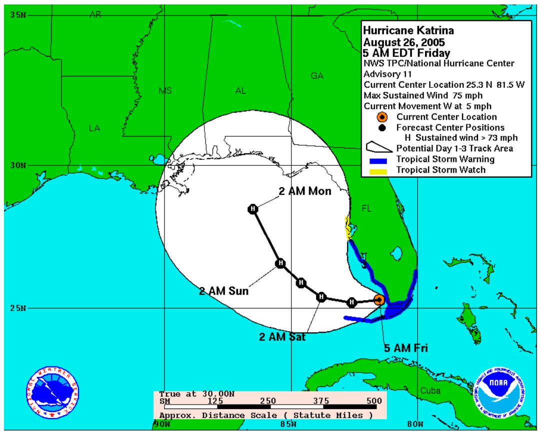
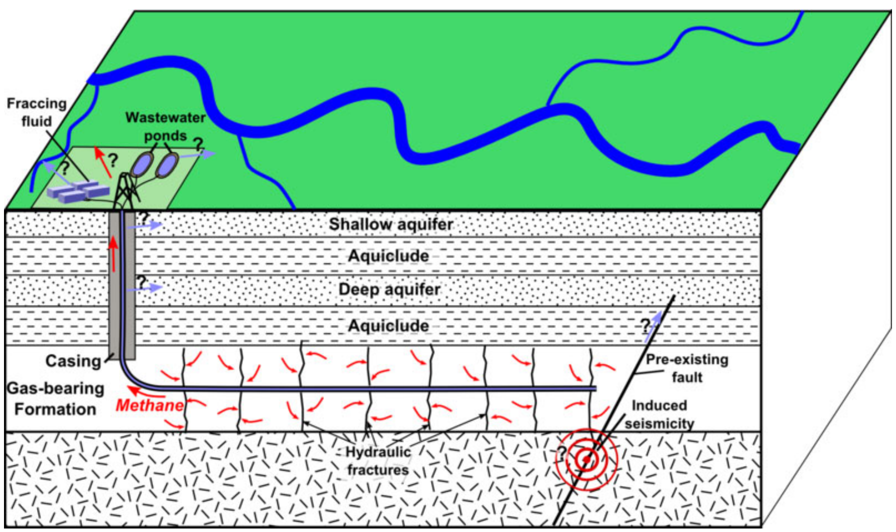
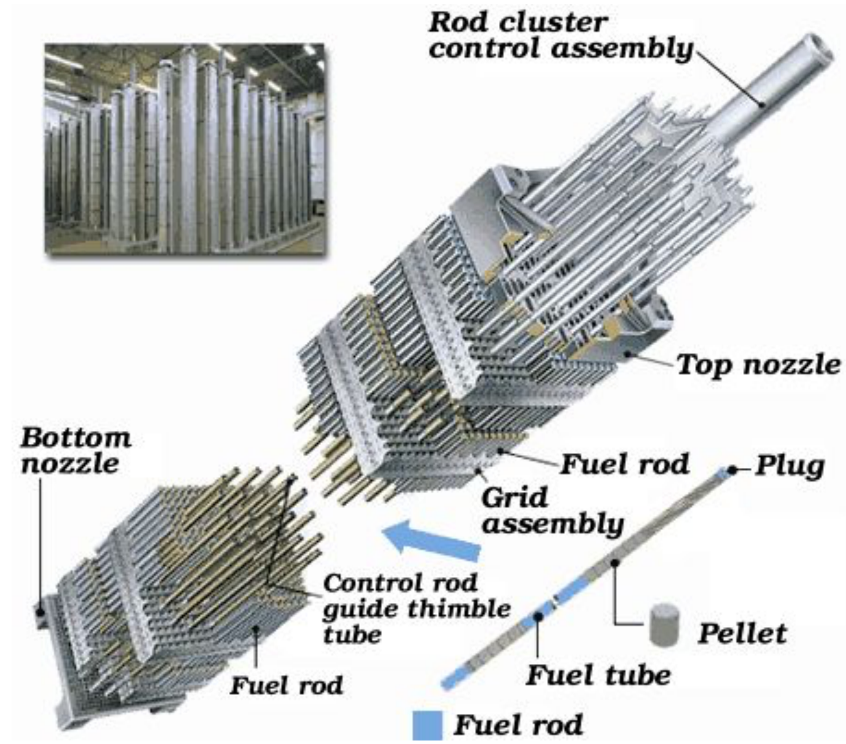
Lecture 1: Motivation and Prototypical Examples

"Essentially all models are wrong, but some are useful,"

George E.P. Box, Industrial Statistician



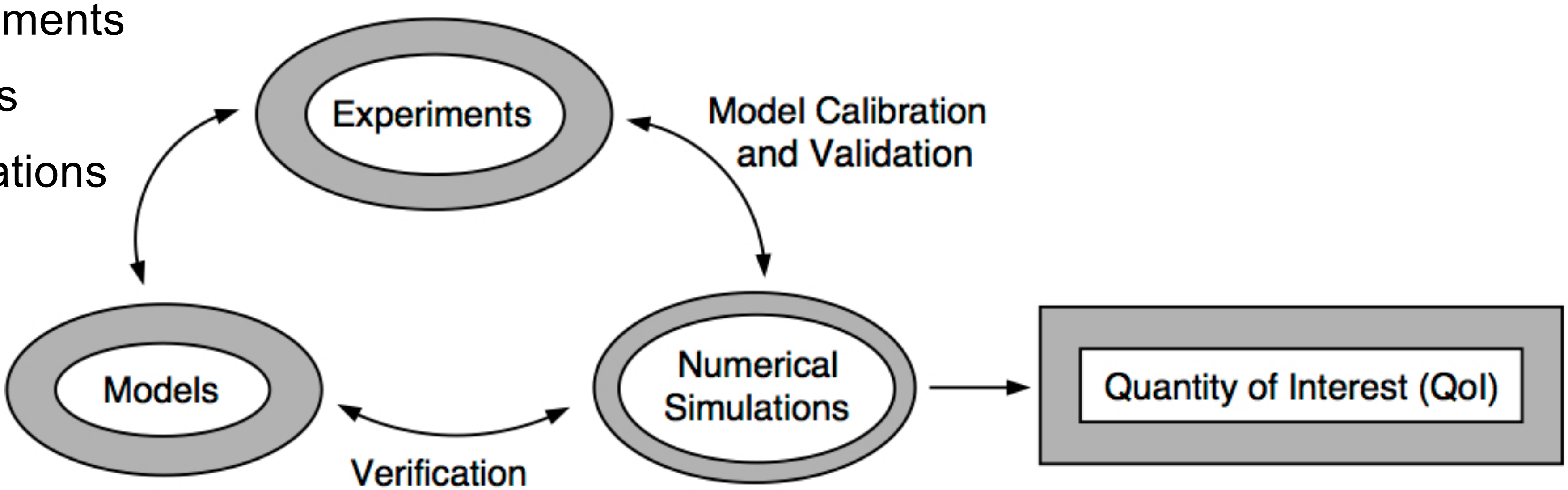




Predictive Science

Components: All involve uncertainty

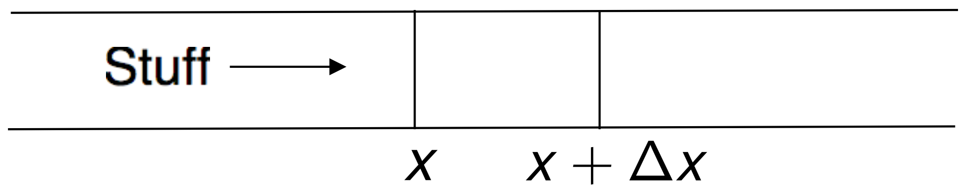
- Experiments
- Models
- Simulations



- *Essentially, all models are wrong, but some are useful,* George E.P. Box, Industrial Statistician.
- *Computational results are believed by no one, except the person who wrote the code*, source anonymous, quoted by Max Gunzburger, Florida State University.
- *Experimental results are believed by everyone, except for the person who ran the experiment,* source anonymous, quoted by Max Gunzburger, Florida State University.
- *I have always done uncertainty quantification. The difference now is that it is capitalized.* Bill Browning, Applied Mathematics Incorporated.

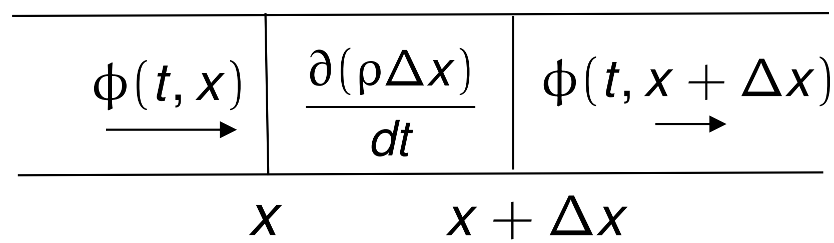
Modeling Strategy

General Strategy: Conservation of stuff



$$\boxed{\frac{dStuff}{dt} = \text{Stuff in - Stuff out + Stuff created - Stuff destroyed}}$$

Continuity Equation:



$$\frac{\partial(\rho \Delta x)}{\partial t} = \phi(t, x) - \phi(t, x + \Delta x)$$

$$\Rightarrow \lim_{\Delta x \to 0} \frac{\partial \rho}{\partial t} = \lim_{\Delta x \to 0} \frac{\phi(t, x) - \phi(t, x + \Delta x)}{\Delta x}$$

$$\boxed{\Rightarrow \frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = 0}$$

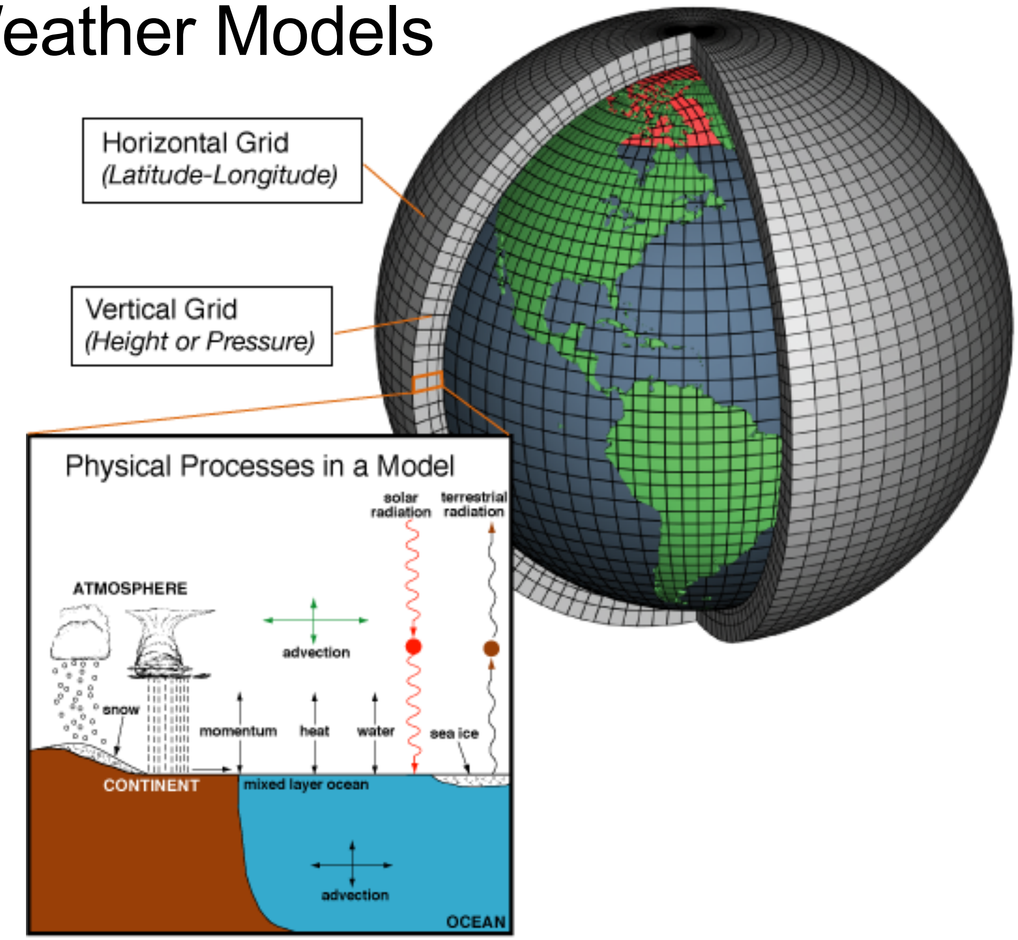
Density: $\rho(t, x)$ - Stuff per unit length or volume

Rate of Flow: $\phi(t, x)$ - Stuff per second

More Generally:

$$\boxed{\Rightarrow \frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = \text{Sources - Sinks}}$$

Example 1: Weather Models

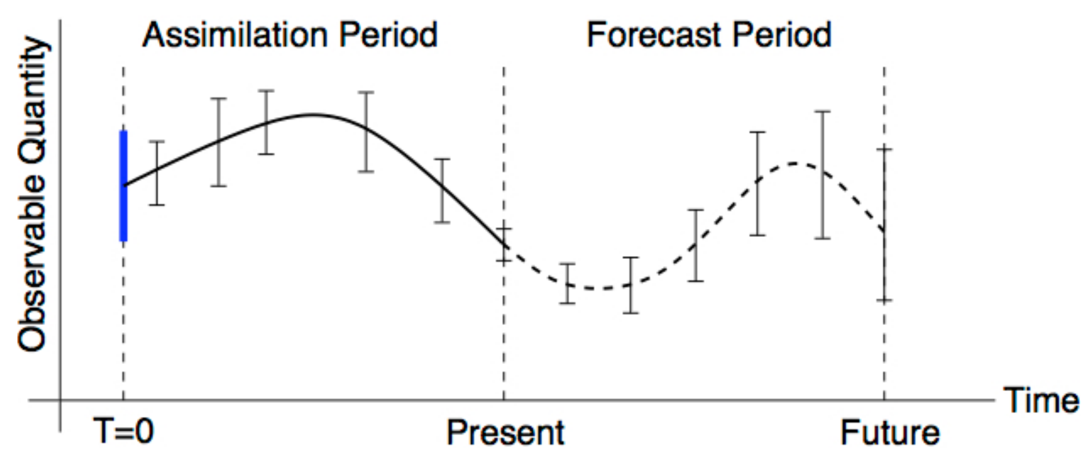


Challenges:

- Coupling between temperature, pressure gradients, precipitation, aerosol, etc.;
- Models and inputs contain uncertainties;
- Numerical grids necessarily larger than many phenomena; e.g., clouds
- Sensors positions may be uncertain; e.g., weather balloons, ocean buoys.

