# WIZER: An Automated Intelligent Tool for Model Improvement of Multi-Agent Social-Network Systems

**Alex Yahja and Dr. Kathleen M. Carley**

Computation, Organizations, and Society Program, Institute for Software Research International,
Center for the Analysis of Social and Organizational Systems, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213
ay@cmu.edu, kathleen.carley@cmu.edu

**Abstract**

Multi-agent social-network models are becoming increasingly used due to their power and flexibility in capturing emergent behaviors in complex socio-technical systems and their ability to link to real data. These models are growing in size and complexity which requires significant time and effort to calibrate, validate, improve the model, and gain insight into model behavior. In this paper, we present our knowledge-based simulation-aided approach for automating model-improvement and our tool to implement this approach (WIZER). WIZER is capable of calibrating and validating multi-agent social-network models, and facilitates model-improvement and understanding. By employing knowledge-based search, causal analysis, and simulation control and inference techniques, WIZER can reduce the number of simulation runs needed to calibrate, validate, and improve a model and improve the focus of these runs. We ran WIZER on BioWar – a city-scale multi-agent social-network model capable of simulating the effects of weaponized biological attacks on a demographically-realistic population against a background of naturally-occurring diseases. The results show the efficacy of WIZER.

## Introduction

A paradigm shift is occurring in how we think about knowledge, individuals, groups, networks, organizations, markets, institutions, and other societal systems due to the developments in computational modeling and analysis (Axelrod 1997; Carley and Prietula 1999; Epstein and Axtell 1996; Prietula et al. 1998). Computational modeling and analysis has emerged as a useful scientific tool for addressing socio-technical problems with complex dynamic interrelated parts, such as natural disaster response and biological attacks, which occur within a context constrained by social, organizational, geographical, regulatory, financial, and other factors.

The use of multi-agent models (Lucena et al. 2004) as well as social network analysis (Wasserman and Faust 1994) to address complex socio-technical problems has increased rapidly. Model assessment – determining how valid, how explainable, and how robust a model is – is becoming a major concern. For example, NATO argued that identifying reliable validation methods for electronic medical surveillance systems is a critical research area (Reifman et al. 2004). Calibration and validation serve as a foundation for model improvement through simulation and inference. However, inherent assumptions/abstractions, changes in reality, and human cognitive limitations make calibration, validation, and model-improvement difficult.

Few multi-agent simulations have exploited the depth and breadth of available knowledge and information for validation that reside in journals, books, websites, human experts, and other sources. Typically, simulation results are designed solely for human analysis and validation is provided by subject matter experts announcing that the model "feels right" (face validity). While this may be sufficient for small-scale simulations, it is woefully inadequate for large-scale simulations designed to inform decision-makers. Thus, automated help for validation and analysis is crucial, but little work to date probes this important aspect of automating validation and analysis. This paper describes our approach for doing knowledge-based simulation-aided validation in multi-agent social-network systems, embodied in a tool called WIZER (**W**hat-**I**f Analy**ZER**). WIZER applies knowledge control of the simulation, inference, and intelligent search to the multi-agent social-network simulations to facilitate the probing of the emergence of social networks and to automate validation and tuning of the simulation model.

The results presented in this paper are based on WIZER runs on BioWar. BioWar is a city-scale multi-agent social-network model capable of simulating the effects of weaponized biological attacks against a background of naturally-occurring diseases within various cities (Carley et al. 2003; Carley et al. 2004). BioWar currently runs 100,000 to several million agents with their social networks. Unlike traditional models that look at hypothetical cities (Epstein et al. 2004; Huang et al. 2004), BioWar is configured to represent real cities by loading census data, school district boundaries, etc. It models the healthy and infected agents as they go about their lives, enabling observation of absenteeism, hospital visits, etc.

