INTRODUCTION TO UNCERTAINTY QUANTIFICATION IN PREDICTIVE COMPUTATIONAL MODELS

Professor: Fernando Rochinha
Mechanical and Nanotechnology Engineering - COPPE
Universidade Federal do Rio de Janeiro

Course Description: The impressive progress experimented by computer technologies and numerical methods has pushed the frontier of predictive modelling. Powerful computer codes built upon physics based sophisticated models are more and more employed in the analysis and design of complex systems within different applied areas. The goal of this course is to give an introduction to the rapidly expanding field of uncertainty quantification which builds on probability, statistics, computation and large scale simulations. There is uncertainty in the mathematical model, in the parameters, and in the initial and boundary data. How do these uncertainties propagate and might hamper the reliability of o the predictions of the computations? In the inverse problem, parameters are determined from measured data. What is the effect of the errors in the data on the parameter estimation? Some basic knowledge of random processes and differential equations is useful. Parameters in mathematical models based on differential equations will be estimated using frequentist and Bayesian techniques. This course is given at graduate level.

Bibliography: R. C. Smith, Uncertainty Quantification, Theory, Implementation and Applications, SIAM, Philadelphia, 2014.

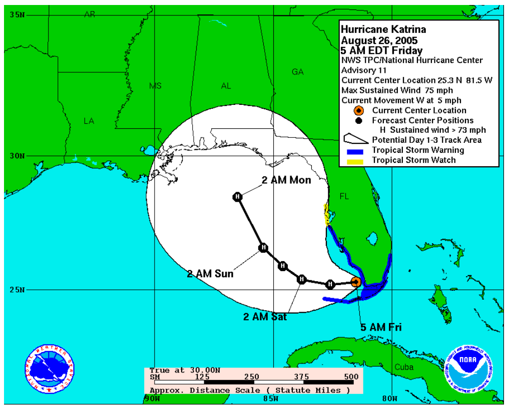
SYLLABUS

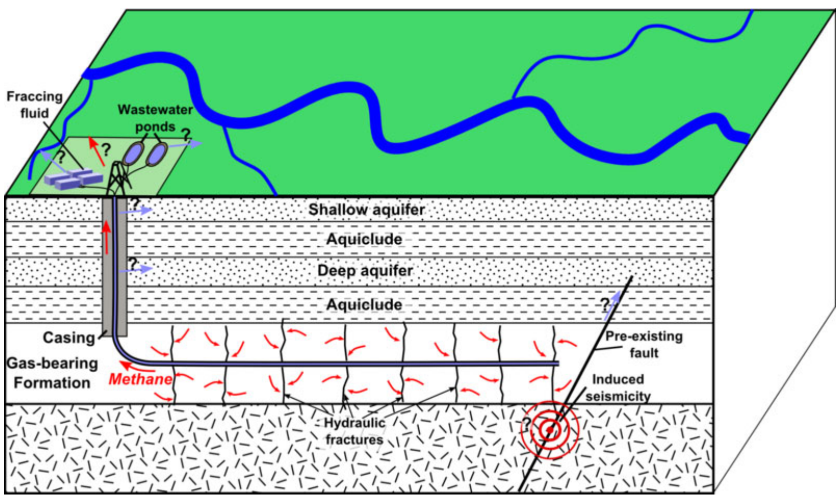
1. Introduction to UQ with examples
2. ~~Fundamentals of probability~~
3. Representation of random inputs
4. ~~Bayesian view of parameter estimation~~
5. Uncertainty propagation in models
6. Applications

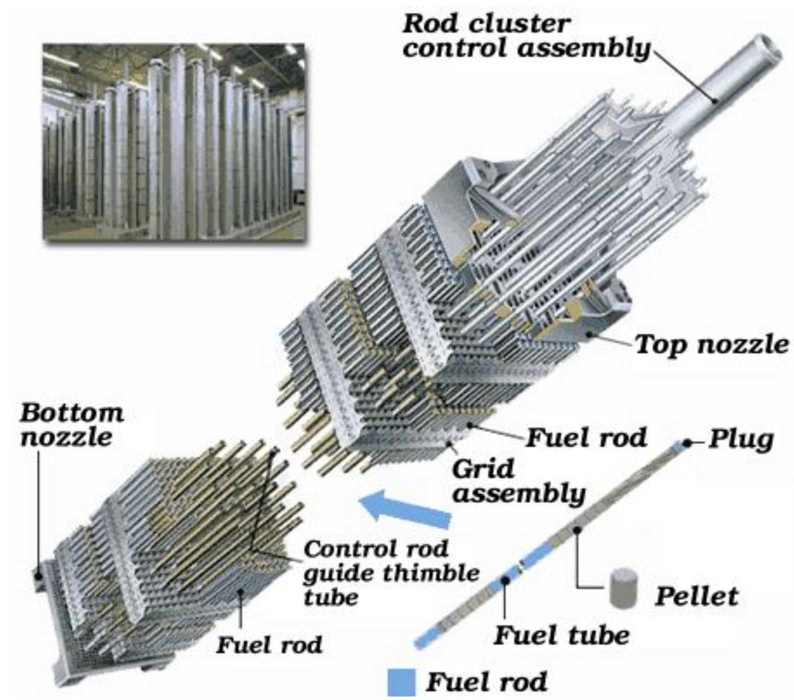
Lecture 1: Motivation and Prototypical Examples

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

George E.P. Box, Industrial Statistician



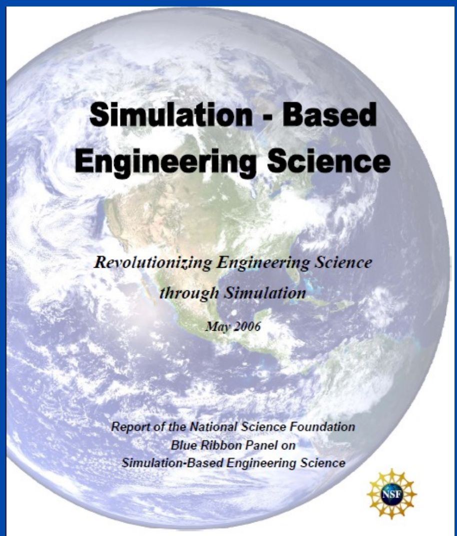




Simulation – Based Engineering Science

"… Can computer predictions be used as a reliable tool bases for crucial decisions? How can one assess the accuracy or validity of a computer prediction? What confidence can be assigned to a computer prediction of a complex event ? "

Ivo Babuska and J.



If an industry is to replace testing with simulation, the simulation tools must undergo robust verification and validation procedures for effectiveness.

Simulation has become indispensable in predictive methods for weather, climate change, and behavior of the atmosphere; and in broad areas of engineering analysis and design.

Why Models in Engineering?

1.Quality Improvement –explore more designs, eliminate early physical model uncertainty, explore variation effects.

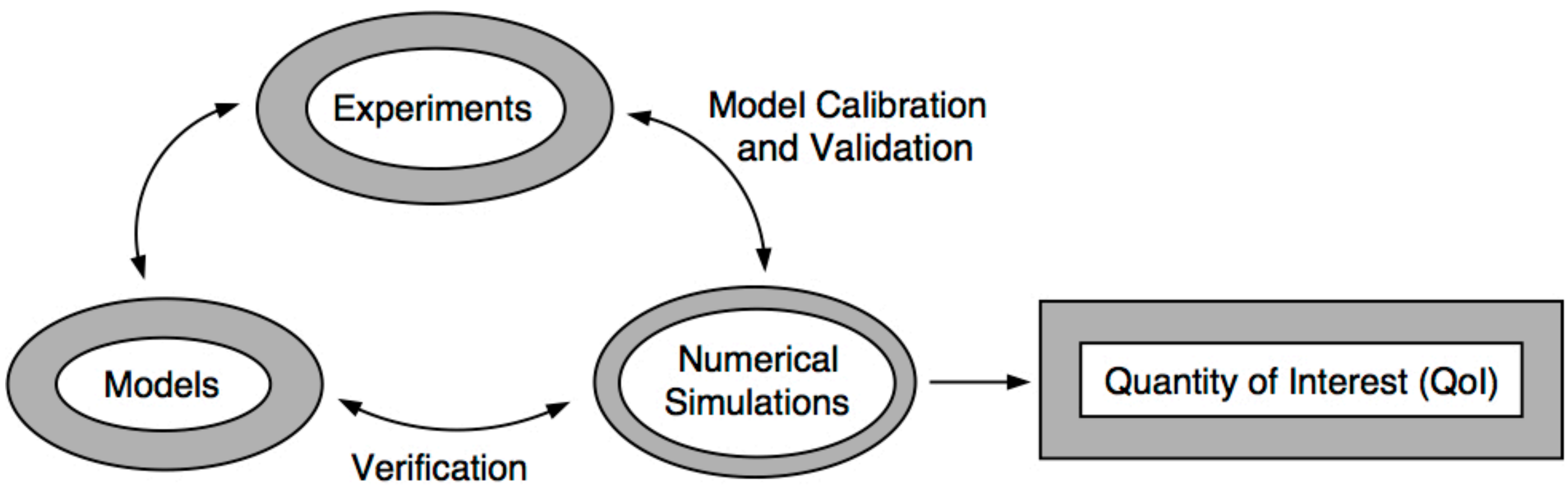
2.Time Constraints/Demand –shorten time from Concept to Showroom

