

Basic Ideas for Testing Models

Zoology 535, Spring 2008

Calibration is the process of fitting a model to data. We have talked a lot about formal fitting methods. Sometimes models are just "fit by eyeball". That can be OK, depending on the intended use of the model.

Uncertainty analysis is the process of determining how the distribution of outputs depends on the distribution of inputs. Suppose the model is our usual

$$Y_{t+1} = f(Y_t, X_t, \theta) + \varepsilon_t \quad [1]$$

perhaps with an observation model for Y and X. Note that f could be a very complicated "black box". The idea of uncertainty analysis is to measure the distribution of Y_{t+1} as determined by the variance of exogenous inputs X_t or parameters θ . For example if we are most interested in variance, using first-order error propagation

$$\begin{aligned} \text{var}(Y_{t+1} | \text{var}(X_t)) &= \left(\frac{\partial f}{\partial X_t} \right)^2 \text{var}(X_t) \\ \text{var}(Y_{t+1} | \text{var}(\theta)) &= \left(\frac{\partial f}{\partial \theta} \right)^2 \text{var}(\theta) \end{aligned} \quad [2]$$

Sensitivity analysis is the process of determining how the outputs change given standard perturbations of the inputs or parameters. The standard inputs are usually small (say 1% of the mean to 10% of the mean). In other words, sensitivity analysis measures $\partial f / \partial X$ or $\partial f / \partial \theta$. So it is a subset of, or first step toward, an uncertainty analysis.

Validation is not so clearly defined. In some cases, it means testing the model with a data set that is different from the one used to calibrate the model. Gardner and Urban (Chapter 10) suggest alternative terms. Adequacy is the extent to which the model output explains observed ecosystem dynamics. Reliability is the extent to which the model output lies within the range of known ecosystem behaviors.

Also consider:

Numerical experiments to ask (1) does the model respond to perturbation in ways that resemble ecosystem behavior? And (2) what is the range of inputs and parameters that give behavior that resembles ecosystems?

Boundaries of behavior: What level of stress will the model tolerate before its behavior changes?

List of potentially important attributes that are not represented in the model, but could be represented in a model.

Repeatability: Is the modeling study explained in enough detail for a knowledgeable scientist to repeat the study?