Information-Theoretic Approaches for Evaluating Complex Adaptive Social Simulation Systems

Auroop R. Ganguly, *Member, IEEE*, Olufemi A. Omitaomu, *Member, IEEE*, and Yu (Cathy) Jiao

***Abstract*—In this paper, we propose information-theoretic approaches for comparing and evaluating complex agent-based models. In information theoretic terms, entropy and mutual information are two measures of system complexity. We used entropy as a measure of the regularity of the number of agents in a social class; and mutual information as a measure of information shared by two social classes. Using our approaches, we compared two analogous agent-based (AB) models developed for regional-scale social-simulation system. The first AB model, called ABM-1, is a complex AB built with 10,000 agents on a desktop environment and used aggregate data; the second AB model, ABM-2, was built with 31 million agents on a high-performance computing framework located at Oak Ridge National Laboratory, and fine-resolution data from the LandScan Global Population Database. The initializations were slightly different, with ABM-1 using samples from a probability distribution and ABM-2 using polling data from Gallop for a deterministic initialization. The geographical and temporal domain was present-day Afghanistan, and the end result was the number of agents with one of three behavioral modes (pro-insurgent, neutral, and pro-government) corresponding to the population mindshare. The theories embedded in each model were identical, and the test simulations focused on a test of three leadership theories – legitimacy, coercion, and representative, and two social mobilization theories – social influence and repression. The theories are tied together using the Cobb-Douglas utility function. Based on our results, the hypothesis that performance measures can be developed to compare and contrast AB models appears to be supported. Furthermore, we observed significant bias in the two models. Even so, further tests and investigations are required not only with a wider class of theories and AB models, but also with additional observed or simulated data and more comprehensive performance measures.**

***Index Terms*—entropy, mutual information, agent-based models, social modeling, performance measures.**

I. Introduction

Agent-based modeling (ABM) has attracted increasing attention in the field of social computing as a main computational approach to social and economic systems simulation. While simulations may give rise to interesting macro-level phenomena, the underlying micro and macro level processes may be far from realistic. Nevertheless, this realism may be important to infer results that are relevant to existing theories of social and economic systems and to policy making. As a result, many new agent-based (AB) models have been proposed with wide applications. Some of the characteristic differences between these models are the resolution of the input data (aggregate vs. fine-resolution data), the computing framework (desktop vs. high-performance computing framework), and the assumptions of the implemented social theories (for instance, the weights assigned to each factor). Therefore, it is important to assess not only the predictive capability of the AB models for Human, Social, Behavioral, Cultural (HSBC) domain but also the ability to quantify and visualize the inherent differences in these models [8].

The critical challenges in systematic evaluation of large-scale social science simulations stem from the inherent multiscale attributes of HSBC processes, models, and theory, as well as from the inadequacy of data and case studies for calibration and validation purposes. The multiscale processes range from psychological profiles of leaders and aggregate crowd behavior to the behavior of institutions or organizations, and of ethnic, geographic, religious, linguistic, and racial groups. The need to adequately handle such processes across scales has spawned a wide range of multiscale social theories, which in turn may be competitive or complementary, and hierarchical or integrated. Also, "surprising" or unusual behavior at one scale may indeed be triggered by minor changes or abnormal behavior at another scale. Existing methods for the evaluation of theories, models, and systems relevant for HSBC or similar domains rely on the exploration of the hypotheses (or parameter) space and on empirical validation. These methods include active nonlinear tests of complex simulation models as well as structural and parametric sensitivity analysis for the evaluation of complex models [14]. These approaches rely on the design of computational experiments [9, 13] and empirical validation [7, 10, 18]. Validation and evaluation in the context of M&S systems for HSBC or similar domains have received some attention from multidisciplinary scientific communities [3, 15, 17].

Given that models are all too imperfect and validation data are inadequate and noisy, traditional calibration and validation approaches are not likely to succeed. Systematic evaluation of models remains useful however, and is perhaps increasing in importance, as decision makers still need to know how to make best use of the available HSBC process understanding, theories and models, as well as how to utilize available data

Manuscript received May 20, 2009. This research was supported in part by the Defense Advanced Research Projects Agency (DARPA) of the US Department of Defense (DoD). The views expressed in this paper are those of the authors and do not necessarily represent the opinions of their institutions or of DoD.

A.R. Ganguly, O.A. Omitaomu, and Yu (Cathy) Jiao are with the Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6017, USA (e-mail: gangulyar@ornl.gov; omitaomuoa@ornl.gov; jiaoy@ornl.gov). Corresponding Author Phone: +1-865-241-1305; fax: +1-865-241-6261.

and computational resources as optimally as possible. In these situations, systematic evaluation may have to take the form of characterization of the space of real-world processes and theoretical simulations. The insights gained may have to be qualitative (e.g., tribal loyalties dominate over individual ideologies in a certain region) or quantitative (e.g., based on structural or parametric sensitivity studies). In this paper, we present information-theoretic approaches for evaluating and comparing complex social simulations.

