Non-intrusive UQ Methods and Software for Engineering Applications

SIAM UQ Conference

April 3, 2012

Charles Tong

Lawrence Livermore National Laboratory

LLNL-PRES-XXXXXX

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Outline

- A few applications
- A UQ Process
- Methods for propagation and analysis of uncertainties
- A UQ software: PSUADE
- Usage examples

Carbon Capture Simulation Initiative

Identify promising concepts → Reduce the time for design & troubleshooting → Quantify the technical risk, to enable reaching larger scales, earlier → Stabilize the cost during commercial deployment

National Labs

Berkeley Lab; Pacific Northwest National Laboratory; NETL; Los Alamos National Laboratory; Lawrence Livermore National Laboratory

Academia

Carnegie Mellon; Princeton University; West Virginia University; Boston University; Institute for Clean and Secure Energy, The University of Utah

Industry

B&W; ADA; Fluor; GE; EPRI Electric Power Research Institute; Air Products; Burns & McDonnell; Duke Energy; URS; Alstom; Southern Company; Southern California Edison; Boeing; Chevron; ExxonMobil; WorleyParsons; Eastman

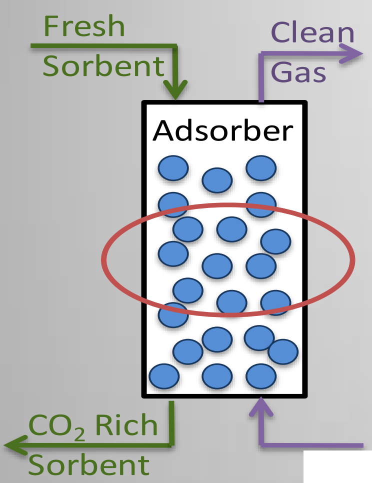
1-31-2012

Essential for accelerating commercial deployment

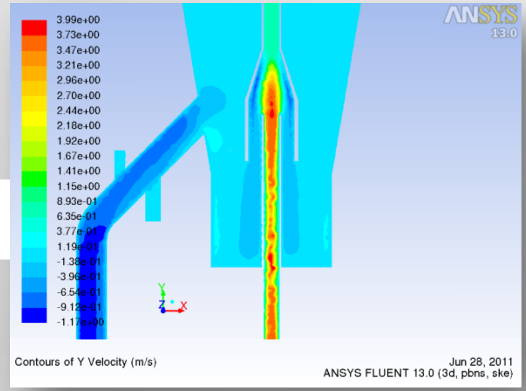
Goal is to develop M&S tools to accelerate the commercialization of Carbon Capture Technologies

Assess different CC technology via simulation and modeling

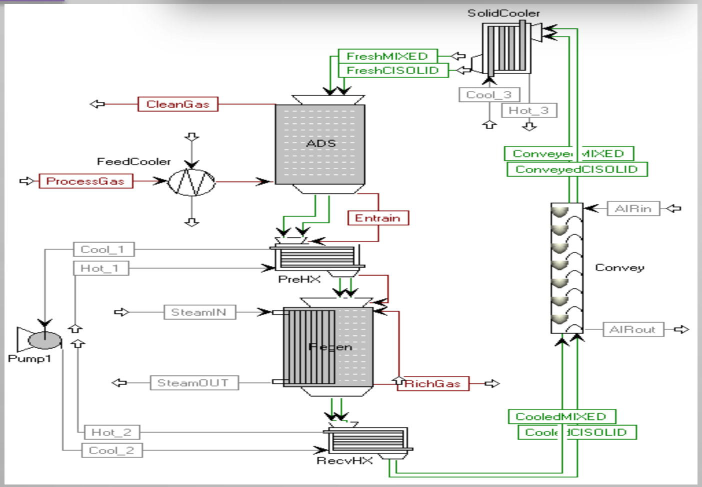
technology



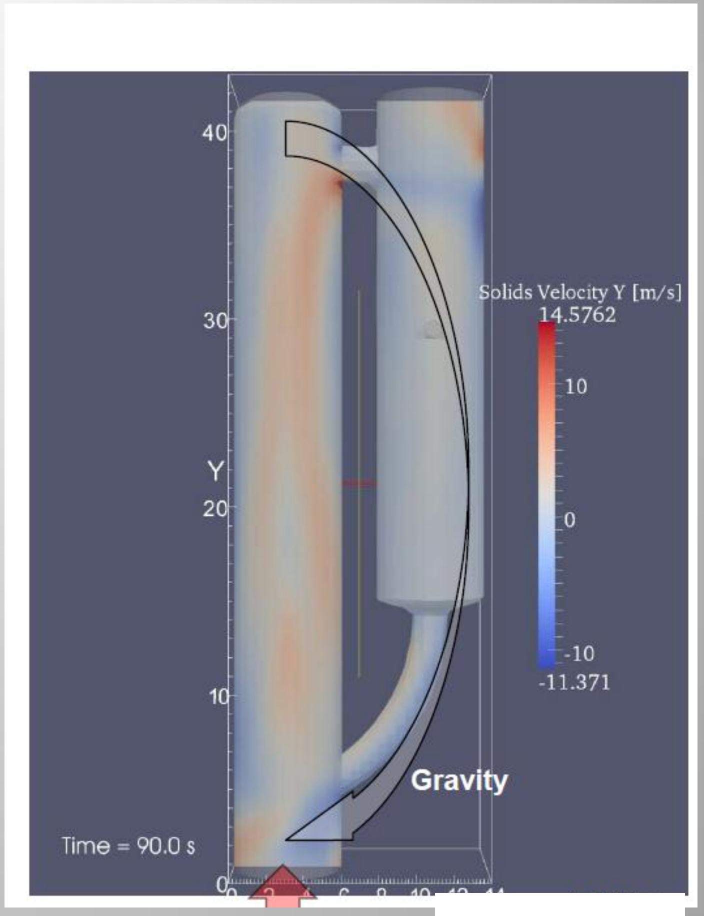
device models



experiment



process models



Prediction with uncertainty on scale-up systems

Source: David Miller's various presentations

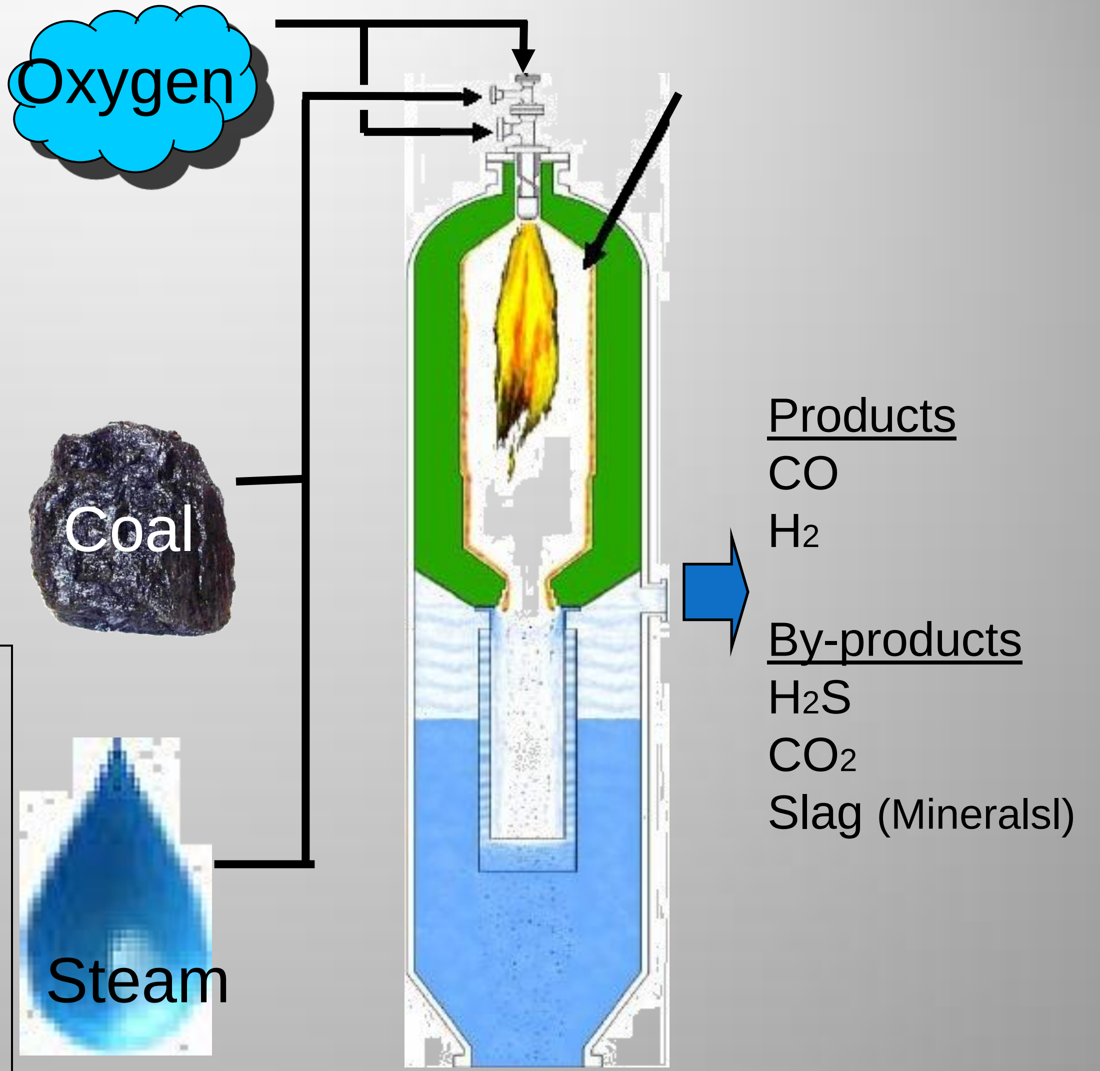
Can we use the world's most abundant and widely distributed fossil fuel source in a different way?