Goal:

- Assimilate data to quantify uncertain initial conditions and parameters;
- Make predictions with quantified uncertainties.



Equations of Atmospheric Physics

Conservation Relations:

$$\boxed{\frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = \text{Sources - Sinks}}$$

Mass $$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) = 0$$

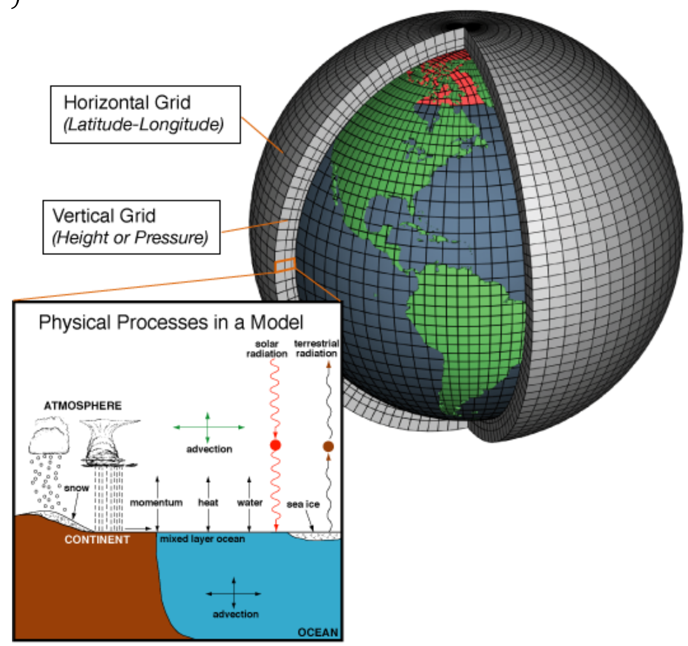
Momentum $$\frac{\partial v}{\partial t} = -v \cdot \nabla v - \frac{1}{\rho}\nabla p - g\hat{k} - 2\Omega \times v$$

Energy $$\rho c_V \frac{\partial T}{\partial t} + p\nabla \cdot v = -\nabla \cdot F + \nabla \cdot (k\nabla T) + \rho \dot{q}(T, p, \rho)$$

$$p = \rho R T$$

Water $$\frac{\partial m_j}{\partial t} = -v \cdot \nabla m_j + S_{m_j}(T, m_j, \chi_j, \rho)\ ,\ j = 1, 2, 3,$$

Aerosol $$\frac{\partial \chi_j}{\partial t} = -v \cdot \nabla \chi_j + S_{\chi_j}(T, \chi_j, \rho)\ ,\ j = 1, \cdots, J,$$



Constitutive Closure Relations: e.g.,

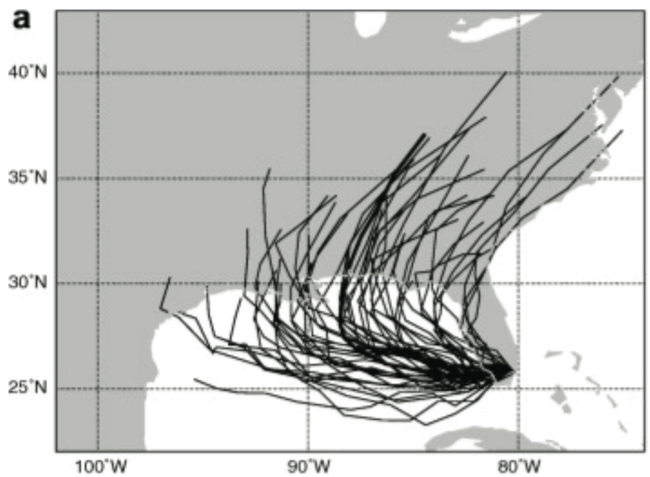
$$S_{m_2} = S_1 + S_2 + S_3 - S_4$$

where

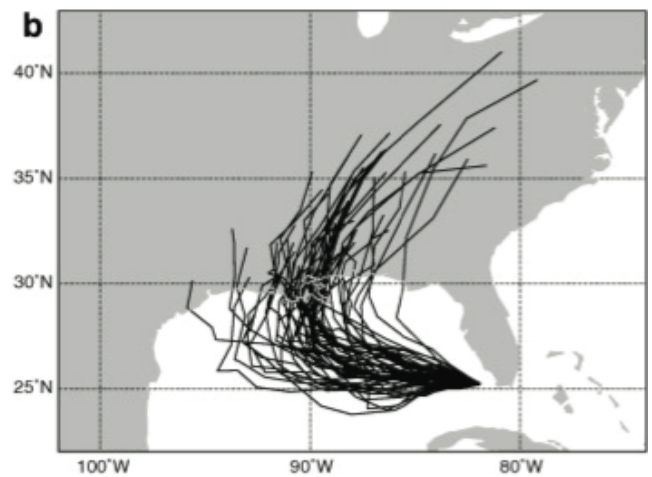
$$S_1 = \bar{\rho}(m_2 - m_2^*)^2 \left[\underline{1.2 \times 10^{-4}} + \left(\underline{1.569 \times 10^{-12}} \frac{n_r}{d_0(m_2 - m_2^*)} \right) \right]^{-1}$$

Ensemble Predictions

Ensemble Predictions:

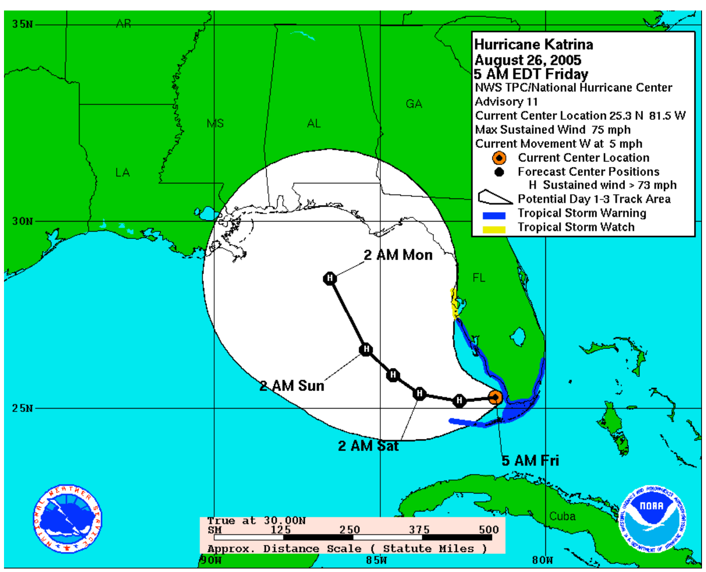


00 UTC on August 26, 2005



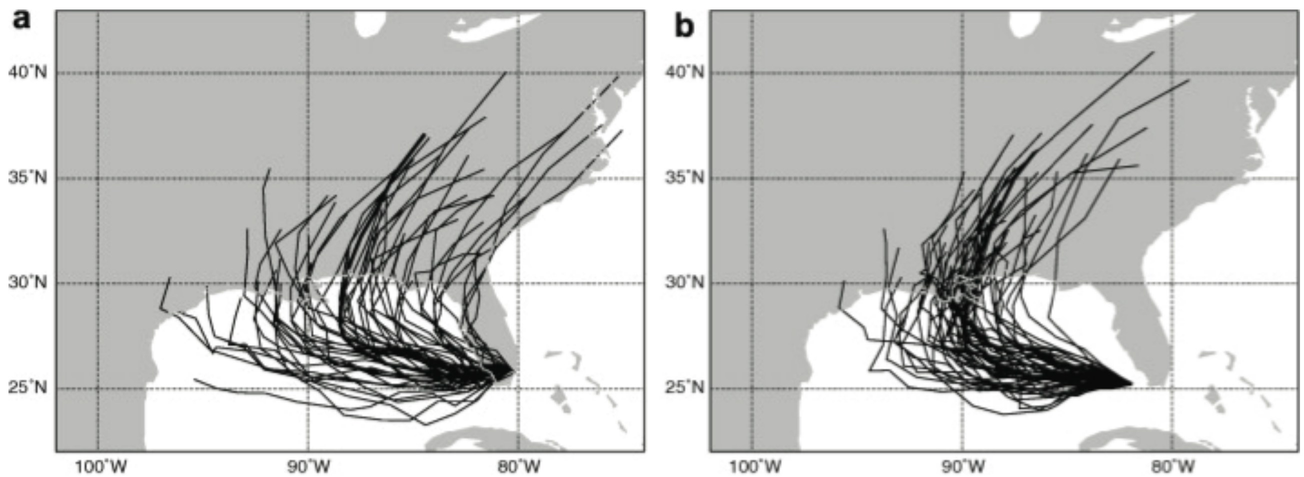
12 UTC on August 26, 2005

Cone of Uncertainty:



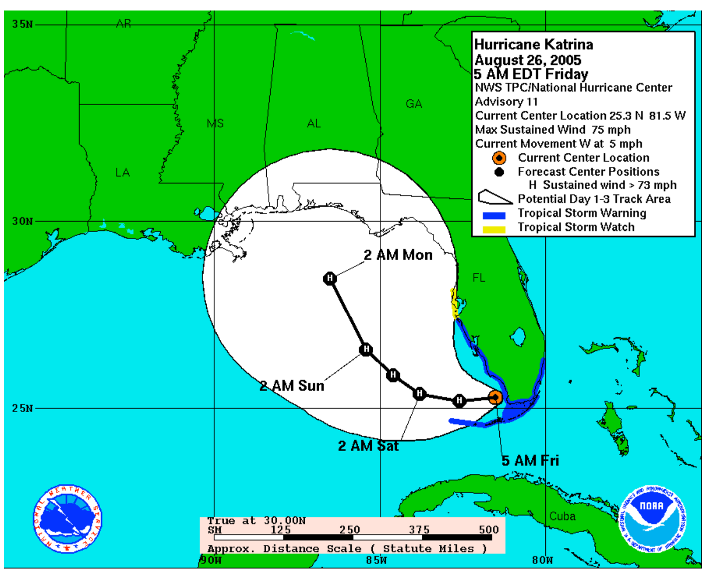
Ensemble Predictions

Ensemble Predictions:



00 UTC on August 26, 2005 12 UTC on August 26, 2005

Cone of Uncertainty:



General Questions:

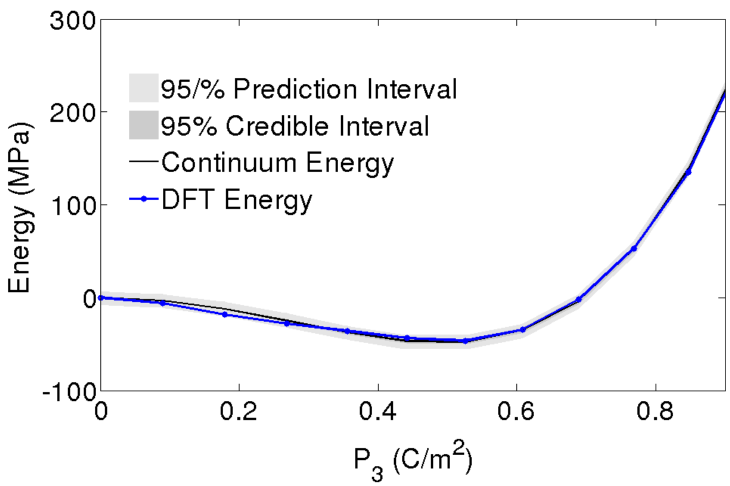
- What is expected rainfall on January 19?
- What are high and low temperatures?
- What is predicted average snow fall?
- Note: Quantities are statistical in nature.