The remainder of this paper is organized as follows. In Section II, we present the theory behind the information-theoretic measures used in this paper for evaluating AB models. In Section III, we review the social theories embedded in each model and describe the computational implementations of the social theories. The results of the comparison of two analogues AB models – ABM-1 and ABM-2 – are discussed in Section IV. A short summary is presented in Section V.

II. Information-theoretic Measures

In this section, we described two information-theoretic measures used in this paper.

A. Entropy

Formally, the entropy for a discrete process X of K classes is defined as:

$$H(X) = \sum_{x \in K} p(x)\log\big(p(x)\big),$$

where $p(x)$ is the probability of x in $\boldsymbol{X}$. Entropy is typical interpreted as the number of bits required to encode and transmit the classification of a data item. If the entropy is smaller the data is more "pure" – all data items belong to the same class, then entropy is zero because there is only one outcome [5]. The entropy is larger for "impure" data. Therefore, entropy has been described as a measure of the rate at which environment appears to produce information. "*The higher the entropy rate, the more information produced, and the more unpredictable the environment appears to be*" [5]. If entropy is used as a measure of the predictability of classes, then the smaller the class entropy, the more predictable the class would be. For example, if all agents belong to one class, then the entropy is zero, and no bits need to be transmitted because the receiver knows that there is only one outcome; therefore, no uncertainty exists and the class predictability appears much higher.

For social simulation outputs, we used entropy as a measure of predictability in the number of agents over time. The smaller the entropy the more predictable the agents over time and the fewer the event types over time.

B. Mutual Information

The mutual information, $I(X;Y)$, of two discrete random variables X and Y can be defined as:

$$I(X;Y) = \sum_{y \in Y}\sum_{x \in X} p(x,y)\log\left(\frac{p(x,y)}{p_1(x)p_2(y)}\right),$$

where $p(x,y)$ is the joint probability distribution function of X and Y, and $p_1(x)$ and $p_2(y)$ are the marginal probability distribution of X and Y respectively. The mutual information measures (MI) the information that X and Y share; that is, it is a measure of how much knowing one of these variables reduces the uncertainty about the other. As an illustration, if X and Y are independent, then information about X does not provide any information about Y and vice versa, so their MI is zero. On the other hand, if X and Y are dependent, then all information provided by X is shared with Y; that is, knowing X determines the value of Y and vice versa. MI is symmetric ($I(X;Y) = I(Y;X)$) and can also be written as:

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

where $H(X)$ and $H(Y)$ are the individual class entropies and $H(X,Y)$ is the entropy of the two classes considered as a joint process. An increase in MI between two classes could be described as an indication that the correlation between the classes is growing. Such a growth has been attributed as an evidence for a phase transition; an important feature for exploring emergent behavior.

In social simulations, we need to determine the effect of one class of agents on the behavior of another class of agents over time; therefore, we could find the "correlations" between events that take place in two classes of agents. The mutual information between two classes of agent X and Y can be used to study "correlations" in social simulation systems. The mutual information, in this case, measures the complete dependence, unlike linear correlations, which are measures of linear associations or rank-based measures that capture only monotonic dependence.

III. The Social Theories And Their Computational Implementations

A prototype experimental test-bed was designed to evaluate two ABMs – ABM-1 and ABM-2 – by maximizing a utility function. The ABM-1 model utilized a desktop environment and aggregated data/models. The ABM-2 model utilized high performance computing and fine-resolution data/models.

Three leadership theories, specifically legitimacy, representative, and coercion, were implemented by assigning appropriate weights to each factor in the utility function. Neighbor interactions were modeled by using two social mobilization theories: (1) social influence and (2) resistance to repression. Four learning theories, each implemented for a change in support for a leader or change in ideology, were developed: socialization, homophily, results-based, and cognitive dissonance. Ninety-six combinations of theories resulted from the nine theories (three for leadership, two for social mobilization, and four for learning or psychological change, where each of the last four can be implemented for leadership or ideology change). The ninety-six theories were implemented for each ABM model along with various heuristics for a case study of Afghanistan. Although the solutions utilized identical theories, the fine-grained implementations required slightly different heuristics.

