SENSITIVITY ANALYSIS

Andrew Ford
Program in Environmental Science and Regional Planning, Washington State University, USA

Keywords: computer simulation, sensitivity analysis, statistical analysis, sample design, Latin Hypercube Sampling, Tacuchi methods, confidence intervals, robust, leverage point, Kaibab model, deer, forage, irruption

Contents

Summary

The parameters in a system dynamics model are normally estimated one at a time based on all the information sources available. Some parameters will be highly uncertain, and supporting information may only be available from informal sources such as personal intuition. It is often useful to include the highly uncertain parameters despite the difficulties in their estimation. This is especially true if the parameters are needed to close a key feedback loop in the system. Most practitioners make their best estimates and proceed with simulating the model. The model is then tested to learn if changes in the parameters lead to important changes in the simulation results. Conducting and interpreting these tests is called sensitivity analysis.

This article describes sensitivity analysis as it is commonly performed in the course of building and testing a model. The article then turns to more formal methods of sensitivity analysis when a model has gained initial acceptance. The article describes a statistical approach that is useful when studying the combined uncertainties of dozens, even hundreds of uncertain parameters. Sensitivity analysis is facilitated by recent advances in software, and this article illustrates with a Vensim model of the deer population on the Kaibab Plateau. The article concludes with pragmatic advice to make formal sensitivity analysis a routine part of the modeling process.

1. Building Models with Uncertain Inputs

System dynamics modeling is an iterative, trial-and-error process. A model is usually built up in steps of increasing complexity until it is capable of replicating the problematical behavior of the system. Some models may be small, with less than a dozen parameters. Other models may contain hundreds, even thousands of parameters. Some parameters may be known with perfect accuracy, as with conversion factors or physical constants. Some parameters may be known to within 10%, but others may be highly uncertain, with ranges of uncertainty at plus or minus 100%.

Analysts may be reluctant to include highly uncertain parameters in a model because of a fear that their inclusion detracts from the credibility of the model. Most analysts have heard the disparaging phrase "garbage in, garbage out," and some might fear that including highly uncertain inputs leaves the modeling study vulnerable to criticism. An experienced system dynamics modeler knows that parameters should not be excluded just because they are uncertain. The more useful approach is to acknowledge the uncertainty and use computer simulation to learn the importance of the uncertain parameters. Rather than "garbage in, garbage out," the serious analyst would speak of "uncertainty in, uncertainty out." The analysis of how uncertainty in model inputs translates into uncertainty in the key outputs is called sensitivity analysis.

Sensitivity analysis is usually conducted after the model has been found to replicate the problematical behavior of the system. The examples below illustrate this practice for the model of the deer population on the Kaibab Plateau (explained in a separate article, number 003388). Figure 1 shows the simulation results of a model to explain the rapid growth and subsequent collapse of the deer population after predators were exterminated. The simulation begins with a population of 4,000 deer, whose numbers are held in check by the 50 predators on the plateau. The predators are reduced to zero during the interval from 1910 to 1920, and the model responds with an irruption in the deer population. The population peaks at around 120,000 shortly after the year 1920 and declines rapidly due to insufficient forage.

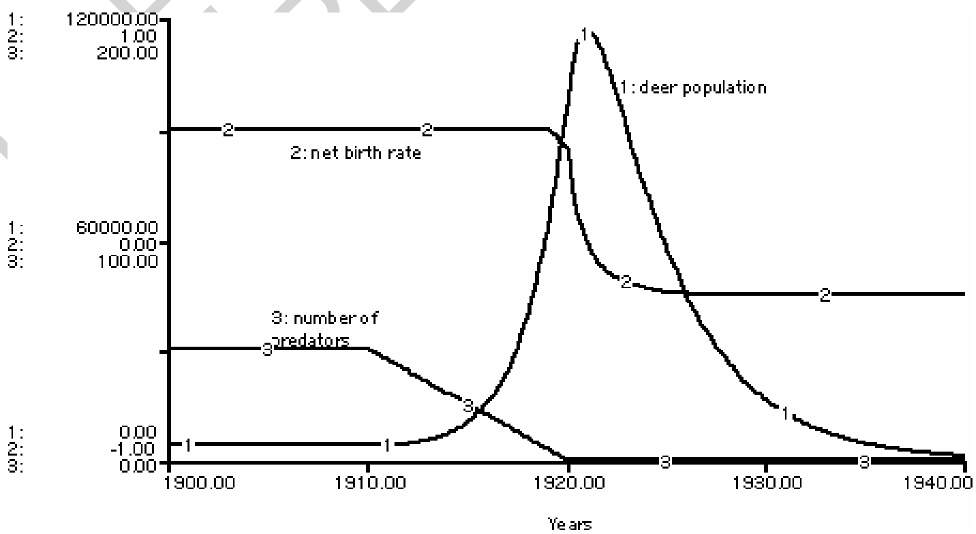

Figure 1. Simulated overshoot and collapse of the Kaibab deer herd.

