Structural Error Quantification in Physical Models

Khachik Sargsyan

Sandia National Laboratories, Livermore, CA

Data Science Workshop II
SNL-CA
September 21, 2017

Sandia National Laboratories

Funded by DOE ASCR/BES/BER, DOD DARPA

Unclassified for Internal Use

Main target: model *structural* error

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

Data meets Models

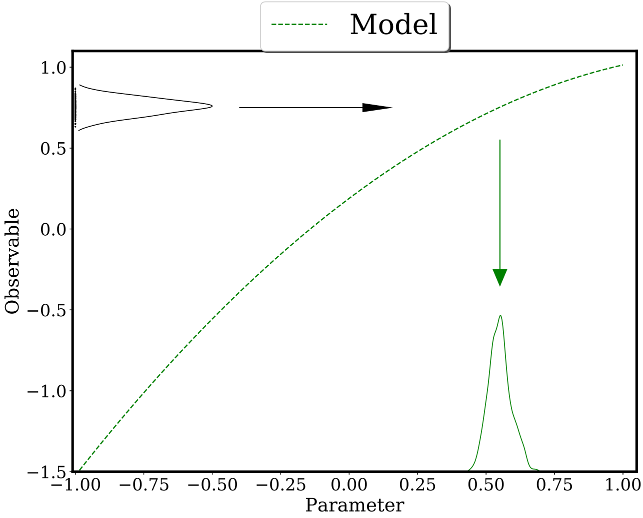
- Inverse modeling context
  - Given experimental or higher-fidelity model data, estimate the model error

---

- Represent and estimate the error associated with
  - Simplifying assumptions, parameterizations
  - Mathematical formulation, theoretical framework
  - Numerical discretization
- ...will be useful for
  - Model validation
  - Model comparison
  - Scientific discovery and model improvement
  - Reliable computational predictions

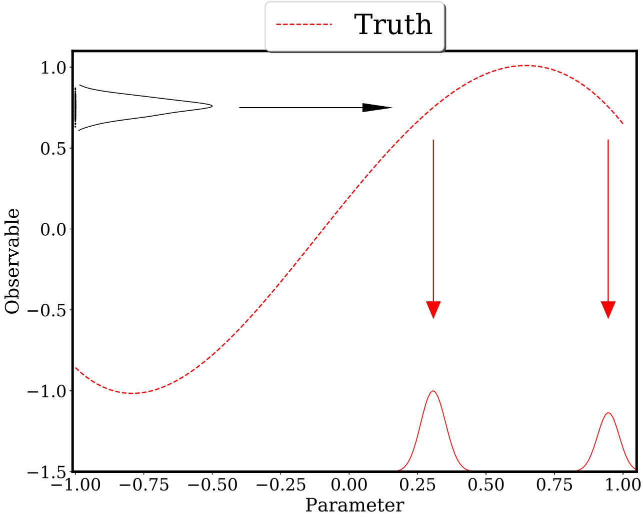
Data informs model parameters:

but what if the model is only an approximation?



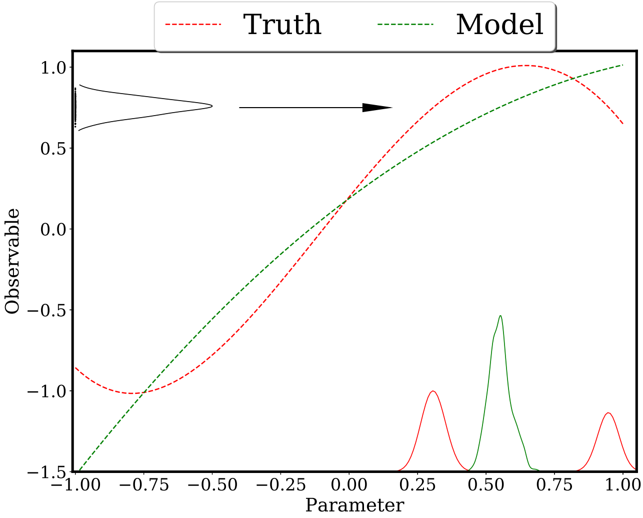
Data informs model parameters:

but what if the model is only an approximation?