Example 2: Quantum-Informed Continuum Models

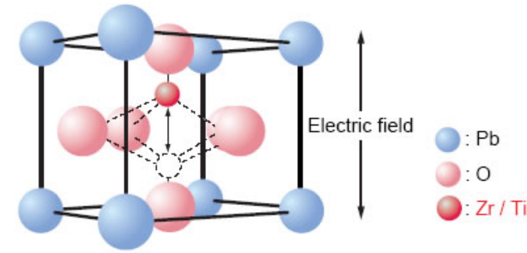
Objectives:

- Employ density function theory (DFT) to construct/calibrate continuum energy relations.
  - e.g., Helmholtz energy

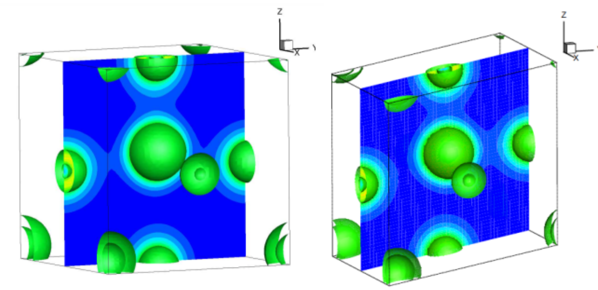
$$\psi(P) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6$$



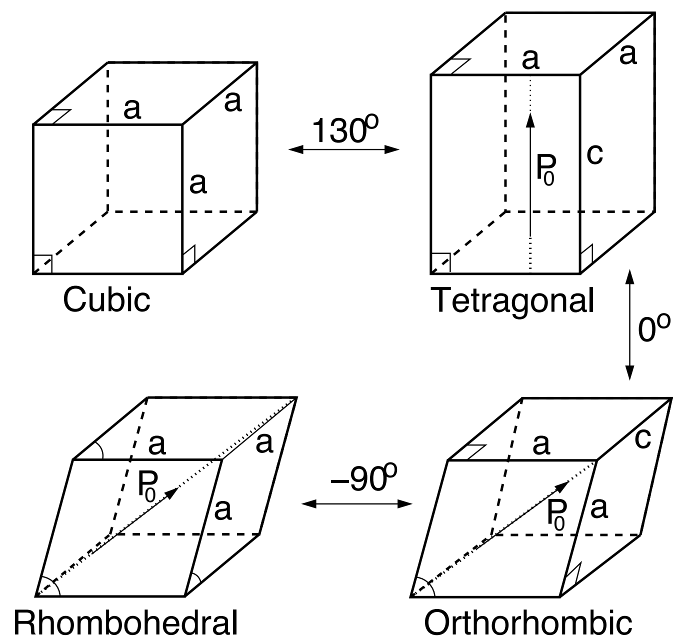
Helmholtz Energy



Lead Titanate Zirconate (PZT)



DFT Electronic Structure Simulation



UQ and SA Issues:

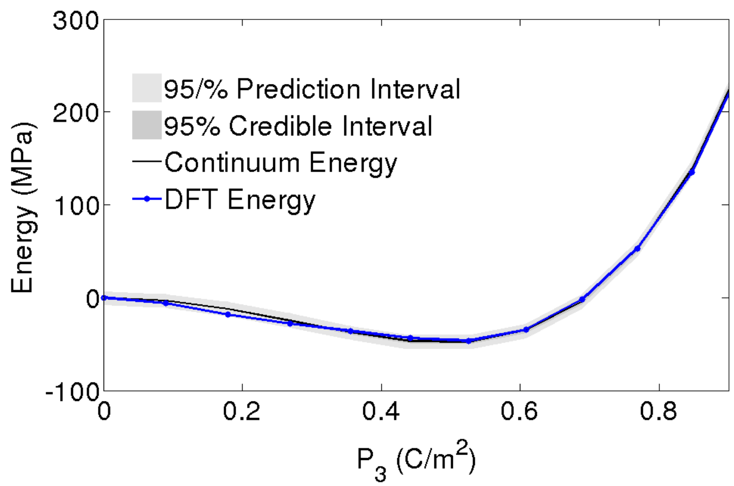
- Is 6th order term required to accurately characterize material behavior?
- **Note:** Determines molecular structure

Quantum-Informed Continuum Models

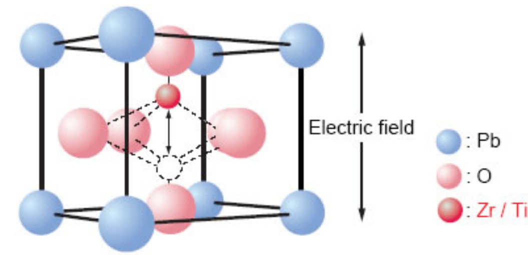
Objectives:

- Employ density function theory (DFT) to construct/calibrate continuum energy relations.
  - e.g., Helmholtz energy

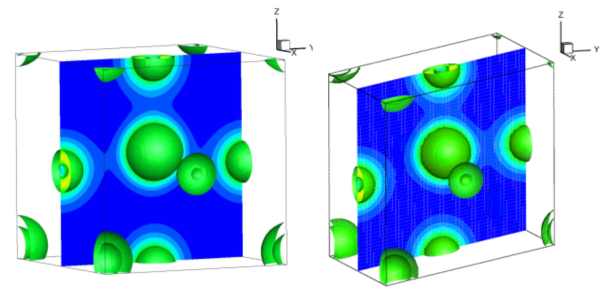
$$\psi(P) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6$$



Helmholtz Energy



Lead Titanate Zirconate (PZT)



DFT Electronic Structure Simulation

UQ and SA Issues:

- Is 6th order term required to accurately characterize material behavior?
- Note: Determines molecular structure

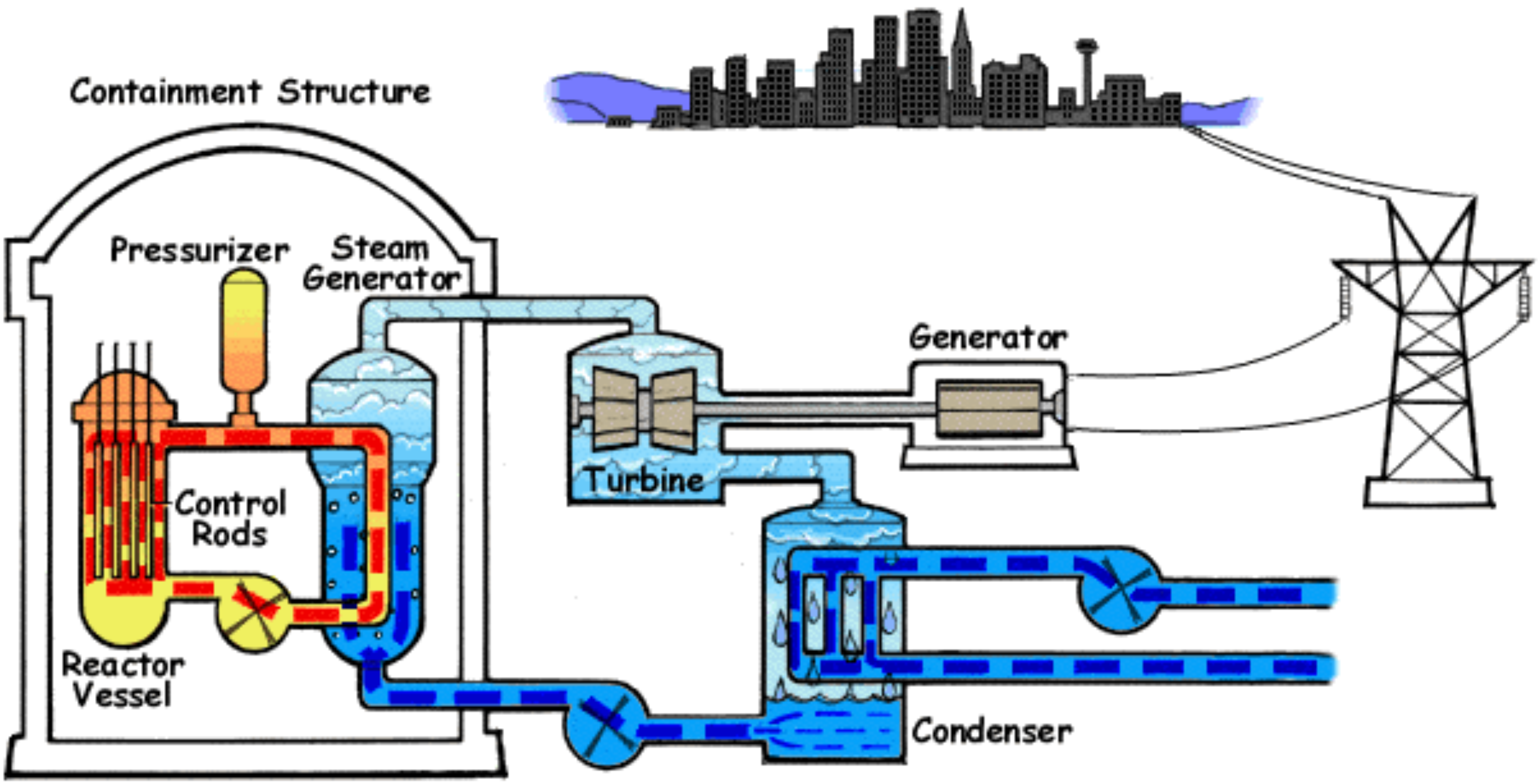
Broad Objective:

- Use UQ/SA to help bridge scales from quantum to system

Note:

- Linearly parameterized

Example 2: Pressurized Water Reactors (PWR)



Models:

•Involve neutron transport, thermal-hydraulics, chemistry.

•Inherently multi-scale, multi-physics.

CRUD Measurements: Consist of low resolution images at limited number of locations.

Example: Pressurized Water Reactors (PWR)

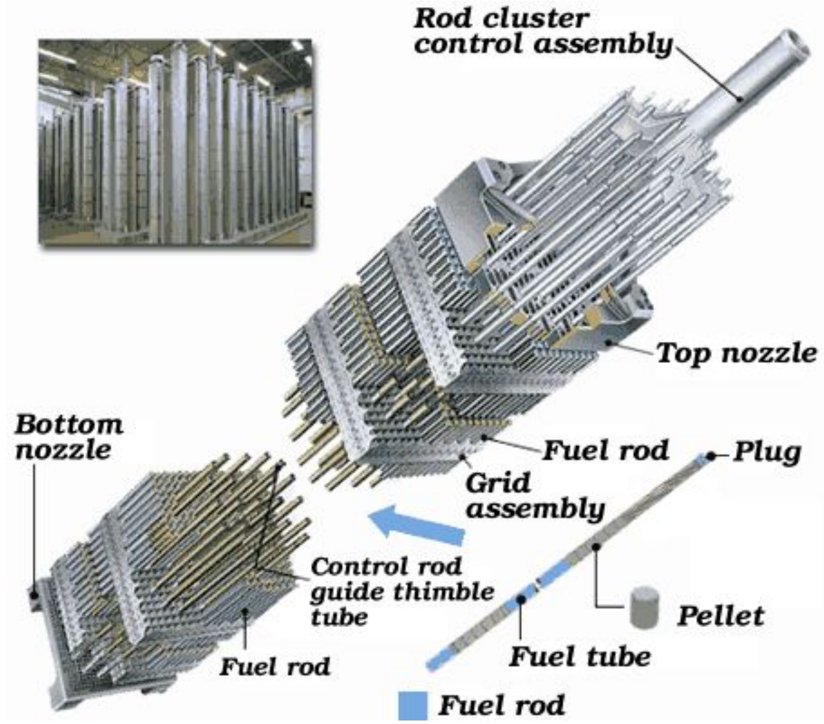
3-D Neutron Transport Equations:

$$\frac{1}{|v|}\frac{\partial \varphi}{\partial t} + \Omega \cdot \nabla \varphi + \Sigma_t(r,E)\varphi(r,E,\Omega,t)$$
$$= \int_{4\pi} d\Omega' \int_0^\infty dE' \Sigma_s(E' \to E, \Omega' \to \Omega)\varphi(r,E',\Omega',t)$$
$$+ \underline{\frac{\chi(E)}{4\pi}} \int_{4\pi} d\Omega' \int_0^\infty dE' \underline{\nu(E')}\,\underline{\Sigma_f(E')}\varphi(r,E',\Omega',t)$$

$$\boxed{\frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = \text{Sources - Sinks}}$$

Challenges:

- Very large number of inputs; e.g., 100,000; Active subspace construction critical.
- ORNL Code SCALE: Can take hours to run.
- Time-dependent surrogate models must accommodate PDE structure.
- Numerical errors often difficult to quantify.
- Predicting future requires extrapolatory or out-of-data predictions; one must address model discrepancy to construct validation intervals.



