The Evolution of Models: Uncovering the Path of Model Improvement

Markus Schwaninger

University of St. Gallen, Dufourstrasse 40a, 9000 St. Gallen, Switzerland
markus.schwaninger@unisg.ch

Abstract. The purpose of this contribution is to learn from the theory of evolution in order to improve the design of modeling processes. The aim of the chapter is to advance the understanding of the evolutionary structural characteristics of modeling processes which lead to better models, and to find design rules for systematically achieving higher model quality. We use a simulation approach, building a System Dynamics model. The simulation-based analysis uncovers several surprising, counterintuitive results and useful insights. These deliver lessons for advancing the design of model-building processes.

Keywords: Modeling and simulation · Model-building · Evolutionary process · Model quality

1 Purpose

The assumption underlying this paper is that the modeling process leading to model quality is driven by the evolutionary "mechanisms" of variation, selection and retention (Aldrich 1999). Assuming that the evolution of models can be influenced, the research questions, then, are how and to what extent this can be done. We intend to learn from the nature of general processes of evolution how we can better design modeling processes.

The principles of evolution have been studied extensively in the biological-ecological realm (Darwin 1859; Jablonka and Lamb 2005). In the technical domain, the principles of evolution have been used to amplify the improvement of technical objects (e.g., Rechenberg 1973; Price et al. 2005). Also, the evolution of social systems – societies, economies and organizations – in both structural and cultural terms, have been explained in terms of the theory of evolution (Campbell 1965; Mesoudi et al. 2006; Aldrich 2008; Dopfer and Potts 2008). Qualitative models of the substantive evolution of models have been published (e.g., Wartick and Cochran 1985), and quantitative model systems for experimental evolution have been built (e.g., Collins et al. 2014).

To the author's knowledge, however, until now modeling processes have not been studied quantitatively in a way that allows one to learn how to improve them by applying the general principles of evolution. We will refer to computer models as used for the purpose of simulation.

To answer the research questions, we are building a simulation model of an idealized modeling process. We are using the System Dynamics (SD) methodology.
developed by Prof. Jay Forrester at the Massachusetts Institute of Technology (Forrester 1961; Sterman 2000). SD models represent complex systems with two kinds of variables, stocks and flows, and their interrelationships. These structures are simulated as continuous processes, with feedbacks and delays. The dynamic behavior of the system under study results from the interactions of the variables, and is nonlinear.

2 A Generic Model of the Modeling Process

The model built here reflects the structure of evolutionary processes. However, we refrain from using the commonly adopted sequential representation. The model consists of causal feedback loops for each one of the evolutionary “mechanisms”. These mechanisms are interdependent, and they converge at the construct “Model Quality” (MQ). MQ is the variable in focus that expresses the degree to which the maturity of a model has evolved. Given its embeddedness in three causal circuits, it can grow, but it can also shrink.

The stock-and-flow diagram, which represents the modeling process, is shown in Fig. 1. The logic of the three causal loops in the diagram is as follows:

1. Variation Loop: The variety of available models is generated by the introduction of new models into the process. The quantity of new models inserted depends on the current level of model quality. These new models are subject to a first validation. The validated models can, but do not necessarily, increase model quality. They also have to stand up to the competition from current models-in-process, i.e., those models which are already in the improvement process (“models-in-process”).

2. Selection Loop: A selection criterion determines the level of errors upon which the decision is made either to use a model-in-process or to adopt one of the new models (from loop 1). Model quality then is changed either by an improved, current model or by the use of a new model. This loop is closed by the connection of model quality, tests and errors.

3. Retention Loop: A comparison between the model quality standard and the actual quality of models-in-process results in a gap. The gap determines the frequency of tests: the greater the gap, the more tests there will be. Errors discovered in these tests and subsequent error corrections cause changes in model quality.

Stocks are represented by boxes and flows by valves. The orthogonal dashes on several arrows denote delays. The arrows marked with negative polarities ("−") represent a causality in which more of the cause leads to less of the effect, and vice versa. All other variables are assumed to have a positive polarity, with more (less) of the cause leading to more (less) of the effect. The overall polarity of a loop is the product of the signs on the arrows constituting that loop. For example, in the case of the Retention Loop: "−" * "+" * "−" * "+" * "−" * "−" * "−" = "−". The simulation model is made up of equations and constants. The accumulation of stocks is captured in differential equations.
Amore parsimonious representation of this quantitative model, at a macro-level, is captured in formula (1):

$$OPT\{M : M \in \text{gen(Processes)}\} = \int_{t=0}^{t_{\text{final}}} S[Coupling(Var, Sel, Ret)]dt,$$

(1)

where $M$ denotes model(s), and $S$ an event-coupling system of cyclic processes. $Var, Sel, Ret$ are dynamically coupled processes, with $Var$ standing for Variation, $Sel$ for Selection, and $Ret$ for Retention.