A. The Social Theories

The individual behavioral choices were modeled with one of two utility functions [4]:

Cobb-Douglas:

$$U = (1-L)^{w_L}(1-C)^{w_C}(1-I)^{w_I}(1-E)^{w_E}(1-V)^{w_V}(1-F)^{w_F}(1-R)^{w_R}$$

Least Squares:

$$U = 1 - W_L L^2 - W_C C^2 - W_I I^2 - W_E E^2 - W_V V^2 - W_F F^2 - W_R R^2$$

The utility functions encompass seven factors: L is the agent's loyalty to the leader ($L \in [-1,1]$), C is the coercion factor ($C \in [-1,1]$), I is ideology ($I \in [-1,1]$), E is economic welfare ($E \in [-1,1]$), V is security against violence($V \in [-1,1]$), F is the influence of "close" associates (geographic or social proximity), and R is repression and social influence for defying repression ($R \in [0,1]$)given as:

$$max\left(0, \left(-sign(A) * sign(B)\right)\right) * max\left((A^2 - \langle B^* \rangle^2), 0\right),$$

where A is the repressive activity in the area and $\langle B^* \rangle$ is the average behavior of agents within a certain region of the focal agent (a larger region than for influence). The weights are required to be non-negative, to be less than or equal to 1, and to sum to unity ($w_L + w_C + w_I + w_E + w_V + w_F + w_R = 1$). The overall computational goal is to identify the behavior value (B) that would allow a citizen agent to maximize the value of her utility function. The seven components considered are:

$$Loyalty: 1 - L = 1 - \eta_1 * abs\,(O - B)/2;$$

$$Coercion: 1 - C = 1 - r_1 * abs\,(O - B)/2;$$

$$Ideology: 1 - I = 1 - \eta_2 * abs\,(P - B)/2;$$

$$Economic\ Welfare: 1 - E;$$

$$Security\ from\ Violence: 1 - V;$$

$$Influence: 1 - |B - \langle B \rangle|/2\,;$$

$$Repression: 1 - R;$$

where B is the considered behavior of the agent to be optimized through the utility function, and O represents an order by the leadership. A variable (e.g., Loyalty: L) is reflected in Cobb-Douglas through the component (e.g., I - L). η_1 and η_2 refer to the agent's support for leadership/ideology, r to the leadership's resources, P to agent's ideology, E to economic dissatisfaction, V to the agent's dissatisfaction with the security situation, and $\langle B \rangle$ to the average behavior of agents within a certain region of the focal agent. Learning theories are implemented by making each of the variables functions of an agent (i) and time point (t), multiplying the utility function by a learning term $\left(\lambda_1(t)\right)$, and allowing both λ and P to be "learned" over time in a prescribed manner. For the detailed description of how this implementation was achieved, please see [8].

B. Computational Implementations

Two analogous AB simulation models were set up. The models implemented the social theories and models described in Section IIIA for contemporary Afghanistan population. The first model, ABM-1, is based on the NetLogo platform [16], which is typically used as a demonstration platform for ABMs. The system considered five types of agents: Afghan government soldiers, coalition forces soldiers, Taliban, leaders, and citizen agents. The citizen agents supportive of the Taliban were called pro-insurgent, whereas citizens who were supportive of soldiers/coalition forces became pro-government. The rest were neutral. Therefore, the end result was the number of agents with one of three behavioral modes (pro-Insurgent, neutral, or pro-government) corresponding to the population mindshare.

The country was divided into six regions, each with multiple "patches" in NetLogo. The purpose of the six regions was to allow multiple leaders for the Pashtun tribe, and to allow each Pashtun leader to have a geographically defined area of influence on Pashtun agents. Therefore, the regions apply exclusively to the Pashtun tribe. Other tribes had only one leader each, and those leaders had influence on their agents across the entire country. The data for agents and their attributes were developed in creative ways. For example, opium production was used as a measure of economic prosperity. A variety of heuristics was used for agent behaviors like geographical movements. The total number of agents was limited to the maximum of 10,000 allowed by the NetLogo environment, which required that the behaviors of citizen agents be modeled at aggregate levels. The data utilized were an aggregate version of the data used for second AB model, ABM-2.

The ABM-2 simulation was developed using the Oak Ridge Mobile Agent Community (ORMAC) platform [11, 12]. The ABM-2 model is an identical model to the ABM-1 model but with fine-grained data and with agents at much higher resolutions. LandScan population data [1] and the relevant geospatial methodologies [6] were used to build a synthetic Afghan population and to geo-locate the 31 million agents corresponding to the 2006 population for Afghanistan. A variety of disparate geospatial sources was utilized to develop and map the agent attributes as well as the theoretical settings. Calibration data were obtained at district levels. The combination of a GIS-based platform with an ABM is by itself a significant step forward [2]. See [8] for detailed description of the data pre-processing methodology for ABM-2. The initializations were slightly different, with ABM-2 using polling data from Gallop for a deterministic initialization and ABM-1 using samples from a probability distribution. The social theories embedded in each model were identical, and the test simulations focused on a test of the leadership theories described in Section IIIA.

IV. Experimental Results

The results of the ABM-1 model were compared with the results of the ABM-2 model to obtain insights into how the fine-grained data and more resolved process models impacted the final results. Furthermore, the evaluation measures were developed to show the feasibility of using these approaches in the HSBC domain.