An artist's rendition depicts the goal of the FutureGen initiative, which aims to build the world's first integrated sequestration and hydrogen production research power plant based on coal gasification.

Source: DOE Office of Fossil Energy

As part of Clean Coal Initiative developing advanced coal technologies based on **GASIFICATION** to achieve zero emission of pollutants (CO2) while still remaining economically competitive

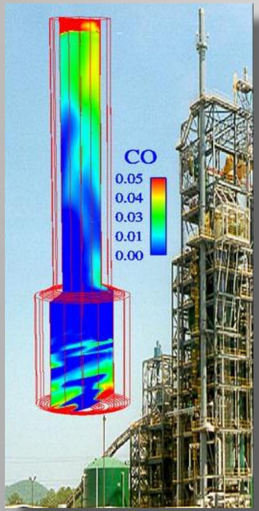
Source: Aytekin Gel's various presentations



Source: Overview of Gasification Technologies, Gary J. Stiegel, NETL (2005)

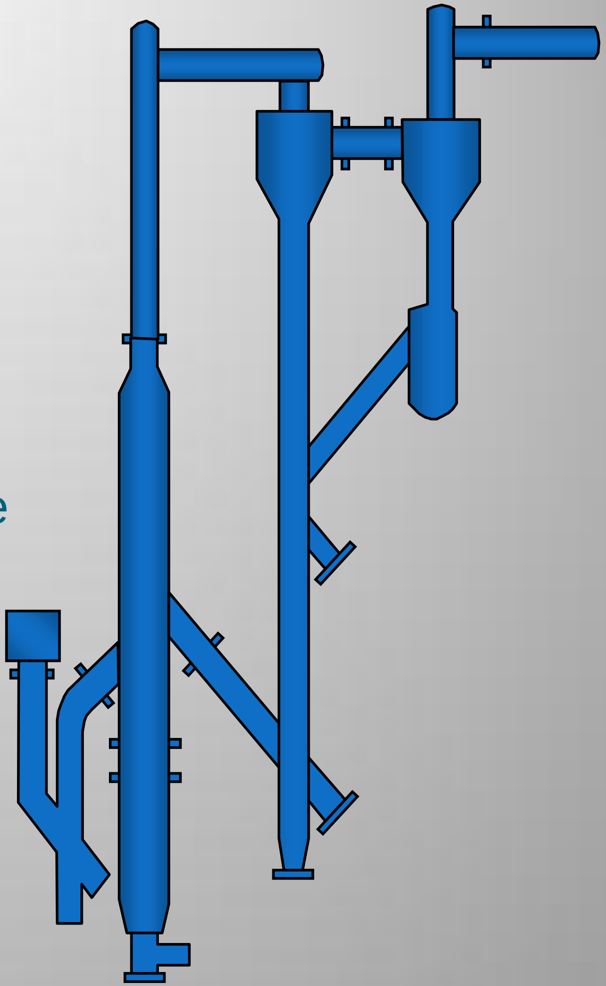
Challenge: How can we design commercial scale gasifiers for optimized operation?

Use validated computer models for answering **scale up** questions



MFIX simulation of pilot scale 13 MW transport gasifier at Wilsonville, AL. Validation of the computer model with prototype system C. Guenther et al (2003)

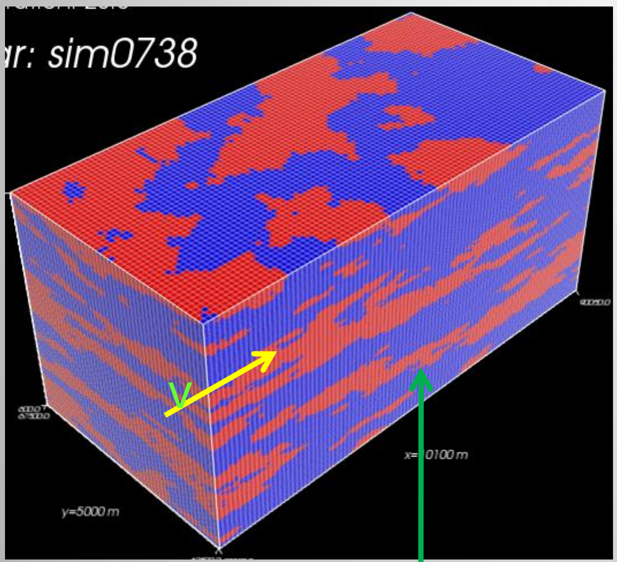
<u>Parametric Study</u>
- Length/Diameter
- Coal feed rate
- Solids circulation rate
- Recycled syngas
- Coal jet penetration



285 MW Commercial Gasifier

Source: Aytekin Gel's various presentations

NRAP: UQ of Geochemical Impact of CO_2 Leakage into Aquifers



Single Source

- 3D Heterogeneous domain of 10000× 10000×240 to represent High Plains aquifer including saturated and unsaturated zones
- Two geologic facies: sand and clay
- Regional groundwater flow is maintained by 0.3% hydraulic gradient
- A single leakage source is at (4500.0, 0.0, 198.41)
- Time steps = 103 for 100 years

Structure Model	NUFT Model	
▪ Sand volume fraction ▪ Correlation length in x ▪ Correlation length in y ▪ Correlation length in z	▪ Sand porosity ▪ Clay porosity ▪ Sand density ▪ Clay density ▪ van Genuchten m in sand ▪ van Genuchten m in clay	▪ van Genuchten α in sand ▪ van Genuchten α in clay ▪ Permeability in sand ▪ Permeability in clay ▪ CO_2 diffusivity ▪ CO_2 leakage flux rate

