Density Estimation Framework for Model Error Assessment

K. Sargsyan[1], H. Najm[1], J. Bender[2], R. Ghanem[3]

[1]Sandia National Laboratories, Livermore, CA
[2]University of Minnesota, MN
[3]University of Southern California, Los Angeles, CA

13th USNCCM
San Diego
July 26-30, 2015

Sandia National Laboratories is a multi-program laboratory operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Density Estimation Framework for Model Error Assessment

K. Sargsyan[1], H. Najm[1], J. Bender[2], R. Ghanem[3]

[1]Sandia National Laboratories, Livermore, CA
[2]University of Minnesota, MN
[3]University of Southern California, Los Angeles, CA

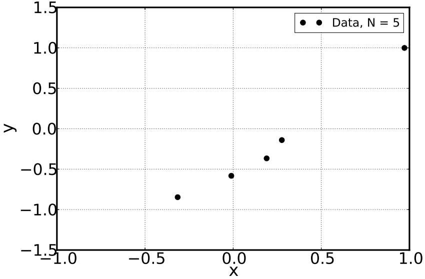
13^{th} USNCCM
San Diego
July 26-30, 2015

- DOE Office of Advanced Scientific Computing Research (ASCR), Scientific Discovery through Advanced Computing (SciDAC)
- DOE Office of Basic Energy Sciences (BES), Div. of Chem. Sci., Geosci., & Biosci.
- SNL Laboratory Directed Research and Development (LDRD)

Motivation

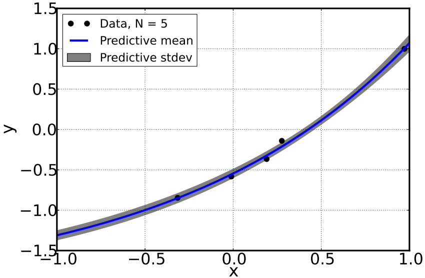
- All models are wrong, but some are useful *[George Box]*
- Models of physical systems rely on
  - Presumed theoretical framework
  - Mathematical formulation
  - Simplifying assumptions, parameterizations
  - Numerical discretization of governing equations
  - Computational software & hardware
- Model error is frequently non-negligible
- Estimating model error is useful for
  - Model validation
  - Model comparison
  - Scientific discovery and model improvement
  - Reliable computational predictions

Motivation

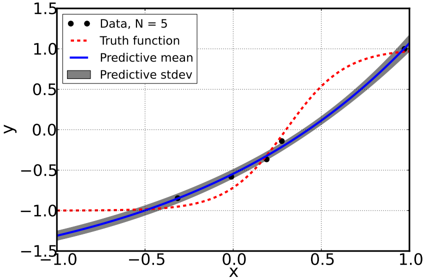


Model-data fit

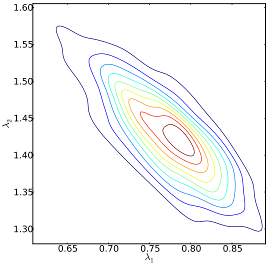
Motivation



Model-data fit

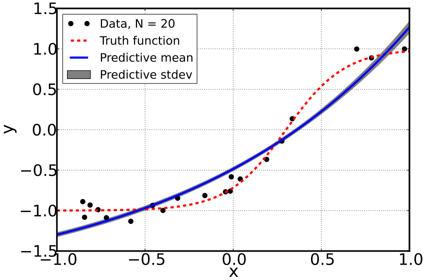


Model-data fit

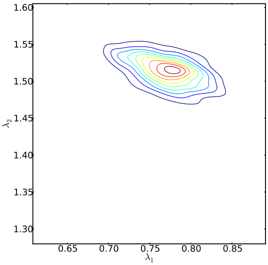


Calibrated parameters

Motivation

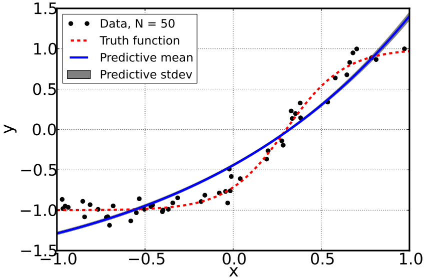


Model-data fit

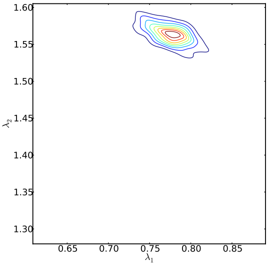


Calibrated parameters

Motivation

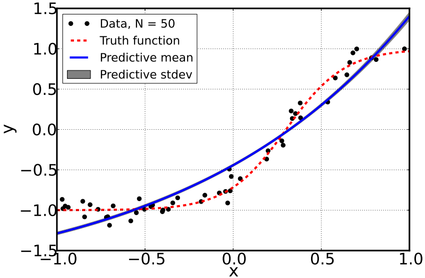


Model-data fit

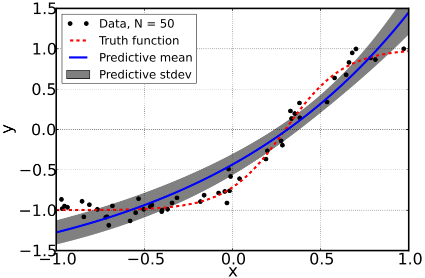


Calibrated parameters

Motivation



Model-data fit



What we want

- If the model has structural errors, more data does not help!
- We target model-vs-truth discrepancy

Data-Model-Truth

- Measurements

$$\overset{\text{data}}{y_i} = \overset{\text{truth}}{g(x_i)} + \overset{\text{data error}}{\epsilon_i^{\mathrm{d}}}$$

- Model

$$\overset{\text{truth}}{g(x_i)} = \overset{\text{model}}{f(x_i;\lambda)} + \overset{\text{model error}}{\delta(x_i)}$$

- Total error budget

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

Statistical modeling of errors in calibrating $f(x;\lambda)$

Data Error: $\epsilon_i^{\mathrm{d}} \sim \mathrm{N}(0, \sigma^2)$

Model Error: $\delta(x) \sim \mathrm{GP}(\mu(x), C(x, x'))$

Estimate model parameters λ along with those of $\delta(x)$, ϵ_i^{d}

State-of-the-art: Issues for physical models

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

- Explicit additive statistical model for model error $\delta(x)$ Kennedy-O'Hagan (2001).
- Calibrated predictive model $\quad y_{mod}(x) = f(x;\lambda) + \delta(x)$
- Potential violation of physical constraints
  - *e.g.* incompressible flow: $\nabla \bullet v = 0$
- Disambiguation of model error $\delta(x_i)$ and data error ϵ_i^{d}
- 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

