Automated Discovery of Polarity from Data in System Dynamics Context

Nefel Tellioğlu
nefel.tellioglu@gmail.com
Boğaziçi University
Industrial Engineering Department
Bebek, Istanbul, Turkey

Yaman Barlas
ybarlas@boun.edu.tr
Boğaziçi University
Industrial Engineering Department
Bebek, Istanbul, Turkey

In initial stages of modeling ('conceptualization'), to decide about the direction of causal effects, System Dynamics modelers generally use *a priori* knowledge (real-life experience or scientific literature), rather than automated data analysis (Sterman 2000; Barlas, 2002). This study investigates to what extent it is possible to derive initial model construction from real data, since model construction time and model subjectivity can be both reduced by the automated analysis of historical data.

We assume that we have real-life (non-experimental) dynamic field data about multiple system variables influencing a given effect variable. In this case, automated polarity discovery becomes a nontrivial, sometimes impossible problem. Our research purpose is to find whether the links are positive or negative between variables by using historical field data under the assumption that we assume/know all the cause variables that have a significant influence on the effect variable (see Figure 1). We also assume that all the causal relations (functions) are either *monotonically* increasing or decreasing, which is typically an accepted best practice in SD formulations (see Sterman 2000 for instance).

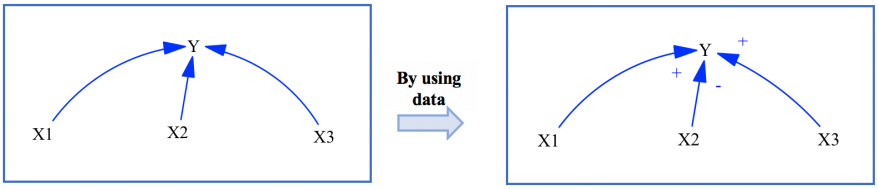

Figure 1: Discovering the Polarity of Relations

In the literature, correlation analysis or a correlation-based analysis such as structural equation modeling (SEM) and generalized additive models (GAM) can be used to extract polarities, since these methods are commonly used to calculate strength and direction of the association between variables (Tarka, 2018; Hauke & Kossowski, 2011). When it comes to the applicability of these methods in SD causal analysis, the assumptions of the methods raise several difficulties: The correlation analysis may fail when there is multicollinearity among variables (Farrar, 1967), strong non-linearities and multiplicative effect relations. Thus, there are no effective approaches to derive the direction of nonlinear causal effects from data, even when the causal variables are assumed to be known.

To solve the problem, we developed a two-phased algorithm that we call *discoverpolarity*. We implemented the algorithm in R programming language and tested it with different datasets that satisfy the data assumptions. Synthetic datasets are generated in R language. Firstly, we specify an underlying structure for each experiment and generate a dataset by using that structure. Then, we assume that we do not have the information about the link polarities. By means of known causal relations and generated dataset, we try to (re)discover the unknown link polarities with the developed algorithm. We used two 'benchmark' methods that require minimum additional information and return the most reliable results in our tests: Shape Constrained Additive Model (SCAM), and Pearson's partial correlation analysis. To apply SCAM, the monotonic shape restrictions (increasing or decreasing) of effect functions must be given. Therefore, we fit $2^{(\text{cause variables})}$ SCAM models having different shape restrictions to our datasets, then, we compare models to each other by two performance measures which are root mean square error and mean absolute percentage error. to obtain monotonic shape restrictions of the best model. We assume that the scope of the variables, the boundary of the system, is well-chosen for the problem and data is noise-eliminated and has no missing values We also assume that we know the range of cause variables (X_i) in which $dfi(Xi)/dX_i$ is significantly different than 0.

The first phase of the algorithm works with proof by contradiction. It compares changes in the effect variable (*Y*) while moving among observed data points in the causal domain. Then, it eliminates the corner points that cannot have the maximum value of *Y*. The algorithm takes the "causal variable" and respective "effect variable" values as main inputs. Then, it compares the differences in causal variables and checks the signs of each difference. Finally, it eliminates the impossible cases from *all the possible polarities list* (given by each *corner* in the causal domain) which has $2^{(\text{cause variables})}$ possibilities. For example, the *list* has 4 values when there are two cause variables: (+, +), (+, -), (-, +), (-, -). The basic process of the algorithm eliminates the *corner* with the same polarity of the causal variable movement while *Y* value is decreasing, and it eliminates the corner with the opposite polarities of the causal variable movement while *Y* value is increasing. Therefore, we analyze all of the paired combinations of the data points. In total, we check the sign of changes in *n*(n-1)/2* points for a dataset composed of n points.