A. *Comparison of Aggregate and Fine-Resolution Models Output*

Each of the two models was run nine times (Table 1), with each run instantiating one of three possible leadership theories (legitimacy, L; representative, R; or coercion, C) and one of three time resolutions (3 days, 7 days, or 14 days per "tick," where a tick corresponds to a clock time).

Table 1: The nine simulation runs

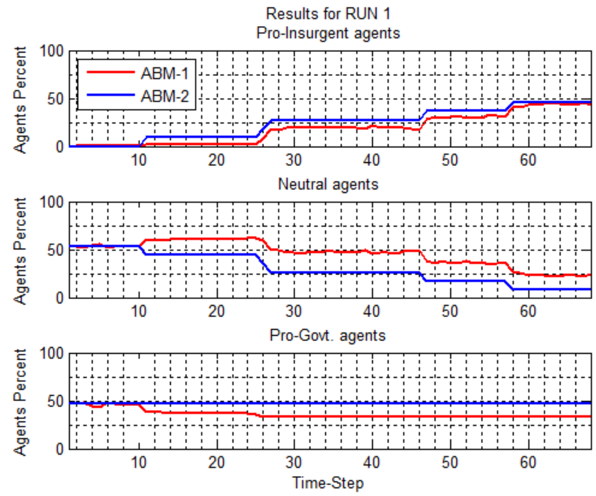
Run Number	1	2	3	4	5	6	7	8	9
Days / Tick	3	3	3	7	7	7	14	14	14
Leadership Theory	L	R	C	L	R	C	L	R	C

Because the focus was on leadership theories, five leader assassinations were introduced to explore the effects of disruptive events (Table 2). The events were based on realistic observations in the region. The different simulations provided us with a test set of outputs to demonstrate the value of statistical distance measures. The nine runs are described in Table 1. A comparison of the percent number of agents in each social class for each model from three of the nine simulation runs are shown in Figure 1.

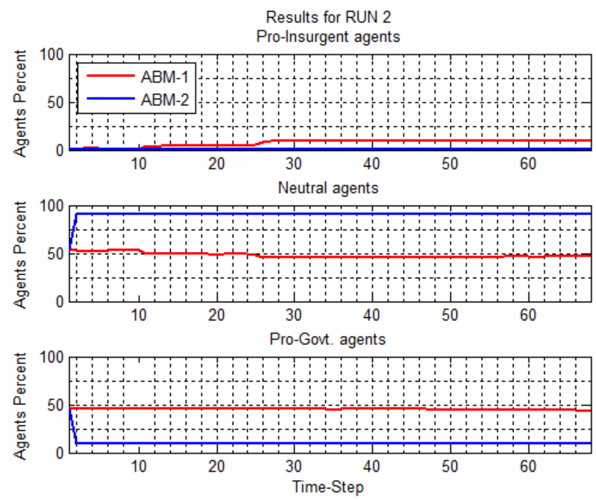
Table 2: Events in the simulations

	Event 1	Event 2	Event 3	Event 4	Event 5	End
3days/tick	11	26	27	47	58	68
7days/tick	5	12	12	21	25	35
14days/tick	3	6	6	13	16	26

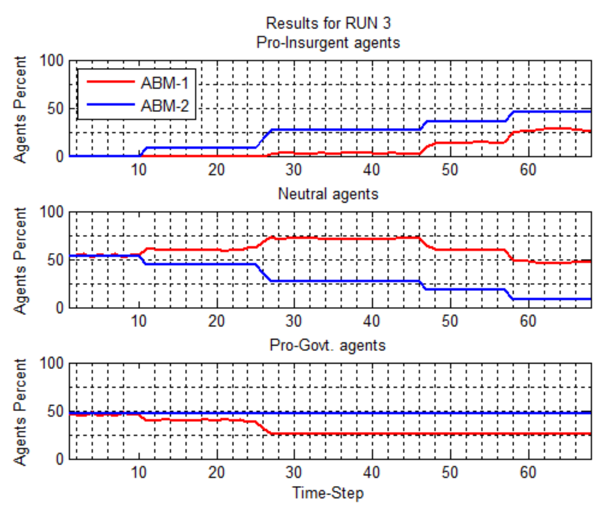
Due to space constraint, only these three outputs are discussed in this paper; however, similar observations are seen in the other six simulation outputs. The significant bias in the ABM-2 versus ABM-1 outputs, even at the end points of the simulations where the outputs at successive time steps appear relatively stable, is obvious from the plots. The only exceptions occur during the instantiation of the legitimacy (L) theory (run 1) and in that case, only for the number of pro-Insurgent agents. The fact that simulations differing primarily in their spatial resolutions result in such large relative biases is cause for concern. Comparisons such as these may help improve model outputs, in this case to correct bias errors.



(a)



(b)



(c)

Fig. 1. Outputs for the first three simulation runs, as shown in Table 1, are shown in (a) to (c) respectively. The outputs generated from ABM-1 (red) and ABM-2 (blue) correspond to the number of agents exhibiting pro-Insurgent (top of each panel), neutral (middle) and pro-government (bottom) behavior.