Example: Pressurized Water Reactors (PWR)

Thermo-Hydraulic Equations: Mass, momentum and energy balance for fluid

$$\frac{\partial}{\partial t}(\alpha_f \rho_f) + \nabla \cdot (\alpha_f \rho_f v_f) = -\Gamma$$

$$\alpha_f \rho_f \frac{\partial v_f}{\partial t} + \alpha_f \rho_f v_f \cdot \nabla v_f + \nabla \cdot \sigma_f^R + \alpha_f \nabla \cdot \sigma + \alpha_f \nabla p_f$$
$$= -F^R - F + \Gamma(v_f - v_g)/2 + \alpha_f \rho_f g$$

$$\frac{\partial}{\partial t}(\alpha_f \rho_f e_f) + \nabla \cdot (\alpha_f \rho_f e_f v_f + Th) = (T_g - T_f)H + T_f \Delta_f$$
$$-T_g(H - \alpha_g \nabla \cdot h) + h \cdot \nabla T - \Gamma[e_f + T_f(s^* - s_f)]$$
$$-p_f \left(\frac{\partial \alpha_f}{\partial t} + \nabla \cdot (\alpha_f v_f) + \frac{\Gamma}{\rho_f} \right)$$

$$\boxed{\frac{\partial \rho}{\partial t} + \frac{\partial \phi}{\partial x} = \text{Sources - Sinks}}$$

Notes:

- Similar relations for gas and bubbly phases
- Surrogate models must conserve mass, energy, and momentum

Example: Shearon Harris outside Raleigh

UQ Questions:

- What is peak operating temperature?
- What is expected level of CRUD buildup?
- What is risk associated with operating regime?
- What is expected profit for new design?

Example 4: SIR Model for Disease Dynamics

SIR Model:

$$\frac{dS}{dt} = \delta N - \delta S - \gamma \underline{kIS} \quad , \; S(0) = S_0 \qquad \text{Susceptible}$$

$$\frac{dI}{dt} = \gamma \underline{kIS} - (r + \delta) I \quad , \; I(0) = I_0 \qquad \text{Infectious}$$

$$\frac{dR}{dt} = rI - \delta R \quad , \; R(0) = R_0 \qquad \text{Recovered}$$

Parameters:

- γ: Infection coefficient
- k: Interaction coefficient
- r: Recovery rate
- δ: Birth/death rate

Response:

$$y = \int_0^5 R(t, q)\, dt$$

Note: Parameters $q = [\gamma, k, r, \delta]$ not uniquely determined by data

Note: Presently employed cholera models have similar form; example this afternoon.

SIR Disease Example

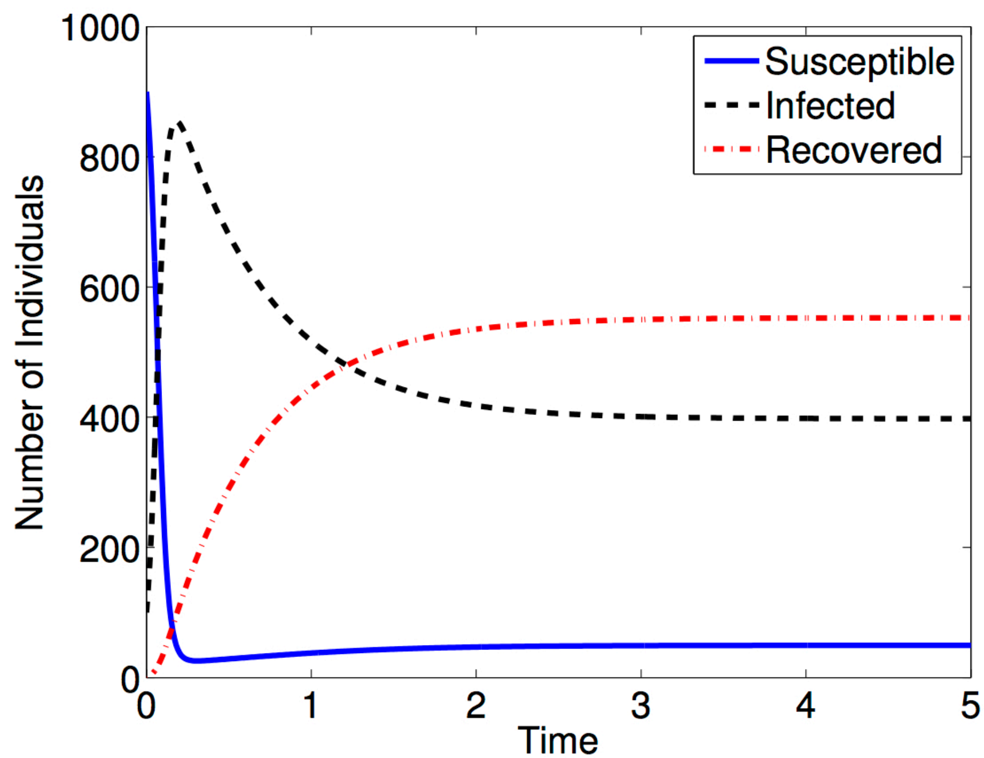
SIR Model:

$$\frac{dS}{dt} = \delta N - \delta S - \gamma k I S \quad , \quad S(0) = S_0 \qquad \text{Susceptible}$$

$$\frac{dI}{dt} = \gamma k I S - (r + \delta) I \quad , \quad I(0) = I_0 \qquad \text{Infectious}$$

$$\frac{dR}{dt} = rI - \delta R \quad , \quad R(0) = R_0 \qquad \text{Recovered}$$

Typical Realization:



SIR Disease Example

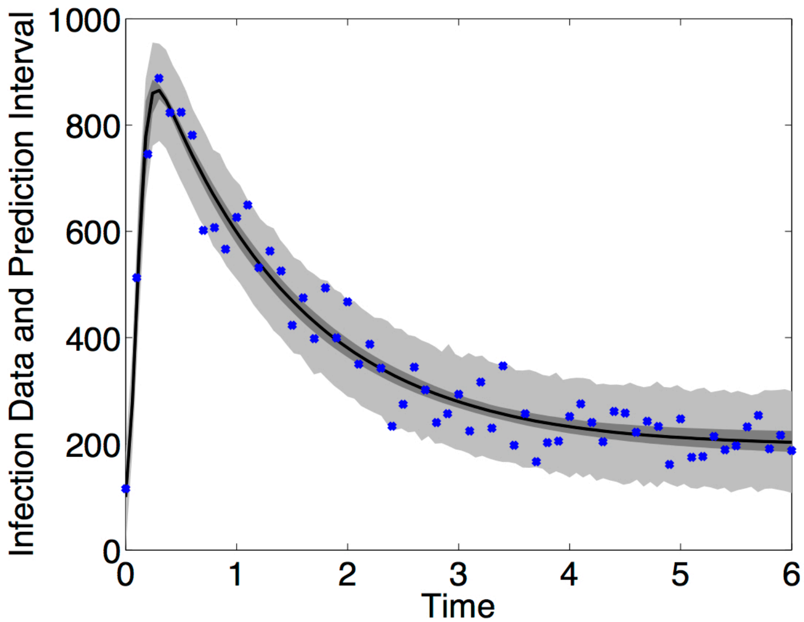
SIR Model:

$$\frac{dS}{dt} = \delta N - \delta S - \underline{\gamma k}IS \quad , \quad S(0) = S_0$$

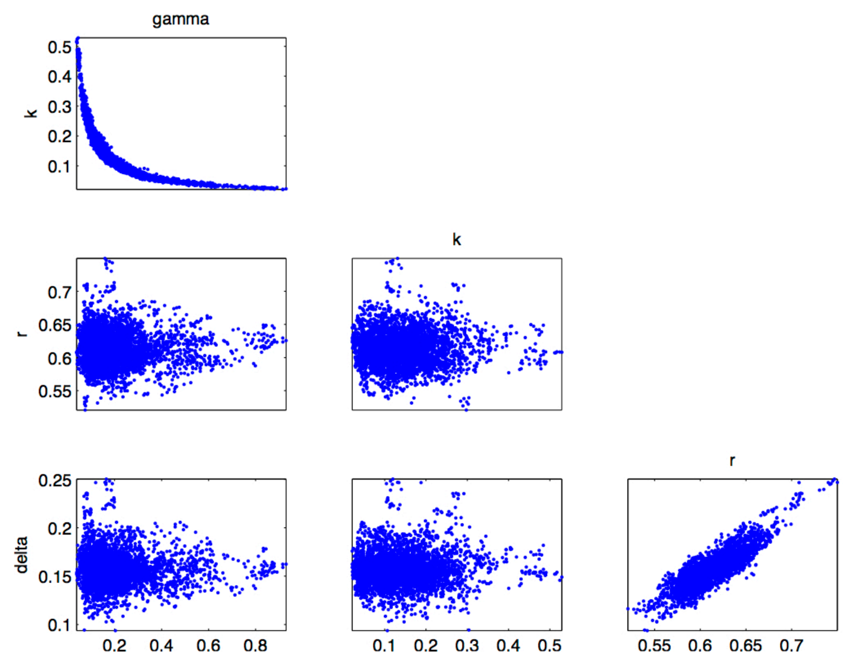
$$\frac{dI}{dt} = \underline{\gamma k}IS - (r + \delta)I \quad , \quad I(0) = I_0$$

$$\frac{dR}{dt} = rI - \delta R \quad , \quad R(0) = R_0$$

UQ Goal: Predict I(t) with uncertainty intervals:



Problem: Cannot uniquely infer parameters



Solution:

- Active subspaces
- Identifiability analysis
- Sensitivity analysis
- Design of experiments

Example 5: HIV Model for Characterization and Control Regimes

HIV Model:

$$\dot{T}_1 = \lambda_1 - d_1 T_1 - (1-\varepsilon)k_1 V T_1$$

$$\dot{T}_2 = \lambda_2 - d_2 T_2 - (1-f\varepsilon)k_2 V T_2$$

$$\dot{T}_1^* = (1-\varepsilon)k_1 V T_1 - \delta T_1^* - m_1 E T_1^*$$

$$\dot{T}_2^* = (1-f\varepsilon)k_2 V T_2 - \delta T_2^* - m_2 E T_2^*$$

$$\dot{V} = N_T \delta (T_1^* + T_2^*) - cV - [(1-\varepsilon)\rho_1 k_1 T_1 + (1-f\varepsilon)\rho_2 k_2 T_2] V$$

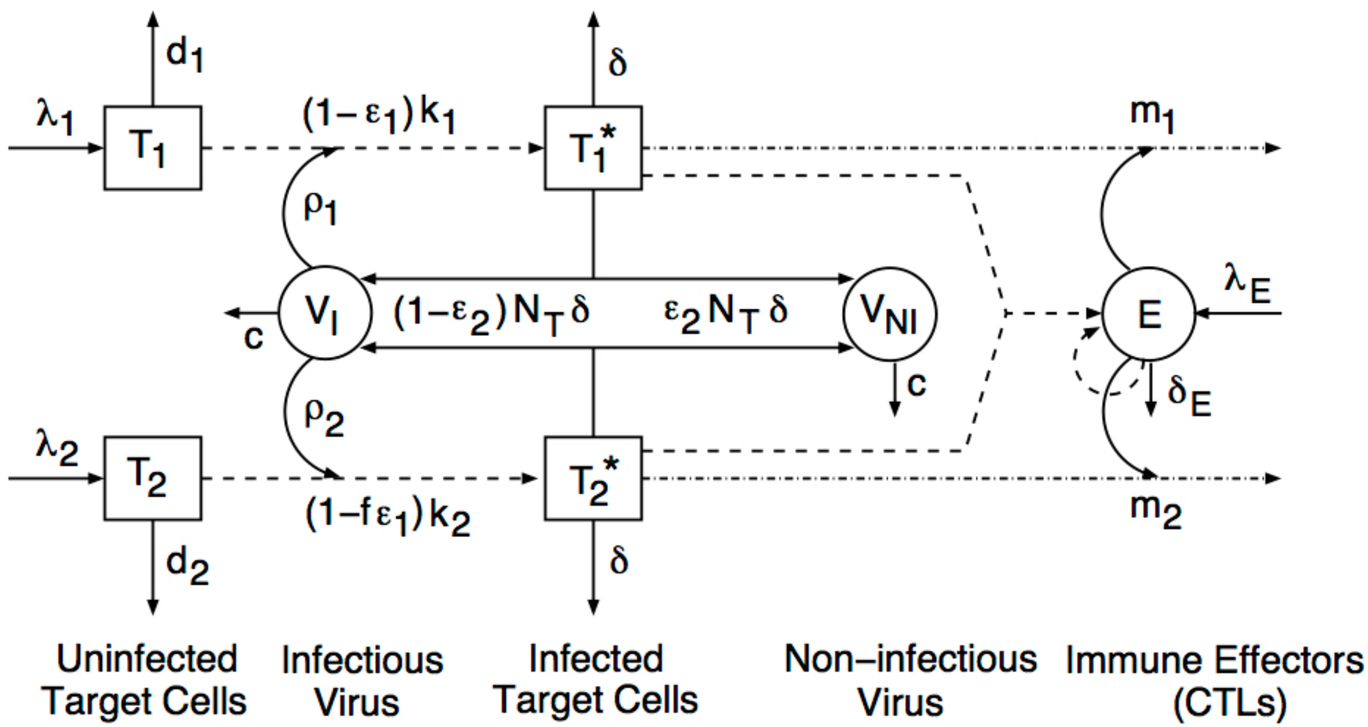
$$\dot{E} = \lambda_E + \frac{b_E(T_1^* + T_2^*)}{T_1^* + T_2^* + K_b} E - \frac{d_E(T_1^* + T_2^*)}{T_1^* + T_2^* + K_d} E - \delta_E E$$