3.Capital Reduction –minimal physical costs for iteration, hardware reduction

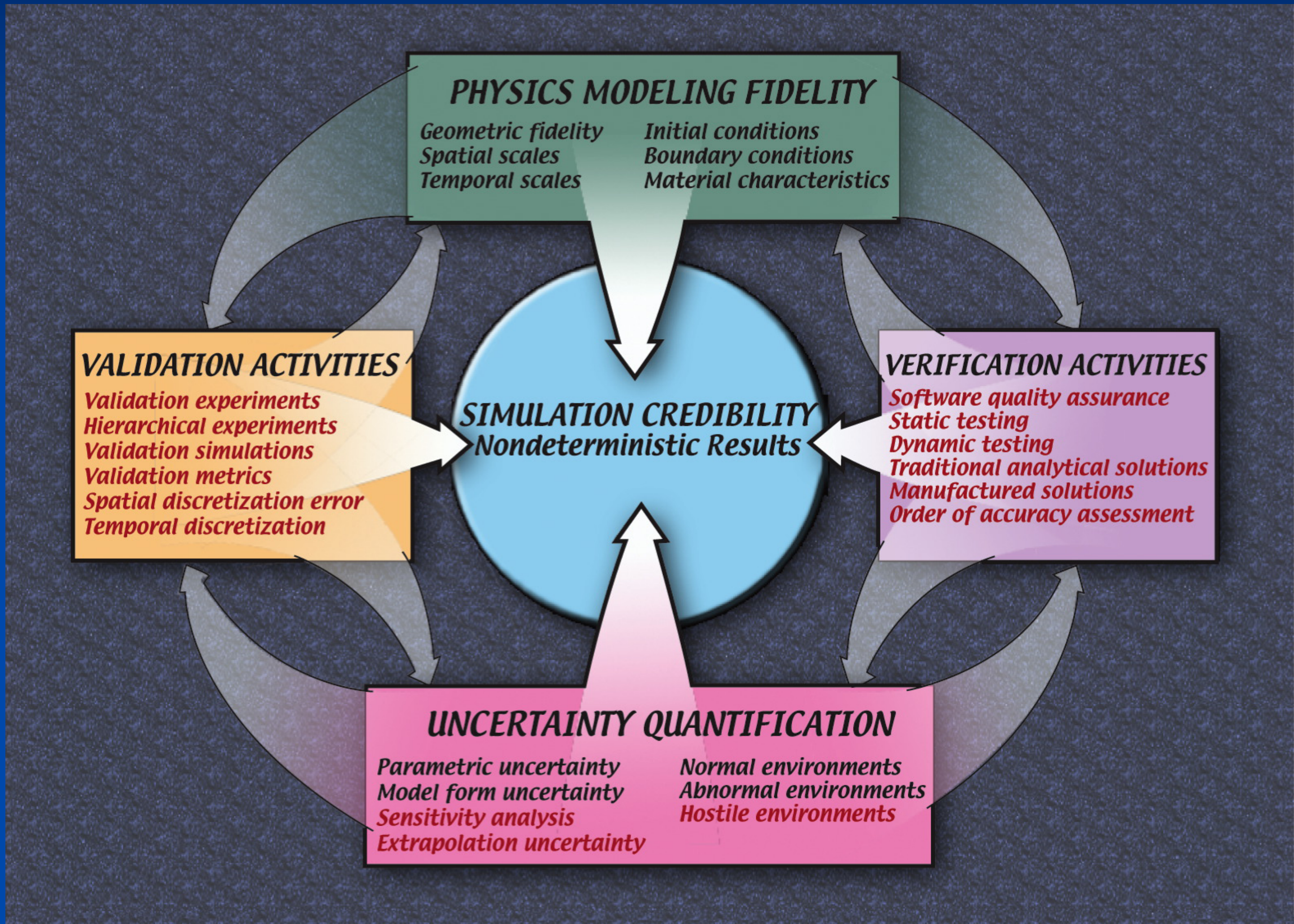
Predictive Science

Components: All involve uncertainty

- Experiments
- Models
- Numerical Simulation
- Quantity of Interest: usually a statistical quantity; e.g., mean



Confiabilidade de Predição Computacional



The Relationship Between Experiments and Simulations is Changing …

• *Old paradigm: Experiments are qualification tests, proof that something does or does not "break". Simulations are used to understand the behavior (generally, after the fact).*

• *New paradigm: Experiments explore the mechanics and validate predictions. Simulations are used to predict, with quantifiable confidence and across the operational space.*

• *Objective: Make decisions based on the predictions from validated science-based simulations. Validation requires an assessment of the sources of uncertainty (including lack-of-knowledge) and their effects on the predictions and decisions.*

General Principles – Science-Based Prediction

- Represent the geometry with the highest possible degree of fidelity.

- Implement models based on "first-principle physics" to describe the materials, initial conditions, boundary conditions, and forcing functions.

- Develop algorithms and numerical solvers based on first-principle physics as opposed to approximations.

- Couple multiple physics and/or multi length-scales.

- Bridge the gap between the continuum (macroscopic) laws of conservation, kinetic theory, molecular dynamics, and general / quantum relativistic physics.

- Propagate the sources of uncertainty (variability, lackof- knowledge) through numerical simulations.

Why is it Difficult?

• Science-based prediction has been enabled by recent advances (within the last 20 years) in programming languages, computer platforms, first-principle physics modeling, numerical algorithms, and visualization.

• The difficulty is therefore shifting from being able to perform complex simulations to validating the models and assessing the degree of credibility of predictions.

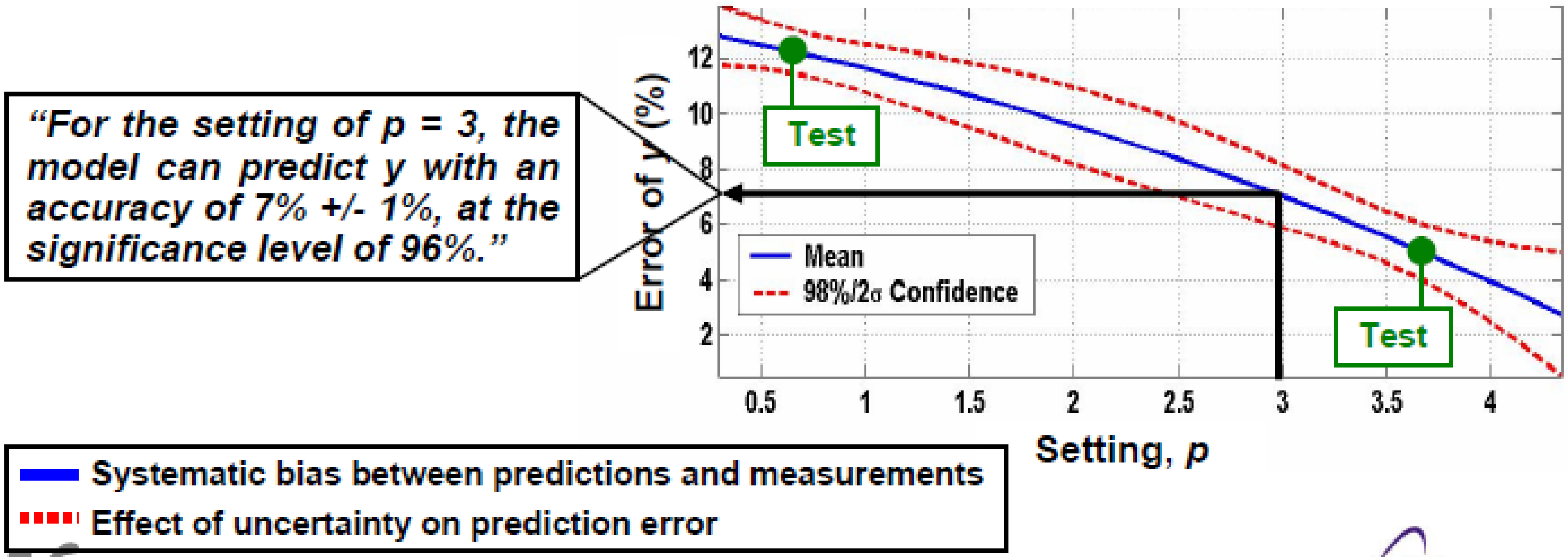
• Science-based prediction is difficult because of our lack of understanding of the sources of variability, uncertainty, lack of knowledge basic phenomena, and their effect on predictions.

• Examples: Material damping, crack growth, radiation hydro-dynamics, weather prediction, pulsars, etc.

Francois Hemez - Engineering Institute Spring 2006 - Los Alamos

What Does it Mean to be *"Predictive"*?

- Models and their predictions are validated when, in addition to assessing fidelity-to-data, the effect on predictions of our lack-of-knowledge is quantified.
- What is very challenging is to assess the effect that all the uncertainty (all of it, not just parameter variability) has on predictions.



What is Model Validation?