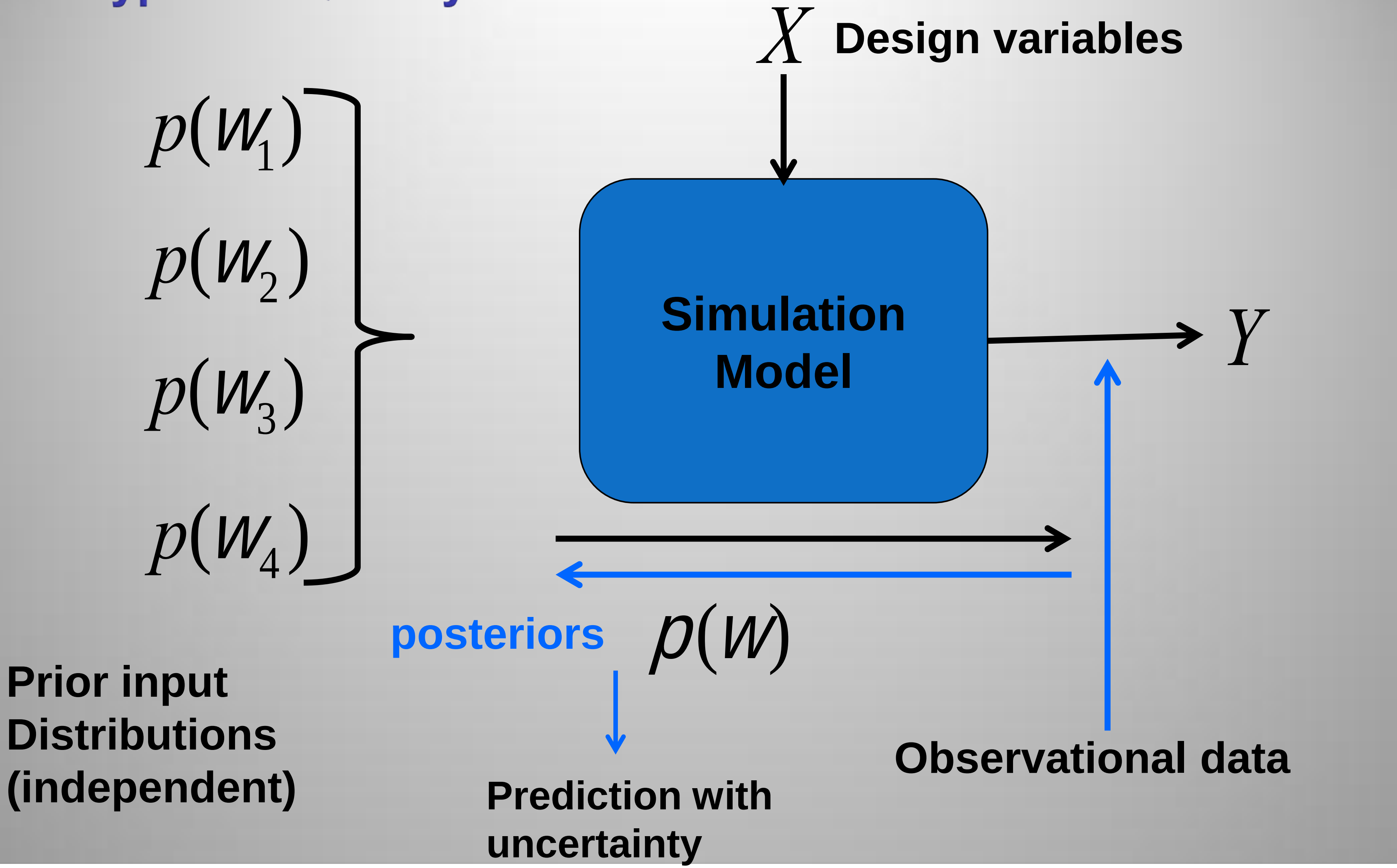
Source: Yunwei Sun’s various presentations

Steps in uncertainty quantification (UQ)

- **Definition (of goals/objectives: what to do?)**
  - simulation models, quantities of interest, assumptions
- **Identification (sources of uncertainties: where?)**
  - model and data uncertainties
- **Characterization (of uncertainty sources, what form?)**
  - parametric distributions (priors) and/or model form
  - Aleatoric/epistemic, continuous/discrete
- **Propagation (of uncertainty through models: how?)**
  - forward and backward
- **Analysis (of impact of uncertainties: so what?)**
  - sensitivity analysis, risk analysis, …

of all relevant uncertainties in simulation models

A typical UQ analysis



Simple uncertainty analysis may turn out to be computationally expensive

mean

$$\mu(Y) = \int_{\Omega} Y(\omega)\, p(\omega)\, d\omega$$

interval

$$\max_{\Omega} Y(\omega) - \min_{\Omega} Y(\omega)$$

variance

$$\sigma^2(Y) = \int_{\Omega} [Y(\omega)]^2\, p(\omega)\, d\omega - \mu^2(Y)$$

+ higher order moments

- Y(ω) may be expensive to evaluate
- Sufficient accuracy may require many integration points
  - Highly nonlinear functions, may be nonsmooth
  - High dimensional uncertain parameter space

➢ Response surface (surrogate, emulator, ROM)

➢ Dimension reduction (parameter screening)

Bayesian inference/Parameter Estimation/Calibration

Continuous Variables

$$\pi(W|D) = p(D|W)p(W) / \int p(D|W)p(W)$$

- Solution method: MCMC-like methods
  - for aleatoric uncertainties → posterior distributions
  - for epistemic uncertainties →feasible subspace
- These methods are computationally very expensive
  - need large sample sizes for burn-in and generating posteriors
- In practice, these are calculated using response surfaces
- More complex cases: mixed continuous/discrete

Global Sensitivity Analysis

First order sensitivity analysis

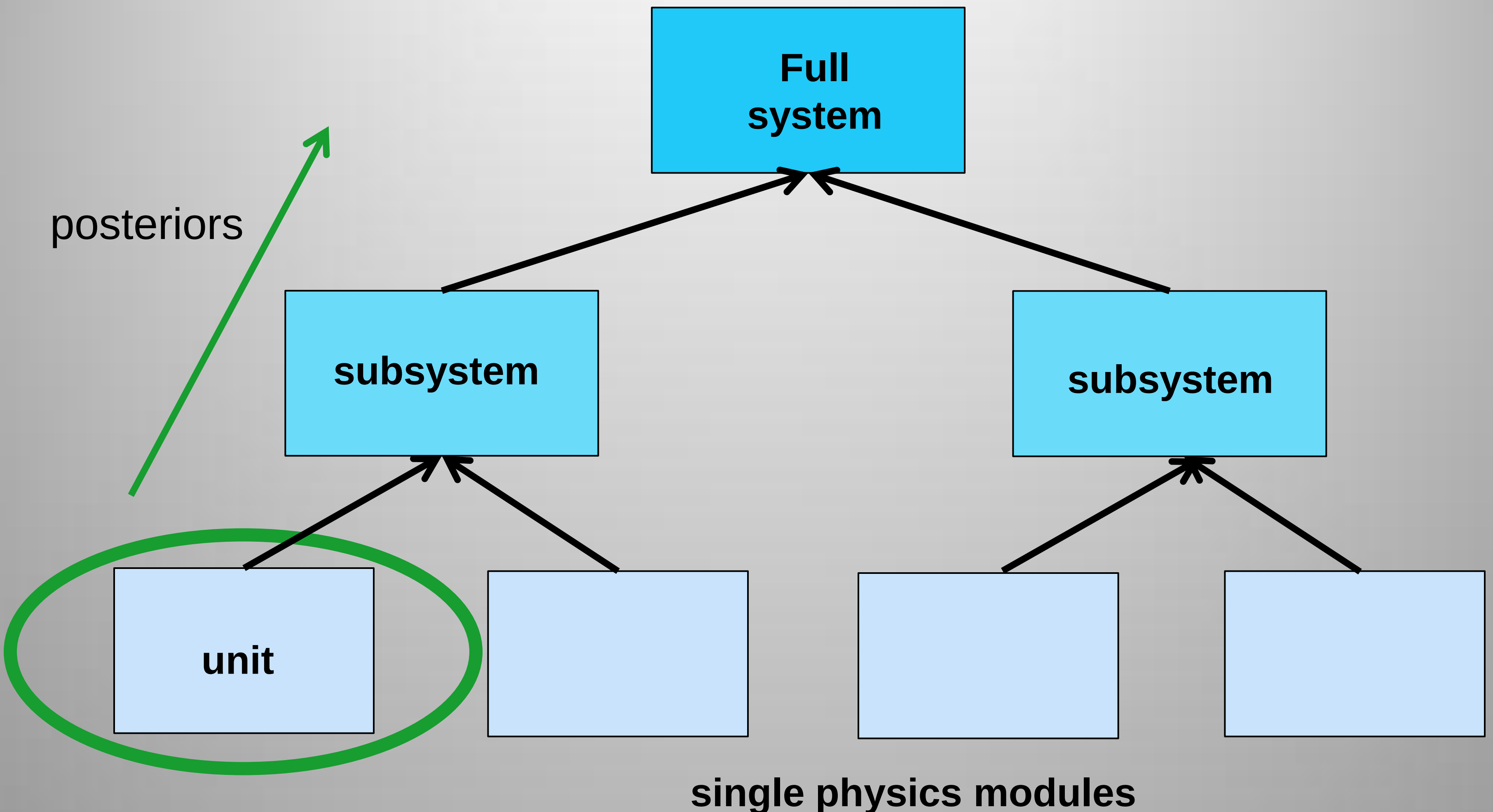
$$\eta_i^2 = \int \left[\int Y(\omega_{\sim i} \,|\, \omega_i) p(\omega_{\sim i} \,|\, \omega_i) d\omega_{\sim i} - \mu(Y) \right]^2 p(\omega_i) d\omega_i \,/\, \sigma^2$$