Notes: 21 parameters

[Adams, Banks et al., 2005, 2007]

Notation: $\dot{E} \equiv \dfrac{dE}{dt}$

Compartments:



Example 3: HIV Model for Characterization and Control Regimes

HIV Model: Used for characterization and control treatment regimes.

$$\dot{T}_1 = \lambda_1 - d_1 T_1 - (1-\varepsilon) k_1 V T_1$$

$$\dot{T}_2 = \lambda_2 - d_2 T_2 - (1-f\varepsilon) k_2 V T_2$$

$$\dot{T}_1^* = (1-\varepsilon) k_1 V T_1 - \delta T_1^* - m_1 E T_1^*$$

$$\dot{T}_2^* = (1-f\varepsilon) k_2 V T_2 - \delta T_2^* - m_2 E T_2^*$$

$$\dot{V} = N_T \delta (T_1^* + T_2^*) - cV - [(1-\varepsilon)\rho_1 k_1 T_1 + (1-f\varepsilon)\rho_2 k_2 T_2] V$$

$$\dot{E} = \lambda_E + \frac{b_E (T_1^* + T_2^*)}{T_1^* + T_2^* + K_b} E - \frac{d_E (T_1^* + T_2^*)}{T_1^* + T_2^* + K_d} E - \delta_E E$$

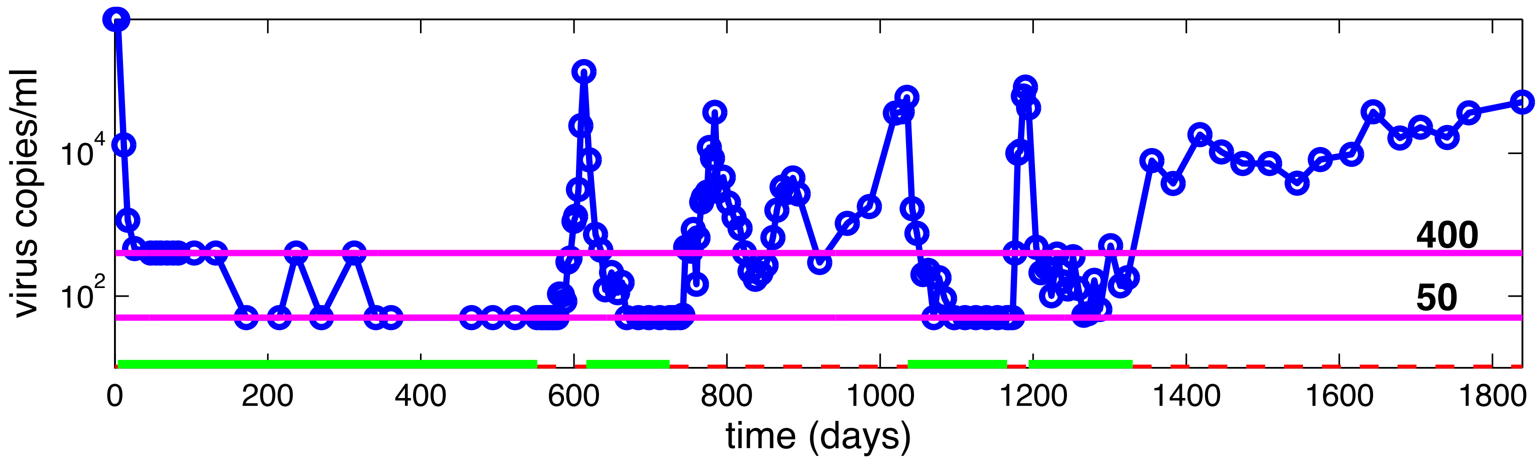
Parameters: Most are unknown and must be estimated from data

λ_1	Target cell 1 production rate	ρ_1	Ave. virions infecting type 1 cell
λ_2	Target cell 2 production rate	ρ_2	Ave. virions infecting type 2 cell
d_1	Target cell 1 death rate	b_E	Max. birth rate immune effectors
d_2	Target cell 2 death rate	d_E	Max. death rate immune effectors
k_1	Population 1 infection rate	K_b	Birth constant, immune effectors
k_2	Population 2 infection rate	K_d	Death constant, immune effectors
c	Virus natural death rate	λ_E	Immune effector production rate
δ	Infected cell death rate	δ_E	Natural death rate, immune effectors
ε	Population 1 treatment efficacy	N_T	Virions produced per infected cell
m_1	Population 1 clearance rate	f	Treatment efficacy reduction
m_2	Population 2 clearance rate		

Example: HIV Model for Characterization and Treatment Regimes

HIV Model: Several sources of uncertainty including viral measurement techniques

Example: Upper and lower limits to assay sensitivity



UQ Questions:

- What are the uncertainties in parameters that cannot be directly measured?
- What is optimal treatment regime that is “safe” for patient?
- What is expected viral load? Issue: very often requires high-dimensional integration!
- e.g., $\mathbb{E}[V(t)] = \int_{\mathbb{R}^{21}} V(t,q)\rho(q)dq$

Experimental results are believed by everyone, except for the person who ran the experiment, source anonymous, quoted by Max Gunzburger, Florida State University.

Example 4: Portfolio Model

Example: Portfolio model

$$Y = c_1 Q_1 + c_2 Q_2$$

Note:

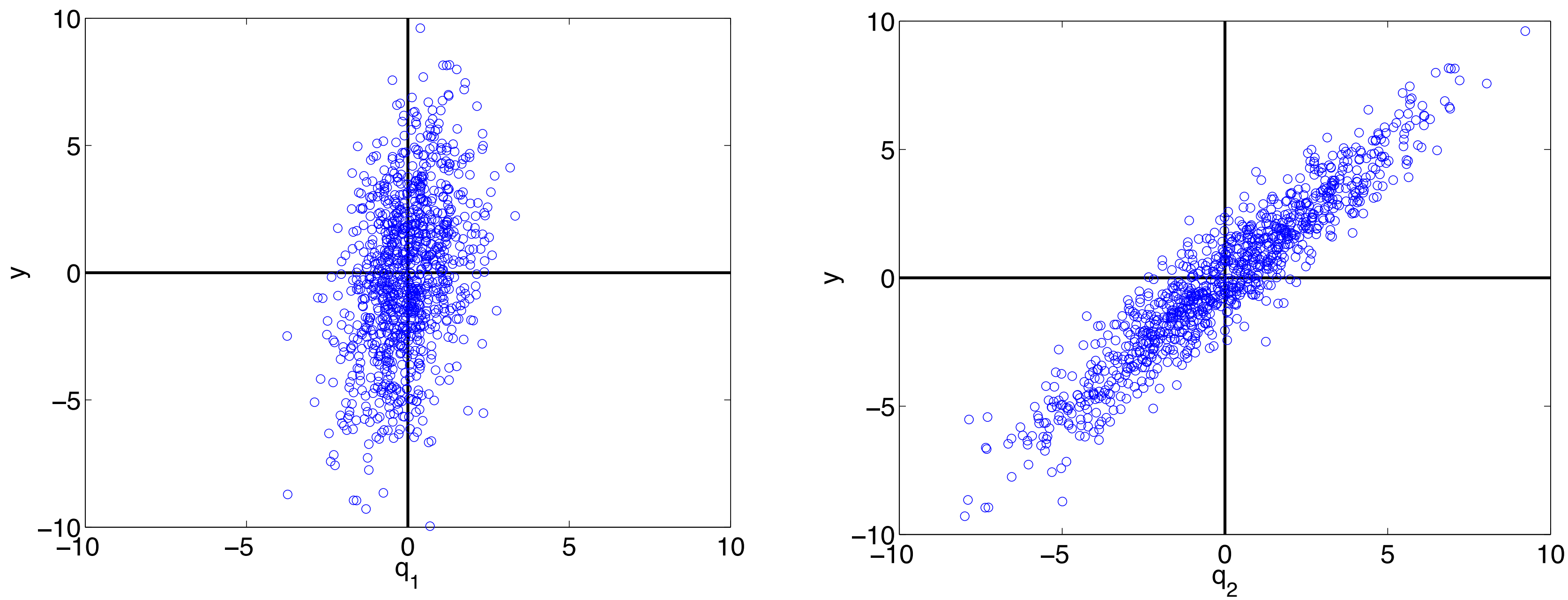
- Q_1 and Q_2 represent hedged portfolios
- c_1 and c_2 amounts invested in each portfolio

Take

$$c_1 = 2 \;,\;\; c_2 = 1$$

$Q_1 \sim N(0, \sigma_1^2)$ with $\sigma_1 = 1$

$Q_2 \sim N(0, \sigma_2^2)$ with $\sigma_2 = 3$



UQ Questions:

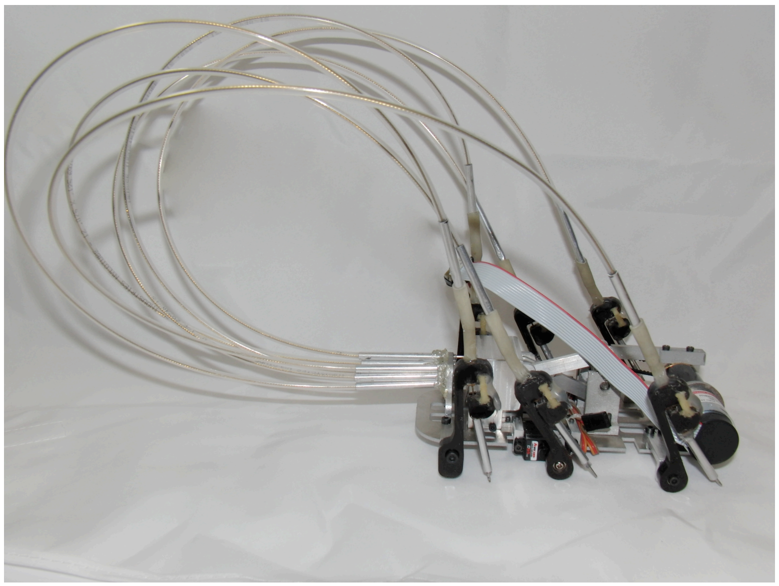
•What is expected investment return?

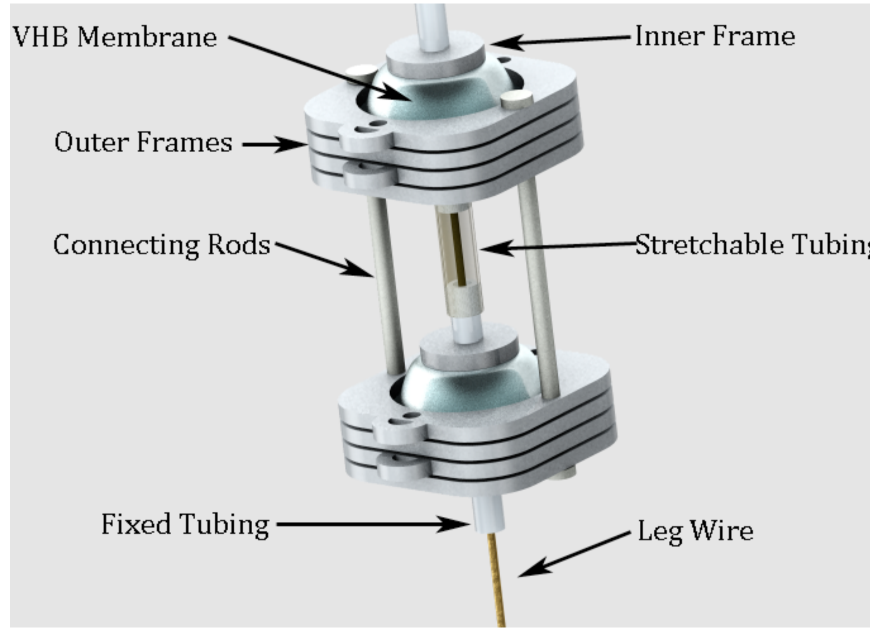
•What is impact of market uncertainty on investment return?