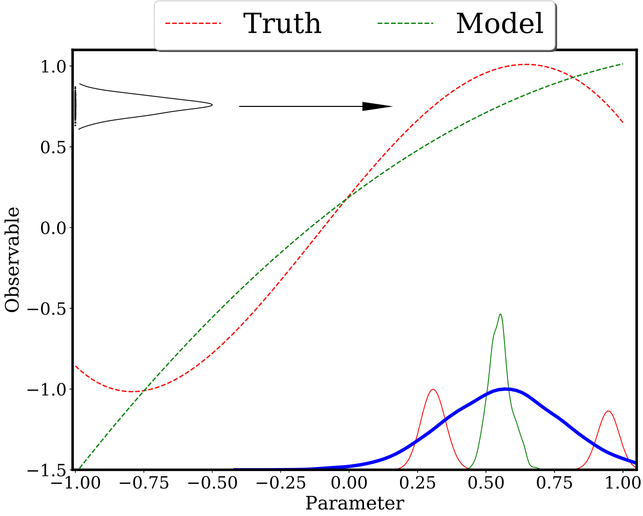
Data informs model parameters:

but what if the model is only an approximation?



Data informs model parameters:

but what if the model is only an approximation?



Calibrate $f(x;\lambda)$, given data $g(x)$

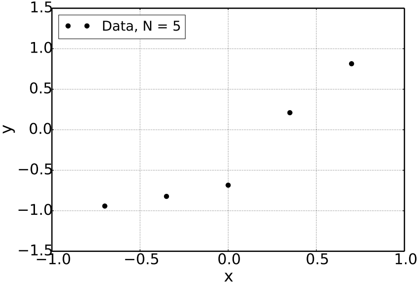
x are operating conditions
λ are model parameters to be inferred/calibrated

- **Default:** Ignore model errors: $g(x) = f(x;\lambda) + \epsilon$
  - Biased or overconfident physical parameters
  - Wrong model predictions

- **Conventional:** Correct for model errors: $g(x) = f(x;\lambda) + \delta(x) + \epsilon$
  - Physical parameters are ok
  - Wrong model predictions (data-specific corrections)
  - Model and data errors mixed up

- **What we do:** Correct *inside* the model: $g(x) = f(x;\lambda + \delta(x)) + \epsilon$
  - Embedded model error
  - Preserves model structure and physical constraints
  - Disambiguates model and data errors
  - Allows meaningful extrapolation

