Validation techniques of agent based modelling for geospatial simulations

Mehdi Darvishi [a, *], Gholamreza Ahmadi[b]

[a] PhD Candidate in Remote Sensing, Department of Remote Sensing and GIS, Geography Faculty, Tehran University, Tehran, Iran
mehdidarvishi@yahoo.com
[b] PhD Candidate in meteorology, Department of Geography, Science and Research Branch, Islamic Azad University, Tehran, Iran
ahmadigholamreza56@yahoo.com

KEY WORDS: Verification, validation, calibration, agent based modelling, Error, Artefact

ABSTRACT:

One of the most interesting aspects of modelling and simulation study is to describe the real world phenomena that have specific properties; especially those that are in large scales and have dynamic and complex behaviours. Studying these phenomena in the laboratory is costly and in most cases it is impossible. Therefore, Miniaturization of world phenomena in the framework of a model in order to simulate the real phenomena is a reasonable and scientific approach to understand the world. Agent-based modelling and simulation (ABMS) is a new modelling method comprising of multiple interacting agent. They have been used in the different areas; for instance, geographic information system (GIS), biology, economics, social science and computer science. The emergence of ABM toolkits in GIS software libraries (e.g. ESRI's ArcGIS, OpenMap, GeoTools, etc) for geospatial modelling is an indication of the growing interest of users to use of special capabilities of ABMS. Since ABMS is inherently similar to human cognition, therefore it could be built easily and applicable to wide range applications than a traditional simulation. But a key challenge about ABMS is difficulty in their validation and verification. Because of frequent emergence patterns, strong dynamics in the system and the complex nature of ABMS, it is hard to validate and verify ABMS by conventional validation methods. Therefore, attempt to find appropriate validation techniques for ABM seems to be necessary. In this paper, after reviewing on Principles and Concepts of ABM for and its applications, the validation techniques and challenges of ABM validation are discussed.

1. INTRODUCTION

For the transfer of real world to model, we use abstraction. This abstraction needs to carry concepts from real word to computational models. Since simulation models are abstract representations of real state systems, they help us to increase our abilities to control, forecast or understand the behaviour of the system. Therefore, simulation models are widely being applied to help in decision-making and to solve complex problems. All users of models and decision makers want to know whether the outputs of a model are correct or not. Thus it can be very difficult to be sure if the behavior that we are observing is truly representative of the actual system (Hodges and Dewar, 1992). Actually, the validation process determines how accuracy a simulation model according to the system. Model variation is a process that determines whether the conceptual model is a reasonably accurate representation of the real world (Law and Kelton). Since there is not a unique validation method, therefore the validation process is a difficult work. Numerous approaches and techniques of validation and verification models have been provided so far by researcher. Validation, verification and calibration are common processes that used in accuracy of model simulation. Verification is generally defined as the process of testing whether or not the logic of the model is acceptable (Crooks et al, 2007). Validation refers to the extent that the model adequately represents the system being modeled (Casti, 1997). Calibration is an iterative process adjusting the unmeasured or poorly characterized experimental parameters in the programming implementation (Xiang et al, 2005). Some researchers provided 75 validation

and verification techniques that are widely applied in validating the models of marketing, engineering and industrial applications (Balci, 1996 and Banks et al, 2001). In an overall view, validation approaches are classified into two general categories: 1) quantitative methods, also called statistical methods that use statistical approaches to evaluate the credibility of the simulation model, Docking and historical data validation are examples of these methods, and 2) subjective methods that are based on the judgment of experts. Black-box testing, internal validity, Turing test and face validation are examples of the subjective methods. In another classification, validation methods are categorized into three classes: 1) operational validation, 2) computerized model verification, 3) conceptual model validation (Sargent, 2004). Figure 1 shows this classification and its internal relationships. Recently, some new approaches of ABM validation are proposed for a special application. VOMAS (Virtual Overlay Multi-Agent System) is one of these new validation methods that used to monitor spatial and non-spatial concepts and VOMAS agents can be used for validation. It works based on a scenario-modeling approach (Niazi et al, 2010). In this paper, after a short introduction to the concepts of ABMS, validation techniques of ABMS are investigated in details.
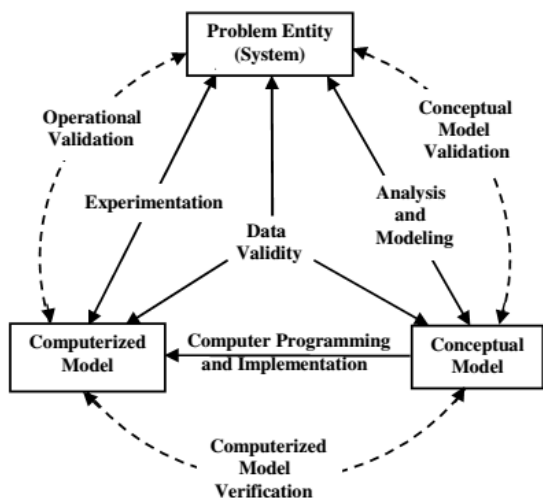
*Corresponding author