Example 5: Viscoelastic Material Models

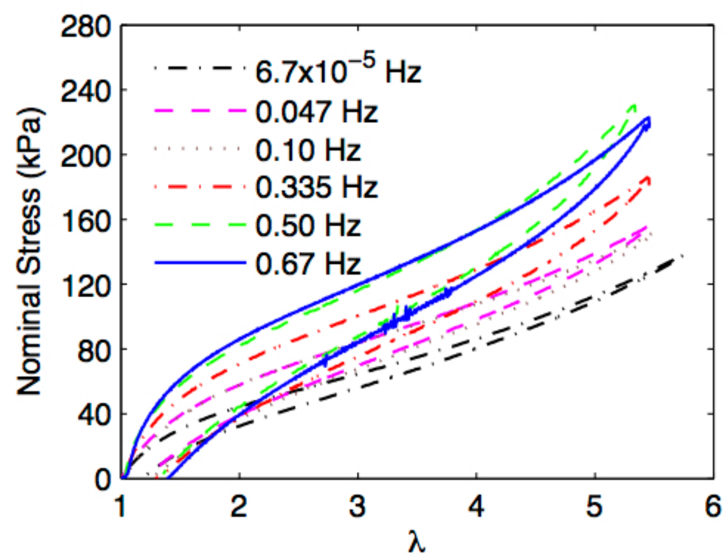
Application: Adaptive materials for legged robotics

- Figure: Billy Oates



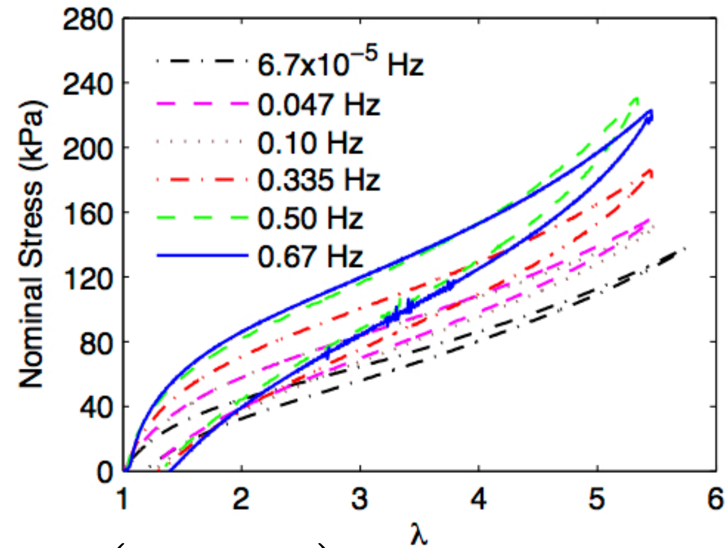


Material Behavior: Significant rate dependence



Example 5: Viscoelastic Material Models

Material Behavior: Significant rate dependence



Finite-Deformation Model:

- Nonlinear non-affine
- Hyperelastic energy function

$$\psi_\infty^N = \frac{1}{6}\underline{G_c} I_1 - \underline{G_c \lambda_{\max}^2} \ln\left(3\underline{\lambda_{\max}^2} - I_1\right) + \underline{G_e} \sum_j \left(\lambda_j + \frac{1}{\lambda_j}\right)$$

Parameters:

$q = [G_e, G_c, \lambda_{\max}, \eta, \beta, \gamma]$

$q = [\eta, \beta, \gamma]$: Viscoelastic parameters

G_c: Crosslink network modulus

G_e: Plateau modulus

$\lambda_{\max}$: Max stretch effective affine tube

Uncertainty Quantification Goals:

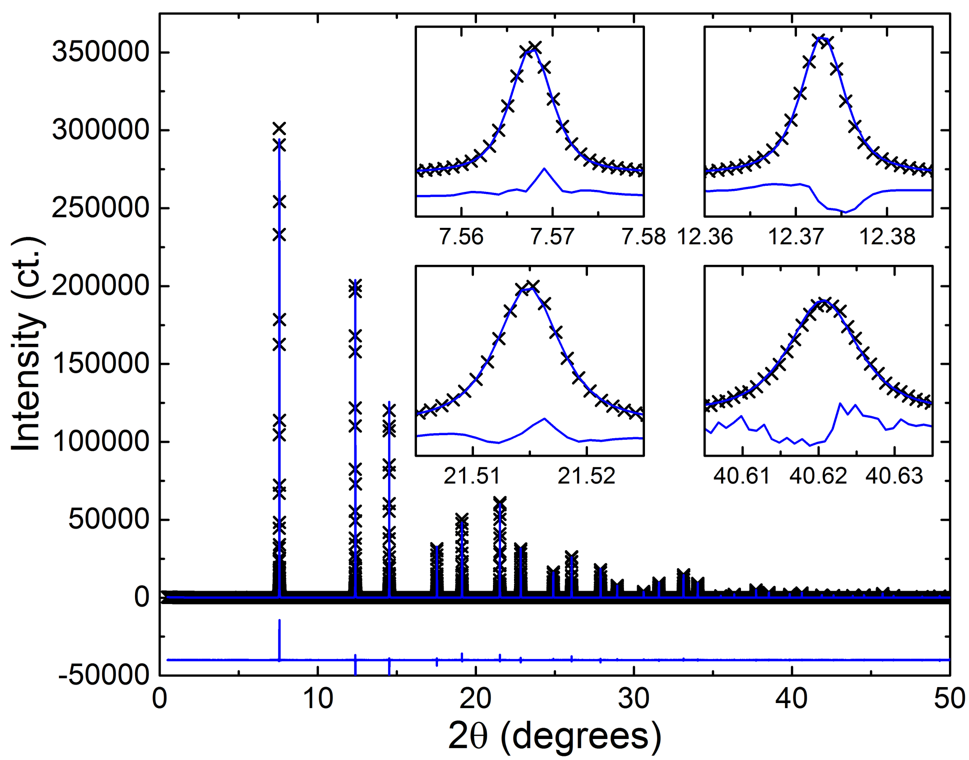
- Quantify measurement errors.
- Quantify uncertainty in parameters.
- Use statistics to quantify accuracy of considered models.

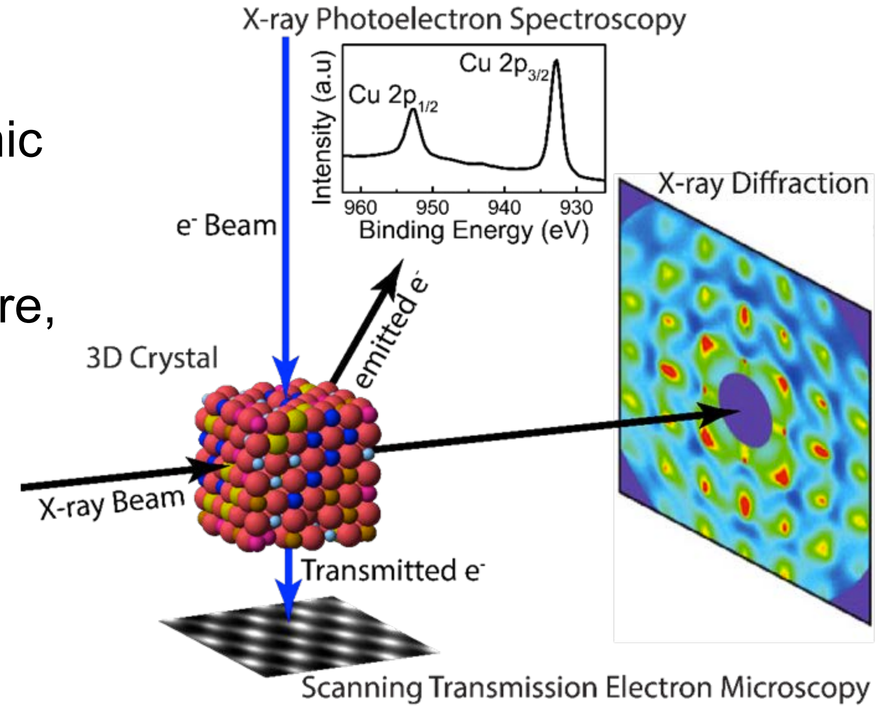
Example 6: X-Ray Crystallography

Properties:

• Reveal relative positions of atoms, their atomic number, types of chemical bonds, etc.

• Applications: determination of of DNA structure, design of pharmaceuticals, etc..





Uncertainty Quantification Goals:

- Use Bayesian analysis to quantify uncertainty associated with Rietveld model and background.
- Quantify heteroskedasticity and correlation of error structure.

Collaborators: Chris Fancher, Zhen Han, Igor Levin, Katherine Page, Brian Reich, Alyson Wilson, Jacob Jones

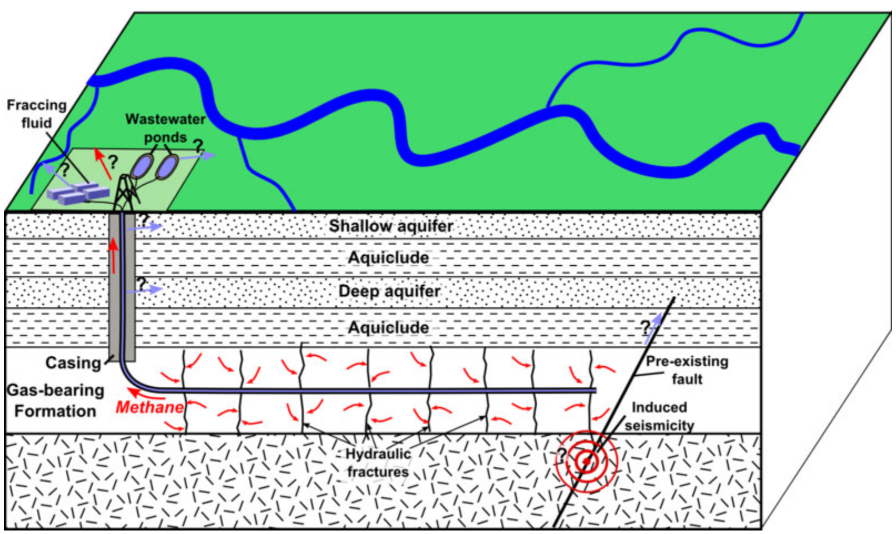
Experimental Uncertainties and Limitations

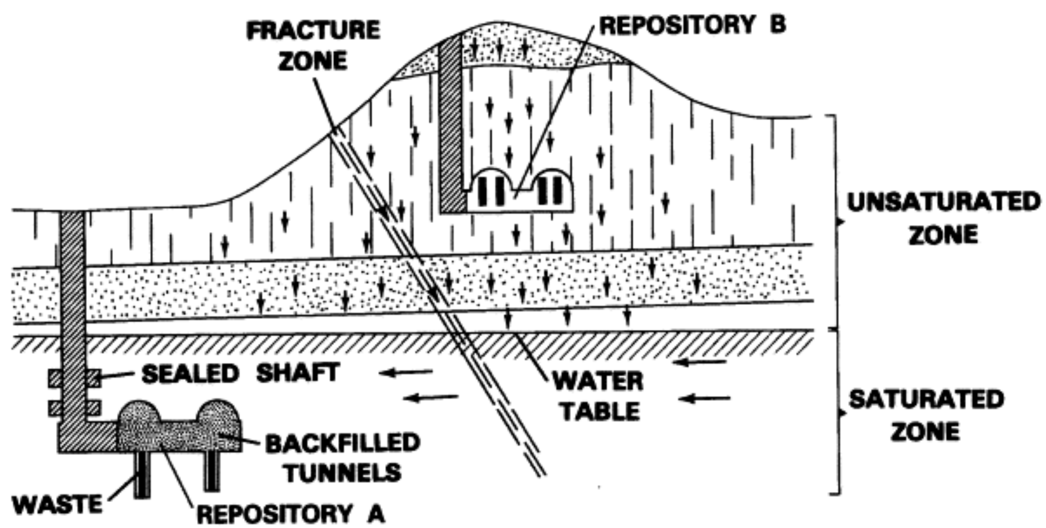
Examples: *Experimental results are believed by everyone, except for the person who ran the experiment*, Max Gunzburger, Florida State University.

• Pharmaceutical and disease treatment strategies often too dangerous or expensive for human tests or large segments of the population.

• Climate scenarios cannot be experimentally tested at the planet scale. Instead, components such as volcanic forcing tested using measurements such as the 1991 Mount Pinatubo data.

• Subsurface hydrology data very limited due to infeasibility of drilling large numbers of wells. Result: significant uncertainty regarding subsurface structures.