The other interesting aspect is that ABM-2 outputs appear more responsive to the disruptive events, while ABM-1 outputs exhibit more random fluctuations but less response to the events. The difference in the response to events versus random fluctuations is more obvious in first differences (not shown in this paper). The response to the events is much clearer in ABM-2, and ABM-1 generates more random fluctuations. Although a definitive explanation may not be possible without further investigation, the dominance of random parameters for initialization in ABM-1 versus the more deterministic initializations and simulations in ABM-2 may be a plausible explanation. In situations where both models exhibit what appears to be "legitimate" (i.e., occurring at the expected time steps) responses to events, ABM-1 responses seem somewhat damped compared to ABM-2 response. This indeed may be caused by resolution effects, even though there are a few exceptions to this empirical "rule." The lack of any response in ABM-2 (even the changes in ABM-1 appear to be no more than random fluctuations) to the disruptive events when the representative (R) theory is instantiated may be worth noting. The lack of a leader may have less immediate effect on followers when the predominant behavior is representative. However such social explanations must be exercised with care given that the simulation runs appear pretty flat in each case when this theory is implemented (runs 2, 5, and 8). This may suggest an artifact of the specific experimental design. Although the causal explanations offered here are only plausible but not proven unless further simulations are performed, the value of simple measure (e.g., bias and first differences) together with visual representations may be apparent from the discussions here.

Traditional statistical distance measures such as correlation coefficients shown in Table 3 further indicate that the increase in the number of days per tick does not significantly improve the dependency between systems with respect to the number of agents in each class. A detailed comparison of a relatively "lumped" or low-resolution model (e.g., ABM-1) with a relatively more spatially "distributed" or high-resolution model (e.g., ABM-2) typically entails one of two approaches: either aggregate the distributed model outputs to the scales of the lumped model and compare at the aggregate scales, or allocate the lumped model outputs to the scales of the distributed model and compare at the higher resolutions.

Table 3: Correlation coefficients between ABM-1 and ABM-2 outputs

Run	Days/tick	Theory	Pro-Insurgent	Neutral	Pro-Govt
1	3	L	0.9750	0.8870	-0.2867
4	7	L	0.9776	0.9463	-0.4136
7	14	L	0.9610	0.9109	-0.6806
2	3	R	-0.2505	-0.2614	0.1203
5	7	R	-0.3966	-0.4101	0.1839
8	14	R	-0.5601	-0.6048	0.2603
3	3	C	0.8479	0.1196	-0.2215
6	7	C	0.9361	0.4923	-0.3166
9	14	C	0.9479	0.5612	-0.4139

In our case, the aggregate "patch" level outputs generated from ABM-1 need to be allocated to the finer grids at which data are obtained from the Geographical Information System (GIS) and which are ultimately used by ABM-2. The simulation results must be compared at scales that matter to decision-makers (e.g., district levels in Afghanistan). The map for the case study region (Ghazni) with one ABM-1 patch, corresponding ABM-2 grids, and the Afghan districts (indicated by identification numbers assigned for the purpose of this simulation) is shown in Figure 2. Specifically, the number of agents for each class and each patch was converted into the corresponding number of agents for each district by multiplying the uniformly distributed number of agents by the number of patches that equal the geographical size of each district. In a sense, this is just an area-weighted allocation strategy. We focused the comparison on "Run 1" (see Table 4) and a few districts for illustrative purposes.

Table 4: Linear correlation between ABM-1 and ABM-2 runs at district levels

RUN	**District ID**	**Pro-Insurgent**	**Neutral**	**Pro-Govt**
1	27713	0.9744	0.8855	0.0000
	27723	0.9759	0.8890	0.0000
	27726	0.9759	0.8897	0.0000
	27731	0.9744	0.8856	-0.2867
	27747	0.9740	0.8846	0.0000

Table 4 shows high correlations between ABM-1 and ABM-2 outputs for pro-insurgent and neutral agents. The zero or negative correlations for pro-government agents may be ignored given that the number of these agents remains relatively constant during the simulation time period (see Figure 1a).

