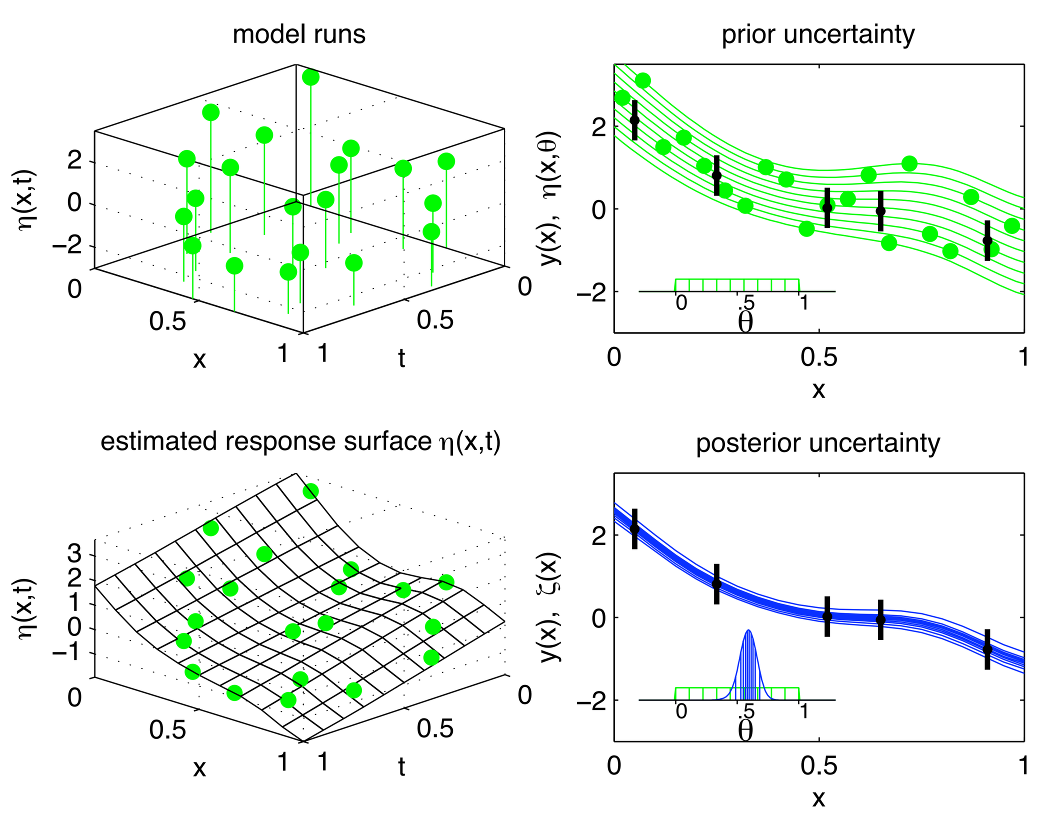
Quantifying Uncertainty with GPM/SA
Recent Work in QUEST Methods

James Gattiker, Dave Higdon
Sham Bhat, Earl Lawrence, Dave Osthus

Los Alamos, Statistical Sciences

Framework for Statistical Calibration & UQ



Important elements include:

- Design of computer experiments
- Response emulation
- Model Discrepancy
- Multivariate modeling with Linear Basis
- Calibration of parameters to observations

x	model or system inputs
θ	model calibration parameters
$\zeta(x)$	true physical system response given inputs x
$\eta(x,\theta)$	simulator response at x and θ.
$y(x)$	experimental observation of the physical system
$\delta(x)$	discrepancy between $\zeta(x)$ and $\eta(x,\theta)$ may be decomposed into numerical error and bias
$e(x)$	observation error of the experimental data