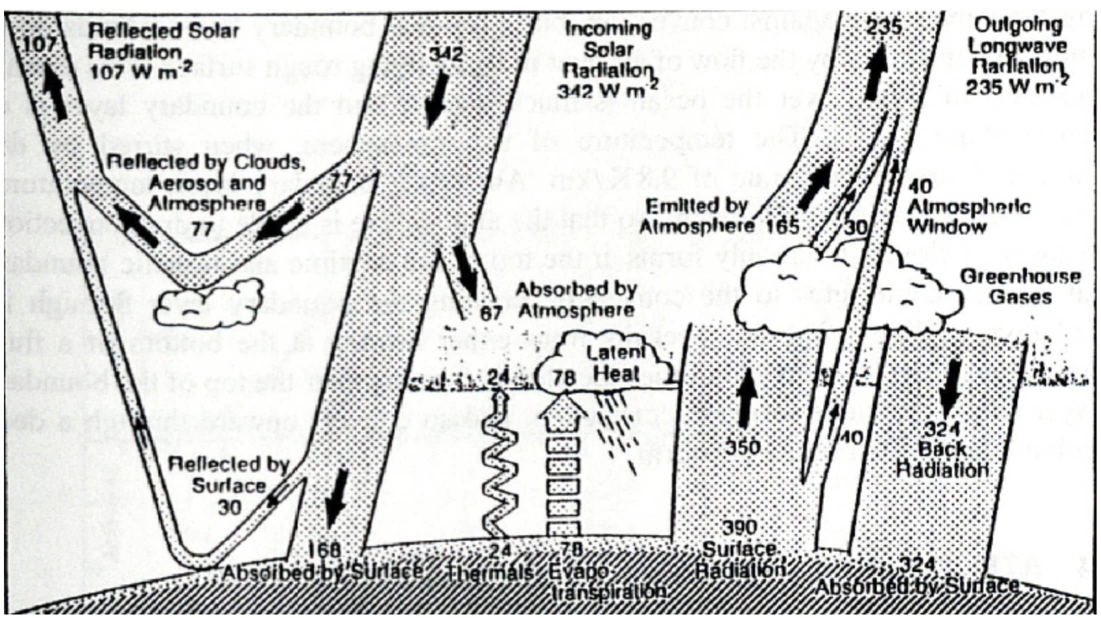


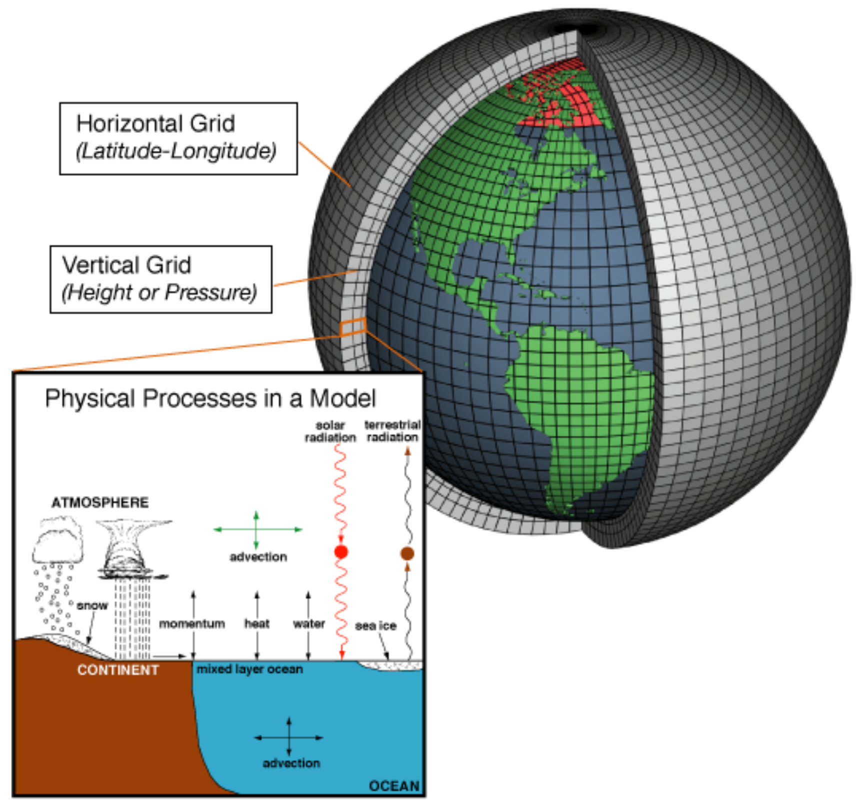
Model Errors

Examples: *Essentially, all models are wrong, but some are useful,* George E.P. Box, Industrial Statistician

• Numerous components of weather and climate models --- e.g., aerosol-induced cloud formation, greenhouse gas processes --- occur on scales that are much smaller than numerical grids used to solve the atmospheric equations of physics. These processes represent highly complex physics that is only partially understood.

• Many biological applications are coupled, complex, highly nonlinear, and driven by poorly understood or stochastic processes.



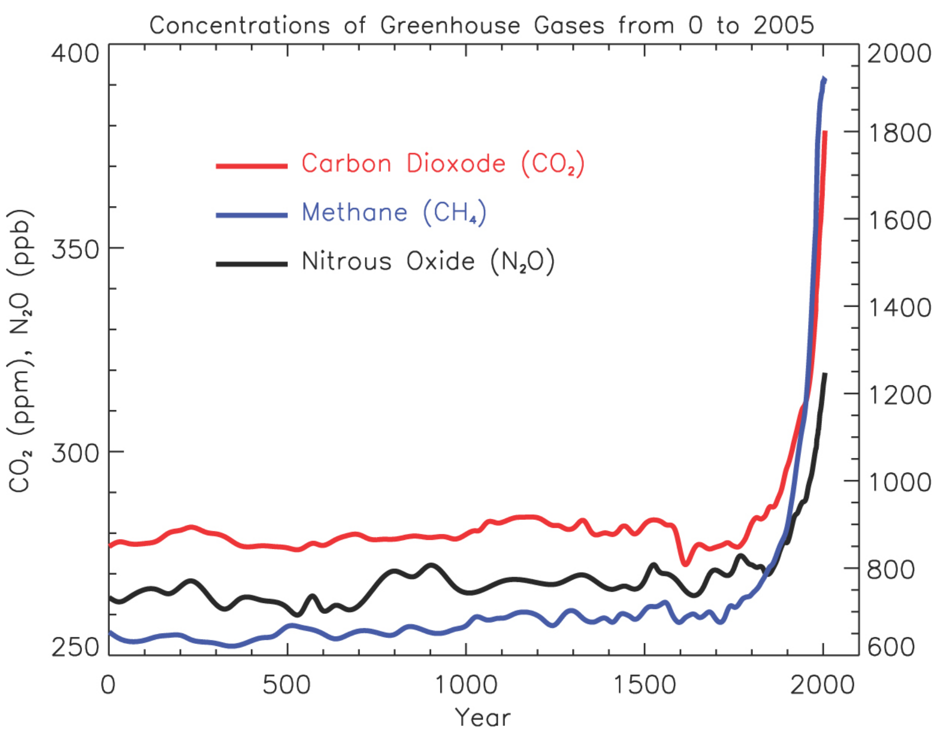


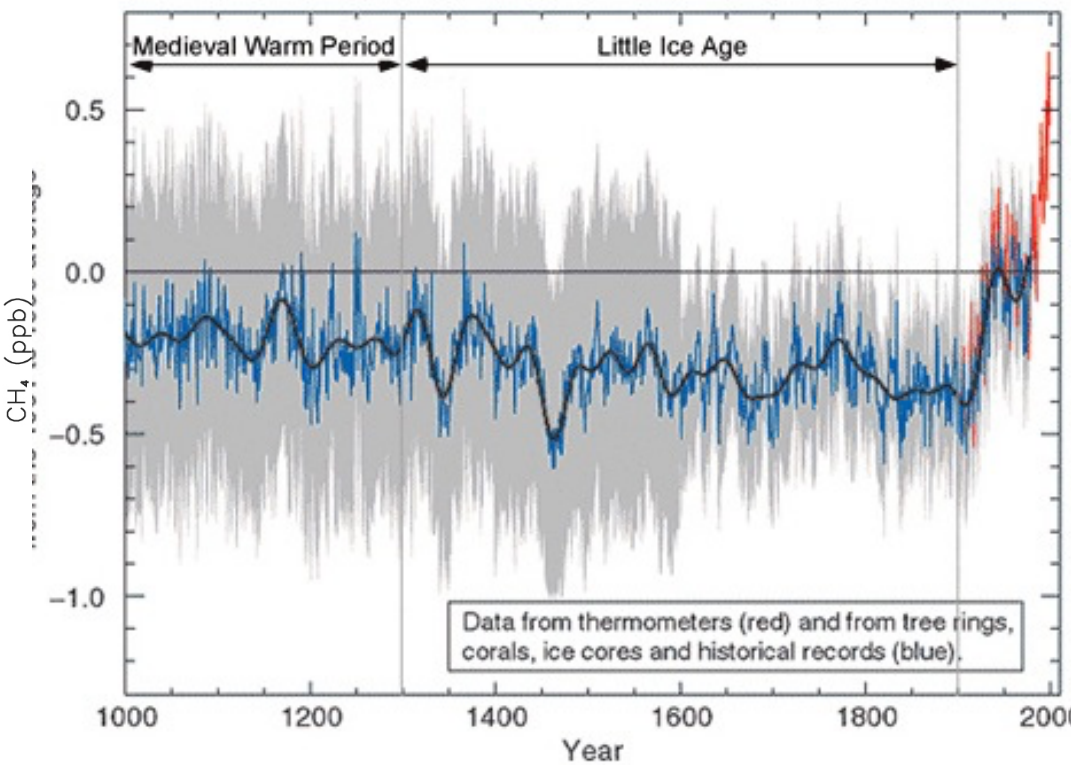
Input Uncertainties

Note: *Essentially, all models are wrong, but some are useful,* George E.P. Box, Industrial Statistician

• Phenomenological models used to represent processes such as turbulence in weather, climate and nuclear reactor models have nonphysical parameters whose values and uncertainties must be determined using measured data.

• Forcing and feedback mechanisms in climate models serve as boundary inputs. These parameterized phenomenological relations introduce both model and parameter uncertainties.





Numerical Errors

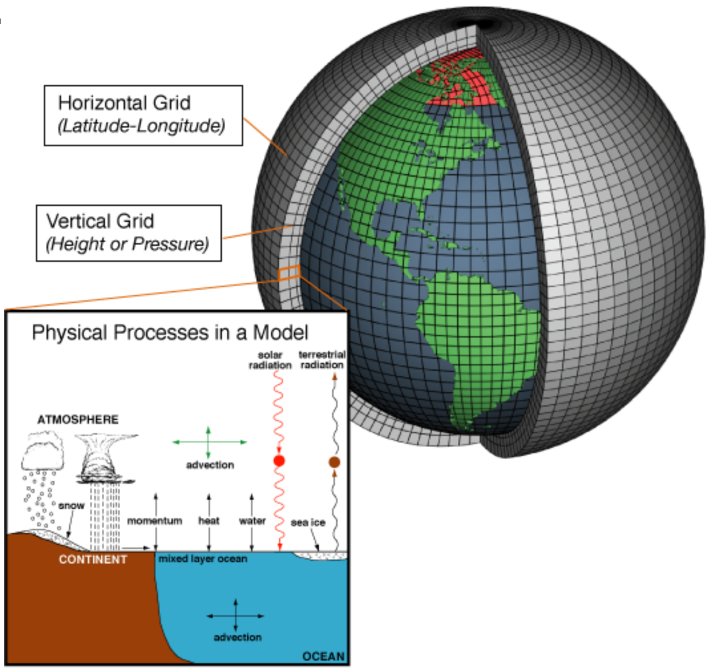
Note: *Computational results are believed by no one, except the person who wrote the code*, Max Gunzburger, Florida State University.

• Roundoff, discretization or approximation errors; e.g., mesh for nuclear subchannel code COBRA-TF is on the order of subchannel between rods.