- *"The process of determining the degree to which a computer simulation is an accurate representation of the real world, from the perspective of the intended uses of the model."*

—DoD Modeling and Simulation
—DoE ASC Program

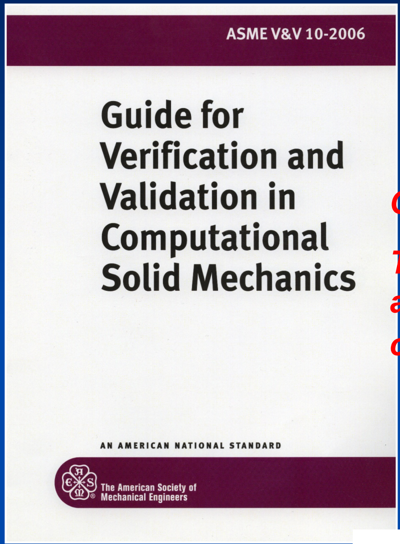
- *"Solving the right equations."*

—Roache (1998)

- *"The substantiation that a model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended applications of the model."*

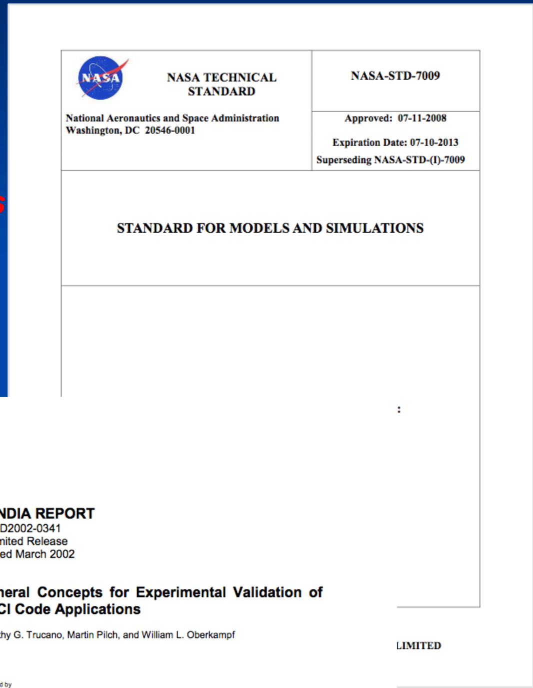
—Schlesinger (1979)

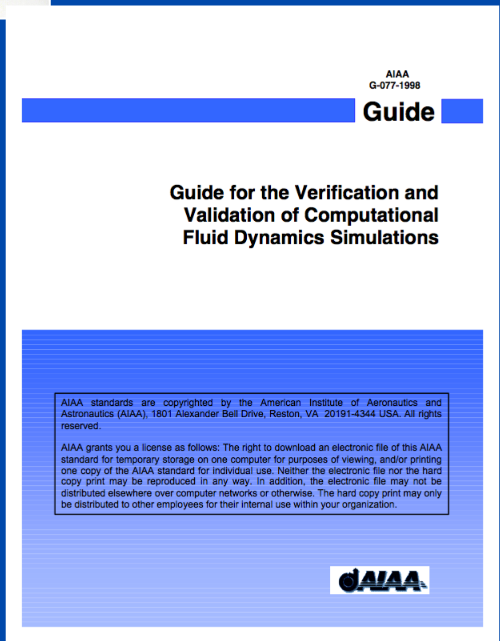
Normas e Guias de Recomendacao

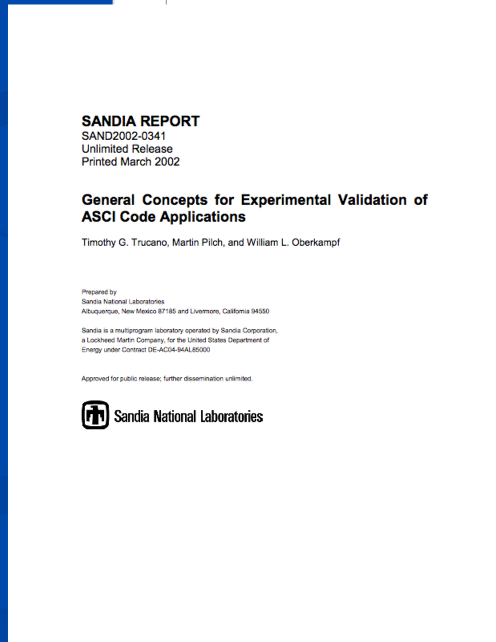


Goal (from ASME Guide to V&V, 2006)

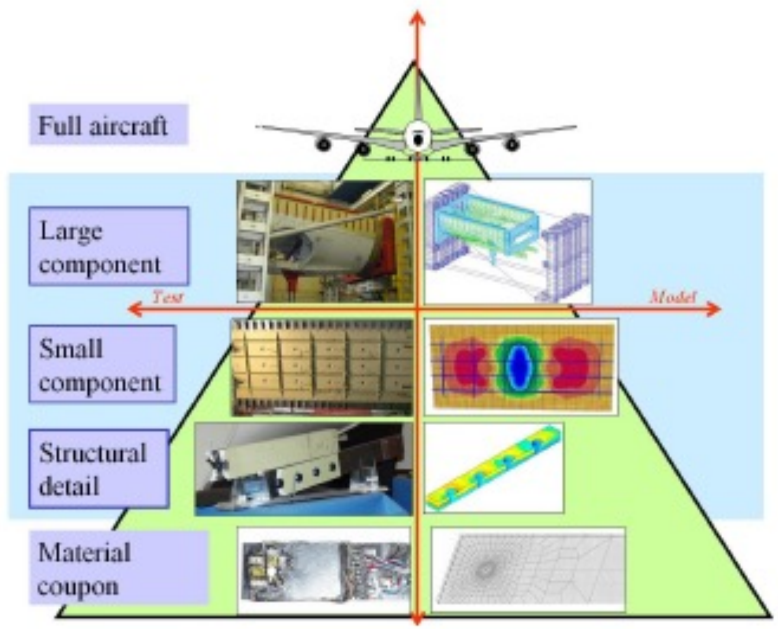
To develop standards for assessing the correctness and credibility of modeling and simulation in computational science.







Validation : a hierarchical approach



[Courtesy of Mr. S. Guinard, EADS Corporate Research]

→ Acreditarão

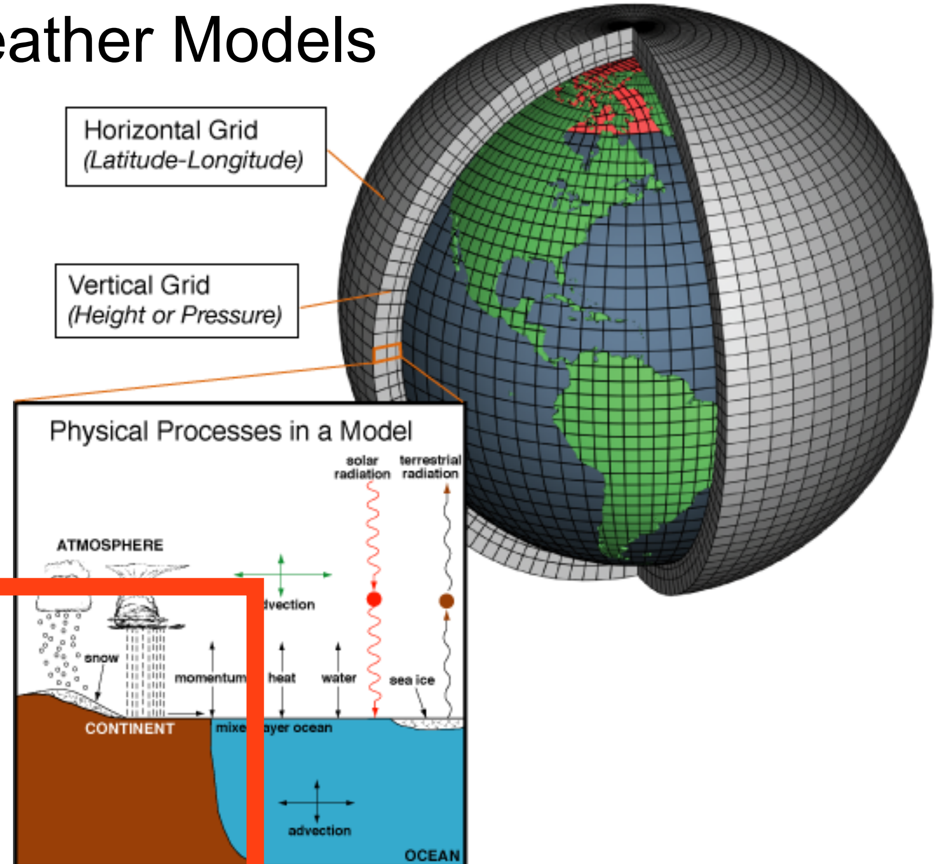
→ Validação

→ Calibração

Example: Weather Models

Physical Processes:

- Temperature
- Precipitation
- Winds
- Chemistry of aerosol species



Equations of Atmospheric Physics:

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

“ First Principles” - Certain

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

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

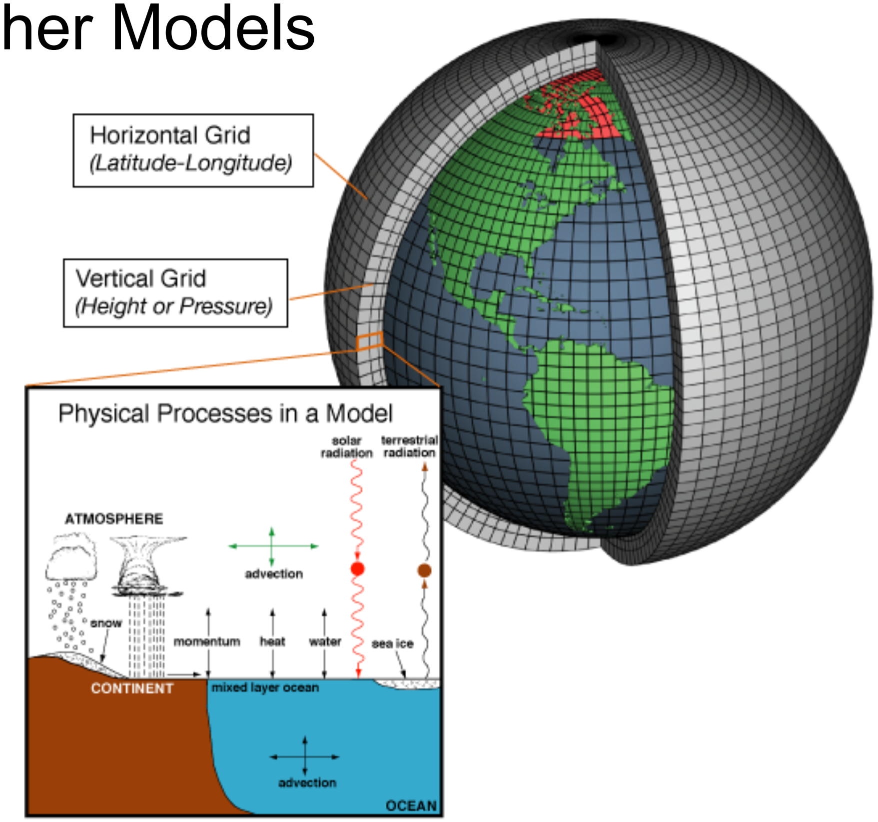
$$p = \rho R T$$

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

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

Closure (phenomenological) equations - different degrees of uncertainty

Example: Weather Models



Equations of Atmospheric Physics:

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

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

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

$$p = \rho R T$$

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

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

Phenomenological Model for Sources:

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

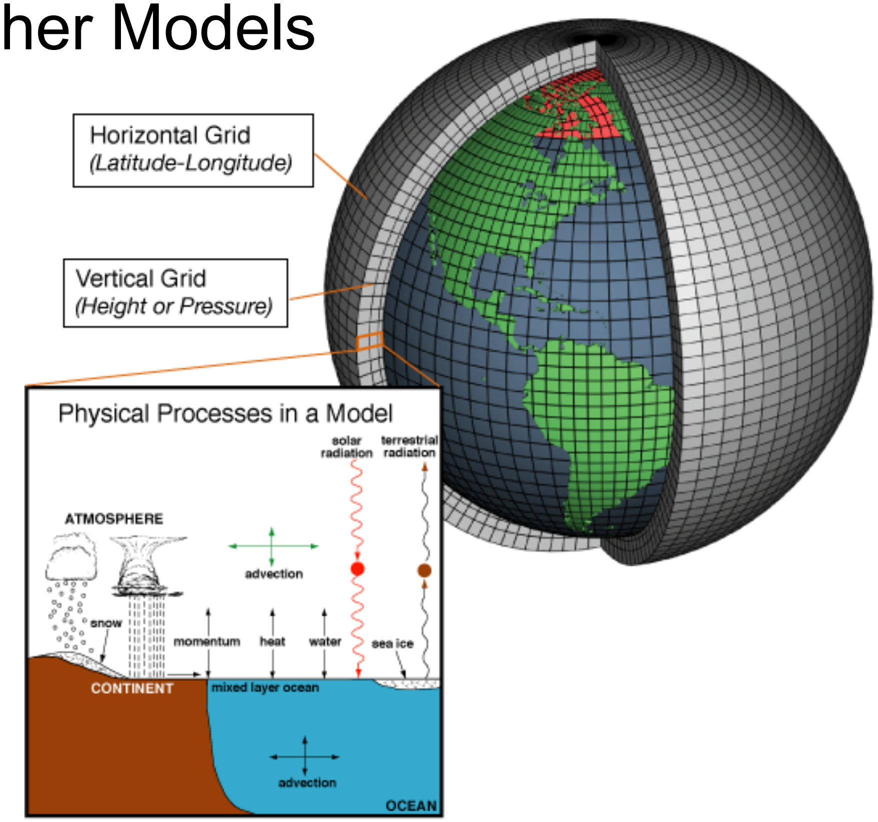
where

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

Example: Weather Models

Sources of Uncertainty:

- Model errors or discrepancies
- Input uncertainties
- Numerical errors and uncertainties
- Measurement errors and uncertainties

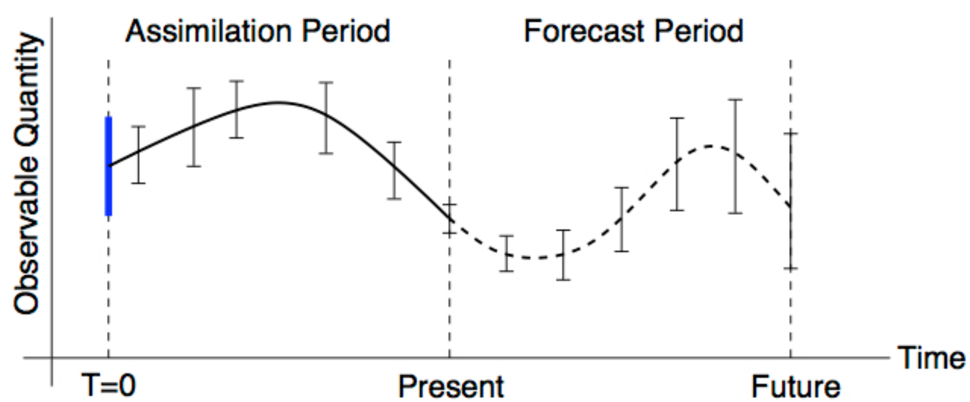


Steps:

• Model Calibration: Involves the assimilation or integration of data to quantify and update input uncertainties.

• Model Prediction: Here one computes the response along with statistics, error bounds, or PDF; extrapolation is important and difficult.

• Estimation of the Validation Regime:

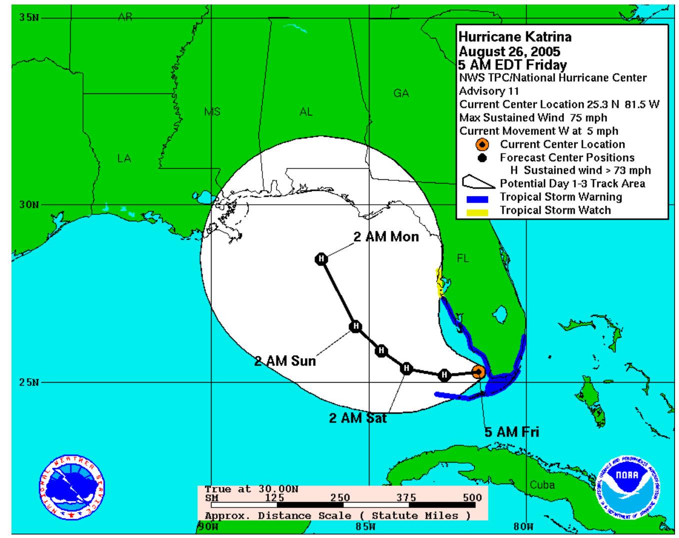


•**Goal:** Construct best estimate parameters and responses or quantities of interest with best estimate reduced uncertainties.

Example: Weather Models

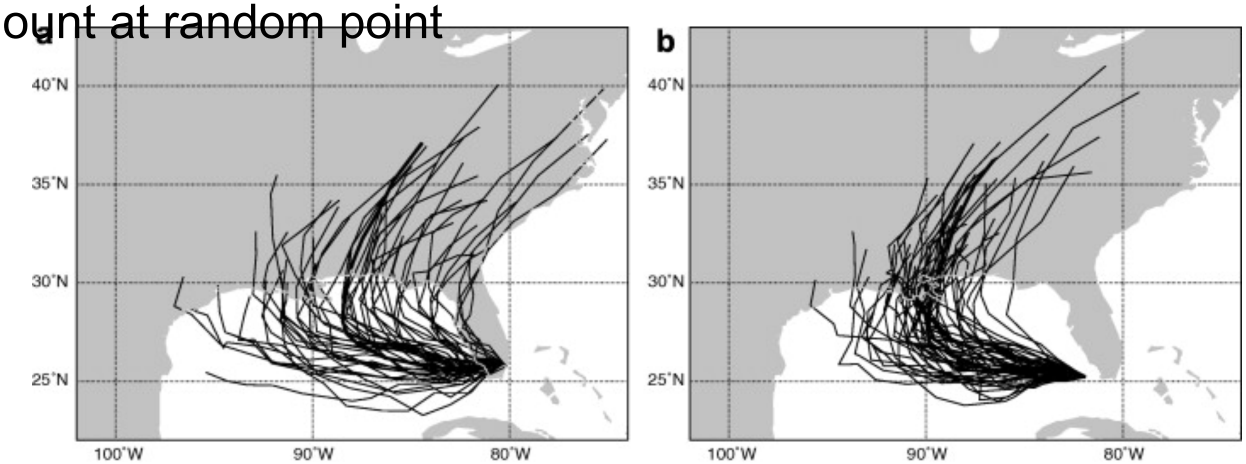
Sources of Uncertainty:

- Model errors or discrepancies
- Input uncertainties
- Numerical errors and uncertainties
- Measurement errors and uncertainties