$$y(x) = \zeta(x) + e(x)$$
$$y(x) = \eta(x,\theta) + \delta(x) + e(x)$$

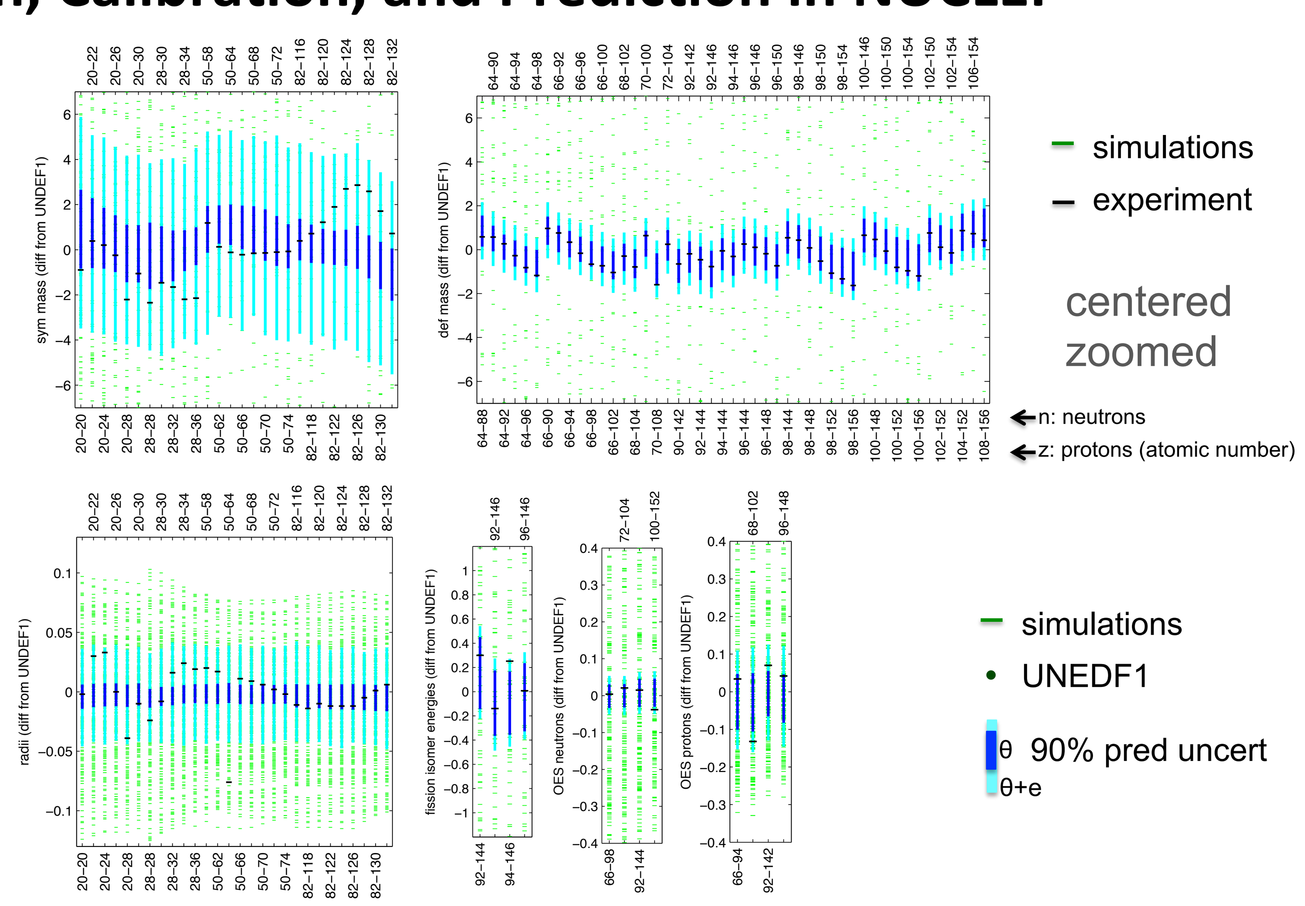
Framework Implemented in GPM/SA code

- Reference Implementation in Matlab, also C++ (GPMSA/Dakota/QUESO)
- Utilities for cross-validation, sensitivity analysis, prediction with uncertainty
- MCMC Sampling calibrates GPM/SA model and parameters:

$$\Sigma = \exp\{-\beta |x_1 - x_2|^p\}$$
$$\pi(\theta, \eta(\cdot,\cdot) \mid y(x)) \propto L(y(x) \mid \eta(x,\theta)) \times \pi(\theta) \times \pi(\eta(\cdot,\cdot))$$
$$L \approx \frac{1}{\sqrt{\det(\Sigma)}} \exp\{-\frac{1}{2}(Y-f)'\Sigma^{-1}(Y-f)\}$$

Emulation, Calibration, and Prediction in NUCLEI

Predicting Measured Atomic Properties



Hierarchical Linked Calibration

Challenges:

Different observation types may have different implications to parameter values and uncertainty
Different models may have different implications for an observation
It's not realistic to expect discrepancy to be known

Upside:

More observation groups/modes give more information about model structural error and uncertainty.

