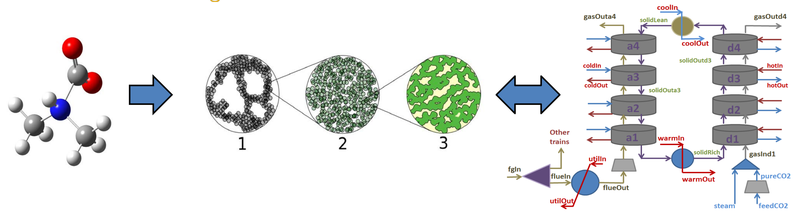
Uncertainty Quantification in Multi-Scale Models of Chemical Processes

David S. Mebane,[1,2] K. Sham Bhat[3] and Curtis B. Storlie[3]

[1]Department of Mechanical and Aerospace Engineering, West Virginia University

[2]National Energy Technology Laboratory

[3]Statistical Sciences Group, Los Alamos National Laboratory



Multi-scale modeling is an inherently statistical problem. Simplifications used in multi-scale approaches introduce uncertainty; a rational goal of a multi-scale paradigm is the quantification and minimization of this uncertainty.

In Bayesian statistics, model parameters are considered to be random variables, and even the functional form of a model can be associated with a probability distribution. A Bayesian approach to model calibration is

$$Y = Z(\theta) + \delta(\xi) + \epsilon$$

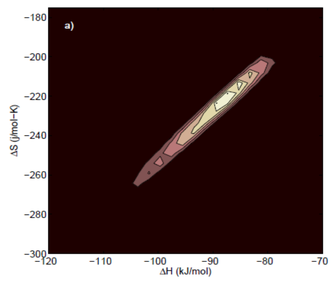
where Y is data, Z is a first-principles (or reduced-order) model, δ is model discrepancy (a stochastic function), and ϵ is observation error. We want to find the set of model functions $Z + \delta$ that are consistent with the data.

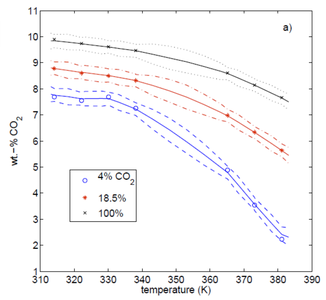
reaction	B3LYP	PBE	MP
CO_2+2MMA→P-COOH:P	-52.72	-76.36	-62.76
CO_2+MMA+DMA→S-COOH:P	-46.86	-70.29	-62.97
CO_2+MMA+DMA→P-COOH:S	-50.21	-72.8	-62.76
CO_2+2DMA→S-COOH:S	-46.86	-70.71	-64.43
CO_2+DETA+EDA→P-COOH:S	-37.66	-69.04	-62.76
CO_2+DETA+EDA→S-COOH:P	-43.93	-68.41	-72.38

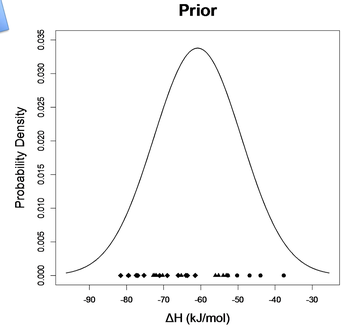
Ab initio calculations can be used to establish prior probability distributions for the physical model parameters θ. These distributions capture uncertainty due to the approximation of the Schrödinger equation as well as approximations to the chemistry.

$$\Omega(\theta, \delta|Y) \propto \mathcal{L}(Y|\theta, \delta)\pi(\theta, \delta)$$

The prior (π) and a likelihood (L) for the data are combined using Bayes' theorem, with the result a posterior distribution (Ω) of models suitable for upscaling.







For an equilibrium process, it's relatively easy to find an appropriate model discrepancy. At left are two marginal distributions of a posterior generated by calibration of a simple equilibrium model of an amine-based CO_2 sorbent to experimental TGA data.

The model-plus discrepancy reproduces TGA data to within observation error, with penalties for interpolation and extrapolation.[1]

Dynamic discrepancies are more difficult, but obviously crucial for the much more common case of chemical reactions not in equilibrium. Here the best approach is to apply the discrepancy function to the rate of reaction:

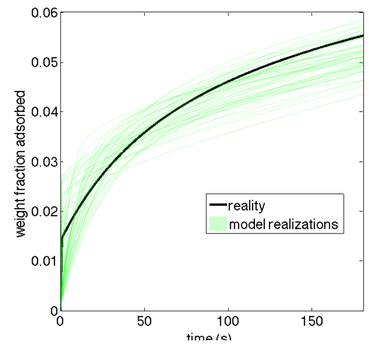
$$\frac{dx}{dt} = f(x, \zeta; \theta) + \delta(x, \zeta; \beta)$$

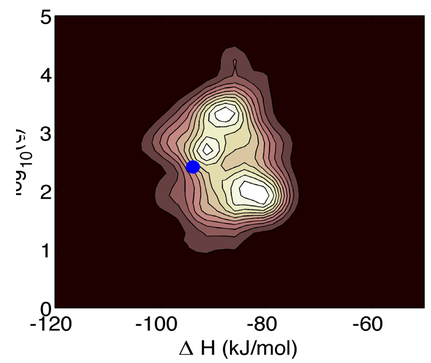
The stochastic differential equation becomes tractable if we use a special form for the discrepancy, called Bayesian smoothing spline analysis of variance. This is a basis expansion of δ – the coefficients are random but the nonlinear basis functions φ are deterministic.[2]

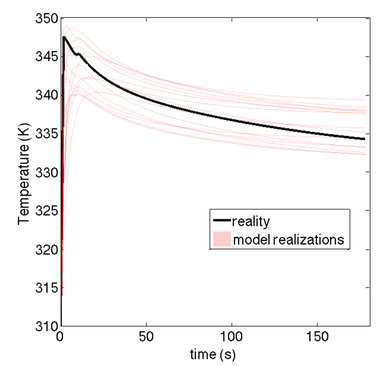
$$\delta(x, \zeta; \beta) = \sum_i \beta_i \varphi_i(x, \zeta)$$

A theorem by Takens[3] makes clear that we can use this method to produce reduced-order models for complex systems of reactions while still capturing the real model behavior.

A hypothetical example appears below. A two-step reaction sequence was modeled by a single-step reaction plus discrepancy.







Simulated TGA data (top left), posterior distribution (above) and "process model" temperature profiles (left) for the hypothetical case. Black lines and blue dots are "reality."

Contact:
David S. Mebane, david.mebane@mail.wvu.edu

References:
[1]D.S. Mebane, K.S. Bhat, J.D. Kress, D.J. Fauth, M.L. Gray, A. Lee and D.C. Miller, Phys. Chem. Chem. Phys. 15 (2013) 4355.
[2]B.J. Reich, C. B. Storlie and H.D. Bondell, Technometrics 51 (2009) 110.
[3]E.R. Deyle and G. Sugihara, PLoS One 6 (2011) e18295.