## Related Work

The code of multi-agent systems is often "validated" by strictly applying requirements engineering. In software engineering terms (Pressman 2001), code-validation means the determination of the correctness of the final program or software produced with respect to the user needs and requirements. In contrast, our concern is with empirical validation (i.e., does the model produce results like the real world and for the "right" reasons). In principle, if the real-world could be specified formally then the formal methods used for code-validation could be applied. However, even the formal methods (Dershowitz 2004) used in software engineering for the control and understanding of complex multi-agent systems lack an effective means of determining if a program fulfills a given formal specification, particularly for very complex problems (Edmonds and Bryson 2004). Societal problems contain "messy" interactions, dynamic processes, and emergent behaviors, and are thus so complex that it is often problematic to apply requirements engineering and/or formal methods.

Another validation method is evolutionary verification and validation or EVV (Shervais et al. 2004; Shervais and Wakeland 2003), which utilizes evolutionary algorithms, including genetic algorithms and scatter search, for verification and validation. While EVV allows testing and exploitation of unusual combinations of parameter values via evolutionary processes, it employs knowledge-poor genetic and evolutionary operators rather than the scientific method, for doing experiments, forming and testing hypotheses, refining models, and inference, precluding non-evolutionary solutions.

Docking or alignment of possibly-different simulation models is another approach to validating multi-agent systems (Axtell et al. 1996). Alignment is used to determine whether two simulation models can produce the same results, which in turn is the basis for experiments and tests of whether one model can subsume another. The more models align, the more they are assumed to be valid, especially if one (or both) of them has been previously validated. The challenges in applying docking are the limited number of previously validated models, the implicit and diverse assumptions incorporated into models and the differences in data and domains among models. Two successful examples of docking are the alignment of the anthrax simulation of BioWar against the Incubation-Prodromal-Fulminant (IPF) mathematical model, a variant of the well-known Susceptible-Infected-Recovered (SIR) epidemiological model (Chen et al. 2003), and the alignment of BioWar against an SIR model of smallpox (Chen et al. 2004). While aligning a multi-agent model with a mathematical model can show the differences and similarities between these two models, the validity it provides is limited by the type and granularity of data the mathematical model uses and by the fact that symbolic (non-numerical) knowledge is not usually taken into consideration.

Validating complex multi-agent simulations by statistical methods alone (Jewell 2003) is problematic because the granularity (at the sample population level) required for the statistical methods to operate properly is distinct from the granularity needed for symbolic knowledge. Furthermore, associations common in statistics characterize static conditions, while causal analysis deals with the dynamics of events under changing conditions.

Human subject matter experts (SMEs) can validate computational models by focusing on the most relevant part of the problem and thinking about the problem intuitively and creatively. Applying learned expertise and intuition, SMEs can exploit hunches and insights, form rules, judge patterns, analyze policies, and assess the extent to which the model and their judgments align. Managed and administered properly, SMEs can be effective as validators. However, such human judgment based validation is subject to many pitfalls such as bounded rationality, biases, implicit reasoning steps, and judgment errors.

Another approach to validation is direct validation with real world data and knowledge. Validation can be viewed as experimentation with data and knowledge, and models as infrastructure or lab equipment for doing computational experiments or simulations (Bankes 2004). Simulation (Law and Kelton 2000; Rasmussen and Barrett 1995) has an advantage over statistics and formal systems as it can model the world as closely as possible (e.g., modeling emergence), free of the artifacts of statistics and formal systems.

Comparison against real data requires that the simulation be used to run a number of "virtual experiments". The results from these experiments are then contrasted with the real data. The quality of the results depends on the scope of the virtual experiments. Two techniques for this are Response Surface Methodology (Myers and Montgomery 2002) and Monte Carlo simulations (Robert and Casella 1999). These two approaches, however, can only be used with numerical data and are limited to a small number of dimensions.

For validating large scale simulation systems designed to characterize complex social problems a new approach is needed. The new approach must be scalable to a large number of variables and a high number of interactions. It must be sufficiently automated that it can be reapplied as new data comes to light and the model is changed. It must be flexible enough to handle data at different levels of granularity, missing data, and otherwise error messy data.