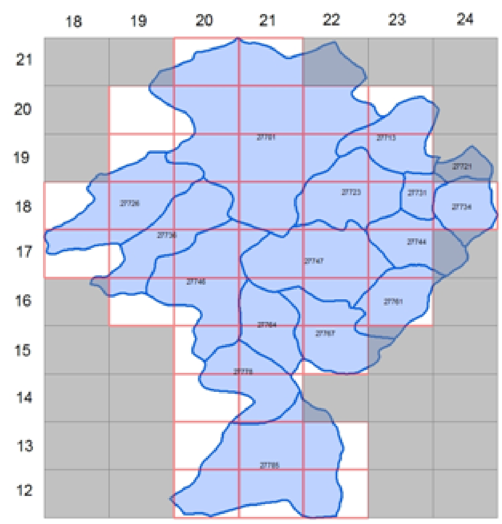


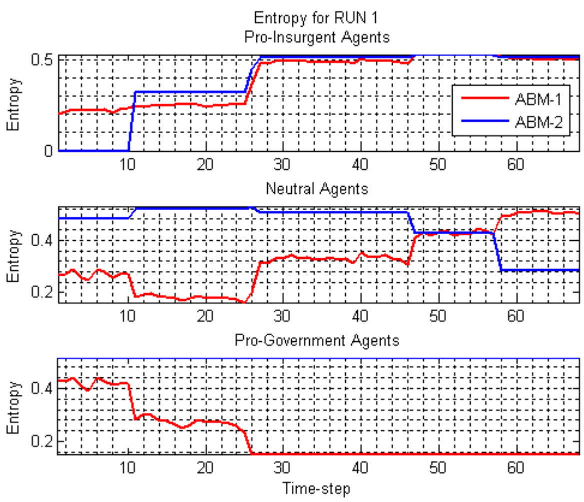
Fig. 2. A map of the Ghazni region in Afghanistan that was used for the case study. The aggregate level ABM-1 patch and the finer resolution ABM-2 grids are indicated. The district boundaries are marked, and each district is assigned an identification number for the purposes of the simulations.

B. *Comparison of Entropy Outputs*

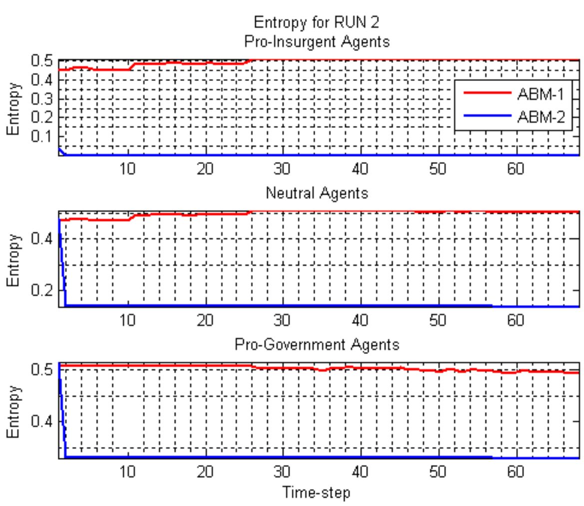
In this section, we compare the entropy of the number of agents in each for each social simulation model. Again, only plots of the entropy of the first three simulations are shown in Figure 3 due to limitation on the number of pages; however, similar conclusions can be drawn from the other six outputs not shown.

One insight from these plots and others not shown in this paper is that in almost all the cases the entropies of the outputs from both models are different over time; except pro-insurgent class for runs 1, 4, and 7.The implication of this observation is that the interpretation of entropy may lead to conflicting conclusions; the determination of which of these models is accurate is a subject for future research. However, the fact that two analogous models could give different entropies is a prerequisite for additional investigations.

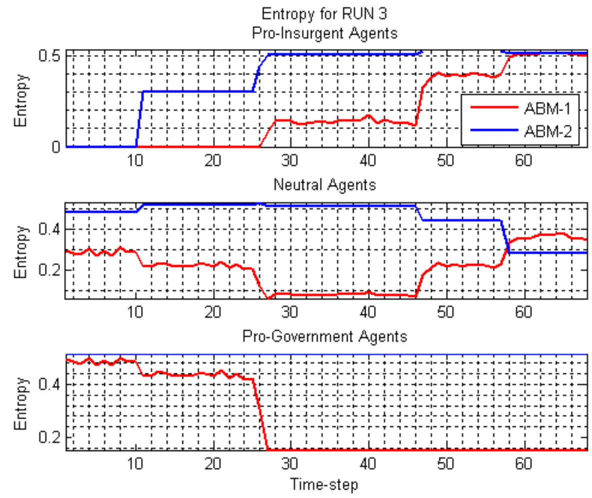
Another insight is that simulations based on L and C theories (Figures 3a and 3c) are reactive to temporal changes; whereas, simulations based on R theory (Figure 3b) are non-reactive to temporal changes. Furthermore, in some cases, both models give similar outcomes with small variations (see Figures 3a and 3c - top plots); whereas in other cases, both models give opposite outcomes (see Figures 3a and 3c - middle and bottom plots). The obvious implication of this insight is that one of the models is not a complete representation of the expected social behavior.