Total order sensitivity indices

$$TS_i = \int \left[\int \left(Y(\omega_i \,|\, \omega_{\sim i}) p(\omega_i \,|\, \omega_{\sim i}) d\omega_i \right)^2 - \mu^2 (Y(\omega_{\sim i})) \right] p(\omega_{\sim i}) d\omega_{\sim i} \,/\, \sigma^2$$

- Variance-based SA is computationally expensive
  - need many integration points for accuracy
- In practice, these are calculated using response surfaces
  - the use of parametric forms reduce computational cost further

In complex applications, these typical UQ analysis will be performed hierarchically



Major Themes in UQ Analysis Methods

- Dimension reduction methods
  - High-dimensional physics parameter space (subset selection)
  - High-dimensional model output space (principal components)
- Response surface (surrogate, emulators) methods
  - Surrogate model selection (splines, GP, regression, ..)
  - Exploit smoothness in the approximate mapping
  - Response surface validation
- Basic uncertainty analysis
  - Statistical moments
  - Correlation analysis, ANOVA
- Calibration/Data integration
  - Bayesian-like inferences (filtering, gates, sets)
- Quantitative sensitivity analysis methods
  - Variance decomposition (1^{st}, 2^{nd}, group, and total order)
- Parameter study/design exploration
- Reliability analysis

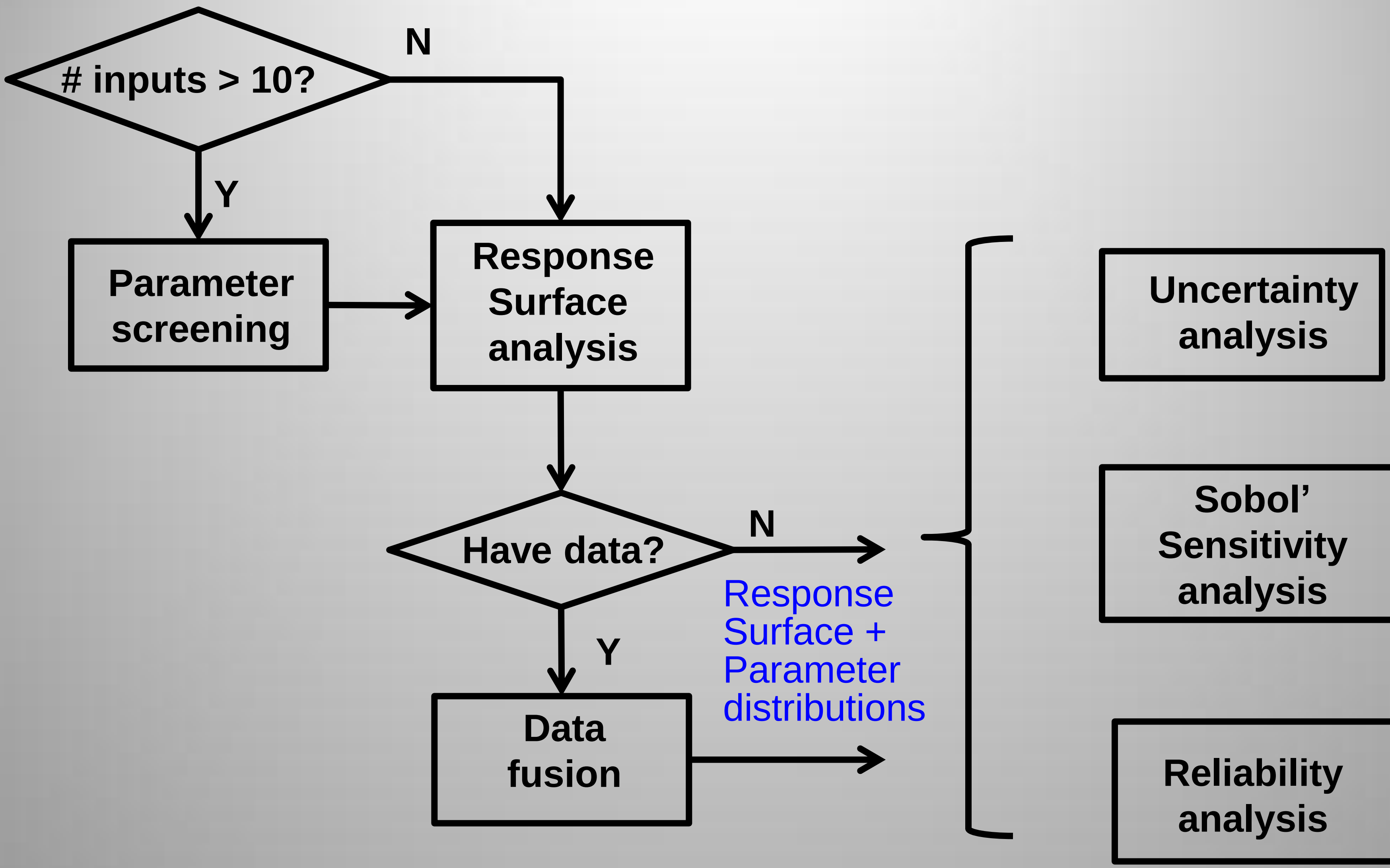
How to make these UQ methods available to users?

Confidence interval analysis
Orthogonal arrays
Cross validation
Correlation analysis
Latin hypercube
Space-filling designs
Gaussian process
Graphical analysis
Nonparametric testing
Kolmogorov Smirnov test
Screening designs
Derivative-based methods
Constrained optimization
Delta test
Principal component analysis
Support vector machines
Probability distributions
Factorial design
Sobol' decomposition
Second order sensitivity indices
Manifold mapping optimization
Multivariate splines
Bayesian Sensitivity analysis
Total sensitivity indices
Main effects
Adaptive response surfaces
Markov chain Monte Carlo
Reliability analysis
Fourier amplitude sampling test
Sparse grids
Regression analysis
Artificial neural networks
Analysis of variance

PSUADE is a software to facilitate UQ analysis

A **P**roblem **S**olving Environment for **U**ncertainty **A**nalysis and **D**esign **E**xploration