## WIZER

WIZER (**W**hat-**I**f Analy**ZER**) is an intertwined inference and simulation engine to do validation against real data. Viewing simulations as knowledge systems, WIZER is designed for controlling and validating them directly with empirical data and knowledge using pattern analyses and knowledge inferences (mimicking those of SMEs) and virtual experiments (mimicking those of RSM). WIZER facilitates knowledge-based simulation control and simulation-assisted inference, enabling reasoning about simulations and simulation-assisted reasoning. It enables the management of model assumptions, contradictory or incomplete data, and increases the speed and accuracy of model validation and analysis. Search in WIZER is performed using simulation virtual-experiments and knowledge inferences.

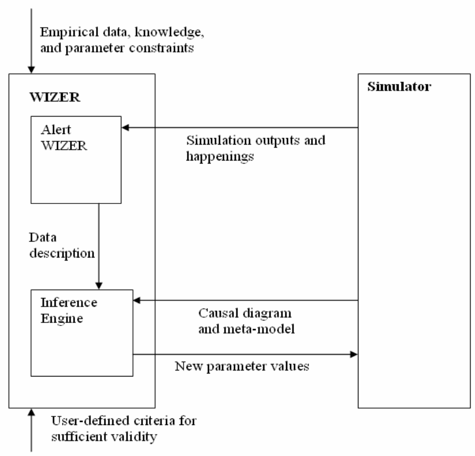


Figure 1. Diagram of WIZER

As shown in Figure 1, WIZER includes Alert WIZER and the WIZER Inference Engine. Alert WIZER determines which data streams do not fall within empirical data value range. The WIZER Inference Engine takes the simulator's causal diagram and the empirical constraints and confidence intervals on parameters to make a judgment on whether and which parameters (and causal links and the model itself if needed) to change and how. This results in new parameters for the next simulation. This cycle repeats until a user-defined validity level (the degree to which the model behaviors and outputs match reality) is achieved.

The knowledge base in the inference engine is populated with the knowledge about the simulator, simulation outcomes, domain facts and knowledge, assumptions, and problem solving strategies, among others. These different types of knowledge are included to enable the inference engine to reason about its reasoning. The emergence of causal links based on low-level interactions can be probed by the inference engine, including probes to see what an individual agent does in its life and what events affected this agent and why, in addition to sample based probes. For sample based probes, WIZER conducts inferences based on the application of its included statistical tests.

The WIZER Inference Engine is based on the rule-based Probabilistic Argumentation System (PAS) (Haenni et al. 1999) for handling assumptions. While the rule-based system (PAS) is sufficient if knowledge engineers are able to check the causal relations inherent in some rules, for large knowledge bases manual checks are cumbersome and prone to errors. Thus there is a need for automated and formal causal checking. Fortunately, causality has been treated mathematically (Pearl 2003). Thus, in WIZER we augment PAS with a causality checker.

Results from social network analysis form one silo of domain knowledge fed into WIZER inference engine. The inference engine in turn provides knowledge-based grounding for the emergence and evolution of social networks from low-level agent actions, behaviors, and interactions. The causal mechanisms encoded in WIZER's causality checker enable formal computation of intervention or action, instead of mere observation.

To account for the probability of causation, the causal model (Pearl 2003; Pearl 2000) specifies the use of Bayesian priors to encode the probability of an event given another event. It does not distinguish between different kinds of uncertainty. It is unable to model ignorance, ignores contradictions, and is incapable of expressing evidential knowledge without the use of the probability distribution format. Pearl's causal model is insufficient for WIZER for two reasons. First, the intended use of WIZER is to do validation in environments with incomplete, contradictory, and uncertain knowledge. Second, WIZER needs to clearly delineate between assumptions and facts. Therefore, we need an improved causal model in WIZER which we are building by borrowing concepts from the Probabilistic Argumentation Systems (PAS). WIZER uses a novel probabilistic argumentation causal system (PACS), which utilizes the probabilistic argumentation (Haenni et al. 1999) in causal analysis (Pearl 2000).