(a)



(b)



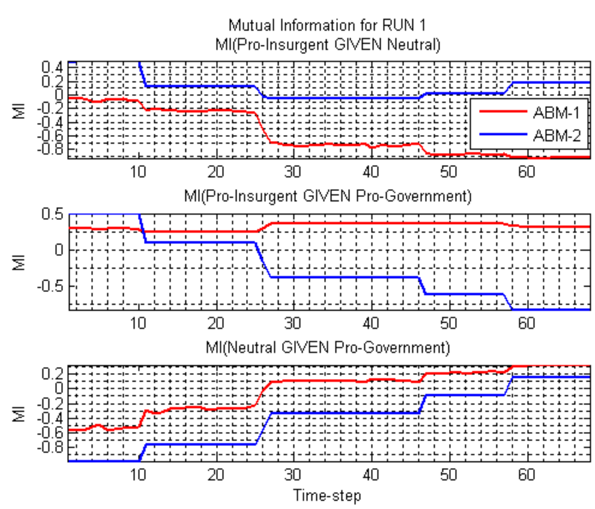
(c)

Fig. 3. Outputs of entropy computation for three of the nine simulation runs are shown in (a) to (c) respectively. The outputs generated from ABM-1 (red) and ABM-2 (blue) correspond to the entropy of the number of agents exhibiting pro-Insurgent (top of each panel), neutral (middle) and pro-government (bottom) behavior.

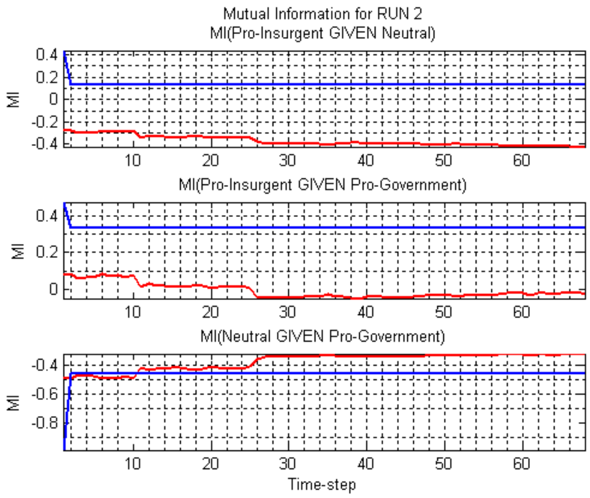
C. Comparison of Mutual Information Outputs

Mutual information is useful for quantifying emergence in complex social systems. Therefore, the first step is to determine how much the knowledge about the number of agents in a particular class could tell about one of the other two classes. To do this successfully, we also compare the mutual information of one class given another class for the two models. The results are shown in Figure 4 for runs 1 to 3. Again from these plots, we notice differences in the MI values over time as we have seen with the entropy computations.

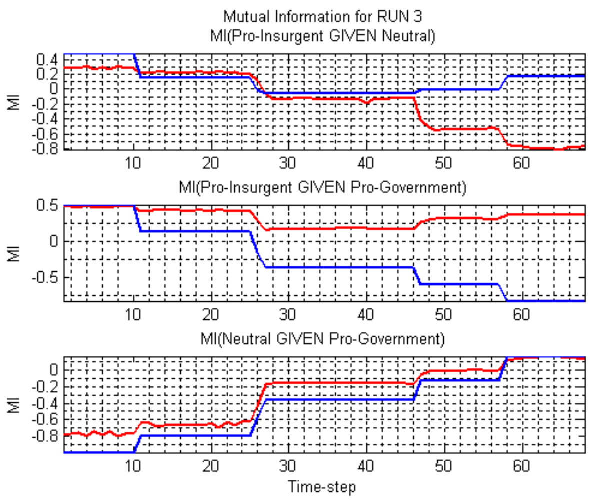
One insight from these results is that the MI (correlation) between classes in all cases ranges from -0.8 and 0.5, which indicates that there is no strong correlation between the classes; hence, they are not overly dependent on each other. The implication then is that the classes are overly independent in most cases.



(a)



(b)



(c)

Fig. 4. Outputs of mutual information computation for three of the nine simulation runs are shown in (a) to (c) respectively. Each plot corresponds to the mutual information of the number of agents in a class GIVEN the number of agents in another class with red plots for ABM-1 and blue plots for ABM-2.

Another insight, considering only ABM-2 outputs, is that for runs 1 and 3, the MI(pro-insurgent GIVEN pro-groverment) decreases over time; whereas, the MI(neutral GIVEN pro-government) increases over time. The increase and decrease in MI over time is an evidence of phase transition; however, we cannot determine if the system is transitioning from a stable system to an unstable system or vice versa. The implication of phase-transition dynamics is that it supports the possibility of emergent behavior; which suggests that MI may be used for quantifying emergent behavior.

V. Conclusions

This paper focuses on systematic evaluation of human, social, cultural, and behavioral (HSBC) modeling and simulation (M&S) systems. Even though the results presented in this paper are preliminary, we believe that an important and promising step, albeit small, has been taken toward achieving the ultimate goal.

The purpose of this paper was to demonstrate, in a preliminary and proof-of-concept fashion, the feasibility of developing distance measures to compare multiple simulation results, as well as to compare simulations with observations (even when such observations are noisy, sparse, partial, or incomplete), with the goal of evaluating performance of HSBC systems in terms of modeling predominant behavior and processes, extreme and surprising behavior, and rare tipping points.

