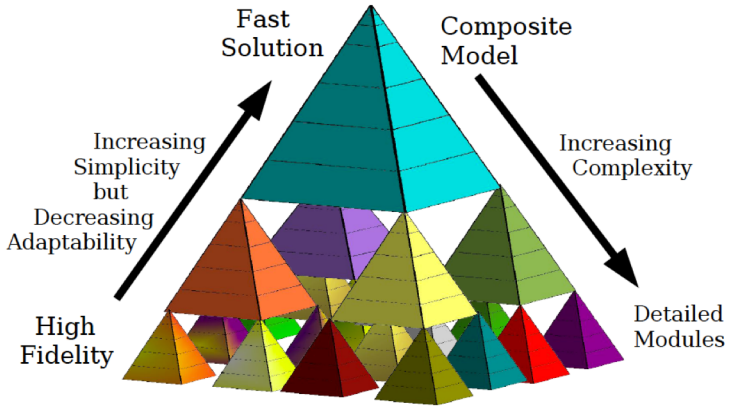
Challenge 5 Raj Vadegepalli / Jim Bassingthwaighte

Reproducible and reusable multiscale models that will be integrated and adopted into model-poor fields (e.g. tissue engineering, regenerative medicine, drug and gene delivery, preventive interventions).

- BREAK THIS INTO Parts to think about it:
- Part 1: A. Define reproducible and reusable modeling (and related terms)
- B. Define the characteristic requirements for computational models: These should be general, not field dependent.
- C. Define A and B for Modules. Single level. Define modularity in technical terms that will permit automation of model construction from modules. Develop examples of automated modular construction from a repertoire of modules. (A function might be captured in a module at three levels, e.g. i. computational simple and fast, ii. a mid-level module with constrained generality, iii. a thermodynamically and chemically detailed, accurate reference "working module" for the field.
- Modules might defined so they only reveal a few variables, the connectors to the system, and a minimal number of variable parameters, leaving the bulk of its innards more or less invisible to the users. Alternatively, keep everything available.

Challenge 5 continued



- Part 2A: Can some ”modules” be multiscale?
- Sometimes: e.g. channels may be multi-level, with gating currents distinct from channel conductance components. An independent module for a currents is not influenced directly by events at other channels or in metabolism.
- Mostly not: the profuse interconnectivity of biochemical networks preclude their isolation, e.g. the TCA cycle is connected in multiple ways at every enzyme and substrate. With just the aspartate-glutamate shuttle the top and bottom of TCA have differing fluxes.
- Part 2B: Can a multiscale model be reconstructed from preconstructed modules at a single keystroke? What does it take to substitute a module?
- If one has for each module of a multiscale model several versions of differing “truthiness” or “speediness”, could one have automated Monte Carlo selection and reconstruct/run 1000 solutions to test for ”structural” uncertainty?

Definition of terms re Repeatability, Replicability, Reproducibility:

- Working Conditions for ***Replicability***:
- Replicability (Different team, same experimental setup.) The measurement can be obtained with stated precision by a different team using the same measurement procedure, the same measuring system, under the same operating conditions, in the same or a different location on multiple trials. An independent group can obtain the same or closely similar result.
- In brief: Independent team, same methods, code, different computational setting, same results. Similar environment.

Definition of terms re Repeatability, Replicability, Reproducibility:

- Working definition for reproducibility while assessing the same problem with "total" independence, starting from the same hypothesis.
- Reproducibility: Different team, different setup. The same output result can be obtained within stated precision by a different team, a different measuring system, in a different location on multiple trials. An independent group can obtain the same result using artifacts which they develop completely independently.
- In brief: Reproducibility implies obtaining same results by different Code developed from the same equations or from the original hypotheses in similar or different language, different solvers and optimizers. There are two levels:
  - (1) Same equations, independent implementations, languages, etc.
  - (2) Same scientific hypothesis, develop the equations independently. Provides assessment of clarity and independent redevelopment, and therefore strong evidence of completeness and accuracy.

A comment on terminology re Repeatability, Replicability, Reproducibility

- These terms have strongly overlapping normal parlance, and we might do better figuring out another set of three terms that start with different letters, or alternatively, use a prefatory term. For example:
  - 1. Repetition variance (e.g. like practicing pipetting -> mean ±SD)
  - 2. Translocation replicability.
  - 3. "Square 1" reproducibility = independent bottom up redevelopment.
- While the exact terms chosen may not be unique, it seems important to use a common set of terms when trying to define standards. This is needed for archiving and classifying or grading what is archived.