Table 1. Assumptions Encoding for Causality

| Type of Knowledge | Logical Representation | Meaning |
|---|---|---|
| A fact | P1 | P1 is true |
| A causation | P1 **causes** P2 | P1 causes P2 |
| An uncertain fact | A1 ➔ P1 | If assumption a1 is true, then P1 is true |
| An uncertain causation | A2 ➔ (P1 **causes** P2) or (P1 ∧ a2) **causes** P2 | If assumption a2 is true, then P1 causes P2 |

Table 1 shows the encoding of facts and causations for causal analysis enhanced with PAS-like assumption management. In the table, let $P_i$ be proposition *i*, $a_i$ be assumption *i*, ***causes*** be the causation operator, and ➔ be the implication operator. We call Table 1's formalism the probabilistic argumentation causal systems (PACS). WIZER includes both rule-based and causal formalisms.

The basic operations of the WIZER Inference Engine are as follows. Let $P = \{p_1, \ldots, p_n\}$ be propositions, $A = \{a_1, \ldots, a_n\}$ be assumptions, $h$ be the hypothesis and $K = c_1 \wedge c_2 \wedge \ldots \wedge c_n$ be the knowledge base of clauses, where $c_i$ is an element of the set of all possible $A$ and $P$ clauses. Let $\alpha$ be the (conjunctive) arguments supporting $h$. We have $\alpha \wedge K$ $|==$ $h$ or equivalently $\alpha$ $|==$ $\sim K \vee h$ or equivalently $\sim(\sim K \vee h)$ $|==$ $\sim\alpha$ and $K \wedge \sim h$ $|==$ $\sim\alpha$. In other words, if we know $K$ and $h$, we can compute the supports, that is, the arguments supporting $h$. The hypothesis $h$ is a clause produced by Alert WIZER after comparing simulation data streams with empirical data. After finding the arguments supporting $h$, the degree of support can be found, defined as

$$dsp(h, K) = prob\text{ (a support } \alpha \text{ of } h \text{ is valid } | \text{ no contradiction, } K)$$

Similarly, the degree of plausibility can be computed as

$$dpl(h, K) = prob\text{(no support of } \sim h \text{ is valid } | \text{ no contradiction, } K)$$

These two measures are used to determine which arguments are the most relevant to the hypothesis at hand, pinpointing which parameter values, causal links, and/or submodels should be changed. In other words, hypothesis $h$ is the input to the WIZER Inference Engine and the arguments supporting $h$ are the output, leading to possible changes in parameter and meta-model values.

The operations described above are performed for both rule-based and causal clauses. Then, for clauses denoted as causal, additional operations are performed to see whether and to what degree the causal relations are empirically correct, partially based on the degree of support and the degree of plausibility and quasi-experimentally via virtual experiments if needed. Sustenance, causal beams and actual cause (Pearl 2000) are also computed.

The intertwining causal computation and virtual experimentation capability of WIZER enhances PACS and is useful in simulations to:

- Provide a formal computational means to convert simulation happenings to user-friendly causal sentences.
- Probe existing and potential causal assumptions and links and examine the robustness of causal links.
- Formally model interventions in simulations.
- Allow symbolic values/events to be considered in determining causal relations.
- Allow experimentation and simulation control. As WIZER modifies, runs, re-modifies, and re-runs simulations, it uses causal mechanisms to keep track of what causes a certain series of modifications to work or fail and to suggest possible next steps.
- Allow better inference by letting the inference engine run simulations in the midst of causal inferences as needed.
- Allow quasi-experimental examination of the empirical claims of causal inferences via simulations.

## Preliminary Results from a Testbed

We are applying WIZER to validate BioWar. As mentioned earlier, BioWar (Carley et al. 2003) is a city-scale multi-agent social-network model capable of bioattack simulations. Alert WIZER and a working prototype of the WIZER Inference Engine have been implemented in C++ and Common LISP.