Figure 1. Simplified Version of the verification models and Modelling Process.

2. CONCEPTS AND PRINCIPLES OF ABMS

Agent-based modelling and simulation is a new technique in modelling methods in complex and dynamic systems that involving interacting, autonomous agents. The behaviour of agents can be expressed by simple rules and interactions between agents can lead to a state that each agent can be affected by other agents, learn from their experiences and adapt their behaviours. The heterogeneity of agents and the self-organization are two main properties that distinct the ABM from other simulation techniques. A typical ABMS often consists of three main elements: a) Agents, with its own characteristics and behaviors, b) topology (relationships and interaction between agents), c) the environment of agents. These key elements are determined by model developers. In figure 2 we can see the elements and their relationships.
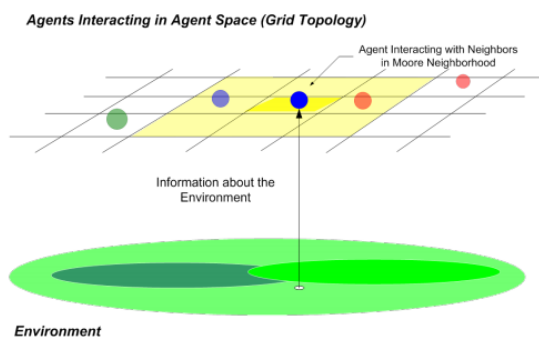

Figure 2. The structure of a typical ABM (Epstein and Axtell, 1996).

In addition to some specifications that mentioned above, regardless of applications, there are several features that are common to most ABM. **Autonomy**, agents are independent units and work without influences of each other and make decision independently. **Heterogeneity,** agents can have different attributes (developed individually) and groups of agents can exist (spawned bottom from the bottom-up). **Active,** agents can independently influence the simulation with following features: a) Pro-active/goal-directed (having goals to achieve), b) reactive/perceptive (sensitive and aware to their environment), c) bounded rationality (agents are rational optimizers with unlimited access to information), d)

interactive/communicative (ability to communicate with other agents), e) mobility, f) adaptation/Learning. Agents can be indicator of any type of autonomous entity such as cars, humans, land segmentations and animals. ABM can be built by use of different platforms, modeling toolkits and object orientated languages. Nowadays, ABMs have been widely used and developed in diverse domains such as modelling economic processes, price validations within stock-market trading, voting behaviors in elections, archaeological reconstruction of ancient civilizations, and growth of bacterial colonies. They have recently been developed for geospatial research to exploit ABM to environmental modelling such as urban growth models, traffic jam, crowd congestion modelling and land-use and land-cover variations modelling. ABMS have three prominent advantages over traditional modeling techniques. They are capturing emergent phenomena, presenting natural environment for the study of systems and having high flexibility. Also it is important to note that ABMs encounter some limitations and considerations. They are sensitive to initial conditions and to small variations in interaction rules (Couclelis, 2002). One of the most challenging issues about ABMs is verification, validation and calibration process, which is the subject of this paper. First of all, it should be clarified the exact meaning of verification, validation and calibration processes. We are addressed these concepts in this paper.