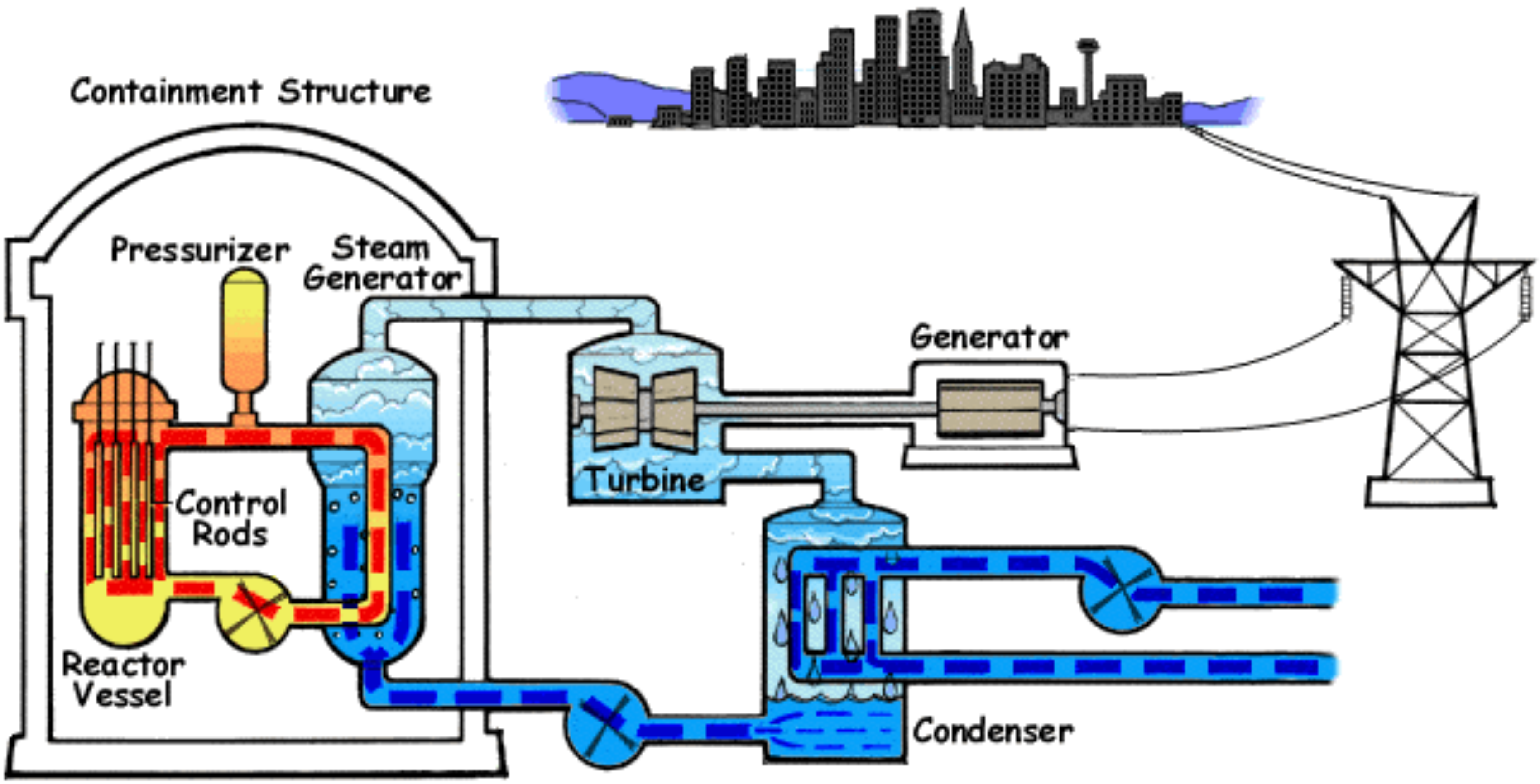


Ensemble Forecasts:

- Run multiple simulations with differing parameter values or initial conditions drawn from appropriate pdf.
- A 50% chance of rain means that given present atmospheric conditions, half of simulations predict measurable rain amount at random point in specified area.



Example 2: Pressurized Water Reactors (PWR)



Models:

- Involve neutron transport, thermal-hydraulics, chemistry
- Inherently multi-scale, multi-physics

Example 2: Pressurized Water Reactors (PWR)

3-D Neutron Transport Equations:

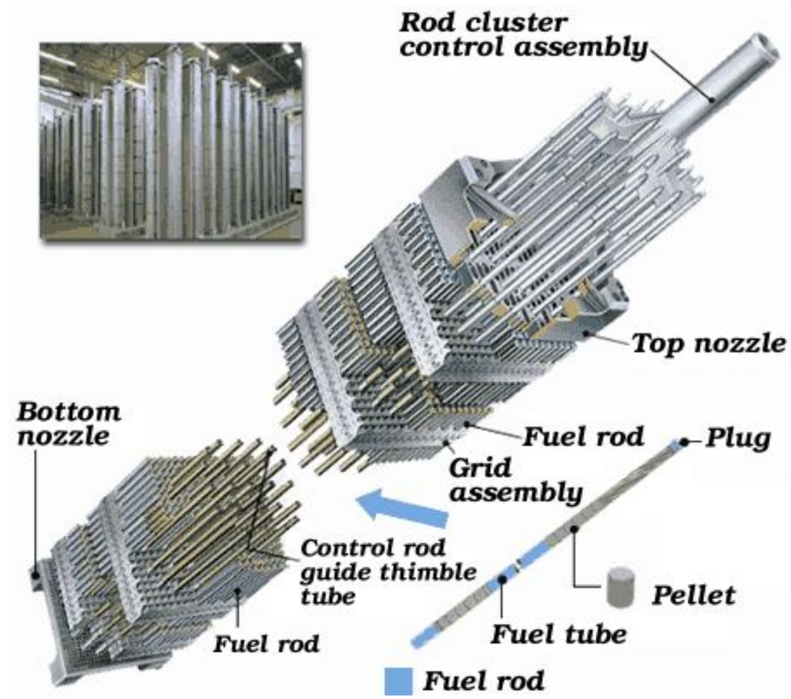
$$\frac{1}{|v|}\frac{\partial \varphi}{\partial t} + \Omega \cdot \nabla \varphi + \Sigma_t(r,E)\varphi(r,E,\Omega,t)$$
$$= \int_{4\pi} d\Omega' \int_0^\infty dE' \Sigma_s(E' \to E, \Omega' \to \Omega)\varphi(r,E',\Omega',t)$$
$$+ \frac{\chi(E)}{4\pi}\int_{4\pi} d\Omega' \int_0^\infty dE' \nu(E')\Sigma_f(E')\varphi(r,E',\Omega',t)$$

Challenges:

- Linear in the state but function of 7 independent variables:

  $$r = x, y, z; E; \Omega = \theta, \phi; t$$

- Very large number of inputs or parameters; e.g., 100,000
- ORNL Code: Denovo;
- Codes can take hours to days to run.



Example 2: Pressurized Water Reactors (PWR)

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

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

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

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

Note: Similar equations for gas

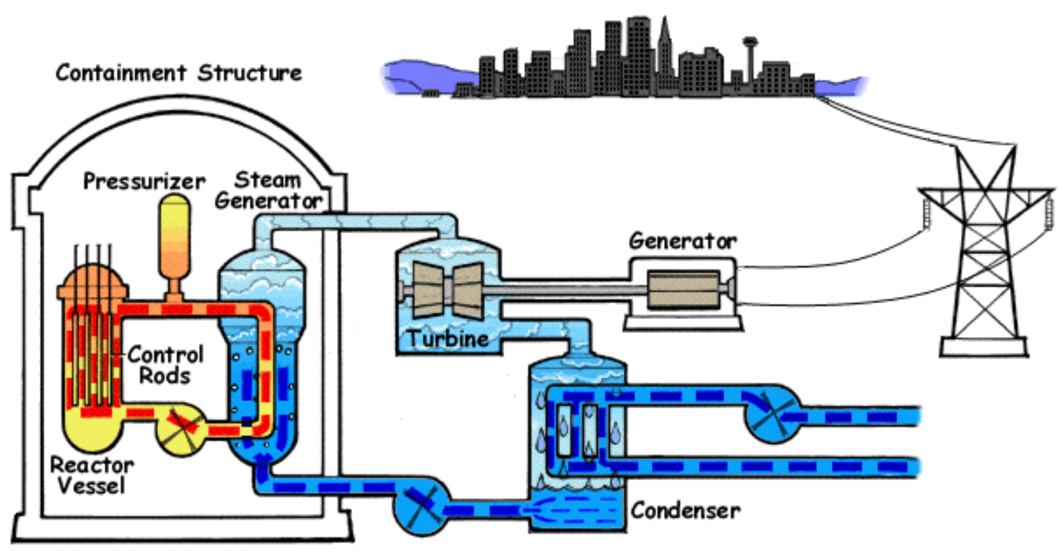
Challenges:

- Nonlinear coupled PDE with nonphysical parameters due to closure relations;
- CASL code: COBRA-TF – Difficult to access primary parameters and inputs.
- Codes can take minutes to days to run.

Example 2: Pressurized Water Reactors (PWR)

UQ Challenges:

- Specify bounds on void fraction distributions and boiling transitions that guarantee specified performance levels and safety margins.
- Specify conditions that limit CRUD on the outside of fuel cladding to within prescribed levels.
- Determine new cladding materials, fuel materials, and fuel pin geometries that provide an average specified improvement in performance and increased resistance to damage.
- Determine conditions that produce specified levels of radiation damage, mechanical thermal fatigue, and corrosion.



Example 3: HIV Model for Characterization and Control Regimes

HIV Model:

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

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

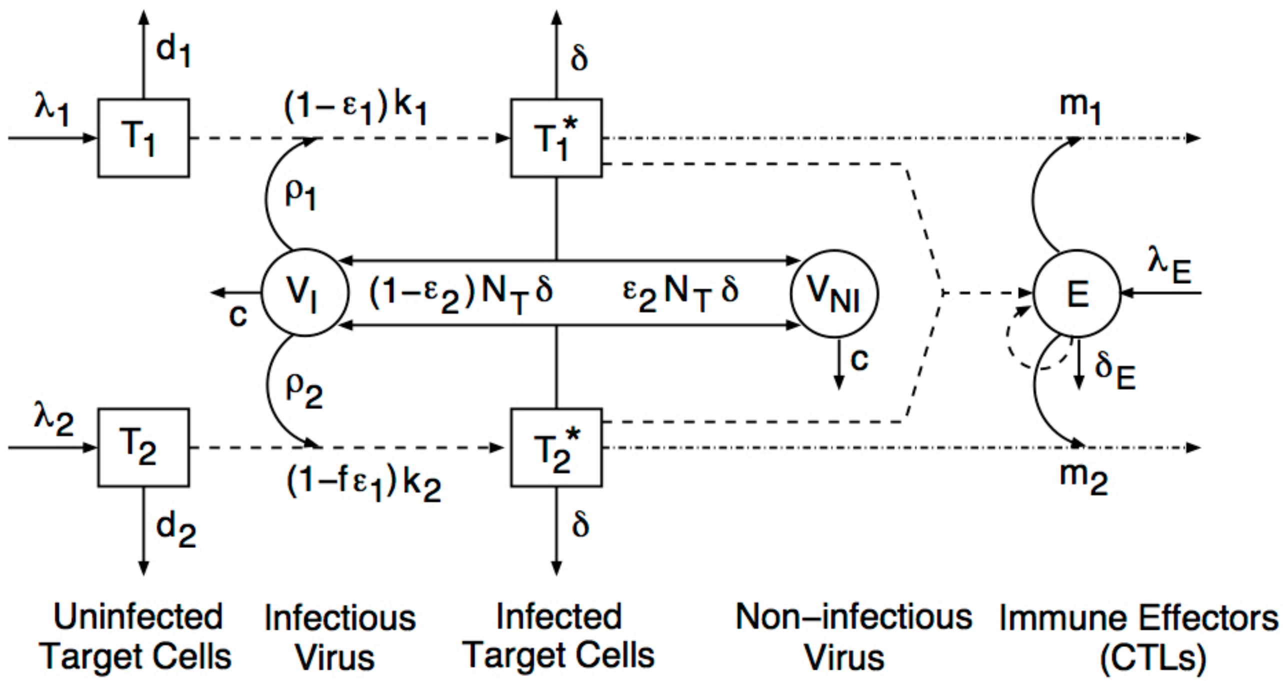
$$\dot{T}_1^* = (1-\varepsilon) k_1 V T_1 - \delta T_1^* - m_1 E T_1^*$$

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

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

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

Compartments:



Example 3: HIV Model for Characterization and Control Regimes

HIV Model: Used for characterization and control treatment regimes.