Alert WIZER takes in the following empirical data:

- NCES Indicator 17, 2002 (Year 2000 data), for calculating school absenteeism http://nces.ed.gov/programs/coe/2002/section3/indicator17.asp
- CDC Advance Data, from Vital and Health Statistics, no. 326, 2002, for calculating emergency room visits http://www.cdc.gov/nchs/data/ad/ad326.pdf
- CDC Advance Data, from Vital and Health Statistics, no. 328, 2002, for calculating doctor visits http://www.cdc.gov/nchs/data/ad/ad328.pdf
- 1997 US Employee Absences by Industry Ranked for determining work absenteeism http://publicpurpose.com/lm-97absr.htm
- Over-the-counter (OTC) Drug Sales from Pittsburgh Supercomputing Center's "FRED" pharmacy sales data.

Alert WIZER also takes in BioWar simulation outputs matching the above empirical data streams.

WIZER was run on the Challenge 3 and Challenge 4 data (Carley et al. 2004) from BioWar. Challenge 3 data used for validation has 4 data streams with 10 simulations for each attack case (no attack, anthrax attack, and smallpox attack) per city for 4 cities. The population size and number of locations were scaled at 20% of actual. The parameters were adjusted following inference engine runs based on a causal diagram of BioWar. First, we present one of the four data stream results from Challenge 3.

Table 2. School Absenteeism

| City (20% scale) | Empirical lower, higher bounds | | No attack (mean) | Anthrax (mean) | Smallpox (mean) |
|---|---|---|---|---|---|
| Norfolk | 3.04% | 5.18% | 3.45% | 3.75% | 3.55% |
| Pittsburgh | 3.04% | 5.18% | 3.52% | 4.67% | 4.46% |
| San Diego | 3.04% | 5.18% | 3.78% | 3.81% | 5.57% |
| Veridian Norfolk | 3.04% | 5.18% | 3.73% | 4.05% | 4.31% |

We see that the simulated means for school absenteeism rate (Table 2) for normal – no attack – simulation cases fall between lower and upper empirical bounds for the simulations of Norfolk, Pittsburgh, San Diego, and "Veridian Norfolk" (a part of Norfolk specified by Veridian, Inc.). For anthrax attack cases, the simulated means are higher than normal means but still lower than the empirical higher bounds. This is plausible as the empirical higher bound contains (contagious) flu cases. For smallpox attack, however, the simulation mean for one city – San Diego – is higher than the empirical higher bound. Smallpox is highly contagious so this is also plausible. For other cities, the simulated means of school absenteeism are within bounds.

Challenge 4 data has 12 data streams: school absenteeism, work absenteeism, doctor visits, emergency room visits, emergency room visits using the Surveillance Data Inc. data, and seven types of drug purchases. Table 3 shows the percentage of validated data streams for six cities for the no attack case.

Table 3. "Challenge 4" Simulation Output Validation

| City | Data Streams Validated |
|---|---|
| San Francisco | 5 out of 12, or 41.67% |
| San Diego | 7 out of 12, or 58.33% |
| Pittsburgh | 7 out of 12, or 58.33% |
| Norfolk | 6 out of 12, or 50.00% |
| Hampton | 4 out of 12, or 33.33% |
| Washington DC | 4 out of 12, or 33.33% |

The WIZER Inference Engine is run on Alert WIZER outputs. One relevant run deals with agents visiting doctor offices. Agents visit doctor offices if they have sufficiently severe symptoms, that is, if the severity exceeds a "going-to-doctor-office" threshold. Alert WIZER checks the simulated doctor visit rate against empirical data to see if it is within bounds and the WIZER Inference Engine takes the resulting information and computes adjustments needed to move the doctor visit rate within bounds. The following shows the relevant knowledge base for this simple case.

assumption_1 → threshold_too_low
threshold_too_low → doctor visit rate is above higher bound
assumption_2 → threshold_too_high
threshold_too_high → doctor visit rate is below lower bound
assumption_3 → ~(threshold_too_high ∨ threshold_too_low)
~(threshold_too_high ∨ threshold_too_low) →
doctor visit rate is within bounds

If Alert WIZER observes that the doctor visit rate is below the empirical lower bound, the WIZER Inference Engine infers that the threshold is too high (the degree of support for assumption_2 is very high) and makes an adjustment to lower the threshold. The above shows a simple result of an Inference Engine run. There are in fact many other possible causes, so other inferences are performed as well.