Model error embedding: key idea

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

- Cast input parameters λ as a random variable Λ
  Black-box

$$y_i = f(x_i; \Lambda) + \epsilon_i^{\mathrm{d}}$$

- More generally, explore additional parameterizations,
  Extra 'physics'

$$y_i = \tilde{f}(x_i; \lambda, \Theta) + \epsilon_i^{\mathrm{d}}$$

Model error embedding: features

$$y_i = f(x_i; \Lambda) + \epsilon_i^{\mathrm{d}}$$

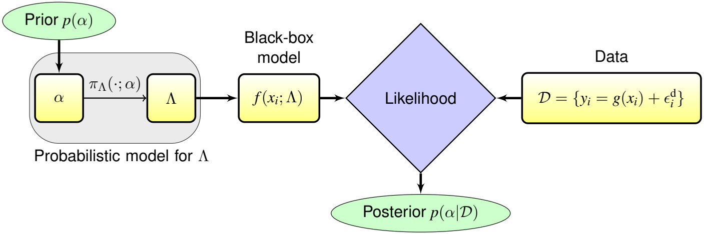
$$y_i = \tilde{f}(x_i; \lambda, \Theta) + \epsilon_i^{\mathrm{d}}$$

- Embed model error in specific submodel phenomenology
  - a modified transport or constitutive law
  - a modified formulation for a material property
- 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

Model error embedding: Bayesian formulation

- In the simplest setting, cast λ as a random variable Λ
  Black-box $$y_i = f(x_i;\Lambda) + \epsilon_i^{\mathrm{d}}$$
- Calibration turns into density estimation for the PDF of Λ
- Polynomial Chaos parameterization for $\Lambda = \sum_{k=0}^{K} \alpha_k \Psi_k(\xi)$
- Back to parameter estimation, now for $\alpha = (\alpha_0, \ldots, \alpha_K)$

$$\underbrace{p(\alpha|\mathcal{D})}_{\text{Posterior}} \propto \underbrace{L_{\mathcal{D}}(\alpha)}_{\text{Likelihood}} \underbrace{p(\alpha)}_{\text{Prior}}$$



Likelihood construction

- Data $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^N$
- Data model (assuming Gauss-Hermite PC for parameters Λ):

$$
\begin{aligned}
y_i &= f(x_i, \Lambda) + \epsilon_i^{\mathrm{d}} = f\left(x_i, \sum_k \alpha_k \Psi_k(\xi_1, \ldots, \xi_d)\right) + \sigma_{\mathcal{D}} \xi_{d+i} \\
&\overset{NISP}{=} \sum_k f_{ik}(\alpha) \Psi_k(\xi_1, \ldots, \xi_d) + \sigma_{\mathcal{D}} \xi_{d+i} = h_i(\alpha, \sigma_{\mathcal{D}}; \boldsymbol{\xi})
\end{aligned}
$$

- Infer $\hat{\alpha} = (\alpha, \sigma_{\mathcal{D}})$.

  Note: for each $\hat{\alpha}$, the data model $\boldsymbol{h}(\hat{\alpha}; \boldsymbol{\xi})$ is a multivariate random variable with easily accessible mean $\boldsymbol{\mu}$ and covariance Σ.

- Full Likelihood: $L(\hat{\alpha}) = p(\mathcal{D}|\hat{\alpha}) = p(y_1, \ldots, y_N|\hat{\alpha}) = \pi_{\boldsymbol{h}}(\boldsymbol{y})$
  - Degenerate if no data noise
  - Requires multivariate KDE
  - Gaussian approximation: $L(\hat{\alpha}) \propto \exp\left(-\frac{1}{2}(\boldsymbol{y} - \boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{y} - \boldsymbol{\mu})\right)$

Likelihood construction

- Data $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^N$
- Data model (assuming Gauss-Hermite PC for parameters Λ):

$$
\begin{aligned}
y_i &= f(x_i, \Lambda) + \epsilon_i^{\mathrm{d}} = f\left(x_i, \sum_k \alpha_k \Psi_k(\xi_1, \ldots, \xi_d)\right) + \sigma_{\mathcal{D}} \xi_{d+i} \\
&\overset{NISP}{=} \sum_k f_{ik}(\alpha) \Psi_k(\xi_1, \ldots, \xi_d) + \sigma_{\mathcal{D}} \xi_{d+i} = h_i(\alpha, \sigma_{\mathcal{D}}; \boldsymbol{\xi})
\end{aligned}
$$

- Infer $\hat{\alpha} = (\alpha, \sigma_{\mathcal{D}})$.

  Note: for each $\hat{\alpha}$, the data model $\boldsymbol{h}(\hat{\alpha}; \boldsymbol{\xi})$ is a multivariate random variable with easily accessible mean $\boldsymbol{\mu}$ and covariance Σ.

- Marginalized Likelihood: $L(\hat{\alpha}) = p(D|\hat{\alpha}) = \prod_{i=1}^N p(y_i|\hat{\alpha}) = \prod_{i=1}^N \pi_{h_i}(y_i)$
  - Gaussian approximation: $L(\hat{\alpha}) \propto \exp\left(-\frac{1}{2} \sum_{i=1}^N \Sigma_{ii}^{-2} (y_i - \mu_i)^2\right)$

Likelihood construction

- Data $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^N$
- Data model (assuming Gauss-Hermite PC for parameters Λ):

$$
\begin{aligned}
y_i &= f(x_i, \Lambda) + \epsilon_i^{\mathrm{d}} = f\left(x_i, \sum_k \alpha_k \Psi_k(\xi_1, \ldots, \xi_d)\right) + \sigma_{\mathcal{D}} \xi_{d+i} \\
&\overset{NISP}{=} \sum_k f_{ik}(\alpha) \Psi_k(\xi_1, \ldots, \xi_d) + \sigma_{\mathcal{D}} \xi_{d+i} = h_i(\alpha, \sigma_{\mathcal{D}}; \boldsymbol{\xi})
\end{aligned}
$$

- Infer $\hat{\alpha} = (\alpha, \sigma_{\mathcal{D}})$.

  Note: for each $\hat{\alpha}$, the data model $\boldsymbol{h}(\hat{\alpha}; \boldsymbol{\xi})$ is a multivariate random variable with easily accessible mean $\boldsymbol{\mu}$ and covariance Σ.