Figure 2 tests the sensitivity of the overshoot pattern to changes in the estimate of the annual forage required per deer per year. The best estimate is 1 metric ton (MT) per deer. This is an approximate estimate based on one expert's observation that a deer consumes around 23 % of the dry matter that would be consumed by a typical cow. Figure 2 illustrates a sensitivity test if it is believed that the forage requirement is known to plus or minus 25%. This graph compares the simulated deer population in three simulations of the model. The previous result is shown in the middle --- the deer population climbs rapidly reaching around 120,000 shortly after the year 1920. If the annual forage requirement is 0.75 MT per deer, the deer herd would attain a much higher population at the peak of the irruption. Figure 2 shows the peak population at around 160,000 instead of 120,000. On the other hand, if the annual forage requirement is 1.25 MT per deer, the population would reach a peak of only around 80,000.

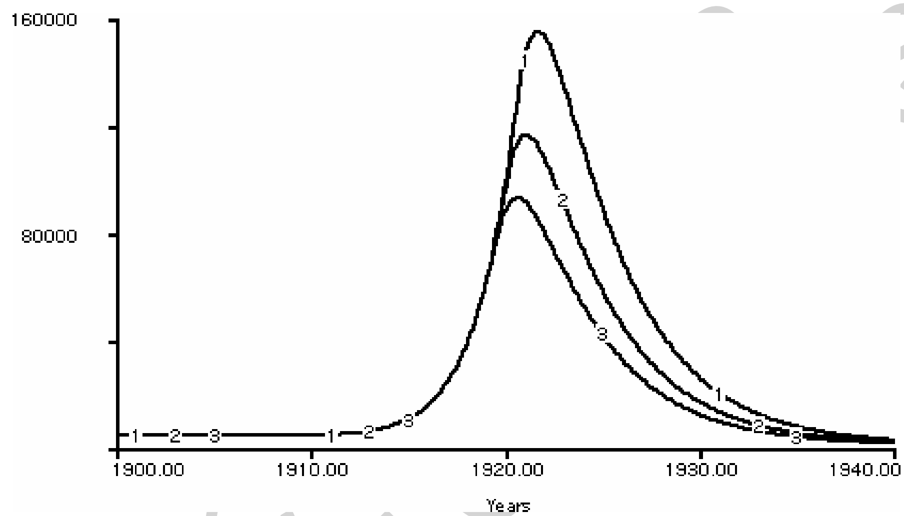

Figure 2. Sensitivity of the simulated deer population to changes in the forage requirement.

2. Robust Results

The numeric results in Figure 2 are certainly sensitive to the estimate of the annual forage requirement. If the goal were to predict the deer population at the peak of the irruption, a 25% variation in the forage requirement changes the peak estimate by almost 100%. One could conclude that the peak population could range from 80,000 to 160,000 so further work on the modeling must stop while analysts go to work to pin down the estimate of the forage requirement. This approach would make sense if the goal were to predict the peak population.

But system dynamics models are constructed to understand general patterns of behavior, not to make point predictions. In the deer example, the goal is to understand how the deer population could undergo explosive growth, followed by a collapse. The focus should be on the general pattern of behavior, not the precise number of deer in a particular year of the simulation.

When one studies the Figure 2 with the general pattern in mind, the interpretation is much different than before. For example, all three simulations show the same behavior during the first decade --- the deer population is held in check at 4,000 by the action of the predators. All three simulations show a rapid growth in the population, reaching a peak in less than ten years. All three simulations show the same result when the peak is reached --- the deer are too numerous to be supported by the biomass on the plateau, and the population declines rapidly. This sensitivity test has revealed that the model is robust. That is, the general pattern of behavior is not altered by changes in the estimates of this parameter.

Robust results are considered desirable because the analyst can proceed with the knowledge that the general pattern of behavior arises from the cause and effect interactions inside the model. Robust results are expected by system dynamics practitioners because they have come to expect that the pattern of behavior will arise from the internal structure of the system, not the precise estimate of a particular parameter. For example, they may expect to see exponential growth when the system is dominated by strong, positive feedback. Or they may expect exponential approach to equilibrium when the system is dominated by strong, negative feedback. The actions of the feedback loop, not the precise values of the parameters create the dynamic patterns.

3. The Structure Causes the Behavior

The expectation that system dynamics models will appear robust in the sensitivity analysis testing is sometimes expressed by the phrase "the structure causes the behavior." This phrase finds common usage in the system dynamics literature because model builders and model users have accumulated their experiences with different models. If their models were well structured, they tended to show robust results during sensitivity testing.

To appreciate why this happens, consider a model to simulate the temperature inside a house. The model will simulate a period of 60-120 minutes to help one learn the pattern of temperature change after the owner returns home from vacation. It's a cold day, and the indoor temperature and the thermostat setting are 60 degrees at the start of the simulation. The owner enters the house and uses the thermostat to set the target temperature to 70 degrees. The furnace will come on more frequently pumping heat into the house and increasing the temperature toward the target. A system dynamics model might include a wide variety of highly uncertain parameters (i.e, heat loss coefficients for the walls, floors, ceilings and windows). If the model is well structured, it will include the negative feedback loop that acts through the measured temperature to control the action of the furnace. Such a model will show the same general pattern of behavior in sensitivity tests with changes in the uncertain inputs. That is, the indoor temperature will climb from 60 degrees toward the 70 degrees in almost a linear fashion. Once the temperature reaches 70 degrees, it will remain close to the target with perhaps some small oscillations if there is a lag in the thermostat mechanism.