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

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

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

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

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

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

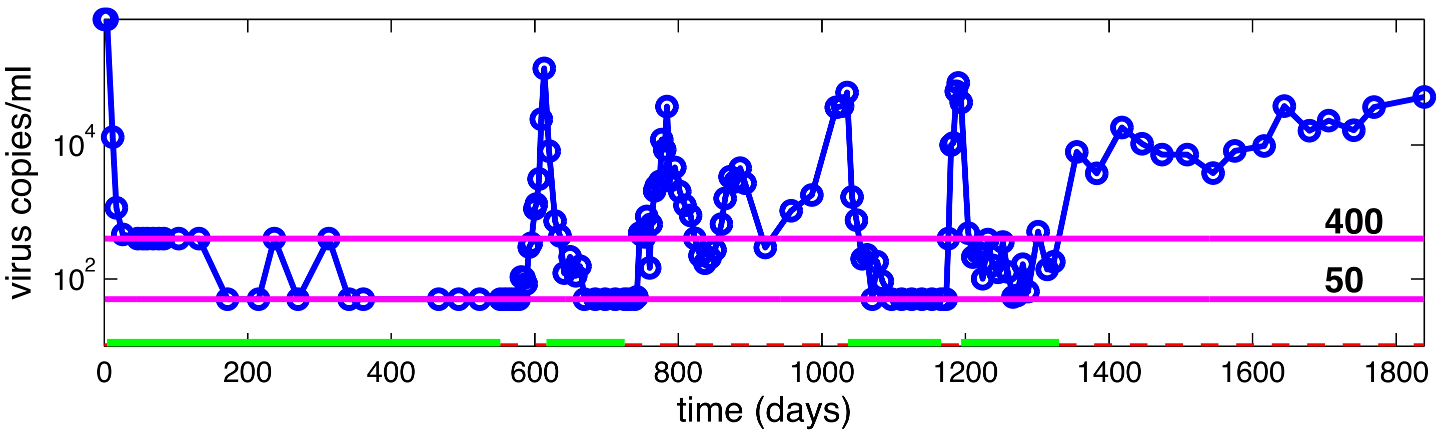
Parameters: Most are unknown and must be estimated from data

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 3: HIV Model for Characterization and Control Regimes

HIV Model: Several sources of uncertainty including viral measurement techniques

Example: Upper and lower limits to assay sensitivity



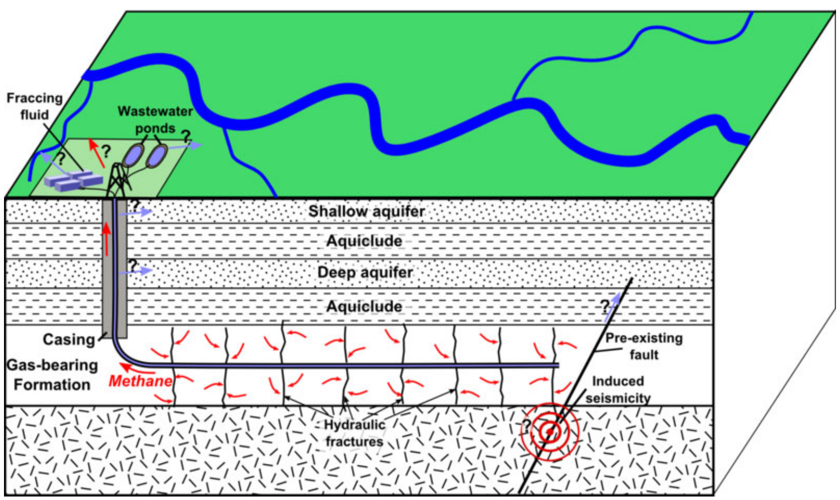
Experimental Uncertainties and Limitations

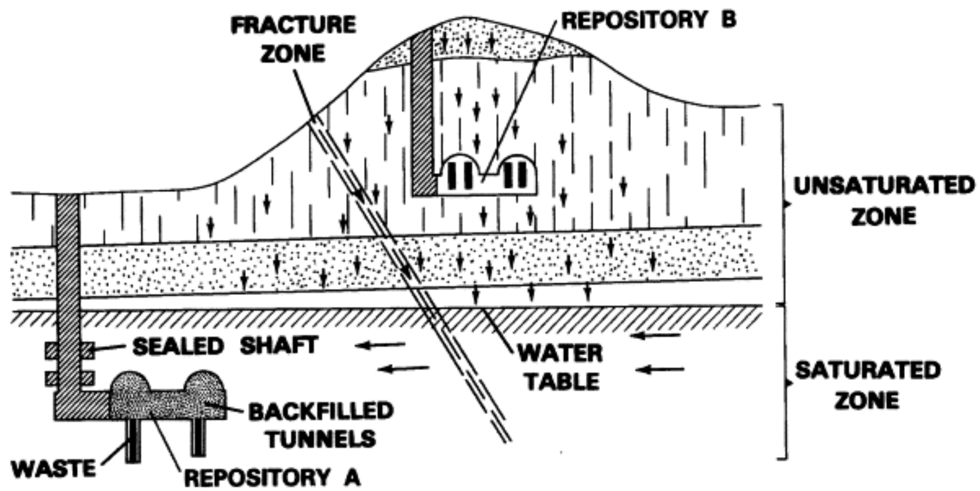
Examples: *Experimental results are believed by everyone, except for the person who ran the experiment*, Max Gunzburger, Florida State University.

• 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.



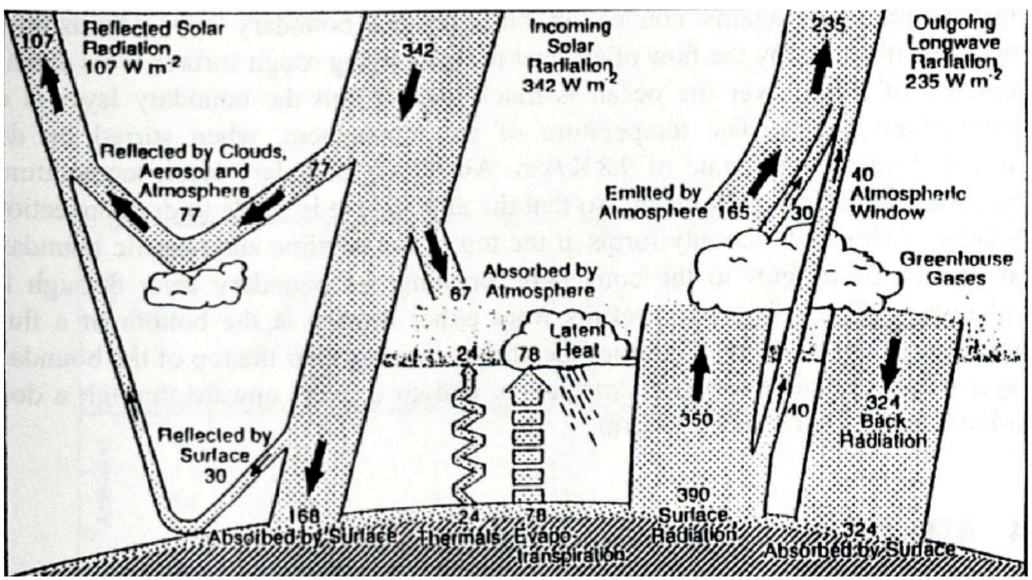


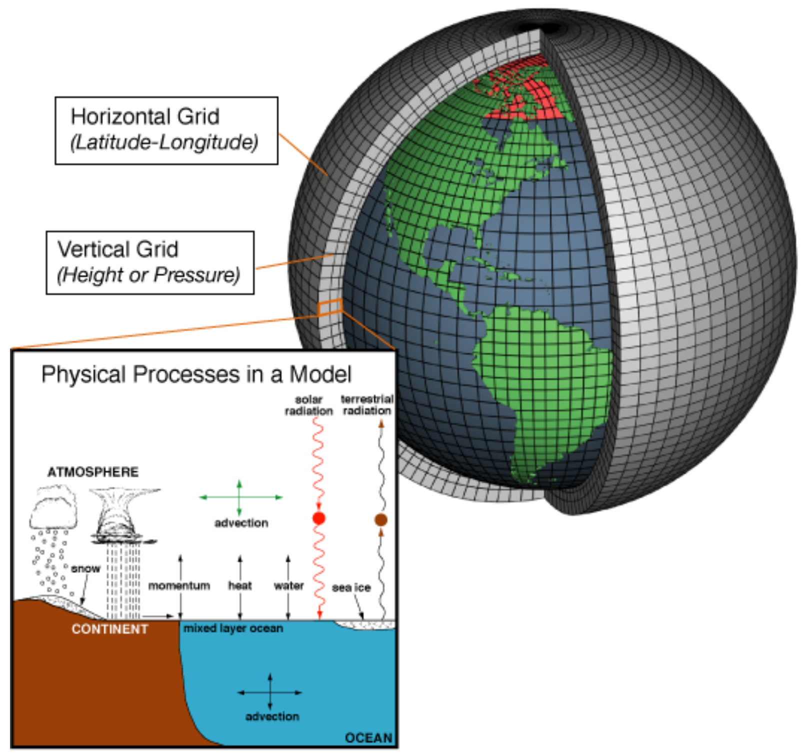
Model Errors

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.



