Homer Jack (Orcid ID: 0000-0003-3690-9892)

Best practices in system dynamics, revisited

Best practices in system dynamics modeling, revisited: a practitioner's view

At the 2001 International SD Conference in Atlanta, Paul Newton and I offered the first Modeling Assistance Workshop, which I co-directed annually through 2016, and which continues to the present as a well-liked, appreciated part of the summer conference. It may be surprising to learn that those seeking assistance are not necessarily beginners, but rather have typically taken courses in system dynamics and may even have built models as part of their work. Their questions are sometimes detailed and technical, but are just as often more basic inquiries about the right approach to model conceptualization, formulation, or testing.

I would sometimes ask myself how it was that people with background in SD still did not have a solid idea of how to proceed effectively through the modeling process. That process was first described in detail by Randers (1973), who introduced the key concepts of reference mode and dynamic hypothesis. It has since been further elaborated in textbooks (Randers 1980, Richardson & Pugh 1981, Sterman 2000, Morecroft 2007, Warren 2008), with many examples and bits of advice given. Yet, system dynamics modeling in practice remains largely an art, and even after all these decades, the range of approaches and quality we see at conferences and in publications is very wide. This "Wild West" irregularity of practice is a concern to many of us in the field, and is surely confusing to newcomers and to people outside the field looking in (Homer 2013).

Some years ago, I participated as one of 20 experts (all former SD Society presidents and award winners) in an extensive study of best practices in system dynamics (Martinez-Moyano and Richardson 2013). We offered lists of ideas on every step of the modeling process and voted on them. We agreed on many things but also apparently disagreed on some. For example, we agreed that the initial conceptualization phase should establish a dynamic hypothesis, but not on whether the process should center more on stock-and-flow diagrams or causal-loop

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/sdr.1630

diagrams. Similarly, we agreed that models should start small and grow in complexity as needed, but not on whether the greater complexity was driven primarily by the need for greater real-life detail or rather to assure robustness (e.g., non-negativity) under extreme conditions.

Reflecting on this study, I would argue those apparent disagreements were more mirage than reality. For example, in developing a dynamic hypothesis, most expert practitioners I know use diagrams that are a hybrid of causal-loop and stock-flow (not just one or the other). Similarly, in considering model expansion, most practitioners would look for both realism and robustness (not just one or the other) and might justify expansion for either reason.

In my opinion, the Best Practices study was a well-intended effort but failed to uncover the substantive differences in modeling practice that really do exist. For example, the expert group strongly agreed with the idea that graphs over time (reference modes) should be identified for central variables. But there was no statement about whether these time graphs should be constructed based on recorded time-series data as opposed to client recollections; nor about the ideal number of such graphs. In my experience, it is exactly on such details that practitioners differ in actual practice. Modeling inevitably involves trade-offs of effort within the time available (Homer 1996), and our expert group was perhaps too polite to get very specific on points about which they might be forced to acknowledge, "Yes, that would be ideal, but I usually don't go that far."

Consequently, the Best Practices study did not really succeed in narrowing down all of the practice recommendations to something more focused and approachable. Instead of identifying a straight and efficient path, we continue to offer a broad menu of options at every step. This failure to "pin it down" or "spell it out" creates confusion and uncertainty—including for many who come to the Modeling Assistance Workshop seeking help.

My purpose here is to offer one such narrower approach, an approach I have most often used during my three-plus decades as a professional SD modeler. It is primarily concerned with the goal of model reliability. This is not to say that other goals, such as group engagement, rapid-cycle learning, or multi-stakeholder consensus (Hovmand 2014) are not important; but rather that when given the choice, I go with model reliability. I have described this approach elsewhere as scientific (Homer 1996) or evidence-based (Homer 2014). Although scientific, I believe this approach is accessible to modelers of all levels and backgrounds, and does not require technical skills that go beyond what is taught in introductory SD courses. It is in the same spirit as the thoroughly rigorous approach John Sterman has laid out recently (Sterman 2018), though less technically demanding.

The approach is as follows:

1. *Problem Definition*

   1a. Get the full historical context from the client, subject matter experts (SME's), and as much related literature as possible. (For a case that has no history, such as the design of a new product or service, identify the broader class to which this case belongs, and compile data and information about relevant members of that class.)

   1b. Assemble a broad array of potentially relevant time-series data, for as many variables as possible, even into the dozens. Some of these data will turn out to be significant and others will not, but you cannot know the difference beforehand, and should not do too much up-front filtering-out. (Client recollections of the past are often unreliable and should not be used in place of data for reference mode construction. Clients are most helpful in describing current structures and policies.)

1c. Based on what you've learned from the information above, define the key problem or question in terms of measured outcomes that define better and worse. List factors that seem central to explaining the problem, and list possible solutions.

1d. Draw a hybrid causal-loop and stock-and-flow diagram that logically connects the central factors, endogenously explains the outcomes, and shows where the policy options may fit in. This diagram is your preliminary dynamic hypothesis (DH), but it is not a complete model, and it should fit easily on a single piece of paper. Core stock-flow structures (e.g., aging chains), and not just causal loops, should be prominent in the diagram; stock-flow dynamics are at least as important as feedback dynamics in well-built SD models (Richmond 1994, Homer 1997).