- Approximate Bayesian Computation (ABC): $L(\hat{\alpha}) = \frac{1}{\epsilon} K\left(\frac{\rho(\mathcal{S}_{\mathcal{M}}, \mathcal{S}_{\mathcal{D}})}{\epsilon}\right)$
  - $p(y|\mathcal{D})$ is "centered" on the data
  - The width of the distribution $p(y|\mathcal{D})$ is consistent with the spread of the data around the nominal model prediction

$$
L(\hat{\alpha}) \propto \exp\left(-\frac{1}{2\epsilon^2} \sum_{i=1}^N \left[(\mu_i(\alpha) - y_i)^2 + (\sigma_i(\alpha) - \gamma|\mu_i(\alpha) - y_i|)^2\right]\right)
$$

Calibrated predictions – general x

For fixed α, e.g. Maximum a Posteriori (MAP) value,

Uncertain prediction:
$$y = f(x, \lambda(\alpha; \boldsymbol{\xi}))$$

Mean:
$$\mu(x,\alpha) = \mathbb{E}_{\boldsymbol{\xi}}[f(x, \lambda(\alpha; \boldsymbol{\xi}))]$$

Variance:
$$\sigma^2(x,\alpha) = \mathbb{V}_{\boldsymbol{\xi}}[f(x, \lambda(\alpha; \boldsymbol{\xi}))]$$

Average over posterior of α

Pushed-forward posterior:
$$p(f|\mathcal{D}) = \int p(f|\alpha)p(\alpha|\mathcal{D})d\alpha = \mathbb{E}_\alpha[p(f|\alpha)]$$

Pushed-forward mean:
$$\mu_{\mathrm{PFP}}(x) = \mathbb{E}_{\boldsymbol{\xi}}\mathbb{E}_\alpha[f(x, \lambda(\alpha; \boldsymbol{\xi}))] = \mathbb{E}_\alpha \mu(x,\alpha)$$

Pushed-forward variance:
$$\sigma^2_{\mathrm{PFP}}(x) = \underbrace{\mathbb{E}_\alpha[\sigma^2(x,\alpha)]}_{\text{model error}} + \underbrace{\mathbb{V}_\alpha[\mu(x,\alpha)]}_{\text{data noise}}$$

Calibrated predictions – compare to data at x_i

For fixed α, e.g. Maximum a Posteriori (MAP) value,
Uncertain prediction:

$$y_i = h(x_i, \lambda(\alpha; \hat{\boldsymbol{\xi}})) = f(x_i, \lambda(\alpha; \boldsymbol{\xi})) + \sigma_{\mathcal{D}} \xi_{d+i}$$

Mean:

$$\mu(x_i, \alpha) = \mathbb{E}_{\boldsymbol{\xi}}[f(x_i, \lambda(\alpha; \boldsymbol{\xi}))]$$

Variance:

$$\sigma^2(x_i, \alpha) = \mathbb{V}_{\boldsymbol{\xi}}[f(x_i, \lambda(\alpha; \boldsymbol{\xi}))] + \sigma_{\mathcal{D}}^2$$

Average over posterior of α

Pushed-forward posterior:

$$p(h|\mathcal{D}) = \int p(h|\alpha, \sigma_{\mathcal{D}}) p(\alpha, \sigma_{\mathcal{D}}|\mathcal{D}) d\alpha d\sigma_{\mathcal{D}} = \mathbb{E}_{\alpha, \sigma_{\mathcal{D}}}[p(h|\alpha, \sigma_{\mathcal{D}})]$$

Pushed-forward mean:

$$\mu_{\mathrm{PFP}}(x_i) = \mathbb{E}_{\boldsymbol{\xi}} \mathbb{E}_{\alpha}[f(x_i, \lambda(\alpha; \boldsymbol{\xi}))] = \mathbb{E}_{\alpha} \mu(x_i, \alpha)$$

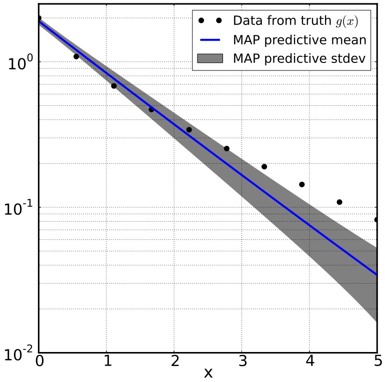
Pushed-forward variance:

$$\sigma_{\mathrm{PFP}}^2(x_i) = \underbrace{\mathbb{E}_{\alpha}[\sigma^2(x_i, \alpha)]}_{\text{model error}} + \underbrace{\mathbb{V}_{\alpha}[\mu(x_i, \alpha)] + \mathbb{E}_{\sigma_{\mathcal{D}}}[\sigma_{\mathcal{D}}^2]}_{\text{data noise}}$$

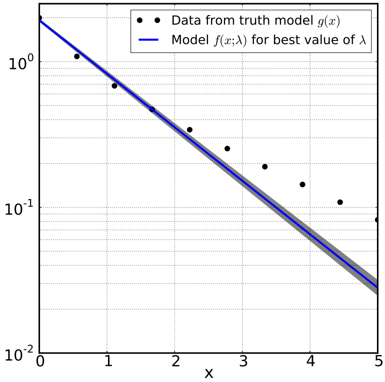
Predictions account for model error

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

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



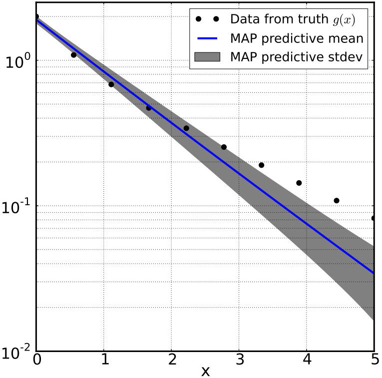
Additive Gaussian error



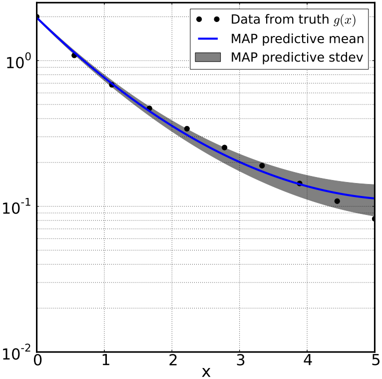
Predictions account for model error

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

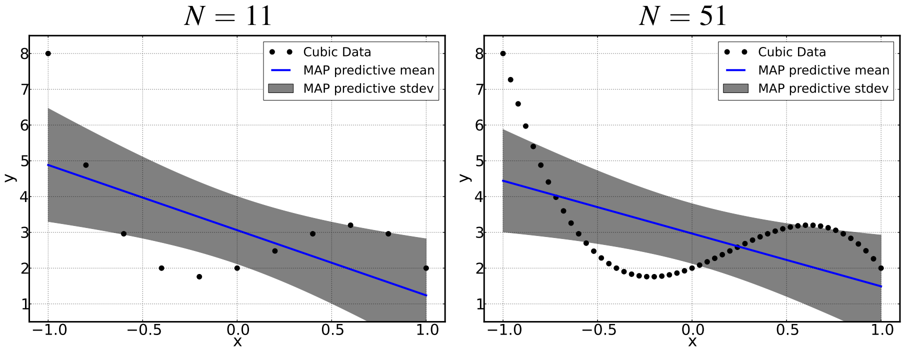
Linear-exponential $f(x, \lambda) = e^{\lambda_1 + \lambda_2 x}$



