Tellus B, 2003]

- 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 $\mathbf{d} \approx \mathbf{R}\mathbf{s}$

$$\mathbf{d} = \mathbf{R}\mathbf{s} + \boldsymbol{\epsilon}_{\mathbf{d}}$$

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

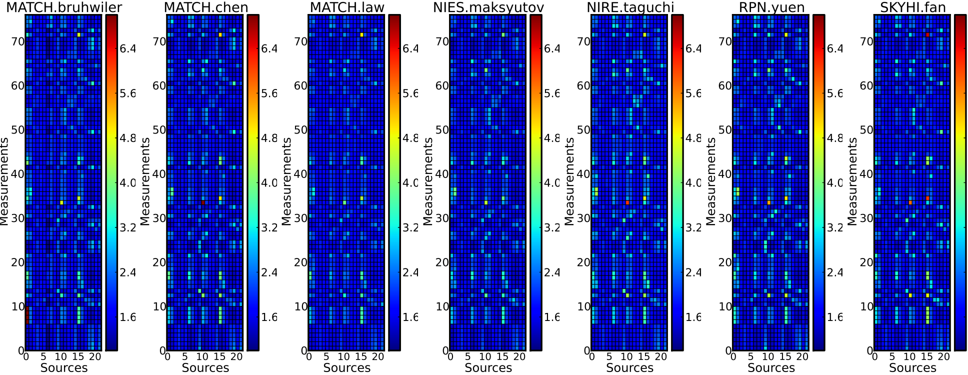
Consider 14 different response models $\mathbf{R}$



Infer fluxes $\mathbf{s}$, given measurements $\mathbf{d}$ to satisfy $\mathbf{d} \approx \mathbf{R}\mathbf{s}$

- Conventional additive Gaussian error (least-squares): $\mathbf{d} = \mathbf{R}\mathbf{s} + \xi$
- Embed probabilistic model for fluxes s: $\mathbf{d} = \mathbf{R}(\mu_\mathbf{s} + \mathbf{C}_\mathbf{s}\xi)$

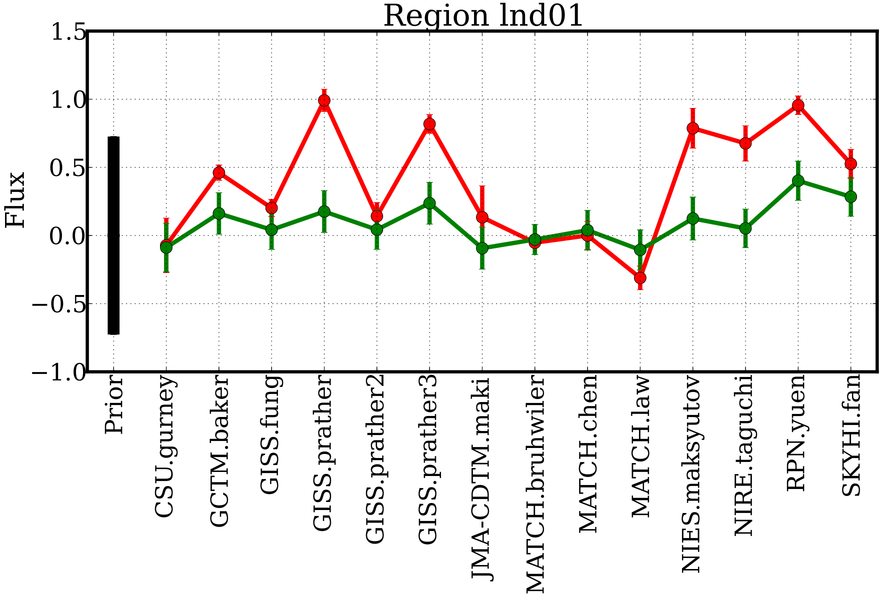
Consider 14 different response models $\mathbf{R}$



