Understanding model behavior using loops that matter.

William Schoenberg, Pål Davidsen, Robert Eberlein

Abstract
The link between structure and behavior is central to System Dynamics, but effective tools for understanding that relationship still elude us. The current state of the art in the field of loop dominance analysis relies on either practitioner intuition and experience or complex algorithmic manipulation in the form of eigenvalue analysis or pathway participation metrics. This paper presents a new and distinct method to find the 'loops that matter' in generating model behavior. This is a numeric method capable of determining the impact for every loop in a model and identifying which dominate behavior at each point in time. The method was inspired by observations on variable value changes during simulations and has been refined using empirical evaluation on a variety of different models. In addition to explaining behavior, the method shows promise for improving visualization and aggregation of simulation results.

The problem
The strong link between structure and behavior is fundamental to system dynamics (Sterman, 2000). A model uses parametric and structural assumptions to derive behavior. A system dynamics practitioner at a general level must go through the following process: create a structure, understand how that structure works, and figure out how to improve that structure. The second step in that process is the focus of this paper and is the key to successfully performing the third step. By referring back to the model structure, the practitioner can explain the reasons why an observed behavior has been produced (Richardson, 1996). Based on that understanding, the practitioner may also propose changes in input values or model structure that will cause a more favorable behavior to be produced.

The current state of the art in the field relies on either practitioner intuition and experience (the art of modeling and model analysis) or complex algorithmic analysis. The former is taught as part of the methodology of model building, while the latter comes from 40 years of work on techniques to derive and explain model behavior based on the analysis of structure (see for example: Graham, 1977; Forrester 1982; Eberlein, 1984; Davidsen, 1991; Mojtahedzadeh, 1996; Ford, 1999; Saleh, 2002; Mojtahedzadeh et al., 2004; Goncalves, 2009; Saleh et al., 2010, Kampmann, 2012; Hayward and Boswell, 2014; Moxnes and Davidsen, 2016; Oliva, 2016 and Hayward and Roach, 2018.

Ford (1999, p.4-5) clearly stated the needs of the system dynamics field as they apply to loop dominance analysis:

> *"To rigorously analyze loop dominance in all but small and simple models and effectively apply analysis results, system dynamicists need at least two things: (1) automated analysis tools applicable to models with many loops and (2) a clear and unambiguous understanding of loop dominance and how it impacts system behavior."*

In this paper we discuss a new loop dominance analysis method and demonstrate its performance on the two tests set out by Ford. For the purposes of this paper, we define one loop as dominant over the others if it is responsible for generating more of the behavior of the stocks in the model at a specific point in time relative to the other loops. Often, there will be one singular dominant loop at each point in time for the model, but there are cases where no single loop will be solely dominant and multiple loops together will form a plurality and therefore together will describe the majority of the behavior of the stocks in a model. This definition holds true for the cases where all stocks are connected to each other and themselves by the network of feedback loops in the model (well connected models). For models where there are stocks which do not share feedback loop relationships, we consider each subcomponent of inter-related feedback loops separately and we refer to each partition of the model's structure along those lines as a behavior origin feedback loop set that we explain in depth in the Defining Loop Scores section of this paper.

Literature review

The current state of the art in the use of mathematical methods for determining loop impact revolves around two methods. The first one is based on eigenvalue elasticity analysis, the second one, uses the pathway participation metric and causal pathways.

Eigenvalues and eigenvectors, specifically eigenvalue elasticity analysis (EEA)

Forrester (1982) was the first to document that eigenvalue elasticities could be used to explain the relative contributions of different loops in models of linear systems. Since then, the formal method of eigenvalue elasticity analysis (EEA) has been further developed and is used to determine how model structure produces the dynamic modes of behavior for the model, specifically those characterizing the state variables, (in SD called stocks) (Saleh, 2002), (Kampmann et al., 2006), (Saleh et al., 2010), (Oliva, 2016). Using EEA, the structure of a model is characterized by the eigenvalues and eigenvectors of that model. It may be demonstrated that the dynamic behavior of a linear systems model may be expressed by a linear combination of behavior modes, each characterized by a specific eigenvalue and weighed by a factor that depends on the eigenvector and the initial state of the model (Saleh et. al., 2010). EEA is applied to examine both link and loop significance with regard to the dynamic behavior of the model. It does so by identifying the relationship, expressed in the form of the elasticity, between the parameters that altogether make up the gains of an individual feedback loop (or link) in the model's structure and the eigenvalues (and eigenvectors) that characterizes the dynamic behavior of the model. The significance of a loop (or link) (relative to other loops (or links)) is expressed by the eigenvalue elasticity of its gain, i.e. how strongly a change in the gain impacts the eigenvalues. Note that this may not only be used to identify the root cause of a model´s behavior, but also the leverage points for controlling the system (policy entry points) if the model is an accurate representation of the system. Kampmann (2012) developed the concept of the independent loop set (ILS) which filters all of the loops in the model into a singular set of independent loops which represent the full behavior of the model so that the analysis can be effectively completed and interpreted. Oliva (2004) extended Kampmann's work on the ILS by developing an ILS composed only of geodetic loops which he termed the shortest independent loop set (SILS) which is the de-facto standard for determining which loops to analyze with EEA.

The purpose of EEA is more encompassing than the other methods discussed in this paper (loops that matter, or the pathway participation metric). EEA is a general method which describes the behavior state space of the model and speaks not only to what behavior the model is producing with a single set of input values, but what behavior modes are capable of being produced using any set of input values. According to Oliva (2016) the EEA approach satisfies Checkland and Scholes (1990) three E's criteria to assess performance. It is efficacious, efficient, and effective.

The downsides of the EEA method are that it is mathematically complex, requires a deep understanding of linear algebra, and may be applied effectively to only a very small subset of models unless they are modified (Saleh et al., 2010). Specifically, models must be linearized to make them well suited for such an analysis, a process that is hard to automate (though that is a problem actively being worked on) and which may change the simulation results. Oliva (2016) when analyzing his service quality model had to, among other changes, remove a stock in order to produce a full rank system matrix which was necessary to perform an EEA analysis and change model equations to ensure the model was continuously differentiable which did have an impact on simulation results.

Pathway participation metric (PPM) and other causal pathway techniques

The pathway participation metric (PPM) approach does not use eigenvalues to describe model structure. Rather, it focuses on the links between variables (Mojtahedzadeh et al, 2004). The starting point in the PPM approach is the behavior of a single variable, typically a stock. The behavior of that single variable is partitioned in time, based on phases where the variable maintains slope and convexity across time with the first and second time derivatives not changing sign (Mojtahedzadeh et al, 2004). This then limits the behavior of the variable at each of these phases to 7 patterns enumerated by Mojtahedzadeh et al, (2004). The PPM approach then determines dominance by tracing along the causal pathways between the stock under study and its ancestor stocks to determine which structure is most influential in explaining the pattern of behavior exhibited by that stock during the selected phase. Mojtahedzadeh et al., (2004) explains that it does this by determining the magnitude of the change in the net flow of the stock under study by making minute changes to that stock. The method then compares these changes in the net flow to determine the change with the largest magnitude in the same direction as the stock under study thereby identifying the most important (dominant) pathway governing the behavior of that stock during that phase.