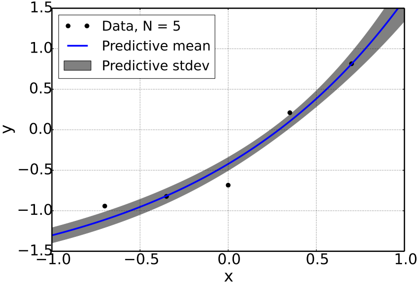
Demo



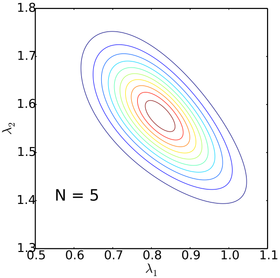
Model-data fit

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

Demo



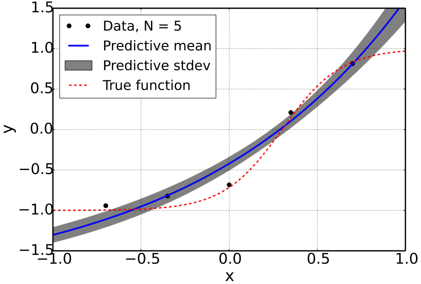
Model-data fit



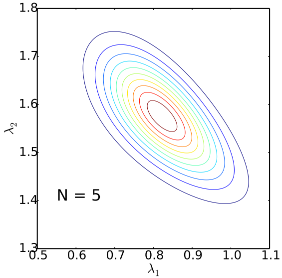
Posterior on parameters

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

Demo



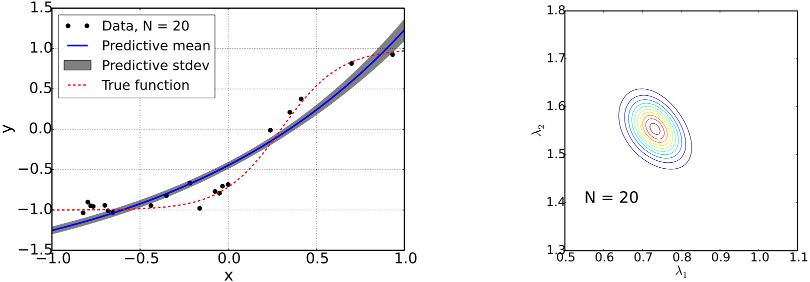
Model-data fit



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 $f(x,\lambda)$

Demo

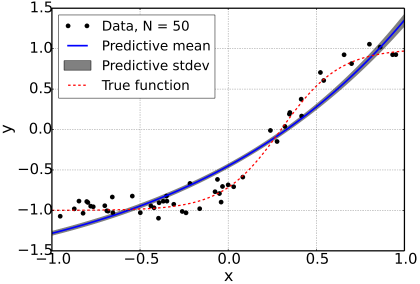


Model-data fit

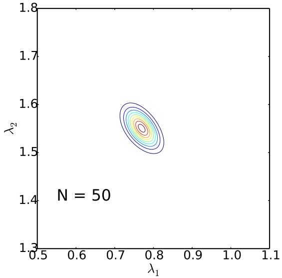
Posterior on parameters

- 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
  - We are increasingly sure about predictions based on the *wrong* model

Demo



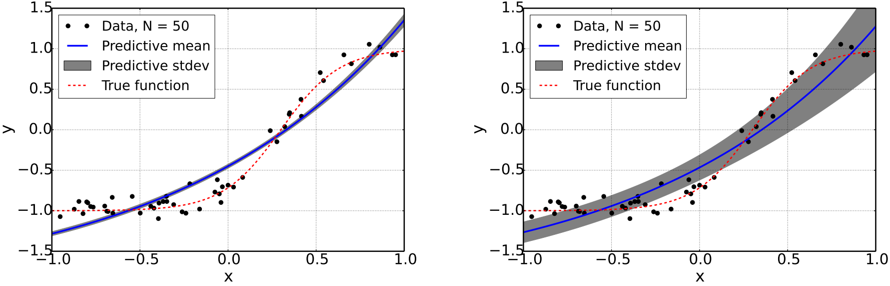
Model-data fit



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
    - We are increasingly sure about predictions based on the *wrong* model

Demo

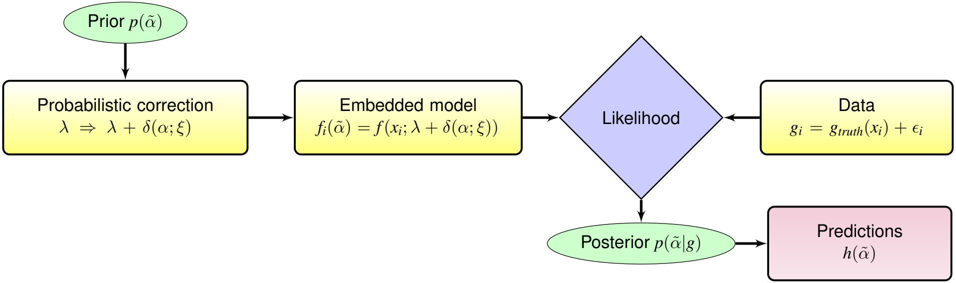


No model error treatment

Embedded model error

- 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 $f(x, \lambda)$
- Embedding model error allows extra uncertainty component to propagate through predictions