Two crucial micro-level equations are in formulas (2) and (3). First, Model Quality is the accumulation of all net flows:

$$Model\ Quality = \int_{t=0}^{t_{\text{final}}} (IncreaseMQ - DecreaseMQ)dt$$

(2)

Second, a formula from the Selection Loop incorporates the focal decision between the options (A) to continue working on the same model, and (B) to use a new model. A Selection Criterion defines the determinant threshold: the minimal number of errors upward from which a new model must be used [Use New Model], instead of carrying on with the correction of errors on the current model. The formula is:

$$Use\ New\ Models = \text{Min}(Valid\ New\ Models/AT, \text{Min}(IF\ THEN\ ELSE(Errors > Selection\ Criterion, 1, 0), INTEGER((valid\ New\ Models)/AT))).$$

(3)

where $AT$ is an adaptation delay.
3 Simulation Results

The variable in focus here is [Model Quality]. Plotted over time, it shows the following pattern in the base run (Fig. 2).

The pattern is oscillatory. In the long term it varies: over 300 weeks the variable MQ [Model Quality] shows a period between 23 and 40 weeks, i.e., an average period of 30 weeks (~7 months) and an amplitude between 0.98 and 1.73 Qual. Qual is the unit of measurement for MQ. Average MQ, in the base run, converges to an equilibrium of 4.7 Qual/week.

![Graph showing oscillatory pattern of MQ over time]

To ascertain the impact of changes in parameter values, elasticities $e_i^j$ are calculated for cumulative MQ, i.e., $Q_{cum}$, and cumulative oscillations of MQ, i.e., $O_{cum}$ over 52 weeks. The formula for the elasticities $e_Q$ and $e_o$ is:

$$ e_i^j = \frac{J_i - J_i^b}{P_i - P_i^b} $$

where $J_i$ are outcomes, either $Q_{cum}$ or $O_{cum}$. $P_i, i = 1, \ldots, 6$ denotes the parameters. $b$ stands for the base value of a parameter, and $j$ stands for parameter changes, - $u$, up, and $d$, down. $j = 1, 2, 3$ indexes the three scenarios for parameter values at $b$, $u$, and $d$.

Oscillations of MQ, $o$ and $O_{cum}$ are calculated as follows:

$$ Q_{cum} = \int_{t=0}^{t_{final}} q \cdot dt, \quad o = \frac{Q_{cum}}{t_{final}}, \quad \bar{q} = \frac{Q_{cum}}{t_{final}}, \quad O_{cum} = \int_{t=0}^{t_{final}} |o| \cdot dt, $$

where variable $q$ stands for MQ, $o$ for oscillations of MQ, and $\bar{q}$ for average MQ.

The parameters $P_i$, analyzed in Table 1, are:

- For the Retention Loop: Selection Criterion (SC), Quality Standard (QS), Tests of Model per Gap (TG), Obsolescence (OB).
## Table 1. Scenario settings, outcomes and elasticities