Relative to the general EEA that yields results covering the entire behavior space (all modes of behavior that may potentially be produced by the model structure), PPM is considerably more specific. PPM, like the loops that matter method, is not aimed at analyzing the entire behavioral space of a model. PPM relies on the specific input currently impacting the model and the specific parameter values characterizing the model structure being analyzed. Using this method, therefore, we can only determine the impact of causal pathways based on the given set of values for inputs and parameters. The only way to determine what behavior modes a model is capable of producing using PPM is to specifically generate each of them (potentially via a Monte Carlo sensitivity analysis) and analyze each one independently.

Among the benefits of the PPM approach, relative to the EEA approach is that, for it to work, it does not require manipulation of the model (in theory), nor is there a need for linearization. Moreover, according to the research by Mojtahedzadeh (1996), repeated applications of the PPM method will cause a convergence on a unique piece of structure as the one most influential with regard to the behavior phase under study. Kampmann and Oliva (2009) state that one of the key benefits to the PPM method is its direct connection between behavior and structure.

Kampmann and Oliva (2009) have criticized PPM for its inability to clearly explain oscillatory systems and also because PPM can fail to identify structure when there are two pathways of similar importance (Kampmann and Oliva, 2009). Hayward and Boswell (2014) have responded to those criticisms by simplifying PPM via the loop impact method. The loop impact method can be implemented in a standard system dynamics model (and software) without any change in the underlying software by adding equations to the model. The key differences of the loop impact method as compared to PPM is that it does not look for dominant pathways, but instead focuses on the direct impacts that one stock has on another (Hayward and Boswell 2014). In addition, the loop impact method identifies instances where multiple loops are required to explain the behavior of a stock.

Expanding on the work done by Hayward and Boswell (2014), Hayward and Roach (2018) have developed a framework around the loop impact method couched in the mathematics of Newtonian physics to explain the model as a series of interacting forces. The stated purpose of this work is to provide a more intuitive and complete understanding of loop dominance in system dynamics models.

The loops that matter method

In this paper, we present the LTM (Loops That Matter) method which determines loop dominance for models of any size, complexity, or dimensionality. LTM is computed using values realized during a simulation and can therefore be used on continuous and discontinuous models without requiring linearization. We compute loop scores at each time in the simulation. The analysis of the relative scores at a particular time identifies which loops are dominating behavior at that time (dependent upon the model input values), and the display of the scores over time builds understanding of why the model behaves the way it does. Loop scores are computed as the product of the link scores for all the links involved in the loop. Because of this, and the definition of link scores, the amount of detail used in defining a model does not change the loop scores. There can be many variables included with simple equations, or a small number of variables with complex equations and the results will be the same.

We use the standard definition of a loop as a set of interconnections between variables in a model that form a closed path from a variable back to itself. The interconnections we refer to as links. Loop scores are computed as products of link scores, and the definition of a link score is tailored to this specific use. In particular, links from variables with unchanging input such as constants, are given a link score of 0 because change in the precedent variable does not contribute to the change in the antecedent variable. Thus, both loop and link scores are completely focused on realized rather than potential influence, and we will return to this point in the conclusion.

The link score computation has been designed for the purpose of determining loop dominance, specifically to be used in the loop power and loop score calculations described later. Link scores are not a general metric to describe the strength or importance of any specific link. The most obvious manifestation of the lack of generality is that a link from an unchanging variable, such as a constant (or even a variable which is temporarily constant), has a score that is definitionally 0 over the time periods where the variable is constant. This is so because the when links in loops do not change, the loop is inactive and therefore not currently of consequence. This is not to say parameters are unimportant. Even though the link score for all links from parameters to variables are 0, the parameter values determine link and loop scores throughout the model. This impact is discussed in depth during the discussion of the inventory workforce model.

We compute link scores for the influence of flows on stocks over time as well as the direct (instantaneous) algebraic influence of a variable on an auxiliary value. Conceptually these will be treated as the same, both being factors (i.e. multiplied) in the determination of a loop score, but they do require a slightly different computation as discussed below.

Defining link scores for auxiliary variables

To simplify the presentation, we will define the link score assuming there are two inputs (x and y) to the dependent variable z characterized by the equation $z = f(x, y)$.

This easily generalizes to the case where there are more (or fewer) inputs to, i.e. links associated with, z.

The link score for the link x → z is:

$$LS_x^z = \begin{cases} \left(\left|\frac{\Delta_x z}{\Delta \mathrm{z}}\right| \cdot sign\left(\frac{\Delta_x z}{\Delta \mathrm{x}}\right)\right), \\ 0, \qquad \Delta \mathrm{z} = 0 \text{ or } \Delta \mathrm{x} = 0 \end{cases} \tag{1}$$

Where Δz is the change in z from the previous time to the current time, Δx is the change in x, and $\Delta_x z$ is the amount z would have changed, conditionally, if x had changed the amount it did, but y had not changed. The first term in this equation represents the magnitude of the contribution, the second is an expression of the polarity.

The exceptions for no change in x or z are included for completeness, but are not important to the end goal of calculating a loop score. In the case of x, the magnitude goes to 0 and the sign, though not computable, is not relevant. In the case of z, any link using z will not be influenced by z, and so any loop going through z will have a loop score of 0.

The magnitude of the effect (force is a good analogy) that x has on z is relative to all of the effects on z. This is a dimensionless quantity, and if all of the effects are in the same direction, it is the fraction of the change in z that originates in a change in x. If the formulation of z is linear, then the values are restricted to the range between 0 and 1. When there are opposing forces, this number may be very large, but this does not harm the overall analysis of loop dominance as we will show below. The absolute value is used because the change in z could be in either direction due to the forces from other variables, regardless of the magnitude of the effect that x has, implying that the polarity can and would be wrong as demonstrated in Table 2.

The polarity of a link is defined as the sign of the partial difference at time t. This formulation is the same as the one used in Richardson 1995, though the formulation there was as a partial derivative, not difference. The polarity numerator is the same as it is for the magnitude, but the denominator is the change in x. When x does not change, the score is definitionally 0, though the magnitude would be 0 in any case. As noted earlier, scores for links emanating from constants will be 0. Nevertheless, the value of the constant may determine other link scores in the model.

We also define the link score to be 0 when z does not change (independent of cases where x does not change). This is both a strength and a weakness of our approach when exhibiting which loops matter. It is a strength because we typically struggle to understand what makes things change. Balancing forces, which would keep a value unchanging, may usually be understood from the steady state characteristics of a model. It is a weakness, because it hides dynamics that might manifest itself, but have not done so in our particular model run. For example, if a model is in an unstable equilibrium, all the link and loop scores would be 0, even though any perturbation would cause dynamics. This is discussed further in the final section of our paper.

Defining link scores for stocks

Stocks change only over time and, more importantly, they change as a result of flows, not changes in flows. This makes the computation of link scores for links going into stocks simpler. Assume the stock equation $s = \int(i - o)$ where s is the stock, i is the inflow, and o is the outflow. We assume a single inflow and outflow for simplicity of presentation, the generalization to multiple inflows and outflows is straightforward.

$$Inflow\colon LS_i^s = \left(\left|\frac{i}{i - o}\right| * 1\right) \quad Outflow\colon LS_o^s = \left(\left|\frac{o}{i - o}\right| * -1\right) \tag{2}$$

We use the same form as we do in previous link score for clarity, and again assume that the link score is 0 if the net flow $(i - o)$ is 0.