A Simple UQ Workflow

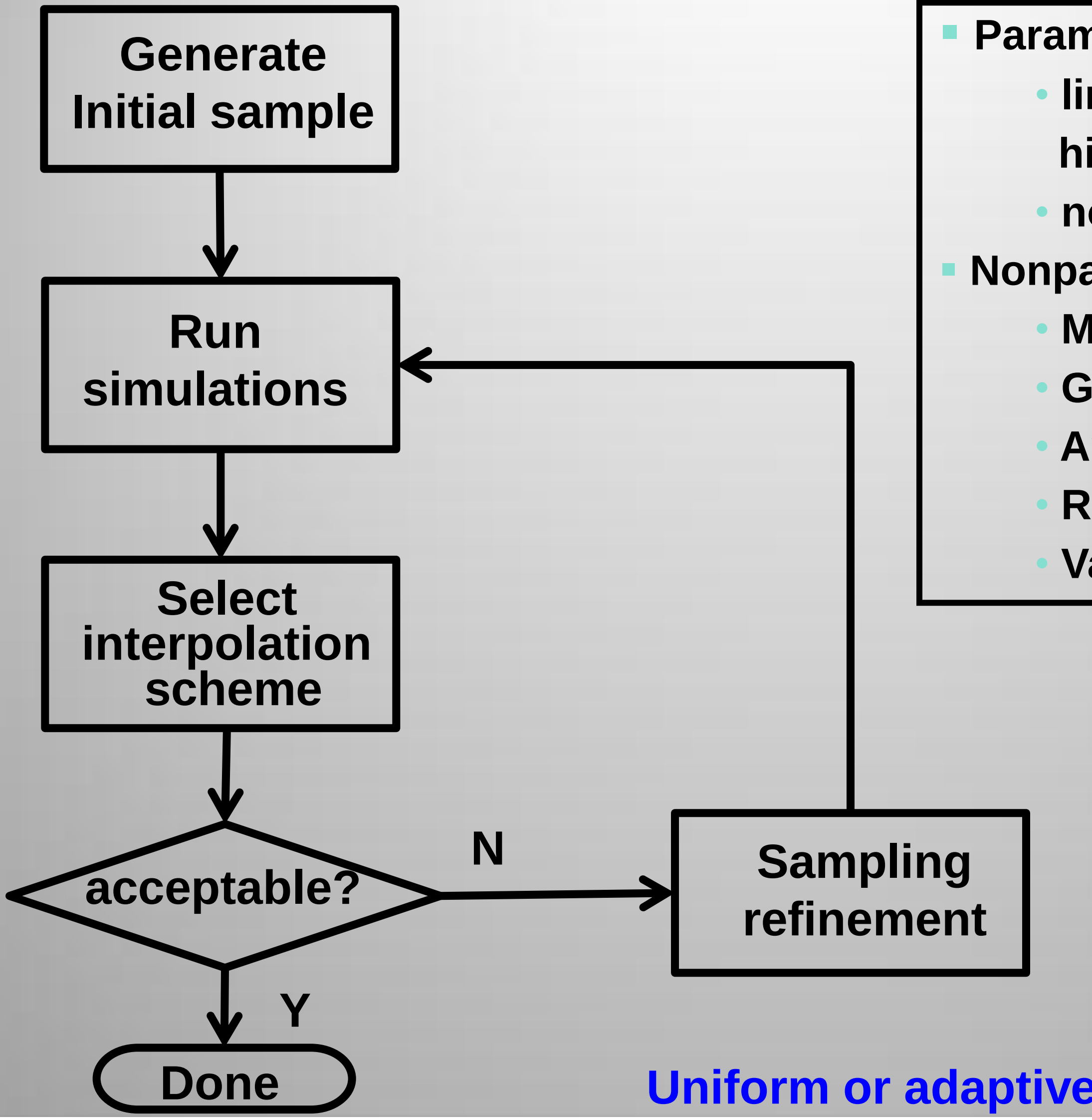


A Suite of Dimension Reduction Methods

- **Parametric methods (e.g. linearity/monotonicity assumptions)**
  - derivative-based methods (local)
  - Standardized regression coefficient (SRC)
  - Spearman rank correlation coefficients (SRRC)
  - Plackett-Burman
  - Fractional factorial
  - Gradient-based Methods (global)
  - Morris method and its variants (e.g. modified Morris)
- **Approximate emulator methods (global)**
  - e.g. splines, Gaussian process, etc.
- **Methods based on nearest neighbors (global)**
  - Delta test, …
- **Tree-based methods (global)**
  - Classification and regression tree (CART) + bootstrap

nonparametric methods are more suitable for multiphysics models

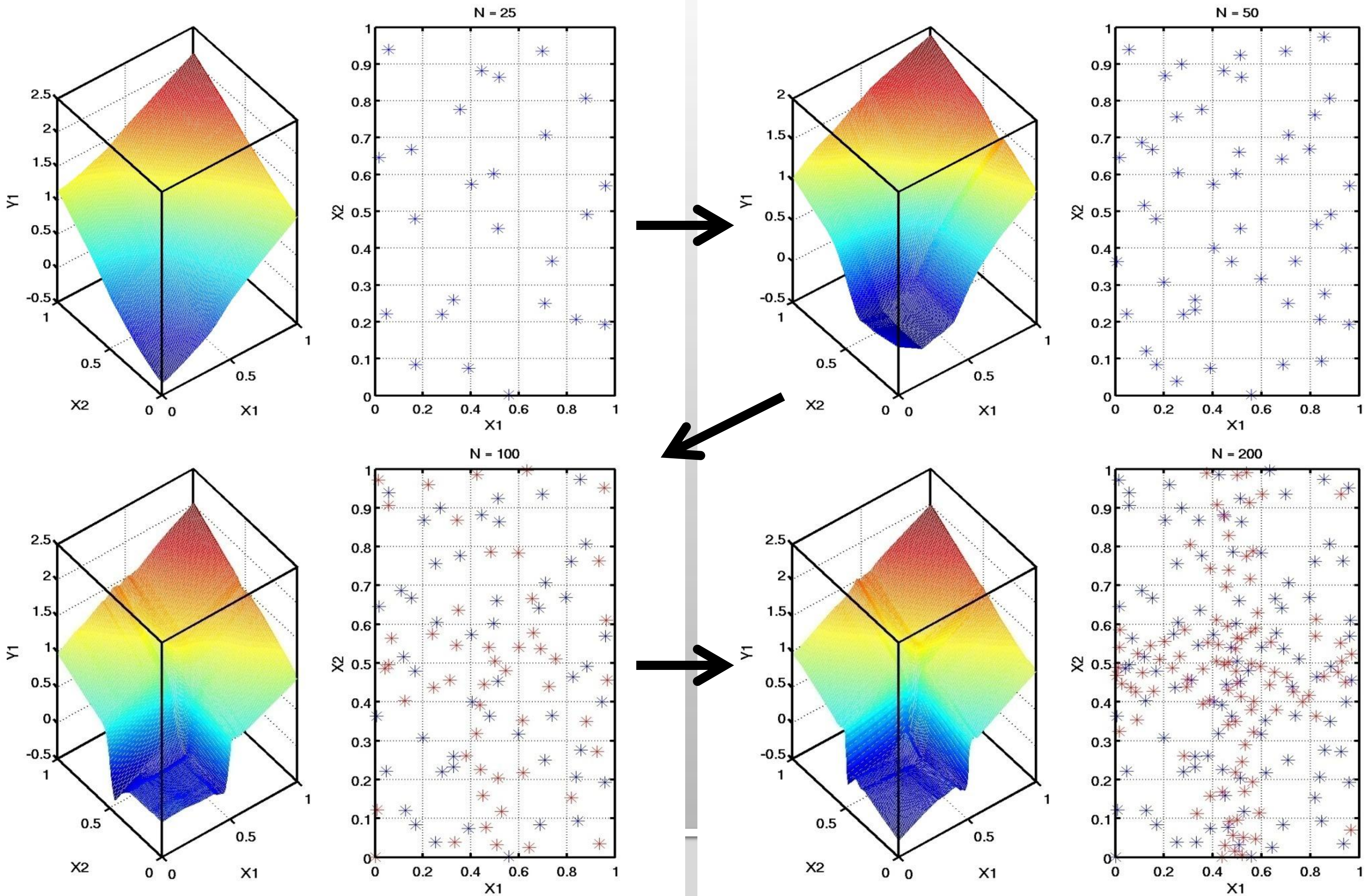
A Suite of Response Surface Methods + Validation + Visualization