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Parameter values</th>
<th>MQ cumulated at $t = 52$</th>
<th>Elasticity for MQ</th>
<th>Oscillations MQ cumulated at $t = 52$</th>
<th>Elasticity for oscillations MQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection criterion</td>
<td>Base SCb</td>
<td>8</td>
<td>249.7</td>
<td>729.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Up SCu</td>
<td>12</td>
<td>260.1</td>
<td>0.08</td>
<td>273.5</td>
</tr>
<tr>
<td></td>
<td>Down SCd</td>
<td>4</td>
<td>231.5</td>
<td>0.15</td>
<td>843.9</td>
</tr>
<tr>
<td>Quality standard</td>
<td>Base QSb</td>
<td>10</td>
<td>249.7</td>
<td>729.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Up QSu</td>
<td>15</td>
<td>231.5</td>
<td>-0.15</td>
<td>843.9</td>
</tr>
<tr>
<td></td>
<td>Down QSd</td>
<td>5</td>
<td>129.4</td>
<td>0.96</td>
<td>106.1</td>
</tr>
<tr>
<td>Model tests per gap</td>
<td>Base TGb</td>
<td>2</td>
<td>249.7</td>
<td>729.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Up TGu</td>
<td>3*</td>
<td>231.5</td>
<td>-0.15</td>
<td>843.9</td>
</tr>
<tr>
<td></td>
<td>Down TGd</td>
<td>1</td>
<td>165</td>
<td>0.68</td>
<td>95.2</td>
</tr>
<tr>
<td>Obsolescence</td>
<td>Base OBb</td>
<td>0.1</td>
<td>249.7</td>
<td>729.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Up OBU</td>
<td>0.15</td>
<td>223.3</td>
<td>-0.21</td>
<td>545.4</td>
</tr>
<tr>
<td></td>
<td>Down OBD</td>
<td>0.05</td>
<td>340.6</td>
<td>-0.73</td>
<td>471.4</td>
</tr>
<tr>
<td>New models made available</td>
<td>Base NMb</td>
<td>4</td>
<td>249.7</td>
<td>729.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Up NMh1</td>
<td>6</td>
<td>259.8</td>
<td>0.08</td>
<td>556.2</td>
</tr>
<tr>
<td></td>
<td>Up NMh2</td>
<td>8</td>
<td>268.2</td>
<td>0.07</td>
<td>557.4</td>
</tr>
<tr>
<td></td>
<td>Up NMh3</td>
<td>10</td>
<td>276.6</td>
<td>0.07</td>
<td>432.6</td>
</tr>
<tr>
<td></td>
<td>Down NMI</td>
<td>2</td>
<td>97.35</td>
<td>1.22</td>
<td>1353</td>
</tr>
<tr>
<td>Frequency of model introduction</td>
<td>Base FLb</td>
<td>1</td>
<td>249.7</td>
<td>729.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Up FLh</td>
<td>1.5</td>
<td>259.8</td>
<td>0.08</td>
<td>556.2</td>
</tr>
<tr>
<td></td>
<td>Down FLi</td>
<td>0.5</td>
<td>97.35</td>
<td>1.22</td>
<td>1353</td>
</tr>
</tbody>
</table>

*same results for TG = 4.00 (saturation)

Correlation MQ cum/Osc cum: -0.36

- For the Variation Loop: New Models Made Available (NM), and Frequency of Model Introduction (FI). The results for these two variables are equal, as changing them has the same effect in both cases. They are kept separate because they represent two distinct policies.
As far as the elasticities for MQ are concerned, if these move in the same direction as parameter changes, then a general assumption would be that most negative values represent unexpected results.

Summarizing the results:

1. It is generally difficult to improve MQ. The only improvements are visible in the small elasticities $e_q$ with SC, NM and FI.
2. MQ and Oscillations are negatively correlated, i.e., there is a tradeoff between the two. Low levels of MQ are often associated with high oscillation levels. There are ways of reducing oscillations, but they are often associated with lower MQ, as in the case of parameters QS, TG, and OB.
3. Quality Standard (QS): Stretching parameter QS too high lowers MQ.¹
4. Testing: Neither an increase nor a decrease in the number of model tests can raise MQ: both damage MQ. We have an inverted-U-relationship here.
5. Obsolescence: Increasing obsolescence should «refresh» MQ, but it does not. What it does is reduce both MQ and Oscillations substantially.

Conclusively, we derive the following insights and implications:

- The best way to reduce MQ is by enhance keeping the number of New Models introduced or the Frequency of Introduction high. With both procedures the necessary redundancy of options is created to enable better models. → Provide redundancy!
- High levels of obsolescence damage MQ. → Do not abandon models-in-process too early!
- Both too many and too few tests are detrimental. → Find a reasonable level of test intensity; avoid over- and under-testing!
- Quality is difficult to increase but easy to lose: Erosion of quality is very dangerous. → Keep Quality Standard high, but do not exaggerate!

4 Conclusion

At the outset, we posed two parallel research questions: if the evolution of models can be influenced, then how and to what extent can this be done? This study has demonstrated that such evolutionary modulation can be accomplished, and that a number of parameters are appropriate to influencing model evolution in both level and direction.

Given a number of unexpected results found via the simulations, certain clues for the improvement of modeling processes have been identified. These findings deliver both conceptual value, further instructing us about the “anatomy” of modeling processes, and also economic potential, in terms of quality and productivity improvements.

The model presented here is structurally validated, but has not been tested yet on the basis of empirical data. For the time being it is a generic model with didactic

¹ $q_{QS}^h$ was already set at a demanding level, but $q_{QS}^l$ was overstretched.
benefits, providing insights about generic patterns of system behavior. Adding more detail could lead to outcomes different in degree but not distinct in kind. Fitting the model to real life could be useful in making it a steering tool for model-builders.

Acknowledgement. The author wishes to thank Profs. Franz Pichler and Stefan Ott for their precious comments and encouragement.

References

Rechenberg, I.: Evolutionsstrategie. Frommann, Stuttgart (1973)
Computer Aided Systems Theory – EUROCAST 2015

15th International Conference
Las Palmas de Gran Canaria, Spain, February 8–13, 2015
Revised Selected Papers