## Discussion

WIZER uses AI techniques to probe the emergence of social networks through multi-agent social-network simulations. It also provides a way to integrate knowledge-intensive and context-sensitive social networks to build effective and humane agents. WIZER interfaces simulation to symbolic inference constrained by empirical data and knowledge. Knowledge-based control of simulation can be viewed as an AI tool to describe individual, emergent, and collective level behaviors.



WIZER indicates that making simulation an integral part of AI can be fruitful, especially when dealing with socio-technical problems which have a high degree of uncertainty and interactions. Based on empirical data and knowledge, simulations can bound the inferences and allow the empirical claims of the inferences to be investigated. At the same time, knowledge-based inference and control of the simulation can reduce the number of search and virtual experiments in the simulation needed.

WIZER uses established theories, empirical knowledge, and empirical data to constrain the degrees of freedom of multi-agent models, facilitating justifications of every change in the agent's and meta-model's parameters. While the lack of empirical data can be a challenge, it also presents an opportunity for WIZER to produce synthetic populations and forecast their behavior based on assumptions. An advantage of WIZER is that it enables validation to proceed by making use of data that is incomplete, from various sources, and at varying levels of granularity. By making sound inferences based on the existing data, socio-technical forecasting is feasible.

More WIZER and simulation runs are needed to get better statistics and to evaluate error margins, the effects of sample choices, search space traversal, and the performance of simulation and knowledge search.

## Acknowledgements

This research was supported, in part, by DARPA for work on Scalable Biosurveillance Systems, the NSF IGERT9972762 in CASOS, the Army Research Labs, and by the Carnegie Mellon Center on Computational Analysis of Social and Organizational Systems (CASOS). The computations were performed on the National Science Foundation Terascale Computing System at the Pittsburgh Supercomputing Center. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of DARPA, the National Science Foundation, the Pittsburgh Supercomputing Center, the Army Research Labs, or the US Government.

## References

Axtell, R., Axelrod, R., Epstein, J., and Cohen, M. 1996. Aligning Simulation Models: A Case Study and Results. *Computational and Mathematical Organization Theory*, 1(2): 123-141.

Axelrod, R. 1997. Advancing the Art of Simulation. In Proceedings of the International Conference on Computer Simulation and the Social Sciences, Cortona, Italy.

Bankes, S. C. 2004. Models as Lab Equipment: Science from Computational Experiments. In Proceedings of North American Association for Computational Social and Organizational Science (NAACSOS) Conference 2004. ISRI, CASOS, CMU, Pittsburgh, PA.

Carley, K. M. and Prietula, M., eds. 1999. *Computational Organizational Theory*. Mahwah, NJ: LEA.

Carley, K. M., Fridsma, D., Casman, E., Altman, N., Chang, J., Kaminsky, B., Nave, D., and Yahja, A. 2003. BioWar: Scalable Multi-Agent Social and Epidemiological Simulation of Bioterrorism Events. In Proceedings of North American Association for Computational Social and Organizational Science (NAACSOS) Conference 2004, Pittsburgh, PA, http://www.casos.ece.cmu.edu/casos_working_paper/carley_2003_biowar.pdf.

Carley, K. M., Altman, N., Kaminsky, B., Nave, D., and Yahja, A. 2004. BioWar: A City-Scale Multi-Agent Model of Weaponized Biological Attacks. CMU-ISRI-04-101 Technical Report, Institute for Software Research International, Carnegie Mellon University, http://reports-archive.adm.cs.cmu.edu/anon/isri2004/CMU-ISRI-04-101.pdf

Chen, L-C, Carley, K.M., Fridsma, D., Kaminsky, B., and Yahja, A. 2003. Model Alignment of Anthrax Attack Simulations. *Decision Support Systems* in the special issue on Intelligence and Security Informatics, in press.