In this case a 0 inflow or outflow will result in a 0 link score. The link scores when a stock has only a small change because the inflow and outflow are nearly balanced will be large, and close in magnitude.

For models in which inflows and outflows are not explicit, but implicitly represented by an equation such as $avg = \int((input - avg)/st)$ for a smooth, we decompose the expression into an explicitly net flow such as $((input - avg)/st)$. The link score for this expression then ends up being the link score that matters, since the stock portion is 1 by definition.

Computational considerations

It is important to remember that link scores are being calculated with the intent of determining loop impact. If the LTM method detects a variable which does not change, the exercise of determining a link score for any link that has that variable as its dependent variable is irrelevant because the next link in the loop will yield a link score of 0, which means any loop that this link is a part of, is inactive for that particular time.

We make our computations as time progresses in the model. The first computation can be made only after the model has been initialized and moved forward in time. In the results we present, we use the model's dt or time step to determine how often to compute link and loop scores, this is most straightforward using the Euler integration method. Conceptually the computation could proceed at a longer or shorter sampling interval allowing it to work with a non-fixed time step integration methods such as Runge-Kutta.

Computing link score magnitudes

To calculate the link score magnitude for an auxiliary (non-stock) variable, z , we decompose its equation to identify its various inputs. For each of the inputs, e.g. x, we calculate $\Delta_x z$, assuming all the other inputs have not changed their values. Only this single input, x, is assumed to take its new value at the current time. If there is only one input, the link score is plus or minus 1 depending on the sign of $\frac{\Delta z}{\Delta x}$. If there are multiple inputs, the equation is recomputed using this one input value from the current time and the other input values from the previous time to yield $\Delta_x z$.

An implication of this is that the equation for z will be computed not just once, but once for every input at every time. This is a proportionally large computational burden, but is not that significant relative to other manipulations such as linearization and dynamics matrix decomposition.

Because the calculation uses only already computed values, the LTM method works on discontinuous as well as on continuous models.

For the example, in equation C=A+B there are two link score magnitudes that must be calculated, one for each link A→C and B→C. To calculate the link score magnitude, two $\Delta_x z$ values must be calculated, $\Delta_A C$ and $\Delta_B C$. The calculation of these can be seen in Table 1:

Table 1: Components necessary to calculate the link score magnitude for the links A→C and B→C based on the equation C = A+B.

Variable	Time 1	Time 2	Δz	$\Delta_x z$	$\frac{\Delta_x z}{\Delta z}$
A	5	7		2	2/3
B	5	6		1	1/3
C	10	13	3		

The absolute value is used in this calculation limiting this calculation's ability to determine polarity. This is done because the sign generated does not accurately report polarity in all cases. An example of this error can be seen in Table 2 where, without the absolute value, the link scores calculated have exactly opposite polarities from what is correct due to the aforementioned effects of the other variables on Δz or in the context of the equation in Table 2, ΔD.

Table 2: Demonstration of wrong polarity when calculating the link score magnitude for the links A→D, B→D, and C→D based on the equation D = (A+B)/C.

Variable	Time 1	Time 2	Δz	$\Delta_x z$	$\frac{\Delta_x z}{\Delta z}$
A	7	10		1	-5
B	2	4		0.67	-3.33
C	3	5		-1.2	6
D	3	2.8	-0.2		

Computing link score polarities

The link score polarity for an auxiliary (non-stock) variable is calculated in the same fashion as the magnitude. We decompose the equation of the variable we are studying to identify the different inputs. For each of the inputs we calculate $\Delta_x z$ by assuming all the other inputs had the values of the previous time, and only the input x has the value at the current time. The difference between the polarity and magnitude calculations is that for polarity we calculate the partial difference as Richardson (1995) does which means instead of putting Δz in the denominator, we put Δx. Table 3 demonstrates how to calculate a link score polarity for the example equation D = (A+B)/C.

Table 3: Demonstration of how to calculate the polarity term of the link score for the links A→D, B→D, and C→D based the equation D = (A+B)/C.

Variable	Time 1	Time 2	Δx	$\Delta_x z$	$sign\left(\frac{\Delta_x z}{\Delta x}\right)$
A	7	10	3	1	+1
B	2	2			
C	3	5	2	-1.2	-1

D	3	2.4

The calculation is continued for the above example shown in Table 3, in Table 4 where we show how this method for determining polarity can capture the shift in polarity when, in this case C goes negative.

Table 4: Continuation from Table 3, a demonstration of how to calculate the polarity term of the link score showing that the technique captures changes in polarity for the links A→D, B→D, and C→D based the equation D = (A+B)/C.

Variable	Time 2	Time 3	$\mathbf{\Delta x}$	$\boldsymbol{\Delta_x z}$	$\boldsymbol{sign\left(\frac{\Delta_x z}{\Delta x}\right)}$
A	10	13	3	1	+1
B	2	2			
C	5	-1	-1	-16	+1
D	2.4	-15			

Defining loop scores

The loop score is a normalized measure taking on a value between -1 and 1. It reports the polarity and instantaneous percentage contribution of a feedback loop to the behavior of all stocks in its behavior origin feedback loop set relative to the other feedback loops in that same subset of the SILS. By comparing loop scores we can determine which loops are dominant in the behavior origin feedback loop set under study.

A behavior origin feedback loop set is the collection of feedback loops where each feedback loop in the set affects at least one stock in a set of stocks where each stock in the set affects itself and all other stocks in that same set. A behavior origin feedback loop set represents a tightly coupled subset of the SILS. Often times there is only a single behavior origin feedback loop set for an entire model and the SILS does not need to be further partitioned (well connected models), but as shown in the inventory workforce case below that is not always true. Behavior origin feedback loop sets are necessary to make sure that we compare loops which affect stocks where the determinants of behavior for those stocks are shared. This is what allows the loop score to describe the percentange contribution of a feedback loop across multiple stocks.

Loop power is the product of all of the link scores in the loop. Note that this multiplies both the magnitude and the sign of the different link scores, with an odd number of negative links yielding a negative loop. The product is used following the chain rule and this also accurately represents the effects of a dead link in an otherwise 'active' loop. This has the consequence of assigning any loop with a dead link a loop power (and consequently a loop score) of 0. The magnitude of loop power represents the force that a loop is exerting to change stock behavior across all stocks in its loop set. A loop score is the result of normalizing loop power values across all loops in a feedback loop subset. The sign of a loop score or a loop power represents the polarity of the feedback loop.

This normalization process is critical to maintaining scores that are easy to work with. Because of the definitions of link scores, loop power values can become very large as an equilibrium is approached. This is shown below for the bass diffusion model. In this case, even though the power of the loops effectively approaches infinity, the transition from positive to negative loop dominance is smooth and clearly visible when using the loop score because it is normalized. The concept of the loop set is important only for this normalization process and ensures that loop power values are safely compared. An example of incomparable loop power values is shown below in the case of the inventory workforce model where the feedback loop B3 is not comparable with the others since it does not share stocks with B1 or B2. Both the loop score and loop power concepts are rigorously defined below in equations (3a and 3b).

$$Loop\ Power(L_x) = \left(LS_{i_1}^{d_1} \cdot LS_{i_2}^{d_2}\ \ldots \cdot LS_{i_n}^{d_n}\right) \qquad (3b)$$

$$Loop\ Score_{L_x} = \left(Loop\ Power(L_{x,})/\textstyle\sum_{y=0}^{n} |Loop\ Power(L_y)|\right) \qquad (3b)$$

Applications of the LTM method to the bass diffusion model

A good model to demonstrate the LTM process is the bass diffusion model depicted in Figure 1 below.