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

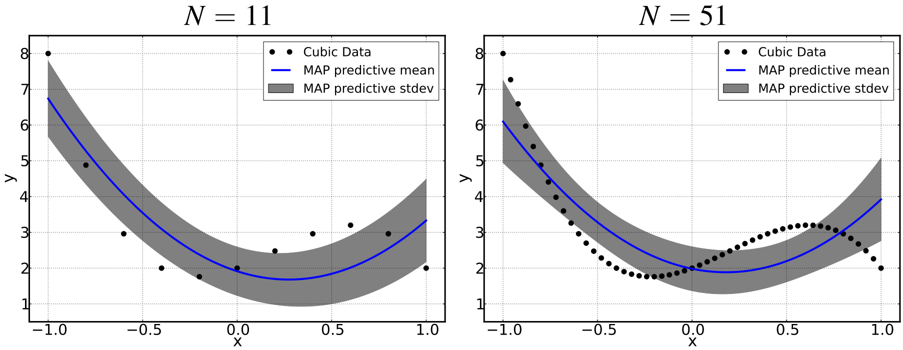


Test problem – Cubic data fit by a line – ABC



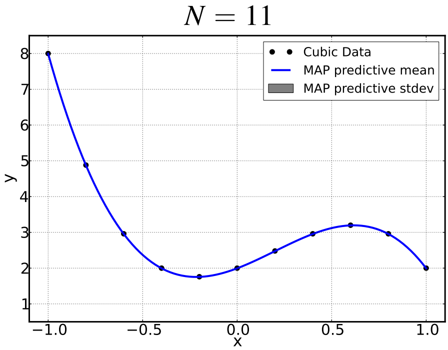
- MAP predictive mean centered on data
- MAP predictive standard deviation captures range of discrepancy
- Increasing number of data points has a small effect on both predictive mean and stdev.

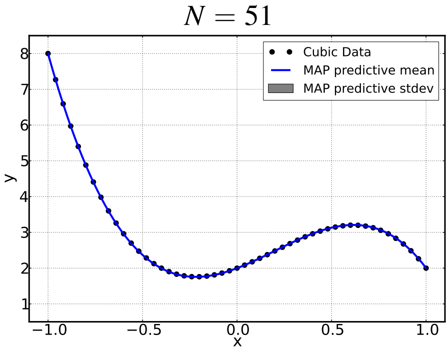
Test problem – Cubic data fit by a quadratic – ABC



- MAP predictive mean centered on data
- MAP predictive standard deviation captures range of discrepancy
- Increasing number of data points has a small effect on both predictive mean and stdev.

Test problem – Cubic data fit by a cubic – ABC

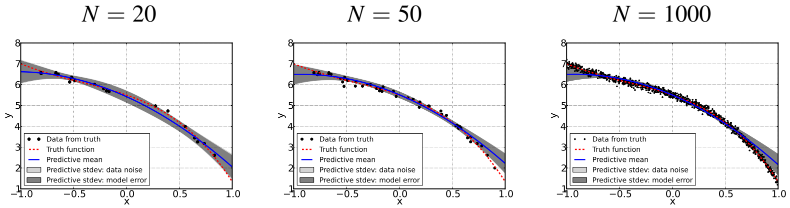




- MAP predictive mean centered on data
- MAP predictive standard deviation captures range of discrepancy
- Increasing number of data points has a small effect on both predictive mean and stdev.

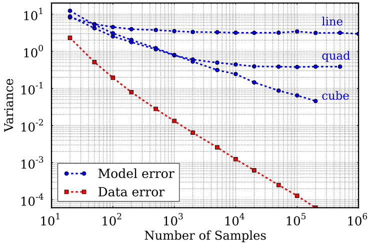
More data leads to 'leftover' model error

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



Chemistry problem – ABC

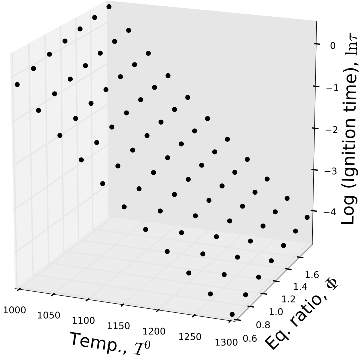
- Homogeneous ignition, methane-air mixture
- Single-step global reaction model calibrated against a detailed chemical kinetic model

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

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$$

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

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

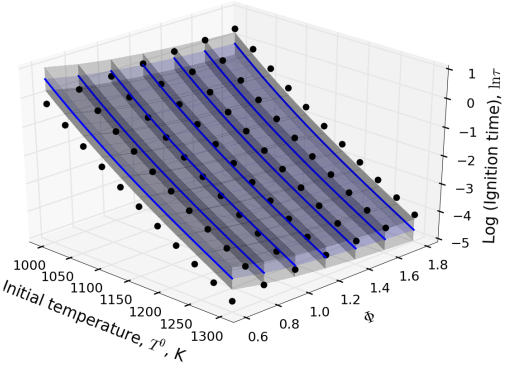


Quality of Uncertain Calibrated Model Predictions

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, HNN, and R. Ghanem
"On the Statistical Calibration of Physical Models"
Int. J. Chem. Kin., 47(4): 246-276, 2015

TransCom3 Experiment of CO_2 Flux Inversion

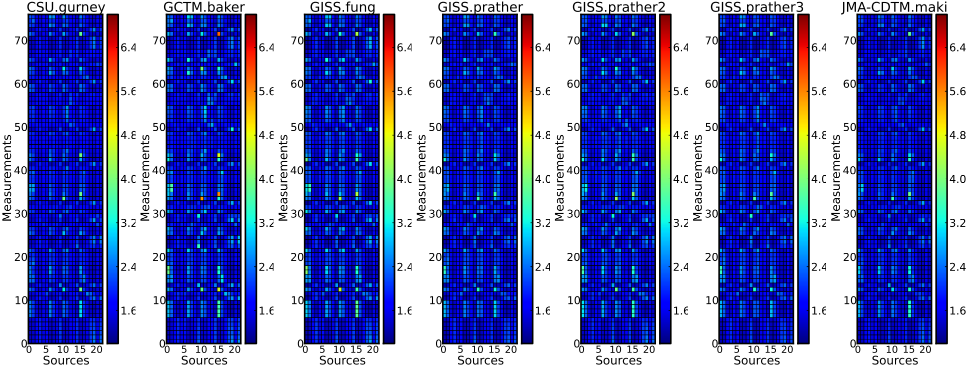
[Gurney *et al.*, Tellus B, 2003]

- 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 $\mathbf{d} \approx \mathbf{Rs}$