A Replicable Package

- An example of Replicable Model Package:
  A JSim Project File to replicate a model plus the data analysis.
- The package, preserved in ascii, stores code, data sets, notes, runs one or several models, preserves the analyses.
  - 8 ODE, 3 PDE solvers, flexible graphing for model solutions and data
  - 8 optimizers for automated parameter adjustment for model fitting to data
  - sensitivity analysis
  - parameter confidence estimates and correlation matrix
  - Monte Carlo for parameter UQ evaluation.
- Numerical / optimization methods retained for each data/solution set.
- QUESTION for IMAG: How to define a platform-independent replication package?

Modeling Stages to VVUQ

Idea, from experience, data, etc.→ The basic hypothesis

Expressing hypothesis in logic and mathematics

Computable code

Code Verification proving that it represents the math

Model Validity testing on data (only after verification)

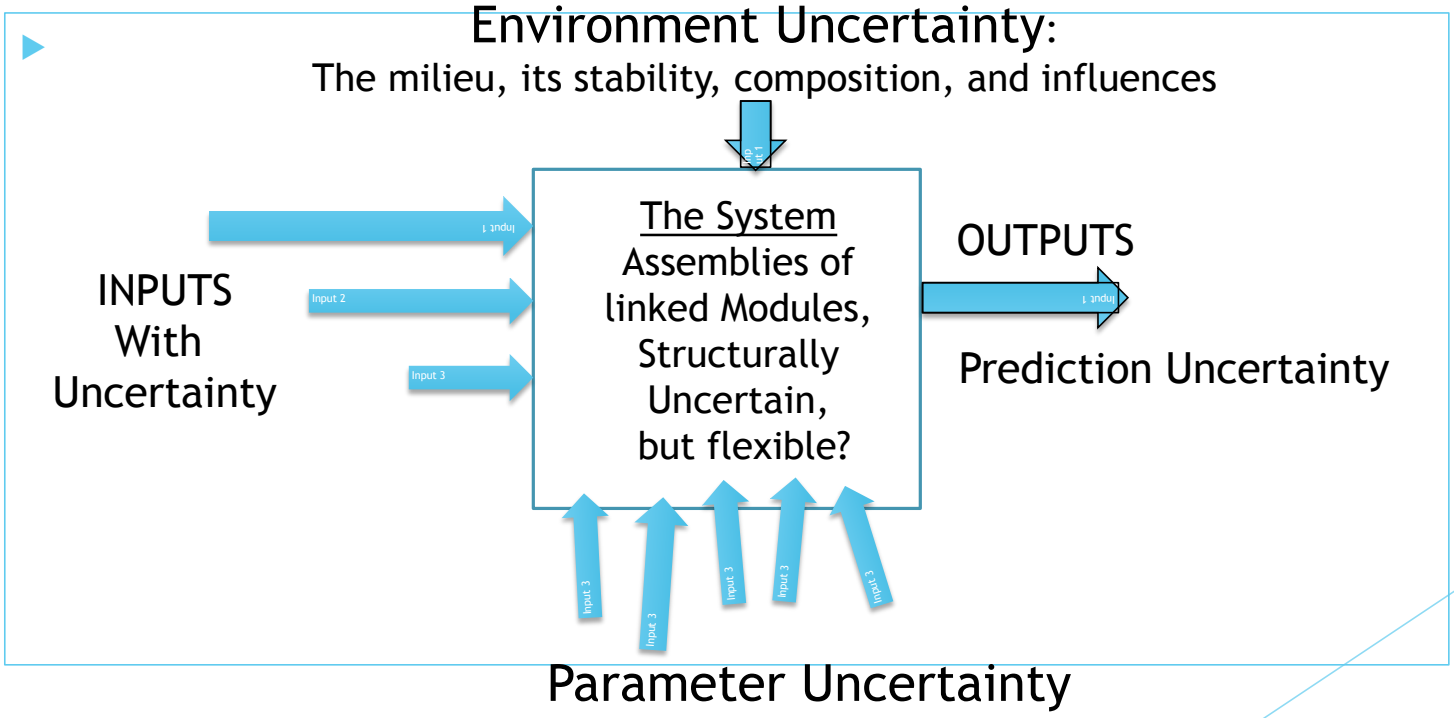
Uncertainty Quantification (how good is the validation)

Aim for *Uncertainty Quantification* as an end in itself for each model!

- Uncertainty types:
  - -parameter
  - -input
  - -environment
  - -structural

Model Uncertainty

Ascertaining OUTPUT UNCERTAINTY
Is the goal of predictive modeling



Definition of terms re Repeatability, Replicability, Reproducibility:

- Working conditions for ***repeatability***: (Same team, same experimental setup.) The measurement can be obtained with stated precision by the same team using the same measurement procedure, the same measuring system, under the same operating conditions, in the same location on multiple trials.
- In brief: Same procedures, operating conditions, location, and replicate measurements on the same objects over a short period of time. Provides estimates of error profile.