- **Parametric:**
  - **linear regression, quadratic, higher order, collocation**
  - **nonlinear regression functions**
- **Nonparametric:**
  - **Multivariate adaptive splines**
  - **Gaussian process**
  - **Artificial neural network**
  - **Radial basis functions**
  - **Variants: with bootstrapping**

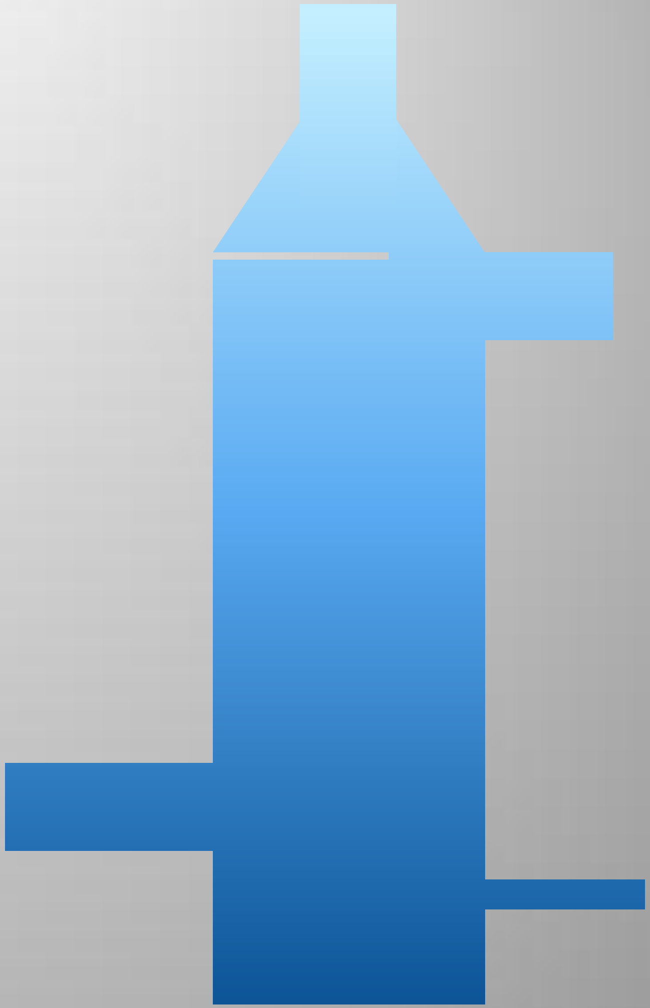
Response surface Validation:
- R-squared
- Resubstitution test
- Test set
- k-fold validation
- visualization

Adaptive Response Surface Analysis



Example: a carbon capture absorber system

Objective: quantify uncertainty in % of CO2 capture



Rate-based reactions

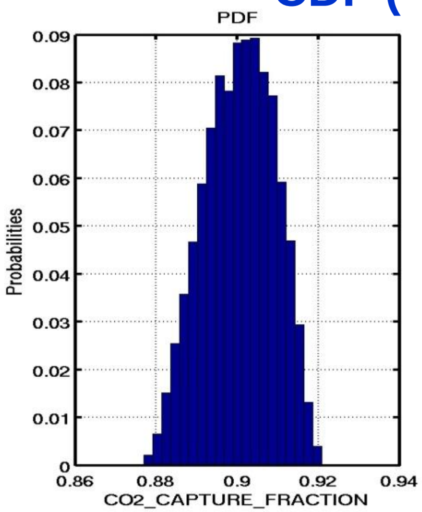
- 7 eqns: 3 equilibrium/4 kinetic
- Equilibrium eqns: ΔT (DT1, DT2)
- Kinetic eqns: rate constant k_n (A4-A7)

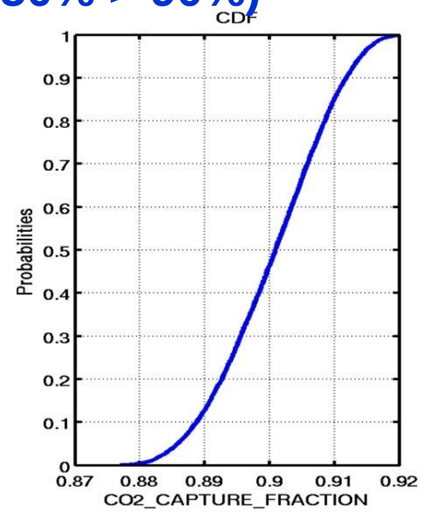
Mass transfer

- Use the Hanley model
- UQ study restricted to tuning the interfacial area factor that affects the mass transfer rates (AE)

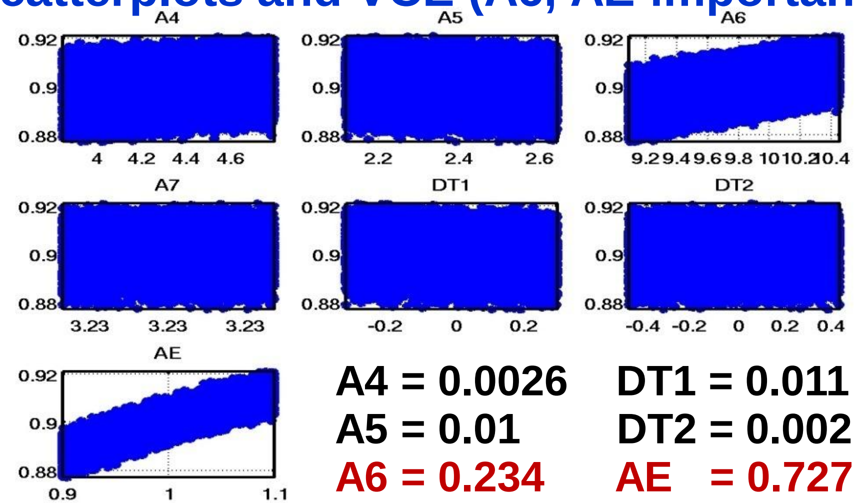
Example: a carbon capture absorber system

CDF (~50% > 90%)



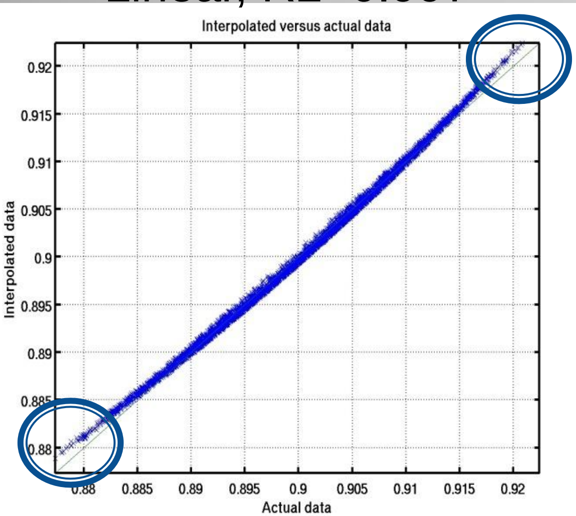


Scatterplots and VCE (A6, AE important)