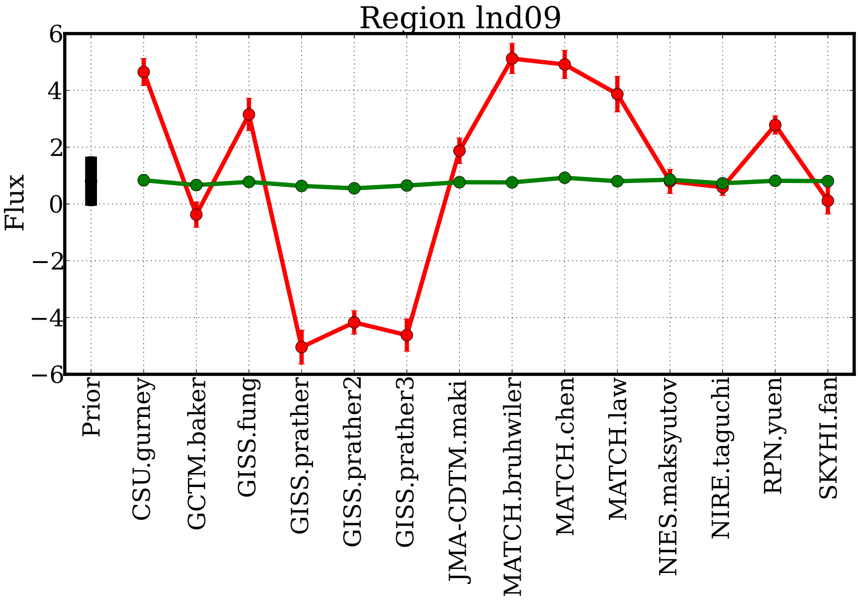
Infer fluxes $\mathbf{s}$, given measurements $\mathbf{d}$ to satisfy $\mathbf{d} \approx \mathbf{Rs}$

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

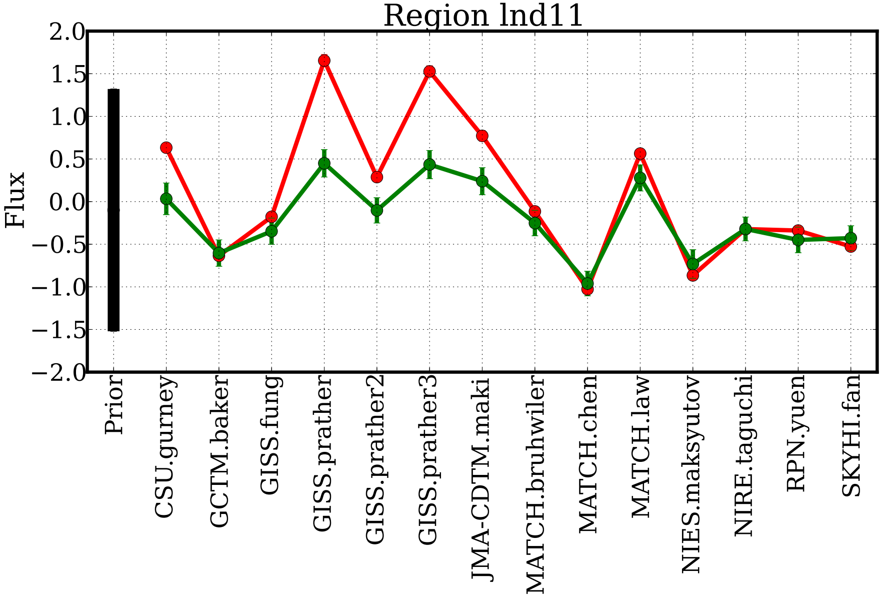
Inferred fluxes show less variability across models



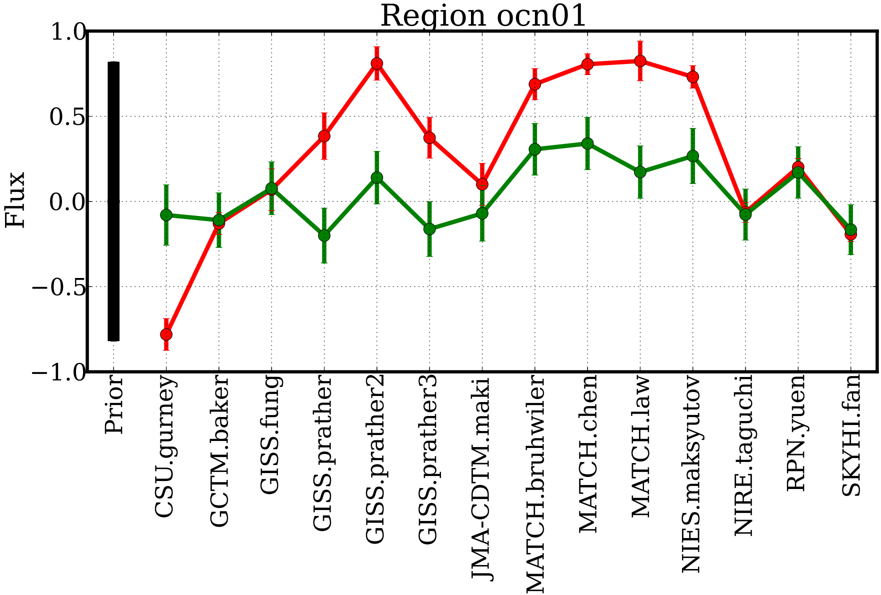
Inferred fluxes show less variability across models