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

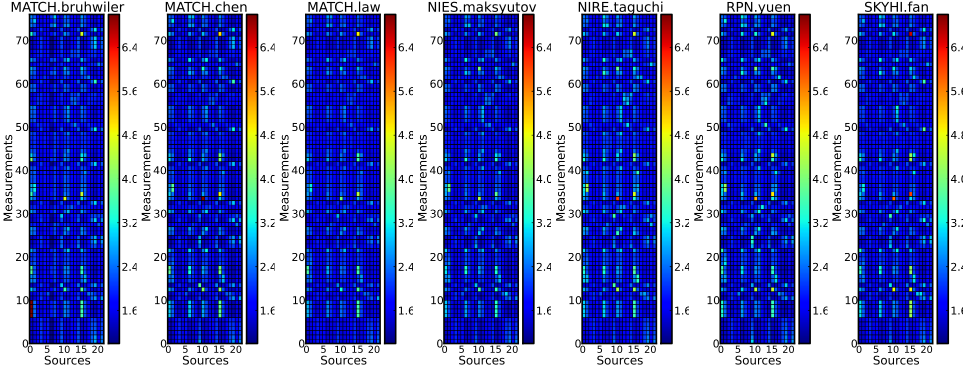
- Model $\mathbf{R}$ is never perfect thus contaminating the inversion
- The inferred values of $\mathbf{s}$ compensate for model deficiencies
- $\epsilon_{\mathbf{d}}$ is meant to capture data errors, but is 'entangled' with model errors

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



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

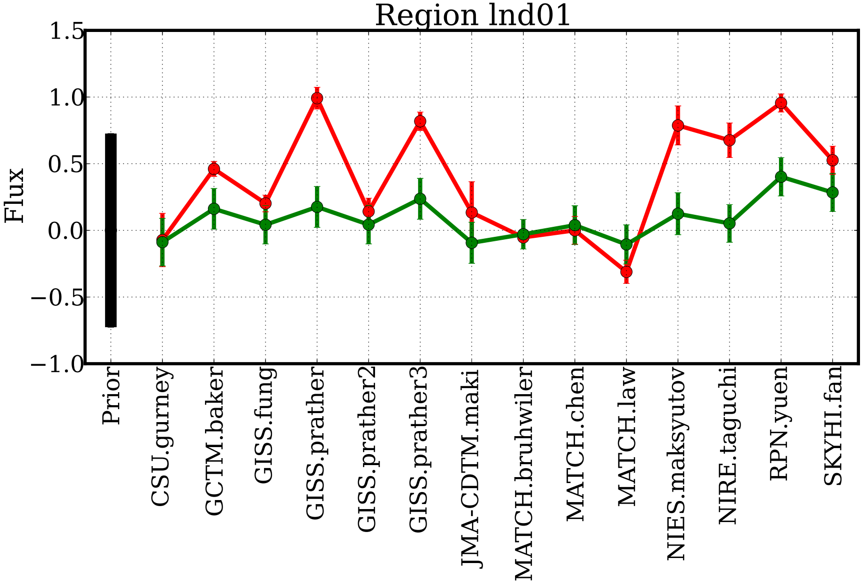
- Conventional additive Gaussian error (least-squares): $\mathbf{d} = \mathbf{R}\mathbf{s} + \xi$
- Embed probabilistic model for fluxes s: $\mathbf{d} = \mathbf{R}(\mu_{\mathbf{s}} + \mathbf{C}_{\mathbf{s}}\xi)$



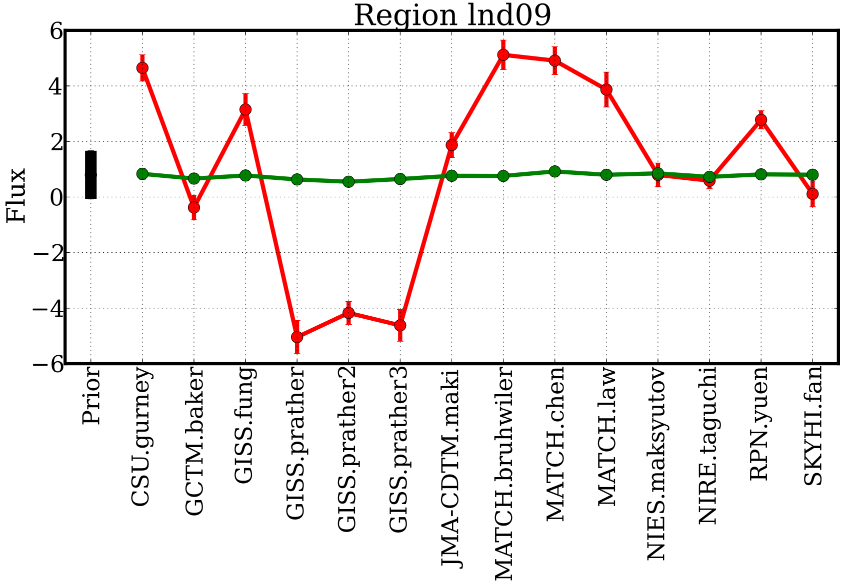
Infer fluxes $\mathbf{s}$, given measurements $\mathbf{d}$ to satisfy $\mathbf{d} \approx \mathbf{Rs}$

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

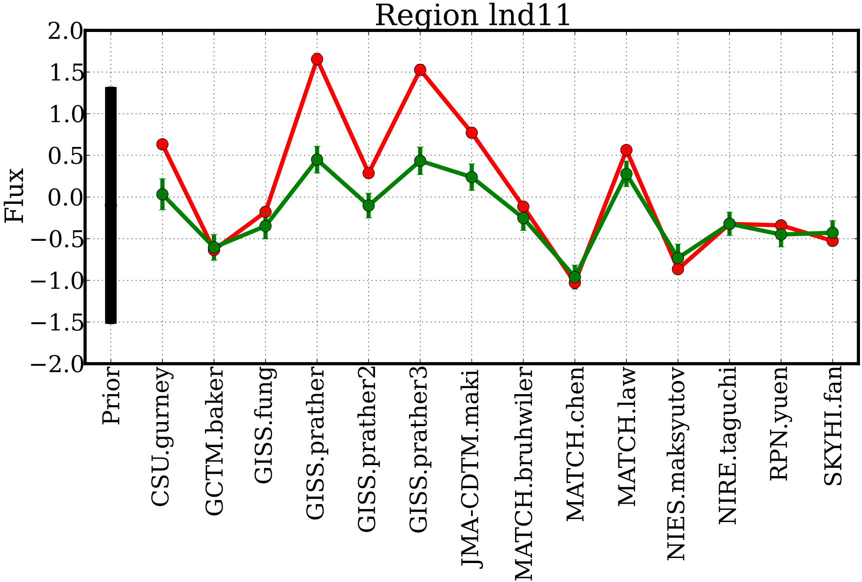
Inferred fluxes show less variability across models