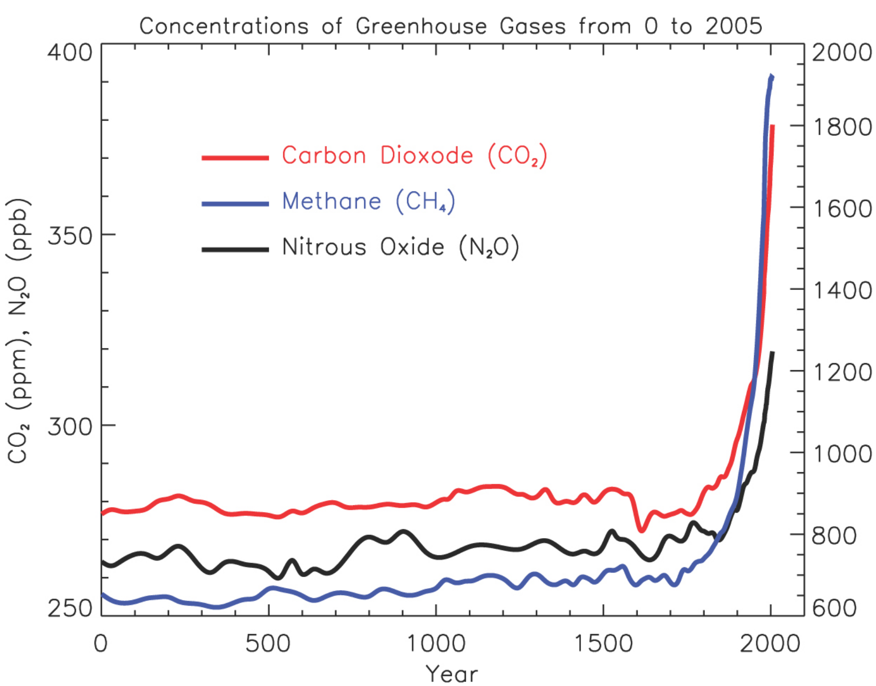


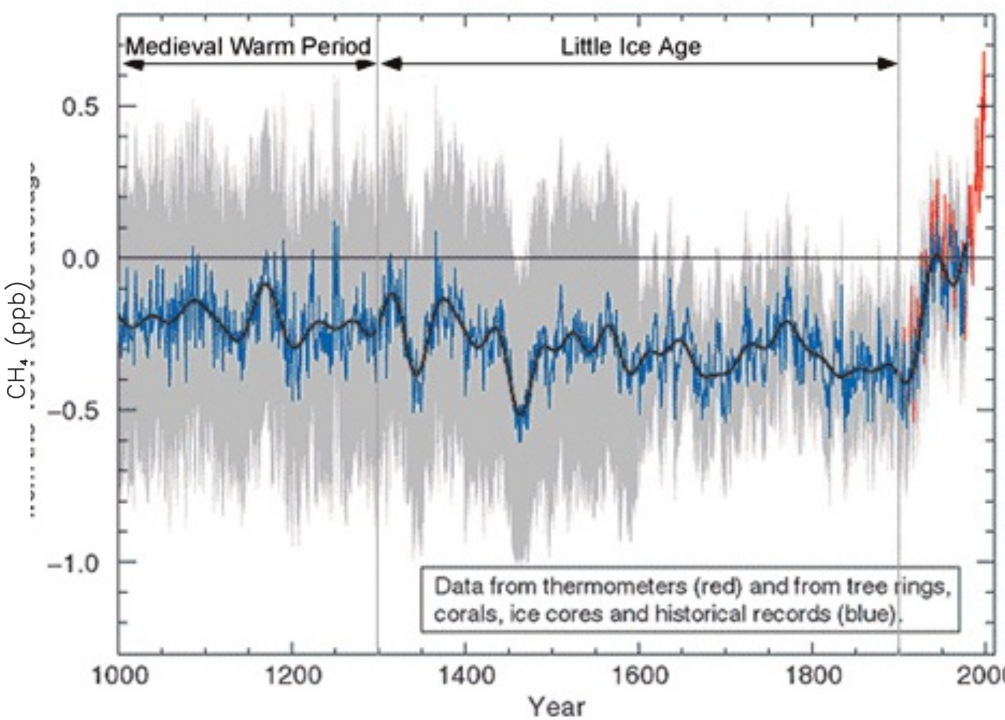
Input Uncertainties

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

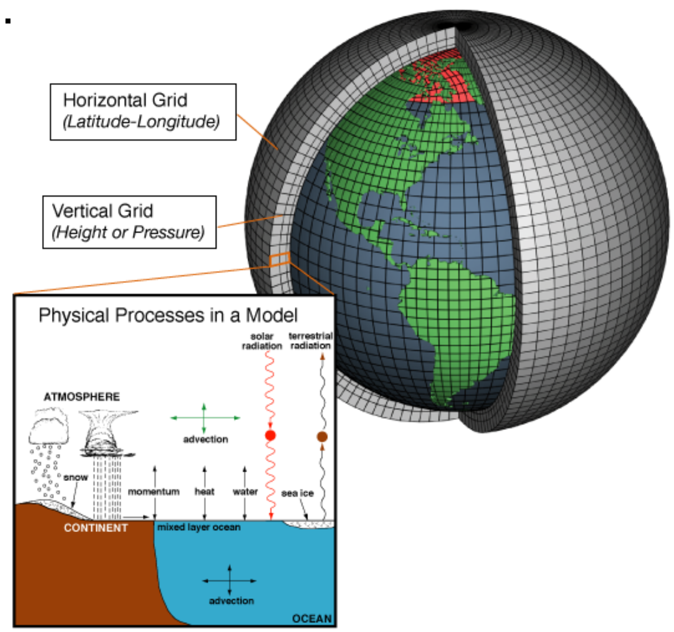
Note: *Computational results are believed by no one, except the person who wrote the code*, Max Gunzburger, Florida State University.

• 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;

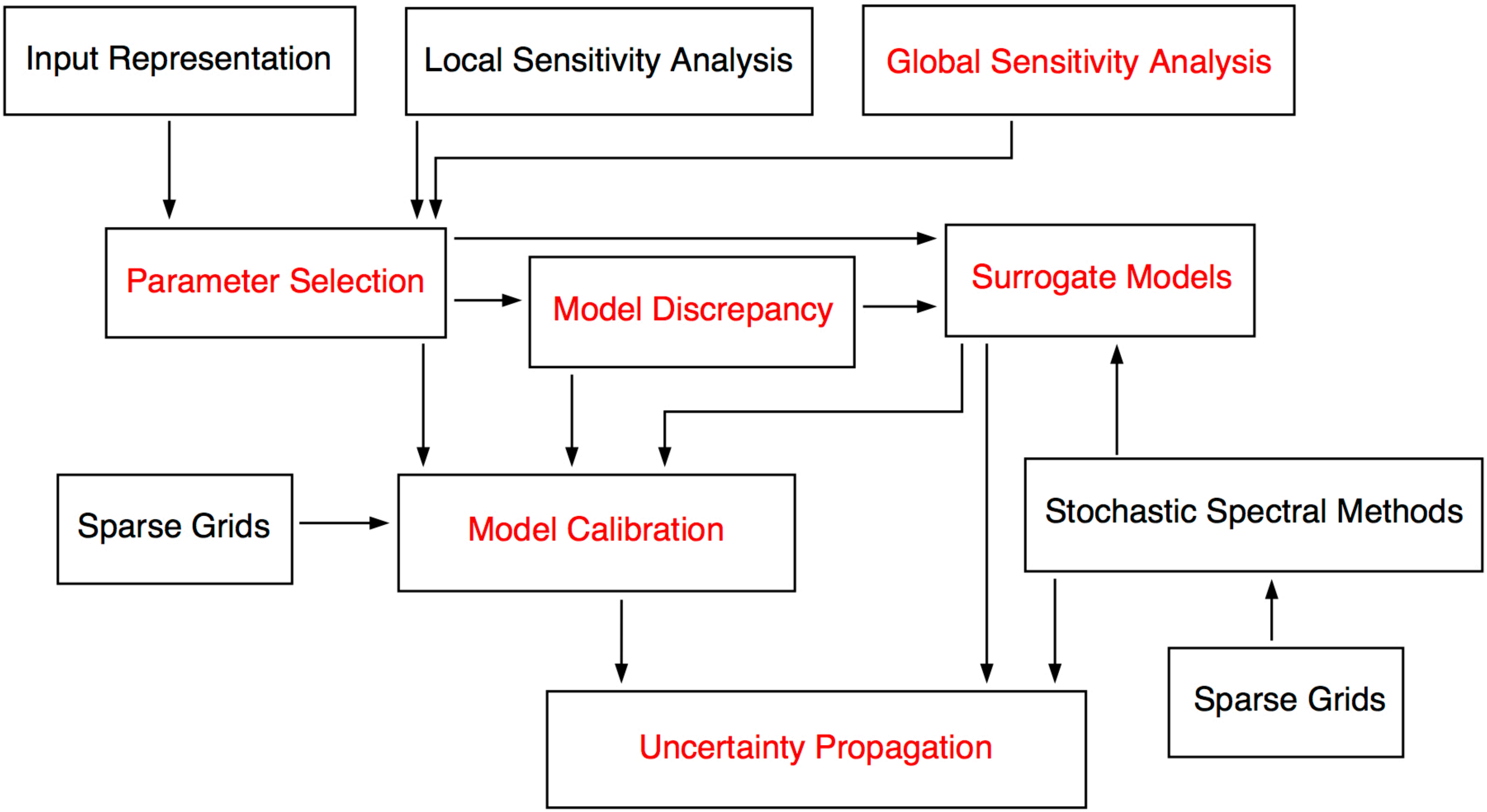
• Bit-flipping, hardware failures and uncertainty associated with future exascale and quantum computing;

• 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).