4. Poorly Structured Models

It's useful to anticipate what might happen if a model is not well structured. In the case of the temperature model, suppose the model builder forgets to close the loop between the measured temperature and the operation of the furnace. For example, suppose the heat production is set at 50% of the furnace capacity because the model builder has observed that the furnace normally operates half the time. Sensitivity testing of this model would reveal that the dynamic patterns would change every time a parameter estimate is changed. With one value of a heat loss coefficient, the model might show linear growth in the temperature. If the coefficient were changed to permit greater heat loss, the model might show that the temperature remains constant. And if the coefficient were changed further, the model might show a linear decline in the temperature. At this point, one might conclude that more time is needed to pin down the estimate of the heat loss coefficient. The system dynamics practitioner draws a different conclusion form these tests. The problem with these results is not the estimate of the heat loss coefficient; it's the failure of the model to close the key negative feedback loop in the system.

5. Leverage Points

As a general rule, the parameters in a system dynamics model do not prove to be critical to the general pattern of behavior. But important exceptions to this rule appear from time to time. When one discovers that an individual parameter makes a lot of difference, one may have found a "leverage point" in the system. A lever converts small movement at one end of the lever into big movement at the other end. A leverage point in a model means that a small change in one parameter can lead to large changes in the overall pattern of behavior. Leverage points are like gold nuggets; modelers keep their eyes open hoping to find one during the sensitivity analysis. When an important parameter is found, the modeler knows that further work on the parameter estimate is truly important. And if the parameter can be shaped by policy, the modeler may have discovered a key factor that could be changed to improve the general pattern of behavior.

An example of a leverage point is the shape of the predation function used in the model of the population of deer and predators on the Kaibab Plateau (see article 003388). In the initial simulations, the model showed growing oscillations that would lead the predators to annihilate their prey after only a few decades. This pattern tends to persist despite many changes in the model's parameters. But there is one key parameter that makes a big difference. It's the values of the predation function at low density of prey. When this parameter is changed (to allow prey refuge from their predators), the pattern of simulated behavior is changed from unstable oscillations to stable oscillations. In this case, the discovery of this leverage point provides important insight for wildlife management. It points to the importance of refuge if the predator and prey populations are to coexist over time.

-
-
-

Bibliography

Bunn, J.; Schneider, A.; Watchel, T. and Brimm, J. (1985). Fluid Therapy in Large Area Burns. *System Dynamics Review*, (1)1. [describes the challenge of formal sensitivity analysis as intractable]

Clemson, B. Tang, Y. Pyne J. and Unal, R. (1995). Efficient Methods for Sensitivity Analysis. *System Dynamics Review*, (11) 1. [reviews the efficiency of Latin Hypercube Sampling and Taguchi methods and uses Taguchi methods to study a model of waste water treatment]

Ford, A. (1990). Estimating the Impact of Efficiency Standards on the Uncertainty of the Northwest Electric System. *Operations Research*, (38) 4. [describes the iterative application of formal methods to a model of the northwest electric system]

Ford, A. (1999). *Modeling the Environment*. Island Press [college text on formal as well as informal methods of sensitivity analysis]

Ford, A. and McKay, M. (1985). Quantifying Uncertainty in Energy Model Forecasts. *Energy Systems and Policy*, (9) 3. [describes the benefit of Latin Hypercube Sampling for a model of an electric utility company]

McKay, M. Beckman, R. and Conover, W. (1979). A Comparison of Three Methods for Setting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*, (21) 2. [the power of Latin Hypercube Sampling is revealed in a comparison of different sampling methods to study the simulated temperature changes in a nuclear reactor]

Palisade Corporation, @Risk: Risk Analysis and Modeling for the PC. [software for spread sheet analysis available at 2189 Elmira Road, Newfield, NY 14867]

Reilly, J, Edmonds, J; Garder R. and Brenkert A. (1987). Uncertainty Analysis of the IEA/ORU Carbon Dioxide Emissions Model. *The Energy Journal*. July 1987 [estimates the benefit of Latin Hypercube sampling as a ten fold reduction in sample size]

Ventana Systems, Inc (1995). Vensim User's Guide, Ver 1.62. [the software used in this article is available from Ventana Systems, 149 Waverly Street, Belmont, MA 02178]

Biographical Sketch

Andrew Ford was born in 1944 in Lone Pine, California. His research deals with energy and environmental problems in the western United States. He is especially interested in the use of simulation modeling for policy analysis in the electric power industry. Dr. Ford completed his doctoral studies in the Program on Public Policy and Technology at Dartmouth College in 1975. He worked in the Energy Policy Group at the Los Alamos National Laboratory and in the Systems Management Department at the University of Southern California. His research on electric power and conservation was honored with the 1996 Jay W. Forrester Award.

Dr. Ford is Professor of Environmental Science at Washington State University where he teaches system dynamics for students interested in systems and the environment. His text on Modeling the Environment was published by Island Press in 1999.