The area of HSBC M&S suffers from models that are poorly understood (relative to models for most physical, built, or natural systems) and data that are inherently noisy, sparse, and incomplete. Thus, validation takes on the form of characterization and systematic evaluation, with the ultimate aim of providing value to end users and stakeholders such as military commanders. The results presented in this paper are a part of the first step in this challenging direction.

Acknowledgment

This research was performed at the Oak Ridge National Laboratory, managed by UT-Battelle, LLC for the U.S. Department of Energy under Contract DE-AC05-00OR22725. The United States Government retains, and the publisher by accepting the article for publication, acknowledges that the United States Government retains, a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

The research is part of a broader effort called "Actionable Capabilities for Social and Economic Systems" funded by DARPA to a team comprising ORNL, University of North Carolina at Charlotte, Georgia Tech and Oak Ridge Associated Universities. The social science theories and computational implementations were developed as part of the broader project [8] and described here for completeness.

References

[1] Bhaduri, B., Bright, E., Coleman, P., and J. Dobson (2002): *LandScan: locating people is what matters*, **Geoinformatics**, 5(2):34–37.

[2] Brown, D.G., Riolo, R., Robinson, D.T., North, M., and W. Rand (2005): *Spatial process and data models: toward integration of agent-based models and GIS*, **Journal of Geographical Systems**, 7(1):25–47.

[3] Bryson, J.J., Ando, Y., and H. Lehmann (2007): *Agent-based modelling as scientific method: a case study analysing primate social behaviour*, **Philosophical Transactions of the Royal Society, B -- Biology**, 362(1485):1685–1698.

[4] Coleman, J.S. (1990): **Foundations of Social Theory**, Cambridge, Massachusetts: Harvard University Press.

[5] Cover, T.M., and J.A. Thomas (2006): **Elements of Information Theory** (2nd ed.). New Jersey: Wiley Interscience.

[6] Dobson, J.E., Bright, E.A., Coleman, P.R., Durfee, R.C., and B.A. Worley (2000): *LandScan: a global population database for estimating populations at risk*, **Photogrammetric Engineering and Remote Sensing**, 66(7):849–857.

[7] Fagiolo, G., Birchenhall, C. and P. Windrum (Eds.) (2007): *Special issue on empirical validation in agent-based models*, **Computational Economics**, 30(3).

[8] Fernandez, S.J. et al. (2008). Actionable Capability for Social and Economic Systems (ACSES). Available at

http://www.osti.gov/bridge/product.biblio.jsp?query_id=0&page=0&osti_id=933058. Accessed on May 5, 2009.

[9] Husslage, B., Rennen, G., van Dam E.R., and D. den Hertog (2006): *Space-filling Latin hypercube designs for computer experiments*, Working Paper No. 2006-18, ISSN 0924-7815, Department of Econometrics and Operations Research, Tilburg University, March 2006.

[10] Marks, R. (2008): *Validating simulation models: a general framework and four applied examples*, **Computational Economics**, 30(3):265–290.

[11] Potok, T., Elmore, M., Reed, J., and F.T. Sheldon (2003): *VIPAR: advanced iInformation agents discovering knowledge in an open and changing environment*, **Proceedings of the 7th World Multiconference on Systemics, Cybernetics and Informatics: Special Session on Agent-Based Computing**, Orlando, Florida, pp. 28–33, July 27–30.

[12] Potok, T.E., Ivezic, N.A, and N.F. Samatova (2000): *Agent-based architecture for flexible lean cell design, analysis, and evaluation*, **Proceedings of the 4th Design of Information Infrastructure Systems for Manufacturing Conference**.

[13] Santner, T.J., Williams, B.J., and W. I. Notz (2003): **The Design and Analysis of Computer Experiments**, Springer Series in Statistics, 283 pages.

[14] Sterman, J., and H. Rahmandad (2008): *Heterogeneity and network structure in the dynamics of diffusion: comparing agent-based and differential equation models*, **Management Science** (forthcoming).

[15] Tesfatsion, L., and K.L. Judd (Eds) (2006): **Handbook of Computational Economics**, **Vol. 2: Agent-Based Computational Economics**, Handbooks in Economics Series, North-Holland/Elsevier, Amsterdam.

[16] Tisue, S., and U. Wilensky (2004): *NetLogo: A simple environment for modeling complexity*, **International Conf. Complex Systems**, Boston, Massachusetts, May 16–21.

[17] Vicsek, T. (2002): *Complexity: the bigger picture*, **Nature**, 418:131.

[18] Windrum, P., Fagiolo, G., and A. Moneta (2007): *Empirical validation of agent-based models: alternatives and prospects*, **J. of Artificial Societies and Social Simulation**, 10(2,8).