Inference model acknowledging bias terms for different observations:

$$y_i = \eta(x, \theta + b_i) + \delta_i(x, \theta + b_i) + \epsilon_i$$

The distribution of the b_i can be estimated by a hierarchical modeling approach:

$$y_i = \eta(x, \theta_i) + \delta_i(x, \theta_i) + \epsilon_i$$

$$\theta_i \propto N(\mu_\theta, \sigma_\theta^2)$$
$$\mu_\theta \propto U(0,1)$$
$$\lambda_\theta \equiv \frac{1}{\sigma_\theta^2} \propto \Gamma(a=1, b \to \infty)$$

Across the observations/models:

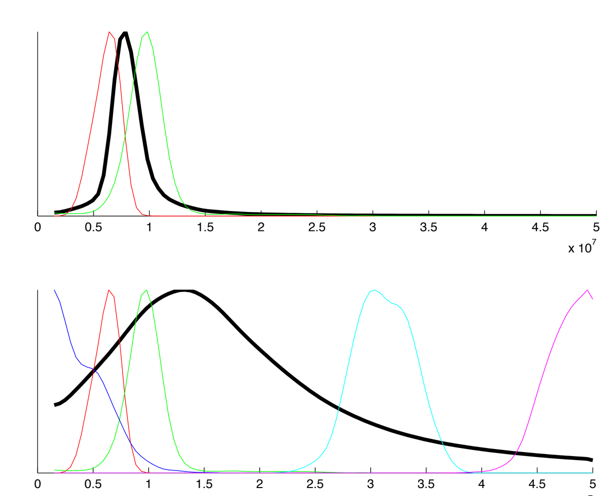
One extreme: parameters are identical
Other extreme: parameters are independent

Generally, this reveals a source of additional uncertainty.

Ocean Model Parameter Calibration:

Inference from red, green, blue measures (in isolation) are strong and consistent, combine as black

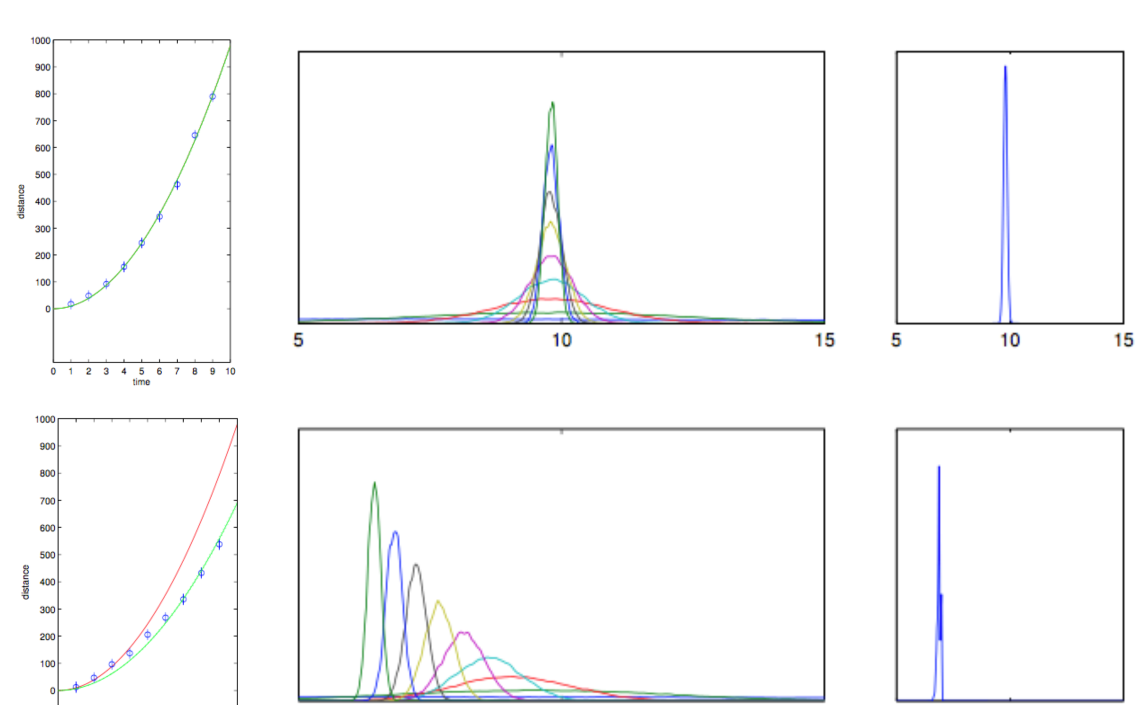
Conflicting inferences from cyan, magenta measures are weak, but not entirely ruled out.