Chen, L-C., Kaminsky, B., Tummino, T., Carley, K. M., Casman, E., Fridsma, D., and Yahja, A. 2004. Aligning Simulation Models of Smallpox Outbreaks. In Proceedings of the Second Symposium on Intelligence and Security Informatics, Tucson, AZ, June 10-11, 2004. Also in *Springer-Verlag Lecture Notes in Computer Science* Vol. 3073.

Dershowitz, N., eds. 2004. *Verification: Theory and Practice*. New York, NY: Springer Verlag.

Edmonds, B. and Bryson, J.J. 2004. The Insufficiency of Formal Design Methods: the Necessity of an Experimental Approach for the Understanding and Control of Complex MAS. In Proceedings of AAMAS 2004, ACM.

Epstein, J.M. and Axtell, R.L. 1996. *Growing Artificial Societies*. Cambridge, Mass.: MIT Press.

Epstein, J.M., Cummings, D.A.T, Chakravarty, S., Singa, R.M., and Burke, D.S. 2004. *Toward a Containment Strategy for Smallpox Bioterror: An Individual-Based Computational Approach*. Washington, DC: Brookings Institution Press. http://www.brook.edu/press/books/towardacontainmentstrategyforsmallpoxbioterror.htm

Haenni, R., Kohlas, J., Lehmann, N. 1999. Probabilistic Argumentation Systems. Technical Report 99-9, Institute of Informatics, University of Fribourg, Fribourg, Switzerland.

Huang, C-Y, Sun, C-T, Hsieh, J-L, and Liu, H. 2004. Simulating SARS: Small-World Epidemiological Modeling and Public Health Policy Assessments. *Journal of Artificial Societies and Social Simulation*, 7(4). http://jasss.soc.surrey.ac.uk/7/4/2.html

Jewell, N.P. 2003. *Statistics for Epidemiology*. Boca Raton, FL: Chapman & Hall/CRC.

Law, A.M. and Kelton, W.D. 2000. *Simulation Modeling and Analysis, 3rd ed.* New York, NY: McGraw-Hill.

Lucena, C., Carcia, A., Romanovsky, A., Castro, J., and Alencar, P., eds. 2004. Software Engineering for Multi-Agent Systems II. *Lecture Notes in Computer Science* 2940. New York, NY: Springer Verlag.

Myers, R.H. and Montgomery, D.C. 2002. *Response Surface Methodology: Process and Product Optimization using Designed Experiments*, *2nd ed.* New York, NY: Wiley.

Pearl, J. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge, UK: Cambridge University Press.

Pearl, J. 2003. Statistics and Causal Inference: A Review. *Test Journal* 12 no. 2 (December): 281-345.

Pressman, R. S. 2001. *Software Engineering*., New York, NY: McGraw-Hill.

Prietula, M.J., Carley, K.M., and Gasser L. 1998. *Simulating Organizations*. Menlo Park, Calif.: AAAI Press/The MIT Press.

Rasmussen, S. and Barrett, C.L. 1995. Elements of a Theory of Simulation. ECAL 95, *Lecture Notes in Computer Science*. Berlin, Germany: Springer Verlag.

Reifman J., Gilbert, G., Parker, M., and Lam, D. 2004. Challenges of Electronic Medical Surveillance Systems. RTO HFM Symposium on NATO Medical Surveillance and Response: Research and Technology Opportunities and Options, Budapest, Hungary. http://www.rta.nato.int-/Pubs/RDP.asp?RDP=RTO-MP-HFM-108

Robert, C. P. and Casella, G. 1999. *Monte Carlo Statistical Methods*. New York, NY: Springer Verlag.

Shervais, S., Wakeland, W., and Raffo, D. 2004. Evolutionary Verification and Validation of Software Process Simulation Models. In the 5th International Workshop on Software Process Simulation and Modeling, extended abstract.

Shervais, S. and Wakeland, W. 2003. Evolutionary Strategies as a Verification and Validation Tool. Portland State University paper, http://www.sysc.pdx.edu/faculty/Wakeland/papers/EvoVandVRevD.pdf

Wasserman, S. and Faust, K. 1994. *Social Network Analysis: Methods and Applications*. Cambridge, UK: Cambridge University Press.