2. *Model Formulation and Reliability Testing*

2a. Assemble a running simulation model based on the DH diagram. To develop a proper simulation model, some of the concepts in the DH diagram will need to be broken out into separate, logically distinct variables. As a result, even a first-version simulation model may end up substantially bigger than the DH diagram. Go back to your information sources (client, SME's, literature) as needed to ensure realism in all formulations. Use other good practices of equation formulation (robustness under extreme conditions, units consistency, etc.) as described in SD textbooks. In some cases, generic structures described in textbooks (e.g., market diffusion, production control, commodity cycles) may be useful, but note that these have many variations and likely need customization for a particular real-world application.

2b. Run the model to determine its ability to reproduce the full array of time-series data corresponding to model variables, and the plausibility of all its base-case outputs. Do sensitivity tests (e.g., neutralizing a causal link, or subjecting the model to plausible extreme conditions) to ensure the model is doing the right things for the right reasons. If such testing reveals logical inconsistencies, go back to revise the model.

2c. Identify all large gaps between model and data, and question your information sources in an attempt to explain these gaps. Additional information gathering is often required. Some gaps may be corrected through better parameter estimation. Other gaps may be the result of systematic flaws in equation formulation or the absence of an important variable—indicating the need for model revision. Yet other gaps may be acceptable deviations corresponding to one-time events or measurement error—for which the model need not be revised.

2d. Revise the model accordingly and repeat the above procedures until all base-case outputs are plausible, and an acceptable fit to data is obtained. If new variables or formulations are added, make sure they are supported by your information sources and integrate properly with the rest of the model without harming its structural coherence.

3. *Policy/Scenario Analysis and Policy Sensitivity Testing*

3a. Test policies and alternative scenarios (individually and in combination) with realistic inputs and determine whether the outputs are plausible. If they are not, this will require a return to Step 2 for model revision.

3b. Initial policy results may spur ideas for other policies and policy structures not included initially, and which may be added at this point.

3c. Test the policies and scenarios also in combination with changes in uncertain parameters. Such sensitivity testing may be done deterministically at first (one or two parameters at a time), saving Monte Carlo testing for later if desired. Make sure model outputs are plausible under all changes in uncertain parameters. If they are not, this will require a return to Step 2 for model revision.

3d. If policy findings are found sensitive to specific uncertain parameters, do one of two things depending on remaining time and budget: (a) move on to final model write-up and presentation reporting the sensitivity as a subject for future research; or (b) proceed to do more research on the area of sensitivity to determine whether the uncertain parameter can be better estimated, or whether it can be replaced with a more certain and evidence-based addition to model structure.

I have described here an approach to the building and testing of models which has worked well with clients, audiences, and scholarly journals throughout my career as an SD modeler. This approach revolves around the careful gathering of evidence for both system structure and behavior. It may not be right for everyone, but I do think it has the advantage of being more streamlined and directed than traditional descriptions of the modeling process, while maintaining essential rigor. It may help people with basic SD training to develop scientifically stronger models, and may also make our approach more understandable for people inside and outside the SD field.

References

Homer JB (1996). Why we iterate: scientific modeling in theory and practice. *SD Review* 12(1):1-19.

Homer JB (1997). Structure, data, and compelling conclusions: notes from the field. *SD Review* 13(4):293-309.

Homer J (2013). The aimless plateau, revisited: why the field of system dynamics needs to establish a more coherent identity. *SD Review* 29(2):124-127.

Homer J (2014). Levels of evidence in system dynamics modeling. *SD Review* 30(1-2):75-80.

Hovmand PS (2014). *Community Based System Dynamics*. Springer: New York.

Martinez-Moyano IJ, Richardson GP (2013). Best practices in system dynamics modeling. *SD Review* 29(2):102-123.

Morecroft J (2007). *Strategic Modelling and Business Dynamics*. Wiley & Sons: Chichester, England.

Randers J (1973). Conceptualizing dynamic models of social systems: lessons from a study of social change. PhD Dissertation, MIT Sloan School of Management.

Randers J (1980). Guidelines for model conceptualization. In: *Elements of the System Dynamics Method*, ed. J Randers. MIT Press: Cambridge, Massachusetts.

Richmond B (1994). Systems thinking/system dynamics: let's just get on with it. *SD Review* 10(2-3):135-157.

Richardson GP, Pugh AL III (1981). *Introduction to System Dynamics Modeling with DYNAMO*. MIT Press: Cambridge, Massachusetts.

Sterman JD (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin McGraw-Hill: Boston, Massachusetts.

Sterman JD (2018). System dynamics at sixty: the path forward. *SD Review* 34(1-2):5-47.

Warren K (2008). *Strategic Management Dynamics*. Wiley & Sons: Chichester, England.

Best practices in system dynamics modeling, revisited: a practitioner's view

Author information:

Jack Homer

72 Station Hill Road, Barrytown, New York, USA 12507

Jack Homer is a system dynamics modeling consultant and has directed Homer Consulting since 1989. He received a PhD from MIT Sloan, and BS and MS degrees in applied mathematics and statistics from Stanford University. He received the SD Society's Forrester Award in 1997, and the Applications Award in 2011. He is the author of the books, "Models That Matter" (2012) and "More Models That Matter" (2017).