If we end up with multiple options at the end of the first phase, the algorithm moves into the second phase. At this point, the algorithm has a default assumption: in the range of the given limits of the variables, the maximum ceteris paribus changes created by all X_i's on *Y* are (approximately) the same. In other words, influence importance weights of X_i's on *Y* are the same. However, if the modeler knows the maximum ceteris paribus impact of each variable which is significantly different than each other, then she must give these impact levels (importance weights) as an input, *varImp*.

In the second phase, the algorithm takes the increasing or decreasing amounts of change of cause and effect variables into account. We already know that a significant change of a cause variable must produce a change (ceteris paribus) in the effect variable. Therefore, when a cause variable X_i does not change much but other cause variables $X_{j,\, j \neq i}$ change significantly, then we can say that the change in the effect variable results from $X_{j,\, j \neq i}$. In this step, the algorithm subtracts the minimum of the variable from the variable value and then divides the variables by their observed ranges (maximum-minimum values). By doing so, it obtains relative changes over a range of (0, 1). Then, if variable importance weights are different, the algorithm multiplies the relative weights (minimum is set as 1) with the difference values. In the new causal domain, the algorithm takes the absolute of the differences of each data point to obtain the relative absolute difference values for each causal variable. Then, the algorithm searches for points close to each other with respect to each cause variable. The points that have relative difference values smaller than *threshold1* (default value is set as 0.15) are considered as 'close' according to a cause variable and are considered in *first candidate indexes*. Then, for each *first candidate index*, the algorithm checks the absolute difference values of other cause variables. If the ratio of "*the absolute difference value of candidate cause variable*/*sum of the absolute difference value of all causal variables in the candidate index*" is smaller than *threshold2* (default value is set as threshold1), then the algorithm sets the difference value of the candidate cause variable at that point to zero. In other words, it concludes that the change in *Y* results from the other cause variables. By doing so, the algorithm eliminates more corners which do not include the maximum Y value.

When the algorithm is applied, the modeler must decide on the threshold values used in the algorithm. This is the trickiest part of the algorithm since the proper values of the thresholds may differ from one dataset to another. *Threshold1* is used to select first candidate indexes for each cause variable to set them to zero. *Threshold2* is used to select indexes from the first candidate indexes list to set their values to 0. *Threshold3* is used to select indexes for the effect variable to set the effect variable to zero. The default threshold values are 0.15. *varImp* is a variable importance list (weights of impact of each cause variable). *varImp* of all the variables is 1 as default.

We generated 7680 synthetic datasets to test the algorithm. The number of data points is 100 in all datasets. In total, 12 different formulation structures (additive or multiplicative formulations with two, three, or four cause variables) are used to generate the datasets. We classified the results according to two questions: "In what fraction of the results, is there the correct polarity in the returned results?" and "In what fraction of the results, the algorithm returned only (uniquely) the correct result?". The answer to these two questions is the same for the benchmark methods since they only return one solution - correct or wrong. In Table 1, the overall results are given. In Table 2, the proportion of correctly returned

results among datasets with the different number of cause variables are given. The histogram of *discoverpolarity* results for the datasets with 3 and 4 cause variables are given in Figure 2.

Table 1: Proportion of Correctly Returned Results in All Datasets

	Fraction of Correct Results	Fraction of Correct & Unique Results
Discover Polarity	0.986	0.862
Pearson's Correlation	0.590	0.590
SCAM	0.865	0.865

Table 2: Proportion of Correctly Returned Results in All Datasets

	Pearson	**SCAM**	***discoverpolarity***	
			All Results	Unique Results
2 cause variables	0.980	0.991	0.999	0.997
3 cause variables	0.442	0.876	0.980	0.864
4 cause variables	0.281	0.689	0.978	0.685

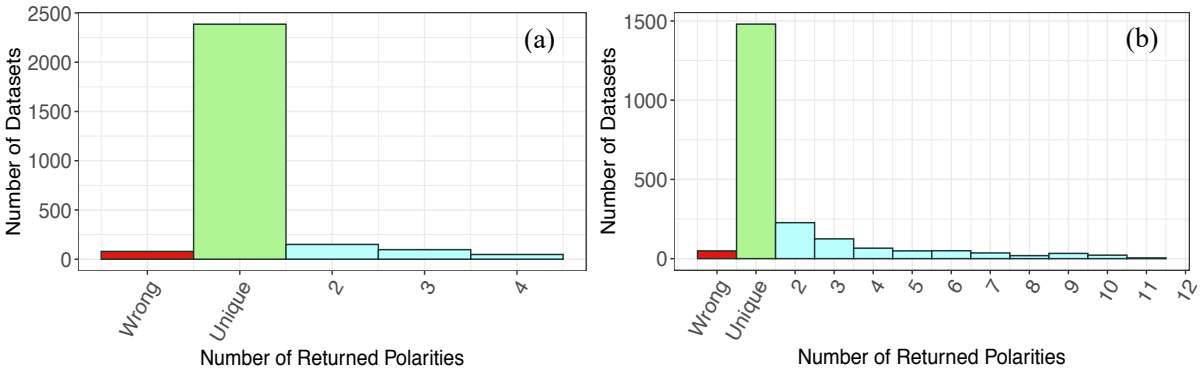

Figure 2: Histogram of *discoverpolarity* results in the datasets with 3 (a) and 4 (b) cause variables

The results show that *discoverpolarity* outperforms partial correlation analysis in terms of "returning correct result" and "returning correct & unique result" when there is collinearity and/or many cause variables in the dataset. Our algorithm outperforms SCAM in datasets with underlying multiplicative structures (see Table 3). This is expected since SCAM does not have a monotonic structure for most of the multiplicative formulations.