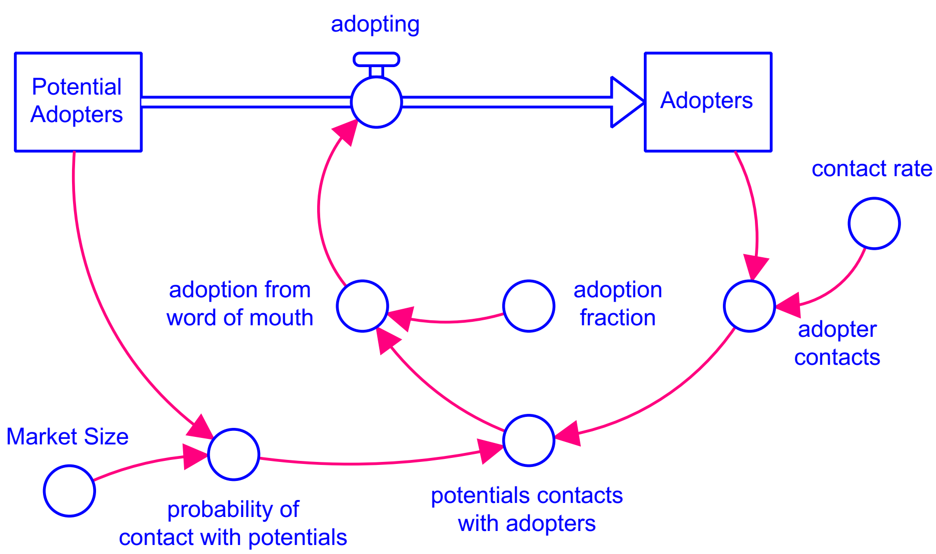

Figure 1: The stock and flow structure of the bass diffusion model analyzed

This version of the bass diffusion model runs from Time 0 to Time 15 with the inflection point reached between time 9.5625 and 9.625. It contains two loops, one balancing and one reinforcing.

- Balancing (B1)
  - probability of contact with potentials
  - potentials contacts with adopters
  - adoption from word of mouth
  - adopting
  - potential adopters

- Reinforcing (R1)
  - adopter contacts
  - potentials contacts with adopters
  - adoptions from word of mouth
  - adopting
  - adopters

In Table 5, the calculation of the loop power of B1 at specific points in time is demonstrated and compared to the loop power of R1.

Table 5: Loop power in the bass diffusion model calculated to 4 significant digits

Link	T_1	$T_{9.5}$	$T_{9.5625}$	$T_{9.625}$	T_{15}
Probability of contact with potentials → potentials contacts with adopters	0.000	9.958	9358	10.91	1.000
Potentials contacts with adopters → adoption from word of mouth	1.000	1.000	1.000	1.000	1.000
Adoption from word of mouth → adopting	1.000	1.000	1.000	1.000	1.000
Adopting → potential adopters	-1.000	-1.000	-1.000	-1.000	-1.000
Potential adopters → probability of contact with potentials	1.000	1.000	1.000	1.000	1.000
B1 Loop Power	0.000	-9.958	-9358	-10.91	-1.000
R1 Loop Power	1.000	11.46	9806	10.41	0.000

The only link score which shows a change (the active link), in the loop B1, is the first one listed, located at the point of the non-linearity, the junction between the reinforcing and balancing feedback loop. Plotting out the loop scores for this model yields the graph portrayed in Figure 2 which demonstrates the common knowledge about the shifting feedback loop dominance in the bass diffusion model.

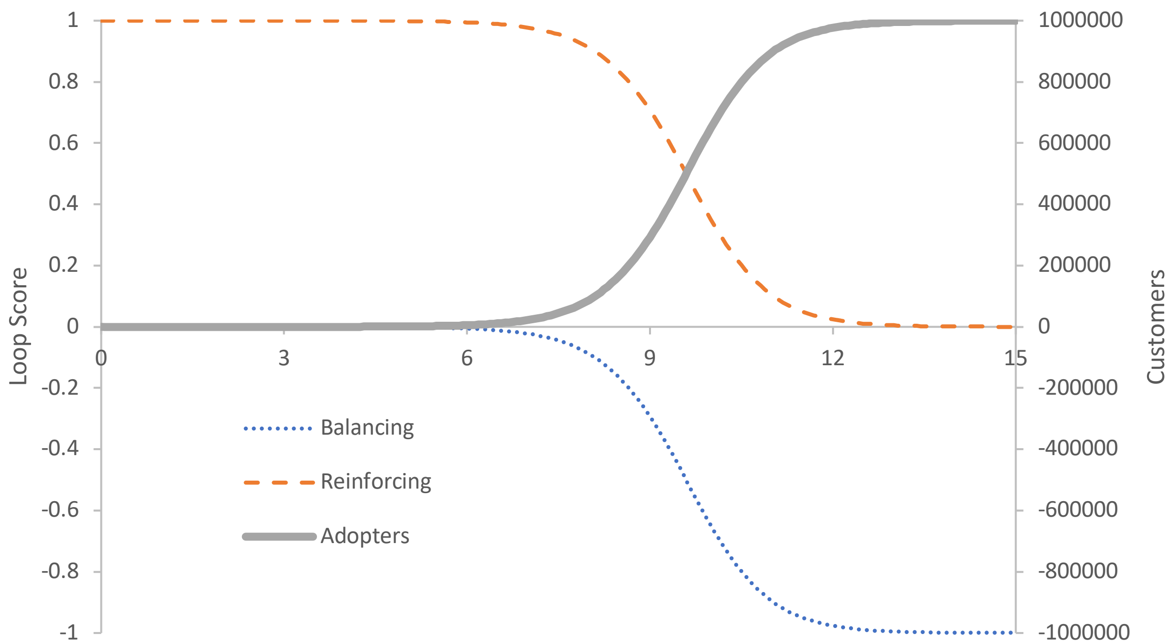

Figure 2: Bass diffusion loop scores plotted over time against Adopters.

Another interesting insight from performing these calculations is that the loop power values of these two feedback loops are highly variable providing a justification for the normalization to loop scores. Figure 3 plots on a logarithmic scale how each of the feedback loops gains and loses power relative to the other over time, producing the shifting feedback loop dominance.