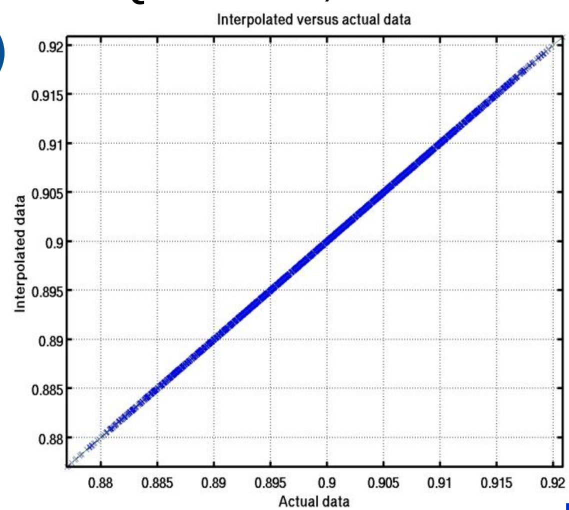


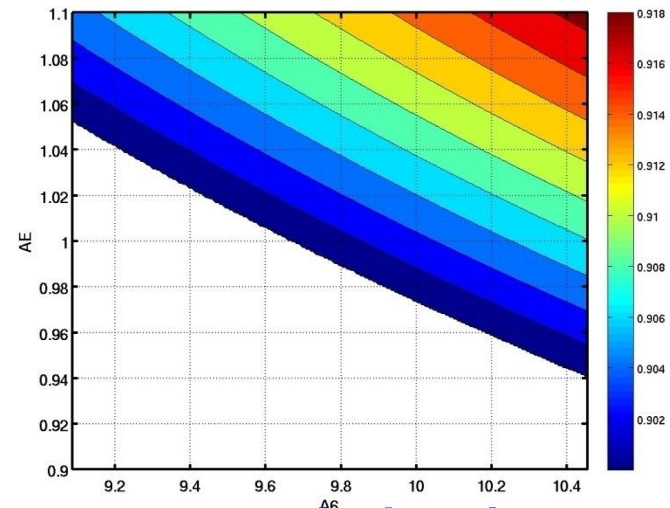
A4 = 0.0026 **DT1 = 0.011**
A5 = 0.01 **DT2 = 0.002**
A6 = 0.234 **AE = 0.727**
A7 = 0.002 (VCE)

Linear, $R2=0.997$



Quadratic, $R2\sim1$

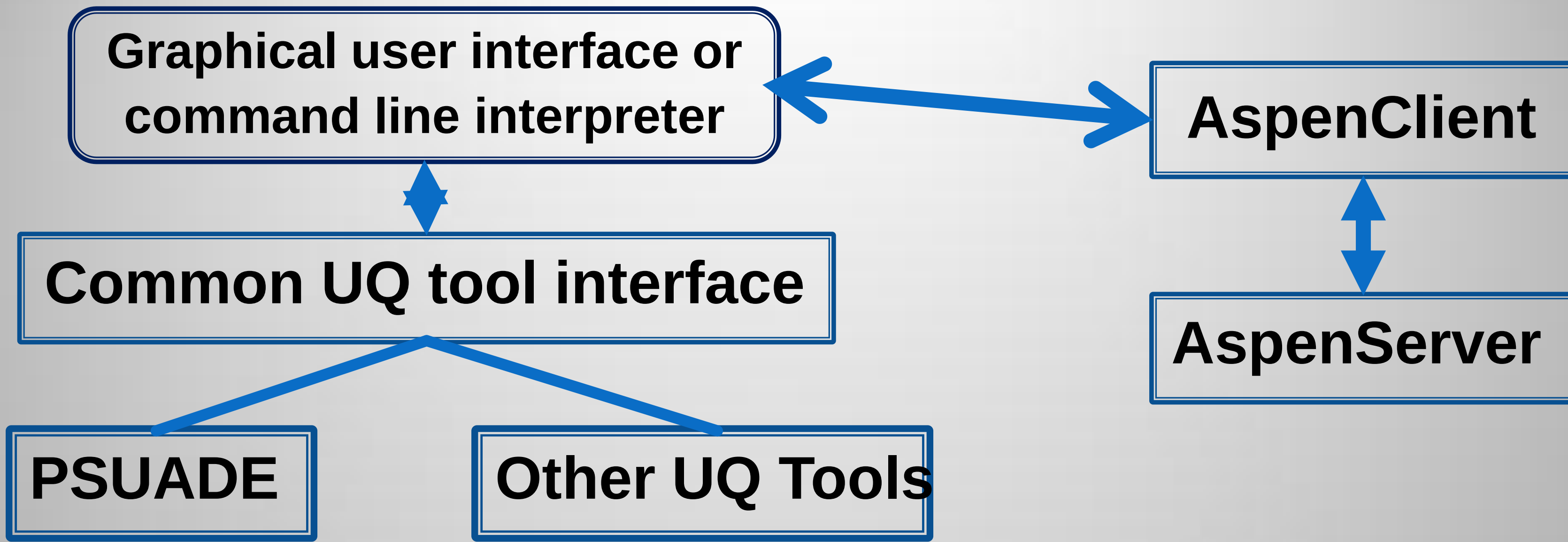


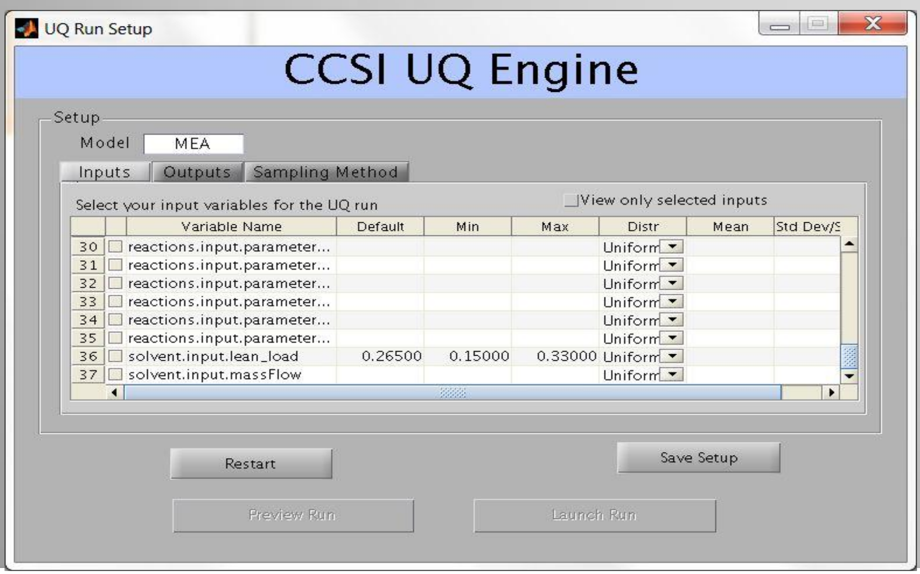


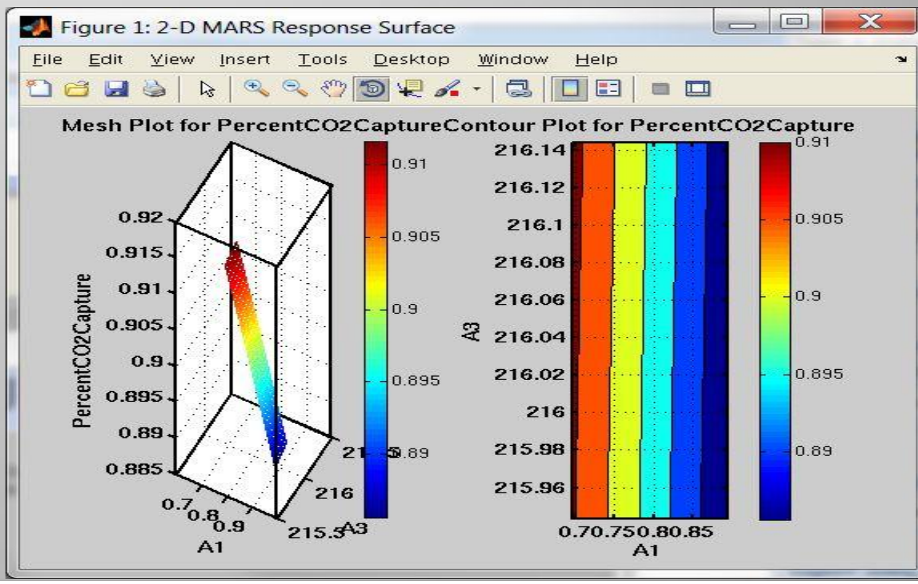
RSA: A6 vs AE (colored: >=90%)

RSA: training and test sets (~5000 each)