Table 3: Proportion of Correctly Returned Results of Datasets with Multiplicative Structure

	Fraction of Correct Results	Fraction of Correct & Unique Results
Discover Polarity	0.975	0.840
Pearson's Correlation	0.587	0.587
SCAM	0.799	0.799

In cases where the algorithm returns multiple possible polarities, the modeler must choose the most plausible of the returned polarities. In addition, the modeler may need to check the link polarities that are eliminated in the second phase of algorithm if the number of the observed differences to support these eliminations are only a few. Another limitation of *discoverpolarity* is that the modeler must decide on the threshold values used in the algorithm. For some datasets, the default values of the thresholds may eliminate all possible polarities, including the true polarity. This may happen if the thresholds used in *discoverpolarity* are high relative to the nature of the data. In such cases, the modeler must decrease the threshold values to obtain a reasonable result. On the other hand, if *discoverpolarity* returns multiple

possible polarities, the modeler may prefer to increase the threshold values. However, usage of too high thresholds can cause the elimination of true polarity because of high non-linearity in the relations.

In further research, we plan to focus on decreasing the sensitivity of *discoverpolarity* to its threshold values and other input parameters to make it more robust. It may also be possible to combine the strengths of *discoverpolarity* and SCAM methods to come up with a method stronger than each. Finally, by using not only the signs of the differences but also the magnitudes of these differences, one may be able to estimate the mathematical forms (additive, multiplicative, or hybrid) of the causal relations.

REFERENCES

Barbiero, A. and Ferrari, P.A., 2017. An R package for the simulation of correlated discrete variables. Communications in Statistics-Simulation and Computation, 46(7), pp.5123-5140.

Barlas, Y., 2002, System dynamics: systemic feedback modeling for policy analysis. In Knowledge for Sustainable Development—An Insight into the Encyclopedia of Life Support Systems, UNESCO-EOLSS: Paris and Oxford, pp. 1131–1175.

Bürkner, P.C., Charpentier, E., 2018, Monotonic Effects: A Principled Approach for Including Ordinal Predictors in Regression Models.

Farrar, D.E. and Glauber, R.R., 1967. Multicollinearity in regression analysis: the problem revisited. The Review of Economic and Statistics, pp.92-107.

Gupta, M., Cotter, A., Pfeifer, J., Voevodski, K., Canini, K., Mangylov, A., Moczydlowski, W., van Esbroeck, A., 2016, Monotonic Calibrated Interpolated Look-Up Tables, Journal of Machine Learning Research 17 1-47.

Hassani, H., Huang, X. and Ghodsi, M., 2018. Big data and causality. Annals of Data Science, 5(2), pp.133-156.

Hauke, J. and Kossowski, T., 2011. Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. Quaestiones geographicae, 30(2), pp.87-93.

Hofner, B., Kneib, T. and Hothorn, T., 2016. A unified framework of constrained regression. Statistics and Computing, 26(1-2), pp.1-14.

Hovmand, P. S., Chalise, N., 2015, Simultaneous Linear Estimation Using Structural Equation Modeling, Analytical Methods for Dynamic Modelers, pp. 71-93, MIT Press.

Kroesen, M., Molin, E.J. and van Wee, B., 2010. Determining the direction of causality between psychological factors and aircraft noise annoyance. Noise and Health, 12(46), p.17.

Malhotra, N. K., M. Peterson, S. B. Kleiser, 1999, Marketing Research: A State-of-the-art Review and Directions for the Twenty-First Century, Journal of the Academy of Marketing Science.

Oliva, R., 2003, Model Calibration as a Testing Strategy for System Dynamics Models, European Journal of Operational Research, Volume (151), pp. 552–568.

Pya, N. and Wood, S.N., 2015. Shape constrained additive models. Statistics and Computing, 25(3), pp.543-559.

Sterman, J.D., 2000. Business dynamics: systems thinking and modeling for a complex world (No. HD30. 2 S7835 2000).

Tarka, P., 2018. An overview of structural equation modeling: its beginnings, historical development, usefulness and controversies in the social sciences. Quality & quantity, 52(1), pp.313-354.