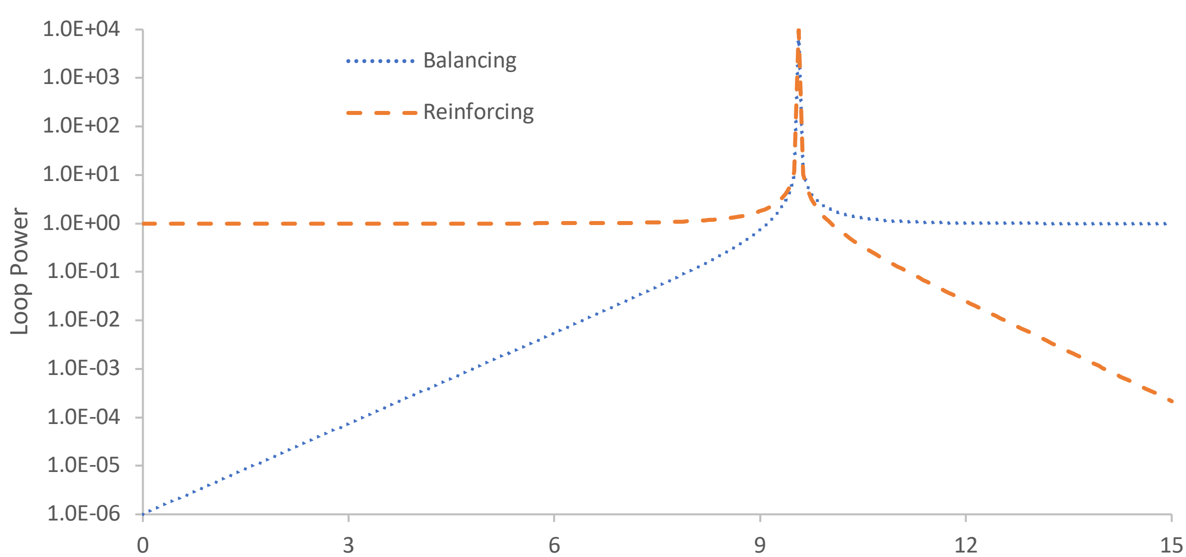

Figure 3: Bass diffusion loop power values

Understanding the yeast alcohol model

The yeast alcohol model has been widely studied using EEA and PPM in spite of the fact that the model contains a conceptual flaw (see below). Therefore it is a good candidate to use to demonstrate the efficacy of the LTM method, to show how it compares to previous research (Saleh, 2002; Güneralp, 2006; Phaff et al., 2006; Mojtahedzadeh, 2008; Hayward and Boswell, 2014), and to reveal flaws in model formulations.

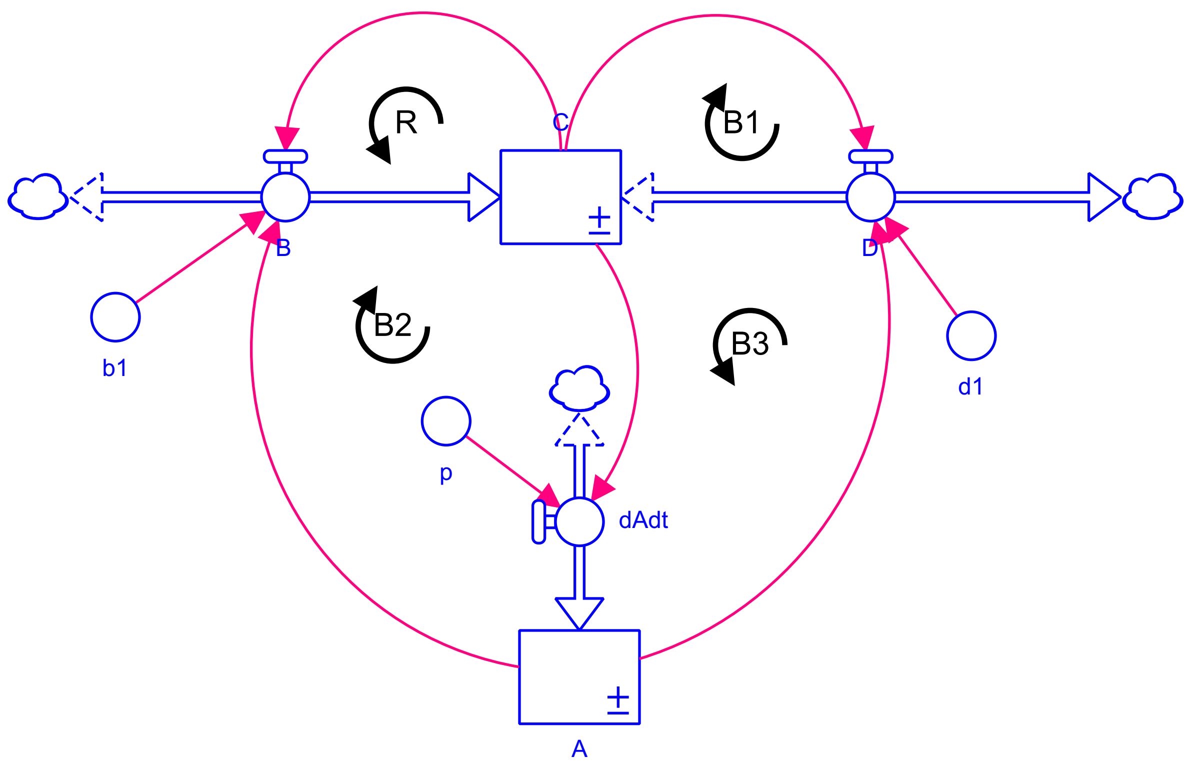

Figure 4: Yeast alcohol model

Figure 4 shows the structure of the model as analyzed which was done so using a DT of .5. The model structure; B = C*(1.1-0.1*A)/b1, D=C*EXP(A-11)/d1, sAdt= p * C , is initialized as such; A=0, B=1, b1=16, d1=30, and p=0.01. It contains 4 loops, all in a single feedback loop subset. Loop R, represents the birth of the cells C, characterized by the fertility, b1. Notice that there is a flaw in the formulation of B in this model causing B to take negative values and the polarity of R to change so that it acts as an additional "deaths loop" under conditions of high levels of alcohol A. Loop B1 represents the natural death of the cells. The main link in Loop B2 represents the slowing of the birth of cells due to the presence of alcohol. The main link in Loop B3 represents the increasing death of cells due to the presence of alcohol. This model produces the overshoot and collapse behavior seen in Figure 5 which matches the behavior generated by Phaff et al. (2006) and Mojtahedzadeh (2008). Hayward and Boswell (2014) use the same parameterization as us and the others, but appears to have used a Stella version of this model in which uniflows were applied for B and D which hides the formulation flaw and causes their results to differ.