• Bugs or coding errors;

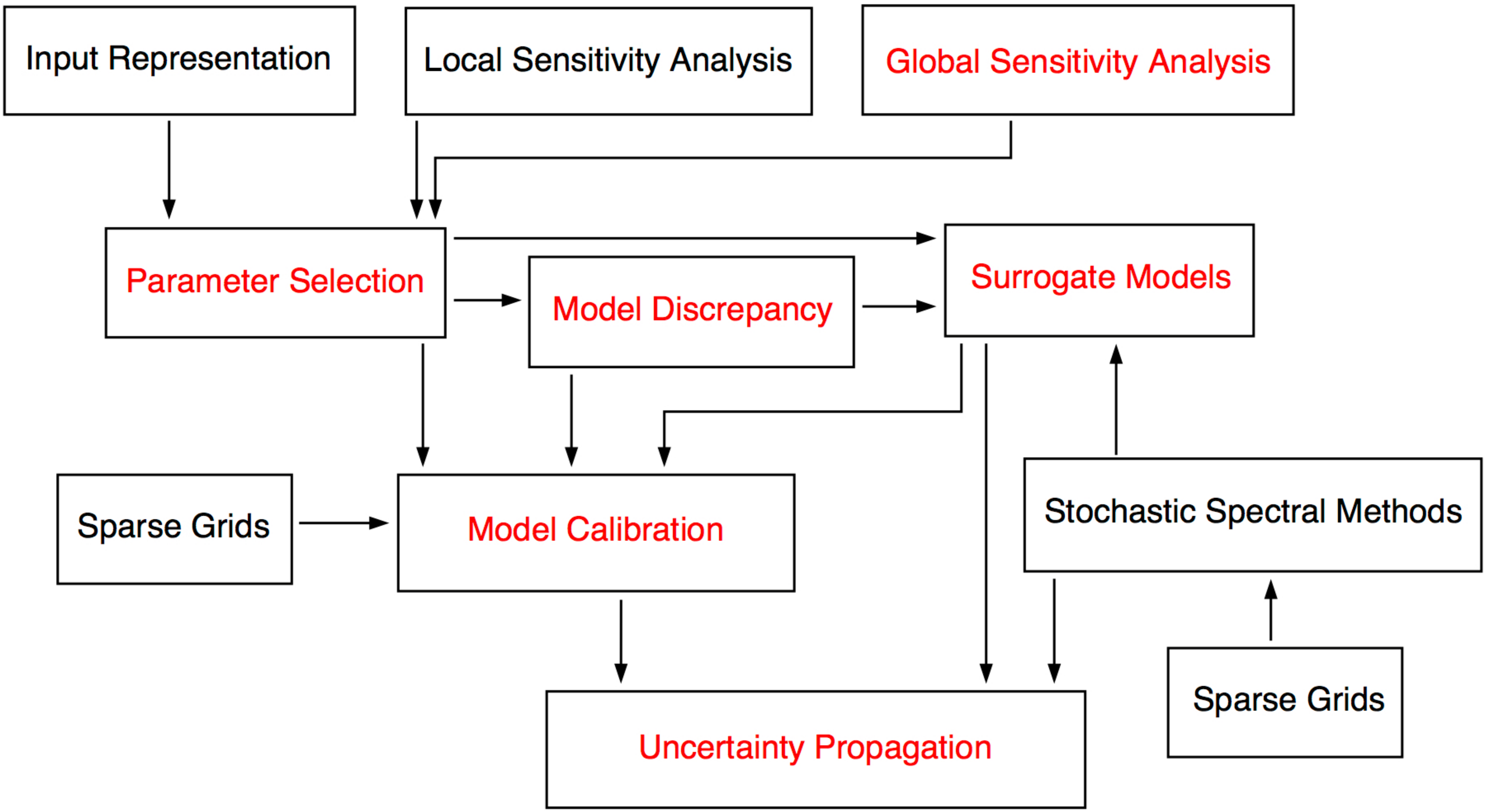
• Bit-flipping, hardware failures and uncertainty associated with future exascale and quantum computing;

• Grids required for numerical solutions of field equations in applications such as weather or climate models (e.g., 50~km) are much larger than the scale of physics being modeled (e.g., turbulence or greenhouse gases).

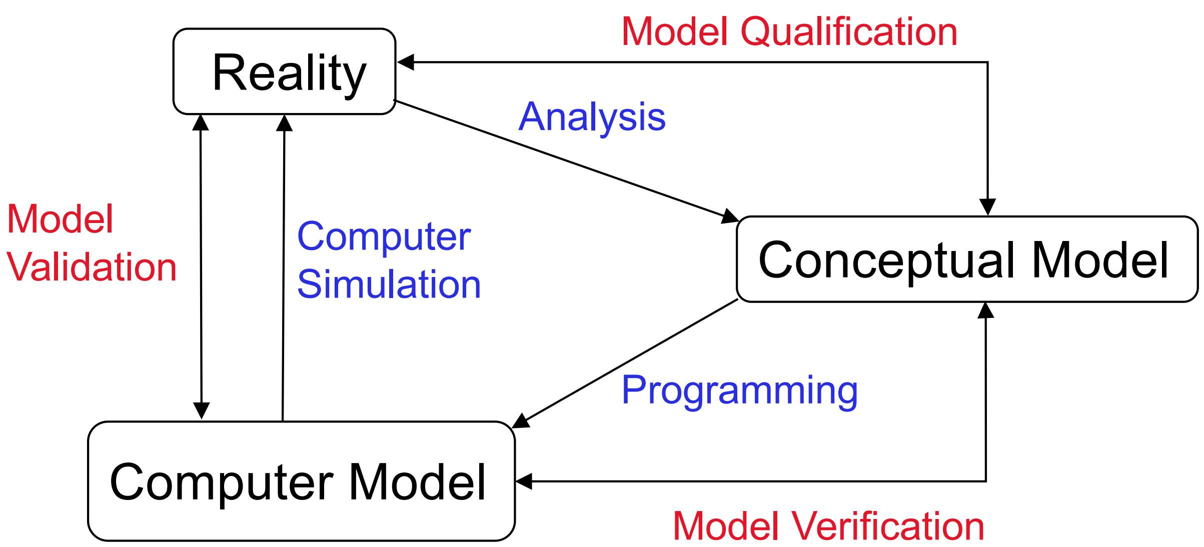


Steps in Uncertainty Quantification

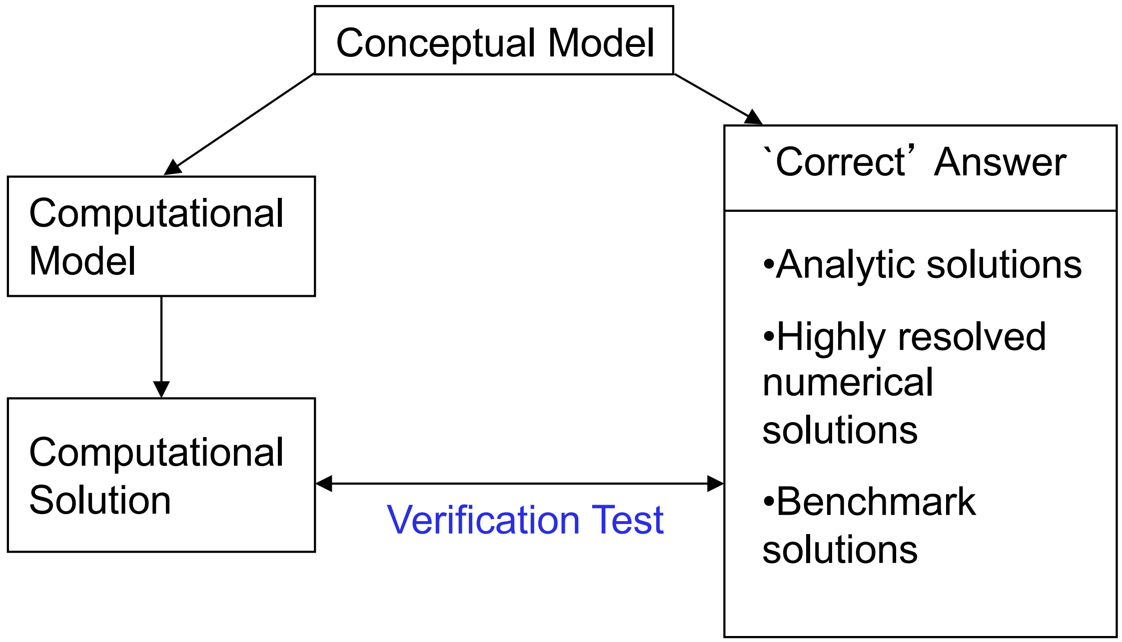
Note: Uncertainty quantification requires synergy between statistics, mathematics and application area.



Modeling Issues



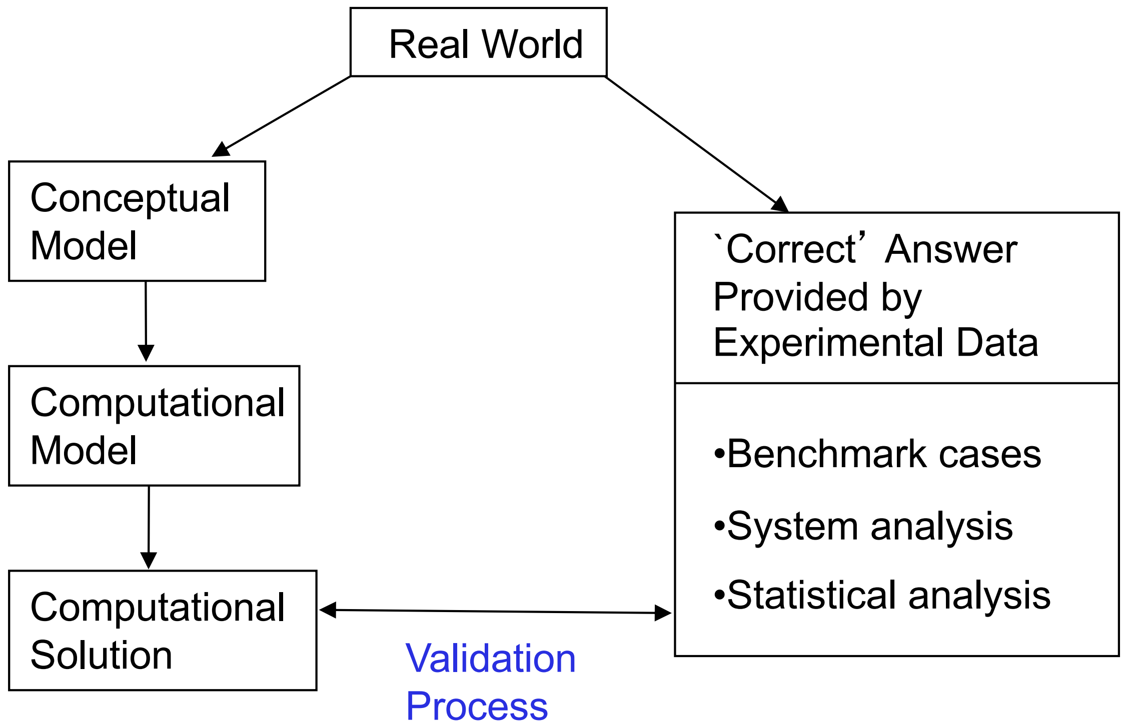
Verification Process



Verification: The process of determining that a model implementation accurately represents the developer's conceptual description of the model and the solution to the model.

Note: Verification deals with mathematics

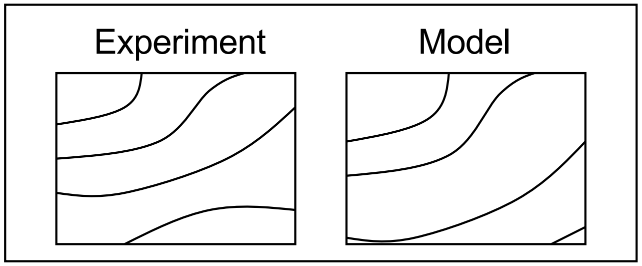
Validation Process



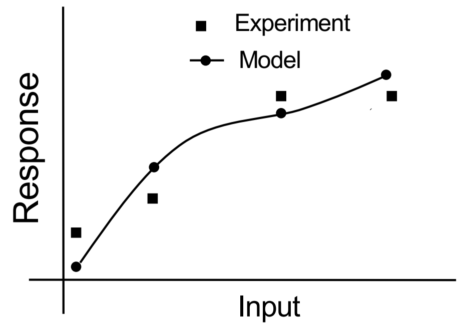
Validation: The process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended model users.

Note: Validation deals with physics and statistics

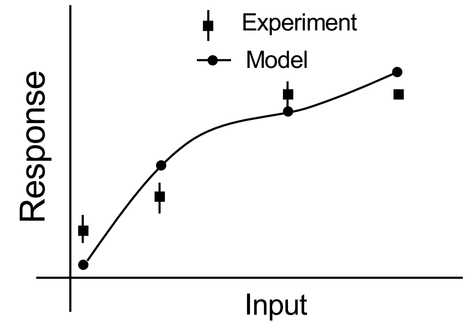
Validation Metrics



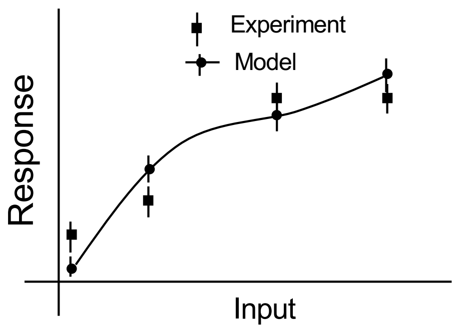
`Viewgraph’ Norm



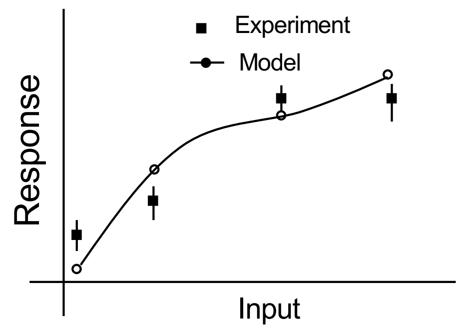
Deterministic



Experimental Uncertainty



Numerical Error



Nondeterministic Computation