Underlying issue
Toy model $x = gt^2 - kv^2t^2$ (falling object with drag)

If generating k=0 and simulation k=0
➔ Expected result

If generating k≠0 simulation k=0 (model deficiency)
➔ Biased *and* confident



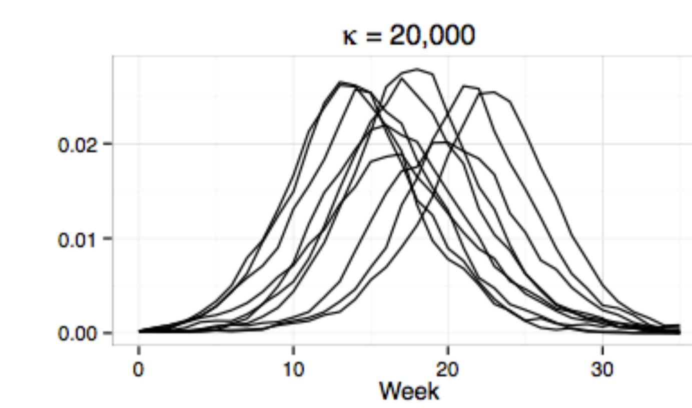
Evidence of each obs. *g* cal.

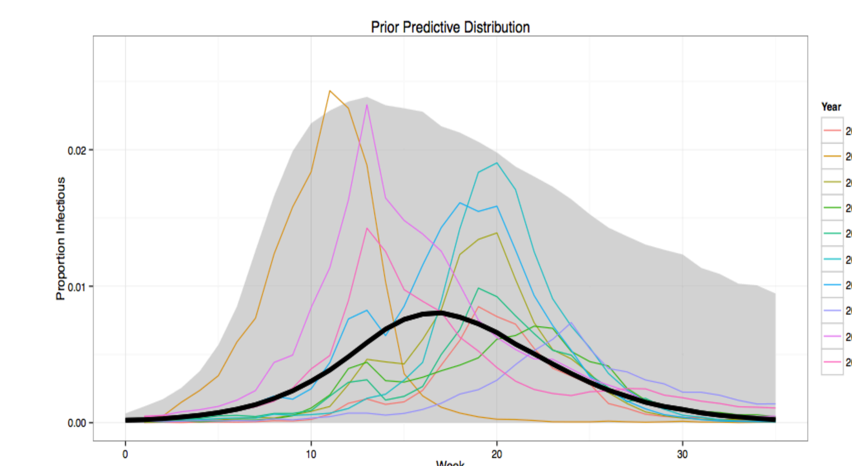
Opening the Model "Black Box"

State-space methods: Epidemic SIR model[1]

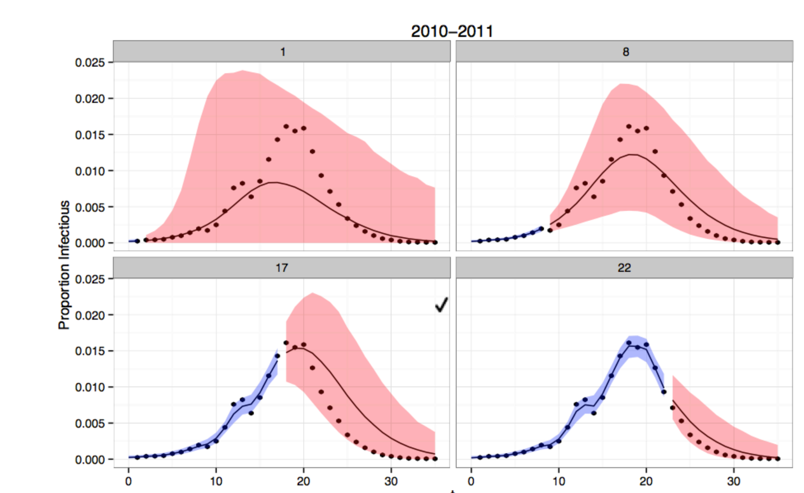
- Simple enough model to identify given flu-season data
- Too simple of a model to fit observations

$[\theta_t \mid \theta_{t-1}, \phi] \stackrel{d}{=} Dirichlet(\kappa\alpha_S(t), \kappa\alpha_I(t), \kappa\alpha_R(t))$





With careful specification of a possible distributions of parameters, the model can predict with good uncertainty partway through the season (as well as probability of peak infected being reached)