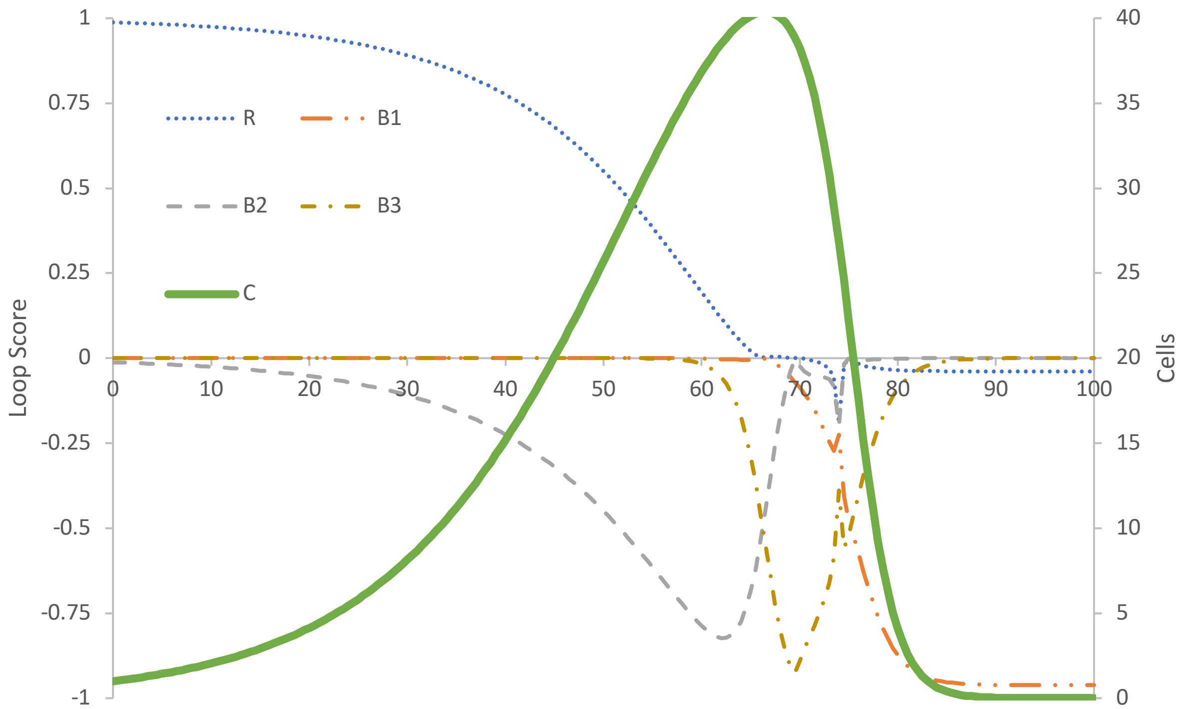

Figure 5: Yeast Alcohol loop scores plotted against C[1]

Table 6: Dominant loops in yeast alcohol model.

Time range	Phase 1: 0-51.5	Phase 2: 52-66	Phase 3: 66.5-75	Phase 4: 75.5-100
Dominant loop	R	B2	B3[1]	B1

In Table 6 the dominant loops for each phase of the model's behavior have been recorded. Comparing these results with Ford's (1999) behavioral approach as applied by Phaff et al. (2006), we identify the same exact 4 phases and agree with the analysis in principal, but as Hayward and Boswell (2014) and Mojtahedzadeh (2008) also point out, Phase 3 is dominated by B3 rather than B2 and B3 together as Phaff et al.'s implementation of Ford's analysis would suggest. From this we can conclude that our analysis of this model matches Ford's behavioral approach with the noted discrepancy. Our results match exactly those of Mojtahedzadeh's 2008 application of the PPM method.

[1] At time 74 no single feedback loop is dominant because this is the point where R is at its strongest as a balancing feedback loop. After time 70 when the birth rate is negative R is acting in a similar fashion as B1. At Time 74 summing the strength of R & B1 yields a loop score which is stronger than B3, but still not over 50%, B3 is the single strongest feedback loop at that exact moment and we therefore consider it alone to be dominant across phase 3.

When we compare our results to Hayward and Boswell's (2014) PPM based loop impact method, we agree in principal (to the extent their analysis matches the others), but a true match cannot be confirmed because of problems that arise due to the differences in the models that we now compare. We posit that their shifts between phases happen at different times because of their use of uniflows. For instance, they note that the shift between B2 and B3 occurs before the peak in C which does not match any of the other analyses. We posit that this is not due to an inherent flaw in their analysis technique, but rather in application because they are using uniflows in their model which changes both structure and behavior. Their analysis does not include the series of complex hidden links and potential feedback loops that in effect dictate C due to the fact that, in Stella, the inclusion of those uniflows creates hidden links between the flows and their stocks. In addition, the use of a uniflow for B does not allow R to switch from a reinforcing to a balancing feedback loop as A increases which does occur with this parameterization during Phases 3 and 4. Though on a macro level, our analyses do agree about the dominance in the four phases, but it is not clear if we agree about the timing of those shifts relative to the changes in behavior of C, and we do not show a combined dominance of B1, B2, and B3 at the inflection point of the growth in C as they report.

Our results in Figure 5 and Table 6 match the EEA analysis of this model performed by Phaff et al. (2006). Phaff et al. conclude that the behavior of Phase 1 is dominated by R with B2 restraining the growth of C. In phase 2 they remark that B2 is now dominant, but R is still a significant factor in explaining C which can be seen in Figure 5 because B2 has a loop score less then -.5 and R is the only other active loop until time ~60 where B3 starts becoming active in preparation for phase 3. In phase 3, they point to B1 and B3 together as describing the behavior of C which is true according to our analysis, but our analysis finds that B3[1] is dominant throughout that time period. They then find that during phase 4 B1 is dominant over B3 which can be seen in Figure 5 as B1 starts growing quickly at the end of phase 3 reaching a loop score of nearly -1 shortly after the start of phase 4.

Using LTM to understand oscillations

Identifying proper oscillatory behavior as the outcome of a negative feedback loop rather than shifting feedback loop dominance is an important benchmark for determining the utility LTM. A model that illustrates our analysis well is the inventory workforce model which appears in Figure 4.

.

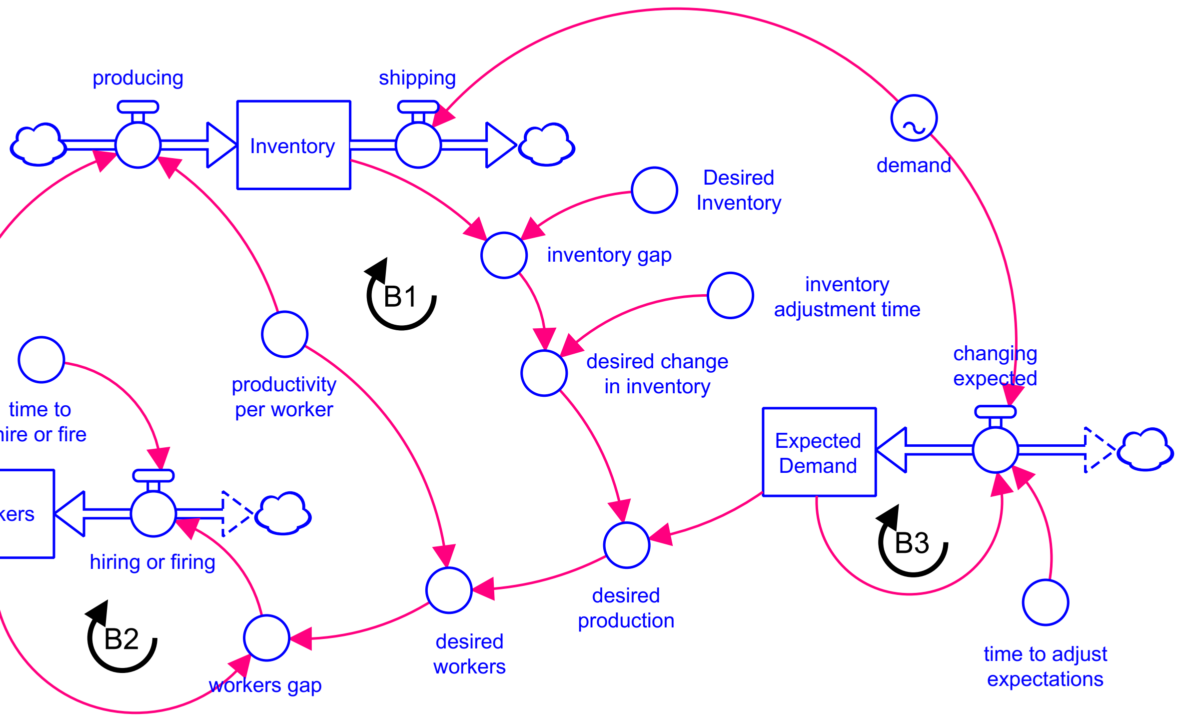

Figure 6: Inventory workforce model

This model runs from Time 0 to Time 60 and has only three balancing feedback loops that appear in two different behavior origin feedback loop sets. The graphical function inside of the 'demand' variable acts like a step function, triggering a single increase in demand between times 1 and 2 which sets off a dampened oscillation in both Workers and Inventory.