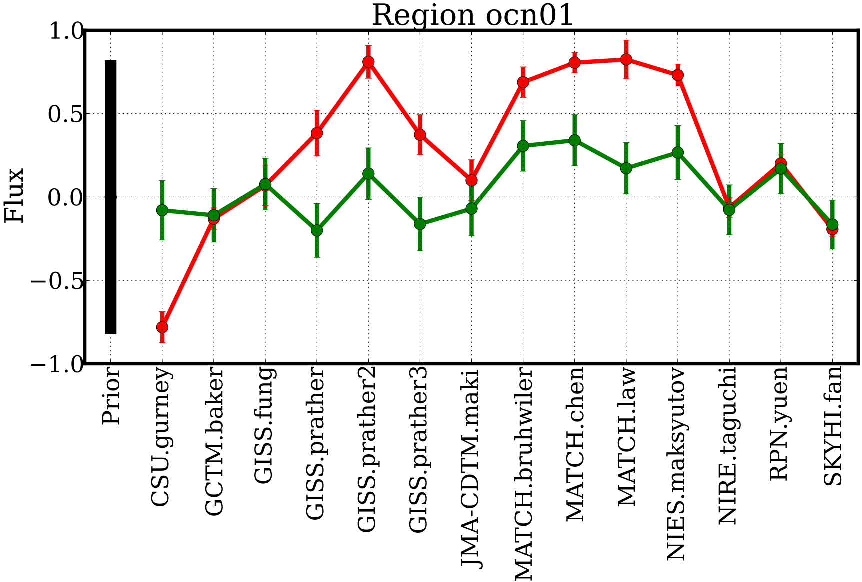
Inferred fluxes show less variability across models



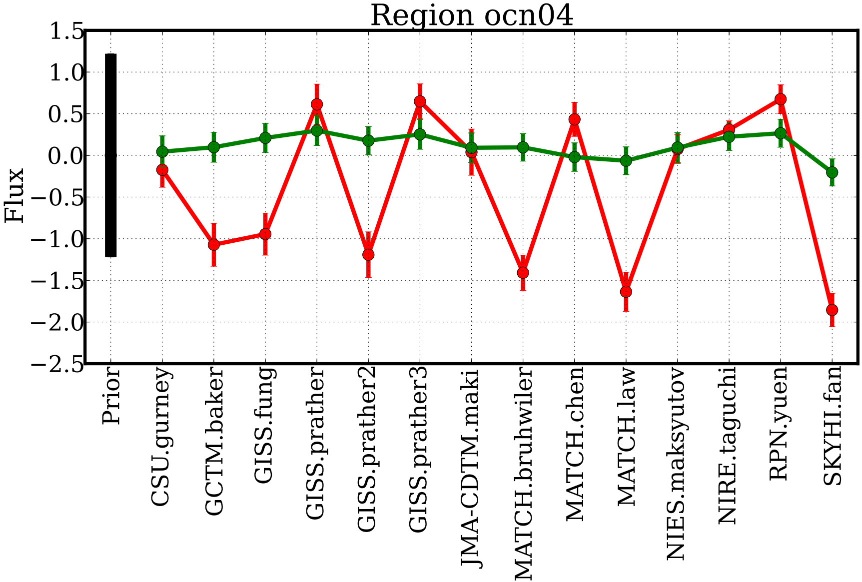
Inferred fluxes show less variability across models



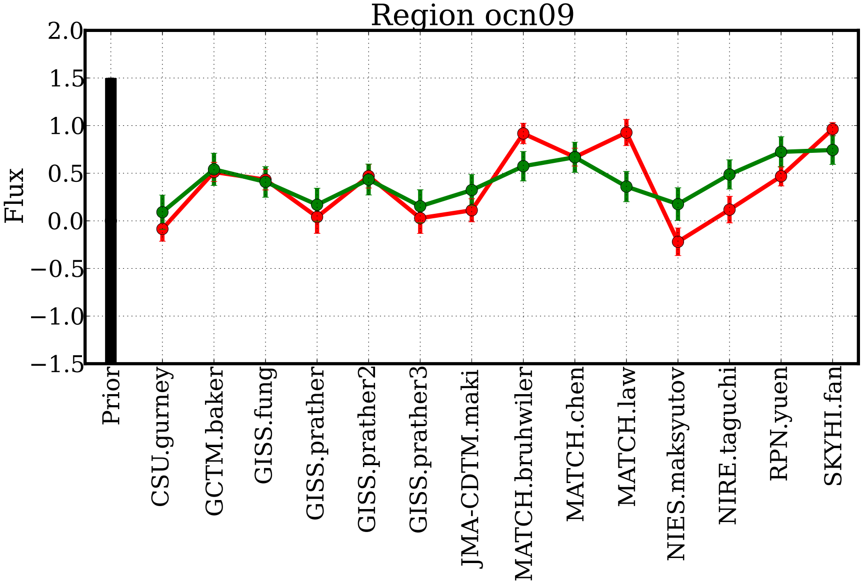
Inferred fluxes show less variability across models



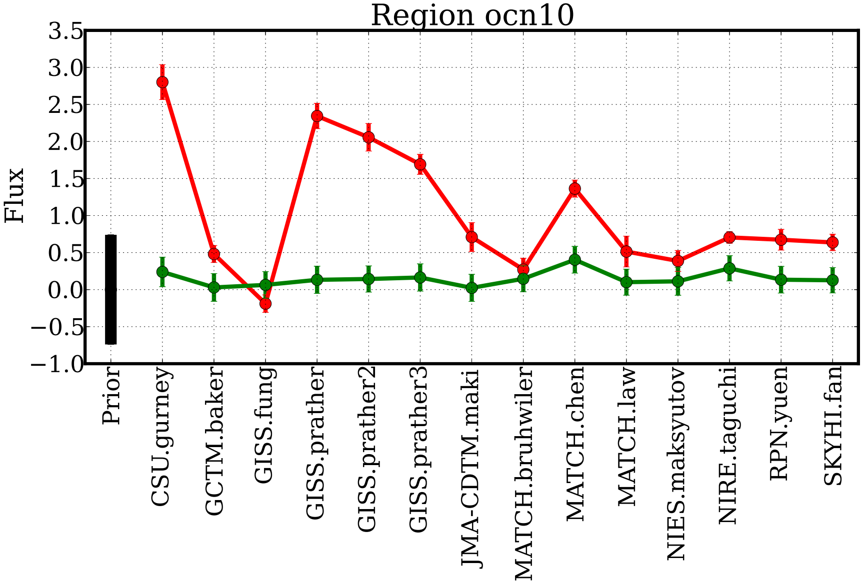
Inferred fluxes show less variability across models



Inferred fluxes show less variability across models



Inferred fluxes show less variability across models



Summary

Thank You

Method

- Targeting model discrepancy from truth
- Reformulate the calibration as a density (PDF) estimation problem
- Bayesian machinery to find PDF parameters
- Likelihood construction targets constraints of interest

Features

- Respects the physical constraints
- Mechanism for model-to-model calibration
- Disambiguates model error (irreducible) and data noise (reducible)
- Meaningful prediction uncertainties

- K. Sargsyan, H. Najm, and R. Ghanem, "On the Statistical Calibration of Physical Models". *International Journal for Chemical Kinetics*, 47(4): pp 246–276, 2015.
- K. Sargsyan *et. al.*, "Bias-Enhanced Bayesian Inference of Atmospheric Trace Gas Sources and Sinks", *in progress, 2015*.