3. VALIDATION METHODS FOR ABM

Validation is a process that helps us to understand whether a model can produce valid and correct outcomes. Before performing the validation step, calibration process must be done. Calibration defines as determining the model parameters using real world data. Both mentioned processes are a big challenge for modellers, due to the vague and debatable definition of validation and different validation provided. Numerous validation techniques have been developed by researchers (Parker et al, 2002), (Windrum et al, 2007) and (Klügl, 2008). In a general classification, validation methods can be divided into three categories:

- **Structural validation**: shows the way that a system works to generate and regenerate the observed system behaviour.
- **Predictive validation**: relating to ability of a model to forecast a behaviour that it has not observed before.
- **Replicative validation**: based on the comparison between the outputs of model and data obtained from the real world.

In the following, we will focus on the structural validation method. Structural validation involves four main processes:

- **Face Validation:** it is performed in the initial step of a simulation process under supervision of conceptual validation. This method consist of three components: a) output assessment, b) immersive assessment, c) animation assessment
- **Calibration**: a process that determines the domain of values for the parameters and fit the model with real data.
- **Sensitivity Analysis**: considering the effects of various parameters on specific behaviour or model outcomes.
- **Output Validation**: matching statistically and spatially the forecast of model's versus real data.

There are two methods in face validation step that experts can judge the accuracy of a model:

- **Animation**: It is a graphical view of the model behaviour during the run-time model. This feature is supported by some simulation software and while the simulation is running, they show the general performance of simulation or even a specific trait in simulation. Swarm and Repast are the examples of this type of software.
- **Graphical representation**: shows the statistical variables of model outputs by some graphs.

In addition to mentioned two main ways above, there are other methods for face validation. **Tracing** is a method the same as Animation that monitors the agents behaviour, especially for isolating strange behaviour, for considering the accuracy of model logic. **Historical data** is applied if there are Historical data. Usually data are divided into two parts, training and test part that are employed in building and testing the model respectively. **Parameter variability** is a technique that is considered the variations of input variables and parameters to determine the effect on the model outputs. By this process the most effective and sensitive parameters will be determined. **Predictive validation** is based on the comparing Actual behavior of the system and the model prediction. **Turing test** is based on the Expert Opinion. The expert considers the output model and real system then he want to know whether there is distinguishable difference between them. If mentioned validation steps above are performed in a model, we ensure that the model produce valid and correct outputs. Sometimes face validation and sensitivity analysis are collectively called as verification (Parker et al, 2002).

3.1 Verification

A process that test whether the logic of a model is acceptable and valid or not. Verification start with considering processes in the model and then comparing the model results versus the real data. As mentioned, verification involves the two steps of face validation and sensitivity analysis. Face validation related to the reliability of plausibility and reasonability of initial outputs. Sensitivity analysis is related to the considering the effects of the model parameters on the outputs. As a result, the parameters that have no a considerable effect on model outputs are omitted from the rest of the process. Monitoring and analyzing the dynamic attributes of the agents and checking the automatic updating of parameters for stability and precision are called the animation and immersive assessments. The next step of face validation is output assessment. Output assessment says whether simulated results are in an acceptable range of real values during the simulation or not. In this step, successive iteration runs are used to ensure that the outputs are located in the range (corresponding to the real world). Finally, the results are statistically compared with real data. The Standard Errors (SE) and the mean values of distribution are compared with each other. If Standard Errors are less than 5%, the mean values are compared with real data. In case of ABMs related to the land use change, since successive iteration runs generate outputs at different scales and the level of details are decreased at lower resolutions, therefore, there will be some differences between model outputs run at different scales. If the difference is not statistically considerable, for decreasing the running time of model, it can be ran the model at the coarser scale. This time reduction will be important, especially when the model is used to a large area. In sensitivity analysis process, before calibration model we need sensitivity analysis for choosing important parameters for the simulation. To perform sensitivity analysis, traditional methods change or manipulate just one parameter at certain time and leaving the other parameters constant. But Traditional method is not to feasible to ABMs. Generally, unstructured approach is applied to sensitivity analysis for ABMs. Design of Experiments (DOE) is a statistical approach that have suggested to sensitivity analysis for ABMs (Kleijnen et al, 2003). In DOE technique, model inputs and outputs are named factors (i) and responses, respectively. $n = i^j$ Indicates that n simulations need to specify the effect of i factors, assuming each factor takes j levels. Of course, DOE can only be used to a limited number of factors, due to the increasing computation time with the addition of other additional factors. Sequential Bifurcation (SB) approach is suggested (Bettonvil, 1997) to deal with the problem of computation time. It is a technique for determining the most significant factors among other factors that affected the efficiency of model. At first the parameters are converted to the binary values (0 or 1), then we obtain the approximation of the output y in following:

$$y = \beta_\circ + \beta_1 x_1 + \cdots + \beta_j x_j + \cdots + \beta_k x_k \qquad (1)$$

$\beta_\circ$ And β_j are overall mean and main effects respectively. y is a first-order polynomial of the variables ($x_1 \cdots x_j \cdots x_k$). There are three assumptions in SB technique: first, errors are assumed to be zero. Second, the sign of each factor effect is known. Third, the factors can be approximated as a first-order polynomial. The operation in SB is an iterative process and each factor is selected based on the results of previous factor that have been already simulated. For more details see (Bettonvil, 1997).

3.2 Calibration

When the sensitivity analysis is done once, the calibration process is performed on the model. Calibration is in order to find suitable values for the model parameters that are the best fit with the real world. For calibration we need to optimize the parameters. A numerous optimization methods exist, but genetic algorithm is one of the suitable methods for optimizing the model parameters. Genetic algorithm with ability of parallel search across a large solution space can well afford to cope. Generally, genetic algorithm is applied to calibration process of ABMs. Validation output is the last step of validation process. In this step we will ensure that the model works correctly. Outputs validation indicates to what extent a model can show the real world. There are some methods that can determine how the model outputs match the real data. ROC[1] is an analysis that uses from a category or forecast plan to determine what cases locate in a specific category. The basic of ROC is based on the ROC curve. This curve is a graphical representation that shows connection between property (x-axis) and sensitivity (y-axis) of a test for thresholds. This curve gives us a tool to properly categorize or forecast households, more details see (Beck ea al, 1986). MRG[2] is another useful analysis for output validation. MRG measures spatial patterns of the output at multiple scales, especially for geographical model outputs. MRG process is based on the following equation:

---

[1] *Relative Operating Characteristic*

[2] *Multiple Resolution Goodness-of-Fit*

$$F_w = \frac{\sum\limits_{s=1}^{t_w} \left[1 - \frac{\sum\limits_{i=1}^{p} |a_{1i} - a_{2i}|}{2w^2} \right]_s}{t_w} \qquad (2)$$

a_{ki} is the number of cells of class i, F_w is the fit for the sampling window size w, p is the number of different classes in the sampling window, s is the sampling window of dimension w by w that moves over the image, t_w overall sampling windows in the image of window size w. MRG is a powerful method of evaluating the adaptability of the spatial model outputs rather common accuracy method in GIS and remote sensing, for example, kappa coefficient and confusion matrix. In addition to the spatial patterns, also it provides the details of the temporal patterns. For more details see (Costanza, 1989).

4. STATISTICAL APPROACH

We examined some validation techniques in section 3. Those techniques are considered as subjective validation category. In contrast to these methods is quantitative technique. Using statistical analyses in model validation causes to improve the reliability of the model. To perform statistical methods, we can take two approaches. First, a statistical comparison between the model outputs and data of the real system, second, we can run one or several model then statistically compare the outputs with each other. Often, the statistical validation techniques use features for quantitative analyses such as means, variances and confidence intervals. The distribution of model outputs can be used to determine an acceptable range of accuracy. More discussion in statistical approach to the validation you can refer to (Kleijnen, 1999).