Model error embedding – schematic



- Infer *both* physical parameters λ and model-error representation α: $\tilde{\alpha} = (\lambda, \alpha)$
- Predictive uncertainty decomposition:

Total Variance =

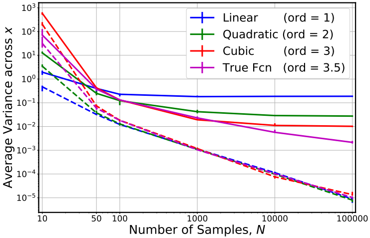
Parametric uncertainty + Data noise + Model error + Surrogate error

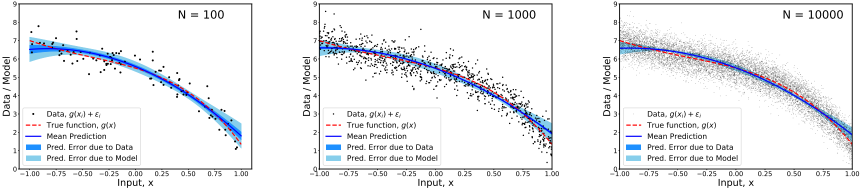
More data leads to ‘leftover’ model error

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 $\sigma = 0.1$.

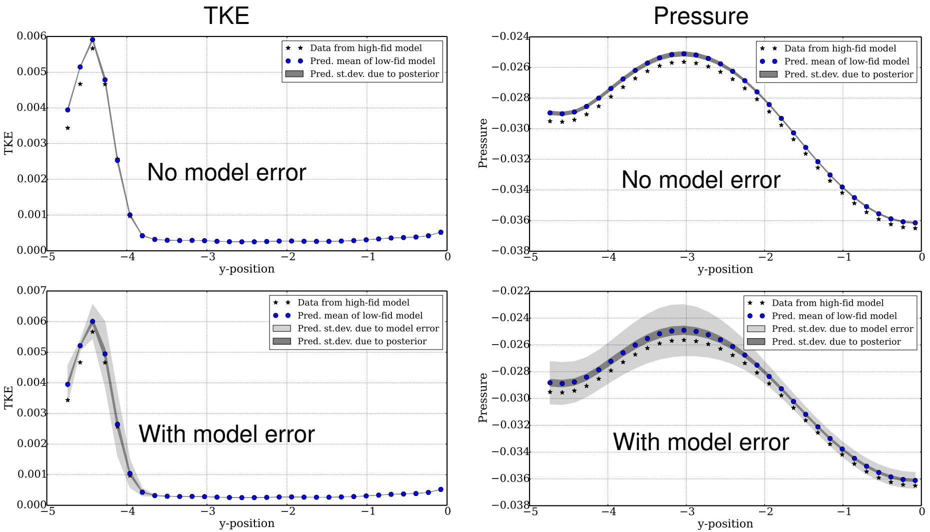
Summary of features:

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





Calibrate with Turbulent Kinetic Energy (TKE) data, predict both TKE and Pressure



Summary Thank You

- Represent, quantify and propagate model structural errors
- Bayesian machinery for simultaneous estimation of physical parameters and structural error
- Differentiates from data noise; allows model-to-model calibration
- Applied in climate land models, transport models, LES, chemistry, fusion.
- Implemented in UQTk (www.sandia.gov/UQToolkit)

- K. Sargsyan, H. Najm, and R. Ghanem. “On the Statistical Calibration of Physical Models”. *International Journal for Chemical Kinetics*, 47(4): 246-276, 2015.
- K. Sargsyan, X. Huan, and H. Najm. “Embedded model error representation for model calibration”. In preparation, 2017.

- Plenty of challenges remaining - best tackled with a driving application
- Open to talk to applications: hierarchy of models, model-vs-data