Full Likelihood

$$L(\alpha) = p(D|\alpha) = \pi_f(y_{\text{data},1}, \ldots, y_{\text{data},N}|\alpha)$$

where:
$\pi_f(\cdot, \alpha)$: N-variate density of the random variable $(f_1, \ldots, f_N)$
with $f_i = f(x_i, \lambda(\alpha))$

Problem: $\pi_f(\cdot)$ is degenerate in general when $N > M$

Consider a case with $M = 1$, $\lambda \sim N(\mu, \sigma^2)$, and $f = \lambda x$
Let $N = 2$, hence $(f_1, f_2) = (\lambda x_1, \lambda x_2)$ for any λ sample
With $f_1/x_1 = f_2/x_2 = \lambda$, (f_1, f_2) are dependent and
$\pi_f(\cdot|\mu, \sigma)$ is non-zero only along the line $f_2 = (x_2/x_1)f_1$

hence $\pi_f(y_{\text{data},1}, y_{\text{data},2}|\mu, \sigma)$ is non-zero only along the line
$y_{\text{data},2}/x_2 = y_{\text{data},1}/x_1$

Marginalized Likelihood

$$L(\alpha) = p(D|\alpha) = \prod_{i=1}^{N} \pi_{f_i}(y_{\text{data},i}|\alpha)$$

where $\pi_{f_i}(\cdot, \alpha)$ is the univariate density of the RV $f_i = f(x_i, \lambda(\alpha))$

Problem: the likelihood has multiple singularities corresponding to α values leading to vanishing marginal variances at each x_i

Gaussian example: Let $f_i \sim \mathrm{N}(\mu_i(\alpha), \sigma_i(\alpha)^2)$, then

$$L(\alpha) = \frac{1}{(2\pi)^{N/2}} \prod_{i=1}^{N} \frac{1}{\sigma_i(\alpha)} \exp\left(\frac{(\mu_i(\alpha) - y_{\text{data},i})^2}{2\sigma_i(\alpha)^2} \right)$$

Multiple singularities, $\sigma_i(\alpha) = 0$, $i = 1, \ldots, N$

Posterior maximization always finds one of these singularities, fitting one point perfectly, while misfitting the rest   ($\Rightarrow$ priors)

Approximate Bayesian Computation (ABC)

Employ a kernel density as a pseudo-likelihood to enforce select constraints:

- Uncertain prediction $p(y|D)$ is centered on the data
  - With $\mu_i(\alpha) = \mathrm{E}_\xi[f(x_i, \lambda(\xi;\alpha))]$:

    minimize $\| \mu_i(\alpha) - y_{\mathrm{data},i} \|_2^2$
- The width of the distribution $p(y|D)$ is consistent with the spread of the data around the nominal model prediction
  - With $\sigma_i^2(\alpha) = \mathrm{V}_\xi[f(x_i, \lambda(\xi,\alpha))]$:

    minimize $\| \sigma_i(\alpha) - \gamma|\mu_i(\alpha) - y_{\mathrm{data},i}| \|_2^2$
  - γ is a factor that specifies the desired match between σ_i and the discrepancy $|\mu_i(\alpha) - y_{\mathrm{data},i}|$, on average

ABC Likelihood

With $\rho(\mathcal{S})$ being a metric of the statistic $\mathcal{S}$, use the kernel function as an ABC likelihood:

$$L_{\mathrm{ABC}}(\alpha) = \frac{1}{\epsilon} K\left(\frac{\rho(\mathcal{S})}{\epsilon}\right)$$

where ϵ controls the severity of the consistency control

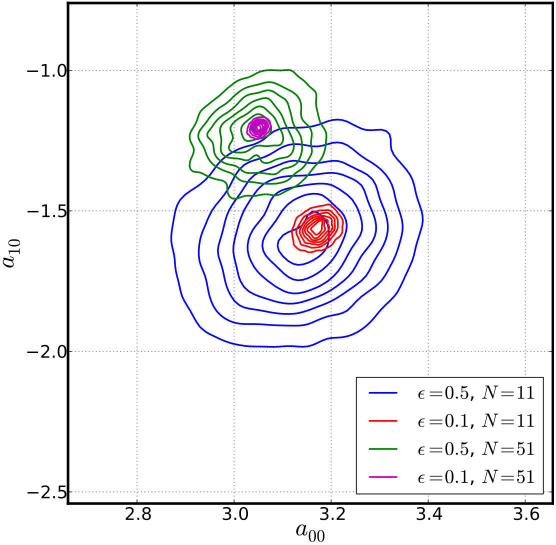
Propose the Gaussian kernel density:

$$L_{\epsilon}(\alpha) = \frac{1}{\epsilon\sqrt{2\pi}} \prod_{i=1}^{N} \exp\left(-\frac{(\mu_i(\alpha) - y_{\mathrm{d},i})^2 + (\sigma_i(\alpha) - \gamma|\mu_i(\alpha) - y_{\mathrm{d},i}|)^2}{2\epsilon^2}\right)$$

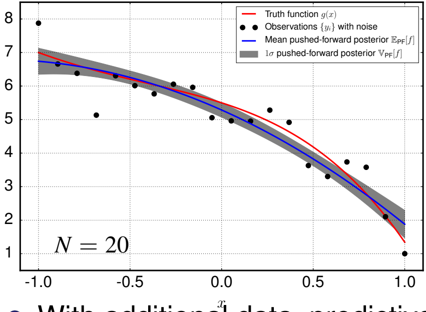
Test problem – Posterior density on α

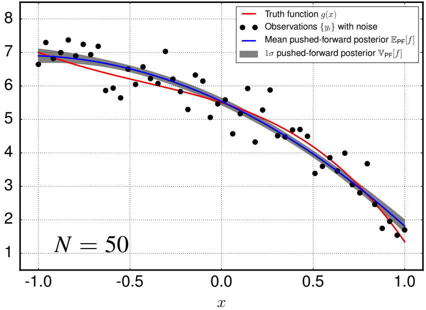
- Cubic data, line-fit
- Joint posterior on two elements of α

- Uncertainty in α is decreased by
  - Increasing N
  - Decreasing ϵ

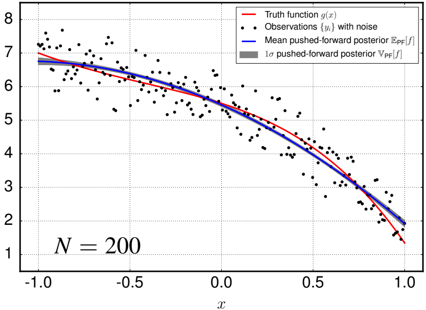


Quadratic-fit – Classical Bayesian likelihood

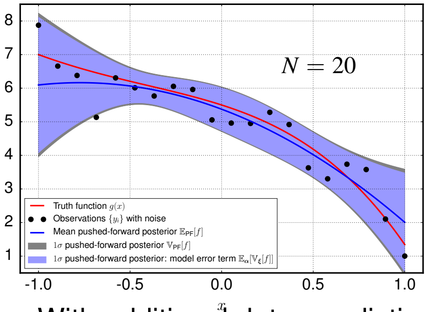


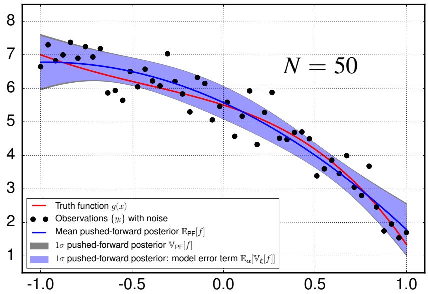


- With additional data, predictive uncertainty around the wrong model is indefinitely reducible
- Predictive uncertainty not indicative of discrepancy from truth