- Loop Set 1
  - Major Balancing (B1)
    - Inventory
    - inventory gap
    - desired change in inventory
    - desired production
    - desired workers
    - workers gap
    - hiring or firing
    - Workers
    - producing
  - Minor Balancing (B2)
    - Workers
    - workers gap
    - hiring or firing
- Loop Set 2
  - Expected Demand Loop (B3)

▪ Expected Demand
▪ changing expected

The two loops in this model that contain the stocks with the oscillatory behavior are B1 and B2 of Loop Set 1. As seen in Figure 7 B2 is responsible for the oscillation, the longer it is active the more pronounced the oscillations are. That tells us that by increasing the delay in time to hire or fire, we increase the cumulative power of the B2 loop causing the oscillations to be more pronounced and to last longer. The loop score of B1 tells us that the dominant mode of behavior in this model is to find a stable equilibrium. These results diverge significantly from those of Mojtahedzadeh (2008) and Hayward and Roach (2018) who explains the behavior of similar inventory workforce models as the shifting feedback loop dominance of B1 and B2. Also of note in Figure 7 is the time period before the shock in demand: Then this model is in equilibrium and, therefore, LTM cannot inform our analyze the model as all link scores are considered 0.

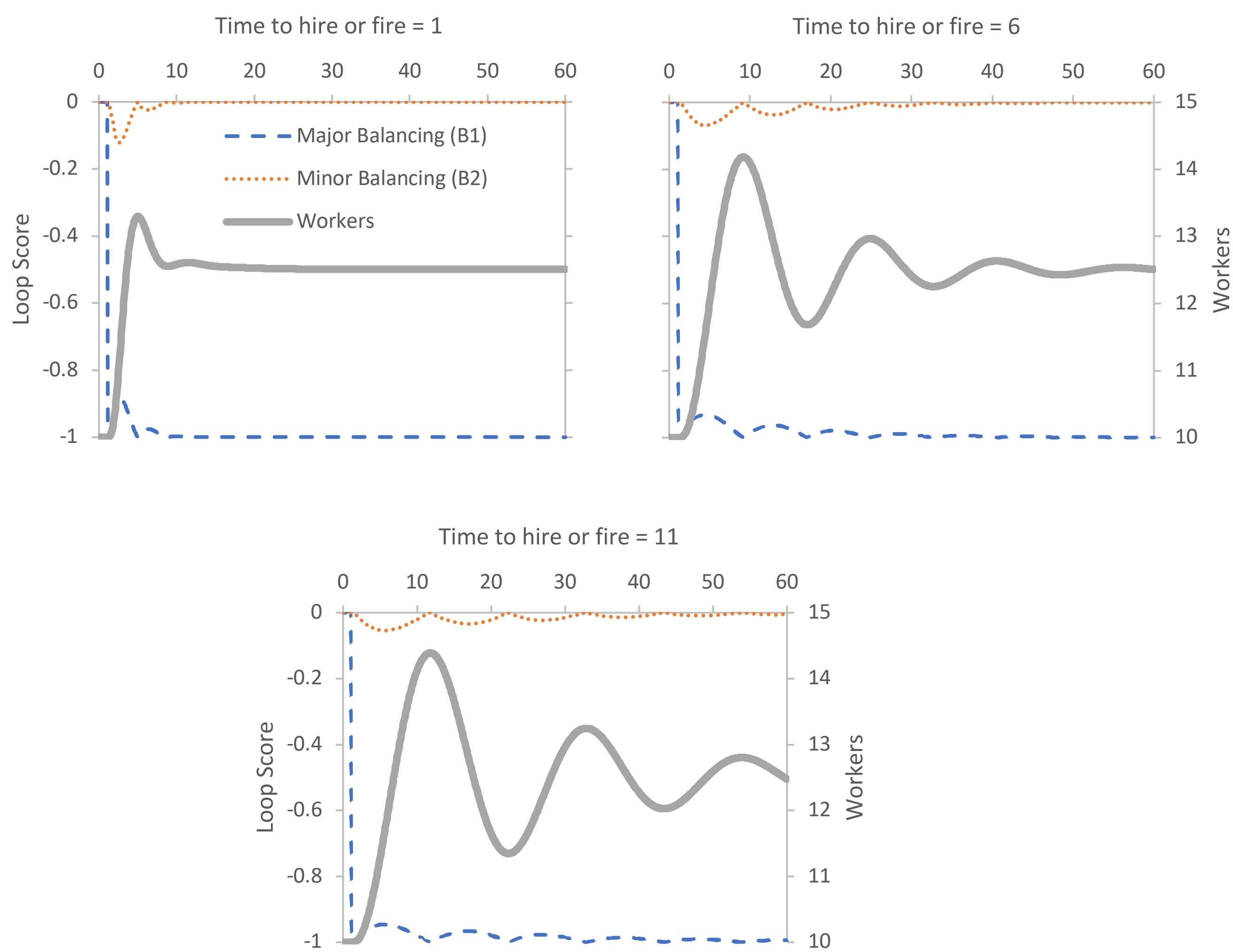

Figure 7: Results of LTM analysis of the Inventory Workforce model showing the effect of time to hire or fire on loop impact and Workers.

Conclusions

As demonstrated, the LTM method provides an easy way to understand and identify, through computation, which feedback loops in a model dominate or in other words describe more of the behavior at each point in time. Dominance is loop set wide, based on the effect on all variables, and is typically driven by the stocks that are changing proportionally most quickly, and the feedback acting in support of that movement. As we have seen in the examples, this measure of dominance correlates well with our structure-based understanding of relatively simple systems. The LTM method has several considerable advantages outlined here:

Most importantly, this method is generally applicable to all models without any manipulation or modification and the format of the results of the analysis are simple, easily interpretable graphs of behavior over time. LTM makes use of the existing skillsets of all modelers and most model consumers and is thus easily accessible. As for its general applicability, at present the current implementation of the LTM method works with any non-arrayed XMILE compliant model containing the most common built in functions and graphical nonlinear relationships. Since the method uses only values calculated by the model as it runs, the structure of the model never needs to change to accommodate LTM analyses. Finally, since the field is so used to mapping out and analyzing behavior over time, it is very beneficial that we conduct our analysis over time in the same format so that it is easier to parse, compare, and understand loop dominance results.

The second key advantage to the LTM method is its relative simplicity. The method as currently developed does not use complex mathematical constructs which are not already in use by the majority of practitioners. From a mathematical perspective the concept of the $\Delta_x z$ is the most difficult part of the method because of its unfamiliar terminology, and not necessarily because of any inherent complexity in the idea itself. The advantage of a simpler method is that it can be understood by all practitioners so that when it comes time to apply the method, practitioners can know 'what it is doing' due to the transparency of the method.