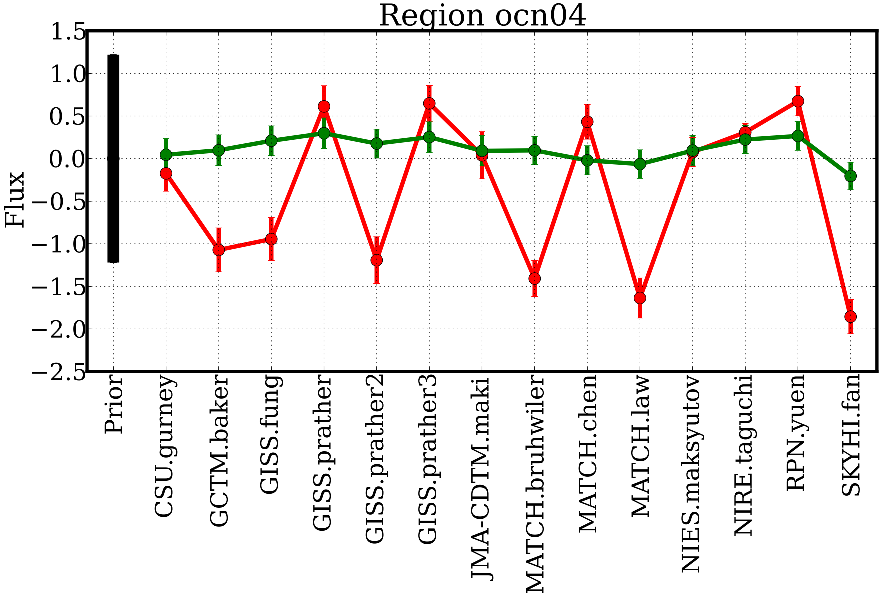
Inferred fluxes show less variability across models



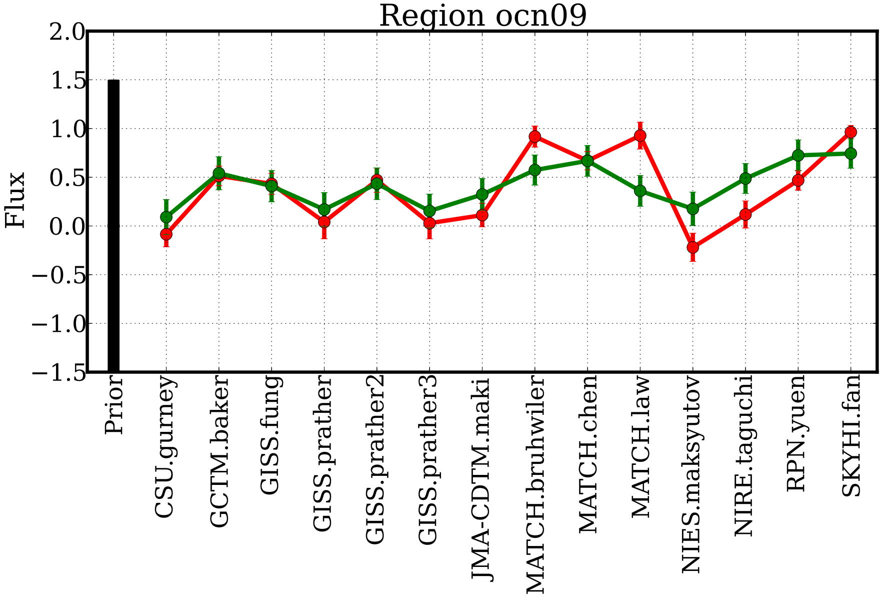
Inferred fluxes show less variability across models



Inferred fluxes show less variability across models



Inferred fluxes show less variability across models



Inferred fluxes show less variability across models