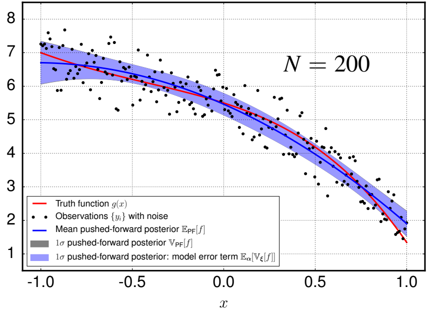


Quadratic-fit – ModErr – MargGauss

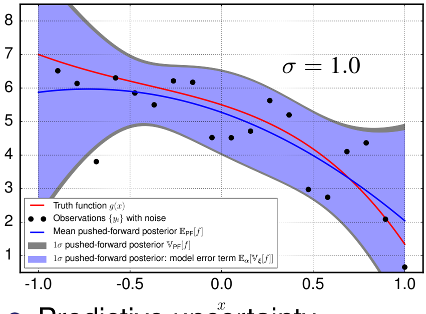


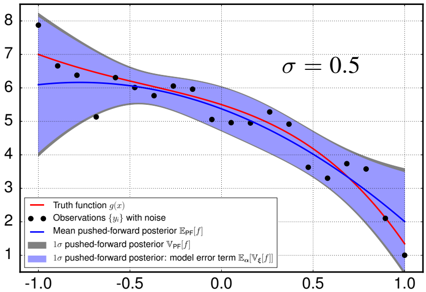


- With additional data, predictive uncertainty due to data noise is reducible
- Predictive uncertainty due to model error is not reducible

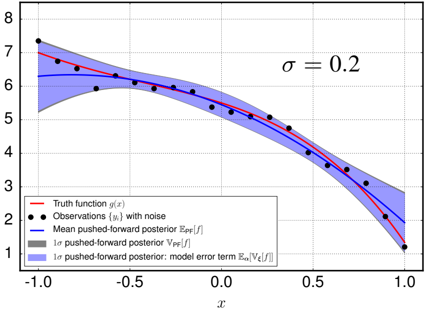


Quadratic-fit – ModErr – MargGauss

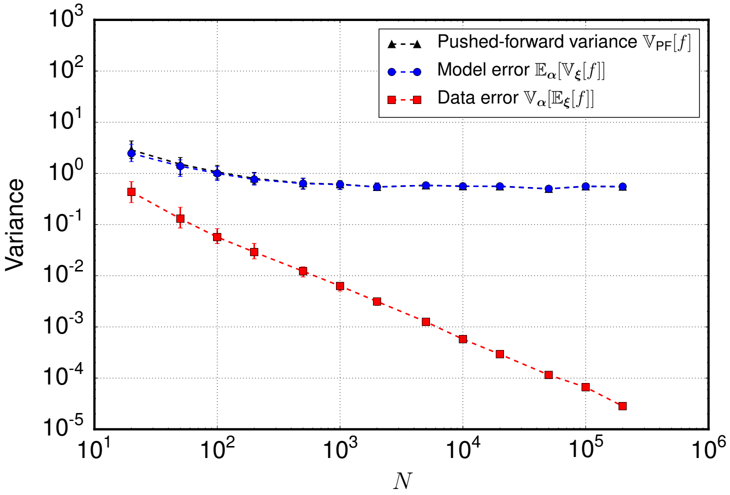




- Predictive uncertainty composed of both model-error and data-noise components
- The data-noise component is reducible with lower-noise in the data



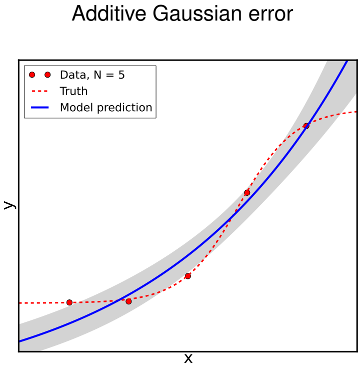
Quadratic-fit – ModErr – MargGauss

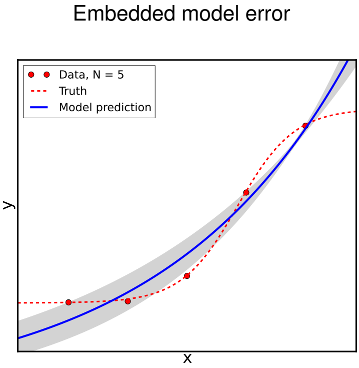


Calibrating a quadratic $f(x)$ w.r.t. $g(x) = 6 + x^2 + 0.5(x+1)^{3.5}$

Predictions account for model error: example 1

Calibrating an exponential model $f(x; \lambda_1, \lambda_2) = \lambda_2 e^{\lambda_1 x} - 2$
with data from a hyperbolic tangent model $g(x) = \tanh(3(x - 0.3))$

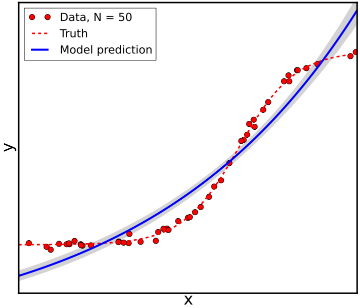




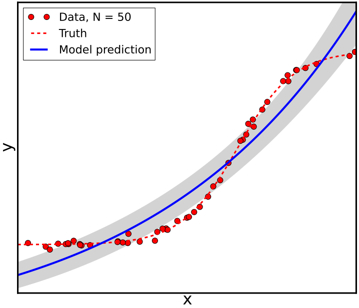
Predictions account for model error: example 1

Calibrating an exponential model $f(x;\lambda_1,\lambda_2)=\lambda_2 e^{\lambda_1 x}-2$
with data from a hyperbolic tangent model $g(x)=\tanh(3(x-0.3))$

Additive Gaussian error



Embedded model error



Challenges/Risks

- Density estimation is a well-known challenging task
  - Inverse problem is ill-posed or intractable

  $\Rightarrow$ Employ empirical likelihoods,
  Approximate Bayesian Computation (ABC)

- Potentially a high-dimensional Bayesian problem
  - Full posterior may be inaccessible

  $\Rightarrow$ Adaptive MCMC algorithms;
  resort to optimization algorithms in no-noise case

- Sparse or noisy data
  - With low information content, calibration may struggle

  $\Rightarrow$ More informative priors/regularization