The third key advantage is that this method is relatively easy to implement in existing simulation engines, especially those that the authors have taken part in constructing. This means its uptake should be relatively painless by software vendors in the field if they so choose. In addition to the key advantages listed above, the LTM method allows for the development of new and exciting visualization tools including animated stock-and-flow diagrams where the links and flows change color and size due to changes in polarity or link strength, - in response to Sterman in Business Dynamics (2000). Going even further, the LTM method allows for the possibility of automated CLD generation and animation. Because the LTM method is able to say on a link-by-link basis which are the key (dynamic) links in the model, it is possible, using the method, to automatically generate a CLD collapsing all of the 'unimportant' static links with scores of 0, +1.0, or -1.0 into links which are conveying a change, which exist at the junction points of the loops. This will allow for the automated generation of structurally correct, minimal CLDs that accurately portray the stuctural components that predominantly produce the dynamics of the model, - laid out according to best practices.

While using only computed variable values is a strength of LTM, it also means that only realized, and never latent, model behavior can be analyzed. Thus, unlike EEA techniques, LTM on its own is unable to provide a general model level of understanding. Some of this understanding, including behavioral sensitivity to parametric changes, can be gained through a combination of sensitivity and LTM analysis such as suggested below.

Additionally, the LTM method is not able to determine loop impact without a change in the model state. As models approach equilibrium, we can see the loop scores balance one another even as they become unbounded, but when a model is in equilibrium all loop scores are 0. Therefore, models in equilibrium cannot be analyzed using LTM. An example of this is a simple "bathtub" population model where the birth fraction equals the death fraction. The limitations of the link score definitionally define the loop power for both loops to be 0 because there is no change across a timestep, dt. A potential solution to this problem from a purely methodological perspective is to start introducing minute changes in these situations in order to measure their impact on loop dominance, but the authors are wary of this approach because of the impacts thas has on discrete and discontinuous models. This is an area which requires future study. Currently, models in equilibrium are much better analyzed using EEA methods like those suggested by Oliva (2016). An alternative approach would be for the model author to offset the model state from its equilibrium using a STEP function or other modeling construct for making constants vary due to exogenous forces.

An additional strength and weakness of the LTM method is that it focuses exclusively on endogenously generated behavior. Such a focus is a hallmark of System Dynamics, but is problematic for models where behavior is driven through external forcing functions that dominate the effects of feedback in the model. Loop score dominance, in this case, may have little to do with behavior generated. Models of this sort are currently much better analyzed using the loop impact method of Hayward and Boswell (2014).

There are a variety of interesting extensions to LTM that combine it with other analysis techniques. The most obvious one is to combine it with sensitivity analysis so that the realized behavior sets encompass the potential behavior sets. For example, using extreme condition testing could, combined with LTM, be used to show that the model is producing the right results for the right reasons. LTM could also be combined with optimization, for example using optimizers to maximize/minimize loop scores. This would allow practitioners to maximize the impact of favorable loops while minimizing the impact of unfavorable loops in order to automatically generate better, more robust policy recommendations. Another area of study would include loop scores in the outputs of Monte-Carlo sensitivity analyses which would allow us to measure the robustness of loop impact to policy or parameter changes. Monte Carlo analysis could also be used to measure the sensitivity of loop power to changes in parameter values.

Finally, it is necessary to testing and analyse larger and more varied models id we are to increase our confidence in the general utility of the LTM method. In general, though, the authors are hopeful that the techniques laid out in this paper, will offer a significant utility and enhance our analysis and understanding of models for years to come.

References

1. Checkland P, Scholes J. 1990. *Soft Systems Methodology in Action*. Wiley: Chichester.
2. Eberlein RL. 1984. Simplifying dynamic models by retaining selected behavior modes. PhD Thesis, Massachusetts Institute of Technology, Cambridge, MA.
3. Davidsen P. 1991. The Structure-Behavior Graph. The System Dynamics Group, MIT, Cambridge, MA.
4. Ford D. 1999. A behavioral approach to feedback loop dominance analysis. *System Dynamics Review* **15**(1): 3–36.
5. Forrester N. 1982. A Dynamic Synthesis of Basic Macroeconomic Theory: Implications for Stabilization Policy Analysis. PhD Thesis, Massachusetts Institute of Technology, Cambridge, MA.
6. Graham AK. 1977. Principles of the Relationship Between Structure and Behavior of Dynamic Systems. PhD Thesis, Massachusetts Institute of Technology, Cambridge, MA.
7. Goncalves P. 2009. Behavior modes, pathways and overall trajectories: eigenvector and eigenvalue analysis of dynamic systems. *System Dynamics Review* **25**(1): 35–62.
8. Güneralp B. 2006. Towards coherent loop dominance analysis: progress in eigenvalue elasticity analysis. *System Dynamics Review* **22**(3): 263-289.
9. Hayward J, Boswell GP. 2014. Model behaviour and the concept of loop impact: a practical method. *System Dynamics Review* **30**(1): 29–57.
10. Hayward, John, and Paul A. Roach. 2017. Newton's laws as an interpretive framework in system dynamics. *System Dynamics Review* **33**(3-4): 183-218.
11. Kampmann CE. 2012. Feedback loop gains and system behaviour (1996). *System Dynamics Review* **28**(4): 370–395.
12. Kampmann CE, Oliva R. 2006: Loop eigenvalue elasticity analysis: Three case studies. *System Dynamics Review*. **22**(2): 141-162.
13. Kampmann CE, Oliva R. 2009. Analytical methods for structural dominance analysis in system dynamics. In Meyers RA (ed.) *Encyclopedia of Complexity and Systems Science* Springer, New York NY, 8948–8967.
14. Mojtahedzadeh MT. 1996. Structural Analysis of the URBAN1 model. University at Albany, SUNY, NY 12222, USA.
15. Mojtahedzadeh M, Andersen D, and Richardson GP. 2004. Using Digest to implement the pathway participation method for detecting influential system structure. *System Dynamics Review* **20**(1): 1-20.
16. Mojtahedzadeh M. 2008. Do parallel lines meet? How can pathway participation metrics and eigenvalue analysis produce similar results?. *System Dynamics Review.* **24**(4): 451-478.
17. Moxnes, Erling, and Pål I. Davidsen. 2016. Intuitive understanding of steady-state and transient behaviors. *System Dynamics Review* **32**(2): 128-153.
18. Oliva, R. (2004). Model structure analysis through graph theory: partition heuristics and feedback structure decomposition. *System Dynamics Review*, **20**(4): 313-336.

19. Oliva R. 2016. Structural dominance analysis of large and stochastic models. *System Dynamics Review* **32**(1): 26-51.
20. Phaff HWG, Slinger JH, Güneralp B, and van Daalen CE. 2006. Investigating model behavioural analysis: A critical examination of two methods. In *Proceedings of the 24th International Conference of the System Dynamics Society*. Nijmegen, The Netherlands, System Dynamics Society.
21. Richardson GP. 1995. Loop polarity, loop dominance, and the concept of dominant polarity. *System Dynamics Review* **11**(1): 67–88.
22. Richardson GP. 1996. Problems for the future of system dynamics. *System Dynamics Review* **12**(2): 141-157.
23. Sterman JD. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World.*Irwin/McGraw-Hill, Boston.
24. Saleh MM. 2002. The Characterization of Model Behavior and its Causal Foundation. PhD Thesis, University of Bergen, Bergen, Norway.
25. Saleh M, Oliva R, Kampmann CE, Davidsen P. 2010. A comprehensive analytical approach for policy analysis of system dynamics models. *European Journal of Operational Research* **203**(3): 673-683.