HOW TO DO STRCUTURAL VALIDITY OF A SYSTEM DYNAMICS SIMULATION MODEL? THE CASE OF AN ENERGY POLICY MODEL

Hassan Qudrat-Ullah, School of Administrative Studies, York University, 4700 Keele Street, Toronto, ON, M3J 1P3, Canada,hassanq@yorku.ca

ABSTRACT

This paper argues that structural validity of the simulation model -right behavior for the right reasons- is a stringent measure to build confidence in a simulation model regardless of how well the model passes behavior validity tests. That leads to an outline of formal structural validity procedures. An illustration of one of a test from a set of six tests for structural validity of a system dynamics model for energy policy analysis follows. Finally, some conclusions on the increased appeal for simulation models for energy policy analysis and design are presented.

INTRODUCTION

Simulation models have been developed and applied to both operational problems and policy issues in energy sector. However, the need of and the evaluation criteria of model validation differs for each case. For instance, in the case of operational problems, the results of a model can be accepted or rejected by exposing the results to a face validity test ([9]; [4]). In a face validity test, experts assess how the model and its results are close to the real system. Model solutions can be tested in real world environments: e.g., efficiency of the oil refinery can be enhanced under the recommended actions ([7]).

In contrast, the majority of policy models such as system dynamics (SD) type model are built for the analysis of policy, exploration of possible future scenarios, and management purposes ([7]; [12]; [10]). From energy policy research perspective, modeling resolutions to important issues simply defy a face validity test. Instead, for policy models, the key issue in validation is deciding (i) if the model is acceptable for its intended use, i.e., does the model mimic the real world well enough for its stated purpose ([5]; [8]; [6]) and (ii) how much confidence to place in model-based inferences about the real system ([1]; [2]; [3]). In order to assess the theoretical content of a policy model, it is imperative to look at the modeling process itself. The crux of SD modeling process is to identify how structure and decision policies help generate the observable patterns of behavior of a system and then identified structures and decision policies be implemented. Therefore, the identification of the appropriate structure is the first step in establishing validity of an SD model. Once the structural validity of an SD model is sufficiently established, behavior validity - how well the model-generated behavior mimics the observed behavior of the real system - is assessed to achieve the overall validity of the model or to build confidence in the model ([7]; [12]). In fact the validation process becomes iterative: structural validity-behavior validity-structural validity. It must be emphasized here that in no way I am discounting the usefulness of behavioral validity of an SD model. Instead, I want to highlight the significance of structural validity, often less explored in SD model validation endeavors.

STRUCTURAL VALIDITY PROCEDURES

Identification of the appropriate structure, responsible for the ‘right’ behavior, is a multidimensional process: problem representation, logical structures, and mathematical and causal relationships. Forrester and Senge discussed some of the tests used for structural validation of an SD model [6]:
Boundary adequacy: Whether the important concepts and structures for addressing the policy

issues are endogenous to the model?
Structure verification: Whether the model structure is consistent with relevant descriptive knowledge of the system being modeled?
Dimensional consistency: Whether each equation in the model dimensionally corresponds to the real system?
Parameter verification: Whether the parameters in the model are consistent with relevant descriptive and numerical knowledge of the system?
Extreme conditions: Whether the model exhibits a logical behavior when selected parameters are assigned extreme values?

Barlas has demonstrated that behavior sensitivity test, originally suggested by Forrester and Senge as a behavior validity test, can detect major structural flaws of the model despite the fact that model can generate highly accurate behavior patterns ([1], [6]). He termed it as a *structurally-oriented behavior test*: Whether the real system would exhibit a similar high sensitivity to those parameters to which model behavior displays high sensitivity.

AN ILLUSTRATION OF STURTCURAL VALIDITY TESTS

We applied all these six tests to an energy policy model: MDESRAP ([11]). However, due to limited space, we will only describe one of these tests: “structurally-oriented behavior test” here.

Structurally-oriented behavior test

In this test, the behavioral sensitivity of MDESRAP is evaluated, i.e. how does the model’s behavior change over time as alternative parameter values are tested in the model, one at a time ([1]). When we applied the test, the fundamental patterns of behavior of production capacity, electricity intensity, and CO_2 intensity proved to be insensitive to most of the system parameters. The parameters were tested with both an increase and a decrease of 50 % to its base value. Changing these parameters affected only the specific numerical values of behaviour characteristics such as an earlier take-off, or shortened doubling time. Such a numercal sensitivity is in line with the anticipated behaviour of the real system under the same conditions. As a result, the more is our confidence in MDESRAP’ structure.

CONCLUSION

Although structural validity tests constitute one of the two general types of tests required to build confidence in an SD type simulation model, these tests nevertheless are the core of SD modeling validation process and have temporal precedence over behavior validity tests. Illustrations provided through the applications of six tests for the structural validity of MDESRAP can help the modelers (and users) in energy policy domain to lend an effective and tangible support to the process of building confidence in their simulation models. Informed by the ‘purpose’ and structurally tested simulation models, should result in the increased appeal for simulation models for energy policy analysis and design. The energy policy issues exist. The simulation models are being built. Validation challenges are being met. Energy policy analysis simulation modeling community owes no apology to those who would only believe in face validity testing alone.

REFRENCES

[1] Barlas, Y., 1989. Multiple tests for validations of system dynamics type of simulation models. European *Journal of Operational Research* 42 (1), 59-87.

[2] Barlas, Y., 1994. Model validation in system dynamics. In: Wolstenholme, E., Monaghan, C. (Eds.), *Proceedings of the 1994 International System Dynamics Conference*. System Dynamics Society, Sterling, Scotland, pp. 1-10.

[3] Curry, G. L., Deuermeyer, B. L. & Feldman, R. M., 1989. *Discrete Simulation*. Holden-Day, Oakland, CA.

[4] Emshoff, J. R. & Sission, R. L. 1970. *Design and use of computer simulation models*. Macmillan, New York.

[5] Forrester, J. W., 1971. The model versus modeling process. *System Dynamics Review* 1(2), 133-134.

[6] Forrester, J. W. & Senge, P. M., 1980. Tests for building confidence in system dynamics models. *TIME Studies in the Management Science* 14, 209-228.

[7] Gass, S. I., 1983. Decision-aiding models: Validation, assessment, and related issues for policy analysis. *Operations Research* 31 (4), 603-631.

[8] Goodall, D. W., 1972. Building and testing ecosystem models. In: Jeffers, J. N. J. (Ed.), *Mathematical Models in Ecology*. Blackwell, Oxford, pp. 173-194.

[9] Hermann, C., 1967. Validation problems in games and simulations. *Behavioral Science* 12, 216-230.

[10] Oliva, R., 2003. Model calibration as a testing strategy for system dynamics models. *European Journal of Operational Research* 151, 552-568.

[11] Qudrat-Ullah, H., 2005. MDESRAP: a model for understanding the dynamics of electricity supply, resources, and pollution. *International Journal of Global Energy Issues* 23(1), 1-14.

[12] Sterman, J. D., 1984. Appropriate summary statistics for evaluating the historical fit of system dynamics models. *Dynamica* 10 (Winter), 51-66.