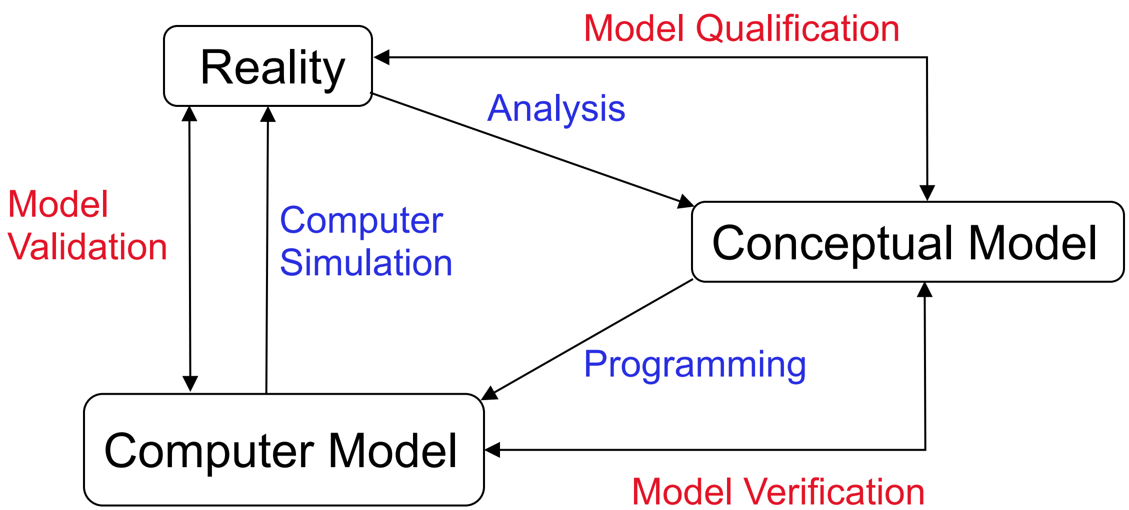


Steps in Uncertainty Quantification

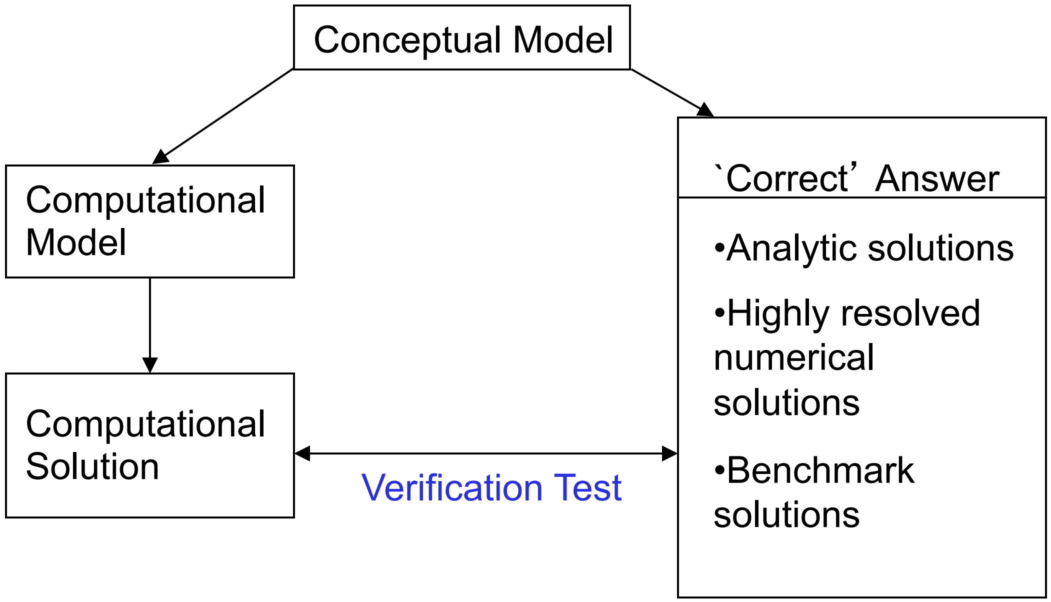
Note: Uncertainty quantification requires synergy between statistics, mathematics and application area.



Modeling Issues



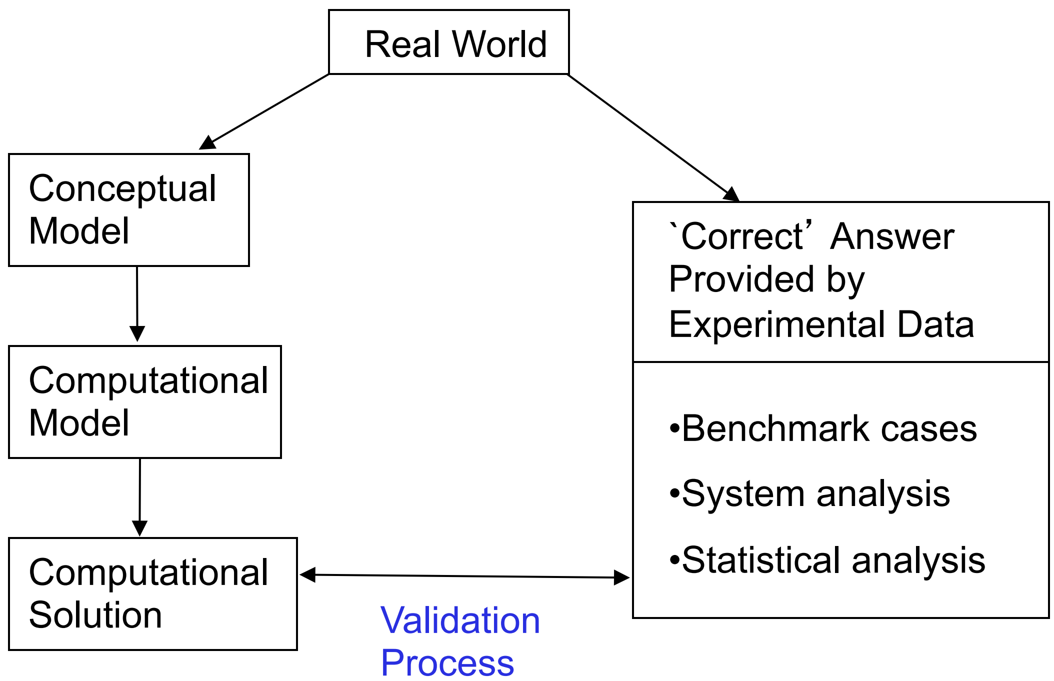
Verification Process



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

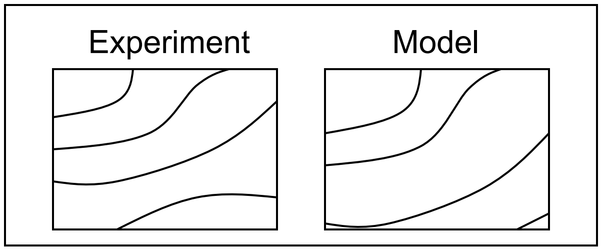
Validation Process



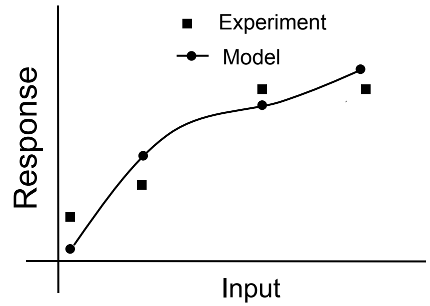
Validation: The process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended model users.

Note: Validation deals with physics and statistics

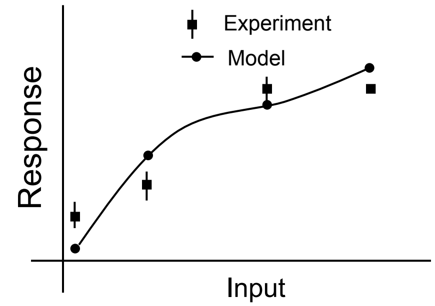
Validation Metrics



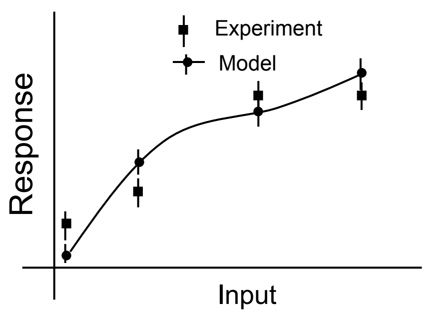
\`Viewgraph’ Norm



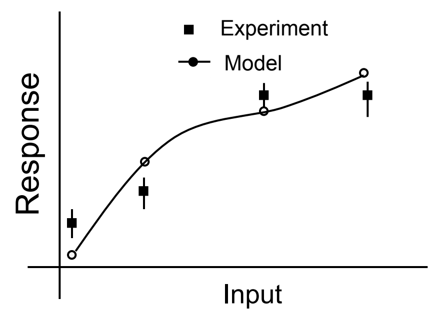
Deterministic



Experimental Uncertainty



Numerical Error



Nondeterministic Computation