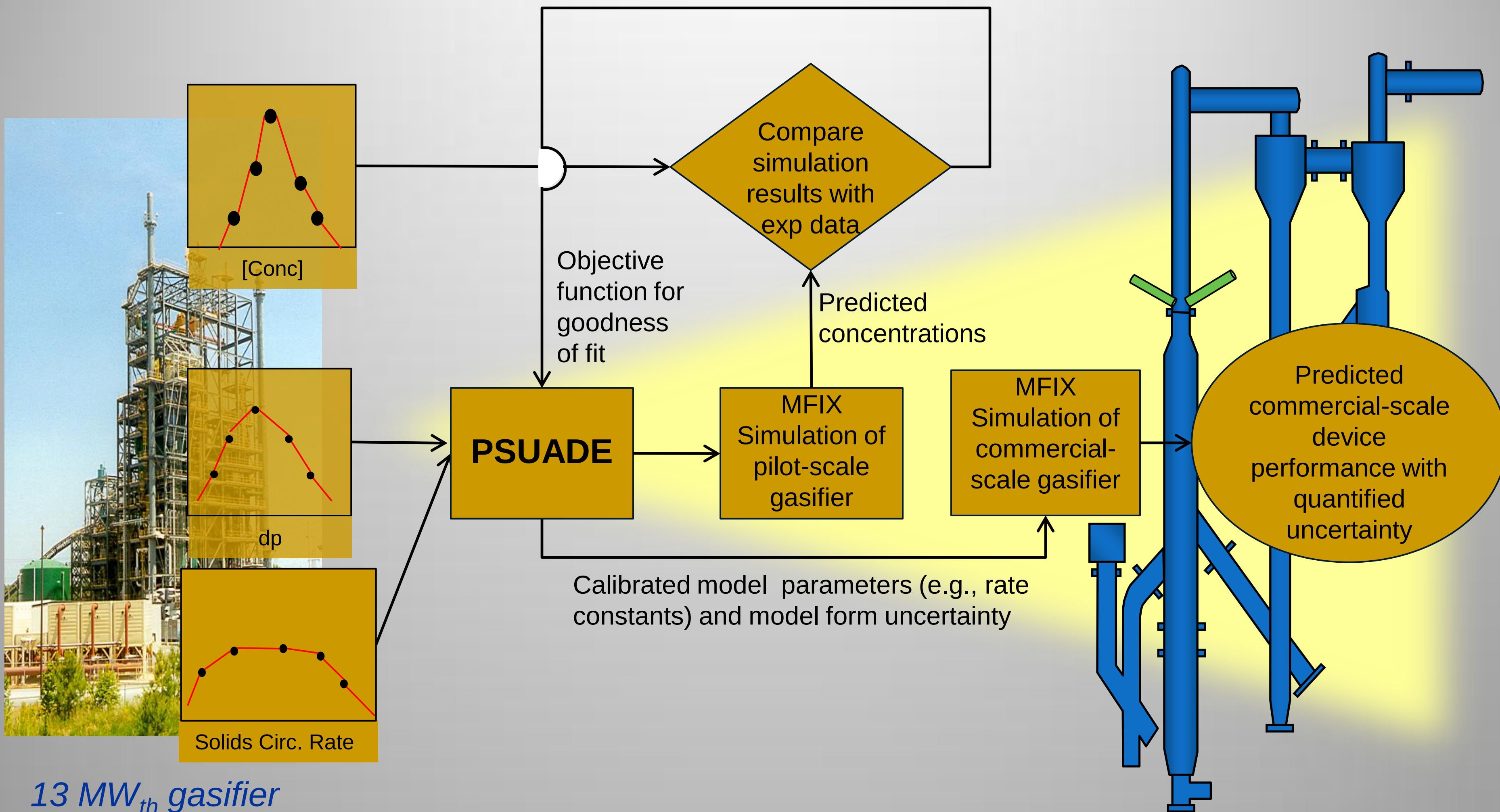
Example: CCSI UQ System Architecture







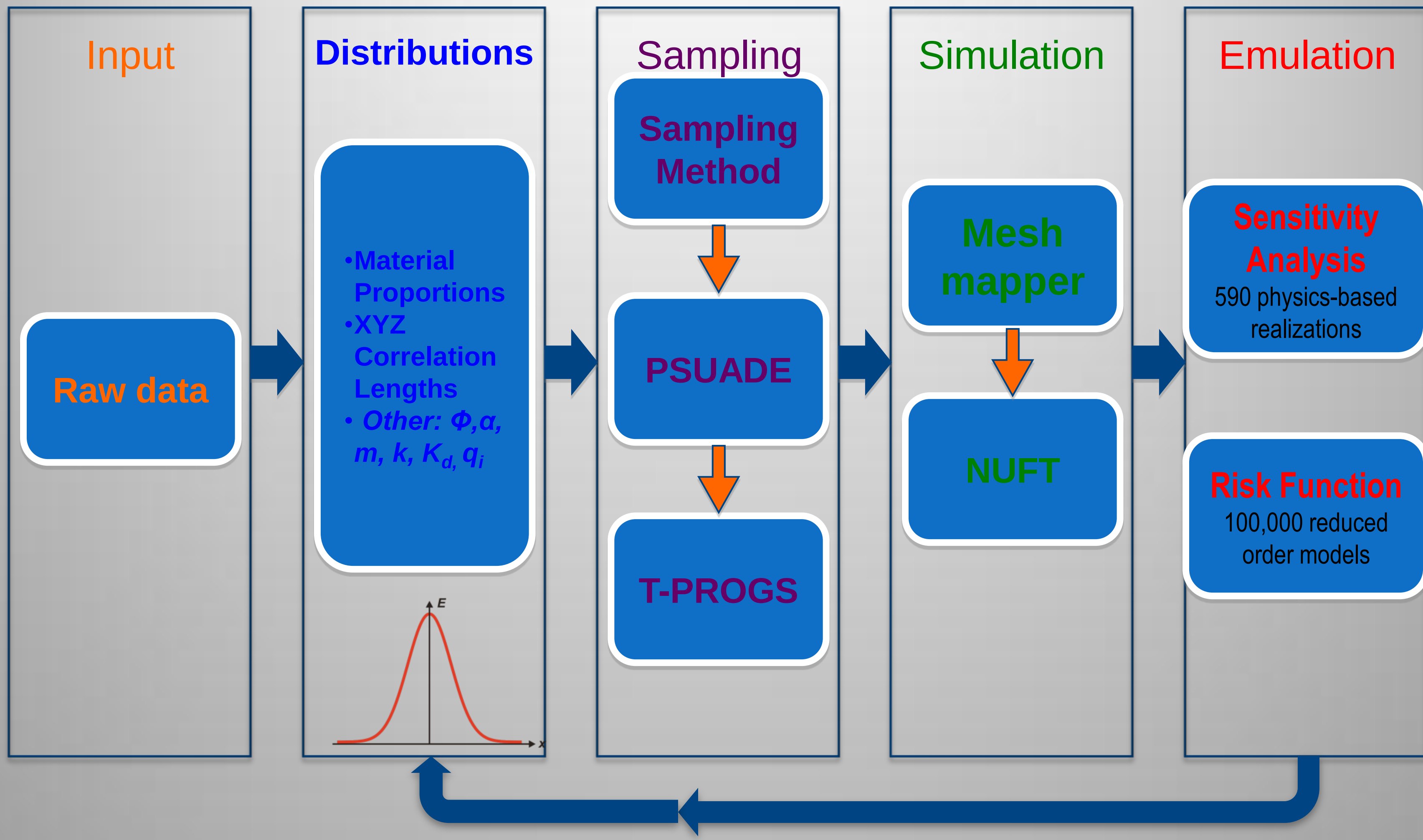
Gasification simulations for Scale-up – with VVUQ



Source:,Sino-US Chemical Engineering Conference, Madhava Syamlal et al., NETL (2011)

285 MW$_e$ Gasifier

Example: UQ Workflow for NRAP



Source: Yunwei Sun's various presentations

PSUADE download site:

https://computation.llnl.gov/casc/uncertainty_quantification

THE END