Dynamic Discrepancy: Adjusting an sequential physics model as it integrates

- Formulate a discrepancy adjustment appropriate to the modeling domain
- Calibrate a discrepancy "update" step iteration, with discrepancy as:

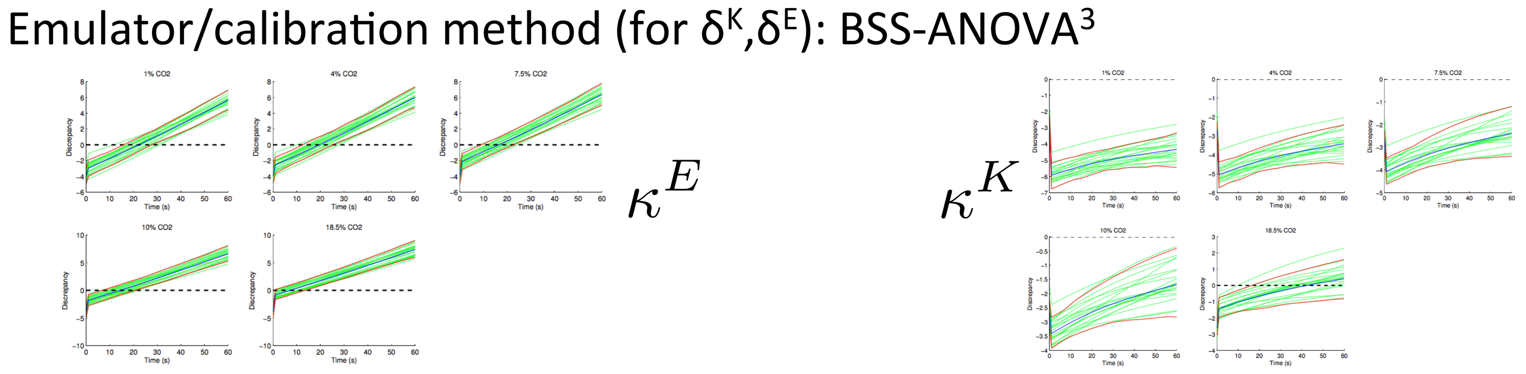
$$\frac{\partial x}{\partial t} = f_s(x, \boldsymbol{\theta}, \boldsymbol{\zeta}(t)) + \delta(x, \boldsymbol{\zeta}(t); \boldsymbol{\beta})$$

Application: Carbon Capture "bubbling fluidized bed" adsorber.

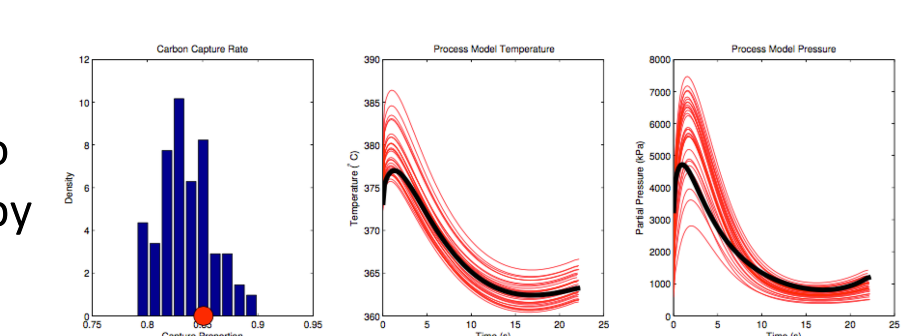
Governing equation: $\frac{\partial x}{\partial t} = \kappa^E \left[(1-2x)^2 p - x^2/\kappa^K\right]$

Enhanced as: $\frac{\partial x}{\partial t} = \kappa^E \exp[\delta^K(x,p,T)]((1-2x)^2 p - x^2/[\kappa^K \exp[\delta^E(p,T)]])$

Emulator/calibration method (for δ^K,δ^E): BSS-ANOVA[3]



Performance on simulation-matching test shows ability to calibrate to target (red dot), by matching response profiles



[1]David Osthus, Forecasting Seasonal Influenza with a State-Space Susceptible-Infectious-Recovered Model
[2]Bhat, Mebane, Storlie, Mahapatra, Upscaling Uncertainty with Dynamic Discrepancy for a Multi-scale Carbon Capture System, in review, JASA Applications.
[3]Storlie, Lane, Ryan, Gattiker, Higdon, Calibration of Computational Models With Categorical Parameters and Correlated Outputs via Bayesian Smoothing Spline ANOVA, JASA vol.110, iss.509,2015.