5. ERRORS AND ARTIFACTS

Errors when occur in a model that it dose not work according to what requirements characteristics applied by its developer. In other words, the errors are maladjustment between what is in the developer's mind and what should really is. Therefore, if a model dose not operates properly, there may be errors in the model. Actually, verification is a process of seeking errors in a model. Some errors may occur in programming time, for example, we want to write a loop for a complete set of agents, but mistakenly we writes a loop for a subset of agents or a situation that we expect the program calculate the computations based on the rules of real arithmetic, but the model generates the outputs based on the floating-point arithmetic (Polhill et al, 2005). Contrary to what was said about errors, in artefacts there is not any maladjustment between what is in the developer's mind and what should really is, of course, artefacts is a maladjustment between the assumptions employed in the model by developer for reaching to the specific event and the assumptions that are really lead to of such event. The assumptions made in model can be divided into two parts, major and minor assumptions. Major assumptions are those whose presences are necessary for running and function of the model. Minor assumptions are those whose presences are not vital part of the model and they cause the model work. Also there is another term that called significant assumptions; they lead to effective and considerable results in outputs of the model. Artefacts are important events that created by minor assumption in the model that are wrongly assumed non-significant assumptions. Actually, artefacts happen when minor assumptions are assumed as non-significant by developer. As

the matter of the fact, due to the changes in some minor assumptions (despite the lack of changes in major assumptions), we will observe different results in output of the model. Artefacts and errors may occur in any different steps of modelling process. For more details refer to (Galan et al, 2005).

6. CONCLUSION

Decision maker for using output of model must be assured of the accuracy of it. Therefore, we need to verify, validate and calibrate the model to ensure the performance and accuracy of the model. Some especial properties of Agent based modeling such as learning process; interactions between other agents, data independence, heterogeneity and complexity of the agents make the validation of ABM difficult. Because of that, nowadays a numerous validation technique has been built to solve the problems of validation process of ABM, the various techniques with different approaches from subjective to statistical methods. Each of these techniques is proportional to a specific application, with a different efficiency, limitations, considerations and accuracies. Therefore, all modelers in all domains of sciences have attempted to develop new methods to improve the efficiency of validation techniques.

REFERENCES

Balci, O., 1998. Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice. Chapter validation, verification and testing, John Wiley & Sons, New York.

Banks, J., Carson, J., Nelson, B., 2001. Discrete-event system simulation. Industrial and systems engineering, Prentice Hall, third edition.

Bettonvil, B., and Kleijnen, J. P. C., 1997. Searching for important factors in simulation models with many factors: Sequential bifurcation. European Journal of Operational Research, 96, 180–194.

Beck, J. R., and Shultz, E., 1986. The use of relative operating characteristic (ROC) curves in test performance evaluation. Archives of Pathology & Laboratory Medicine, 110, 13–20.

Crooks, A., Castle, C., and Batty, M., 2007. Key Challenges in Agent-Based Modelling for Geo-Spatial Simulation. National Centre for Geocomputation (NCG), National University of Ireland, Maynooth, Co. Kildare, Ireland. pp.11.

Costanza, R., 1989. Model goodness of fit: A multiple resolution procedure. Ecological Modelling, 47, 199–215.

Couclelis, H., 2002. Modelling frameworks, paradigms, and approaches. In K. C. Clarke, B. E. Parks, & M. P. Crane (Eds.), Geographic information systems and environmental modelling (pp. 36–50). London, Prentice Hall.

Casti, J. L., 1997. Would-Be-Worlds: How Simulation is Changing the Frontiers of Science. John Wiley & Sons, New York.

Epstein, J.M., and Axtell, R., 1996. Growing Artificial Societies: SocialScience from the Bottom Up. MIT Press: Cambridge, MA.

Galan, J. M., and Izquierdo, L. R., 2005. Appearances Can Be Deceiving: Lessons Learned Re-Implementing Axelrod's

'Evolutionary Approach to Norms'. Journal of Artificial Societies and Social Simulation, 8(3) 2.

Hodges J. S., and Dewar J. A., 1992. Is It You or Your Model Talking? A Framework for Model Validation. RAND Corporation, Santa Monica, California.

Klügl, F., 2008. A validation methodology for agent-based simulations. In Proceedings of the 2008 ACM Symposium on Applied Computing – Advances in Computer Simulation. ACM. New York, NY, USA, 39–43.

Kleijnen, J. P. C., Sanchez, S. M., Lucas, T. W., & Cioppa, T. M., 2003. A user's guide to the brave new world of designing simulation experiments. Tilburg: Tilburg University, Center forEconomic Research.

Kleijnen, J. P. C., 1999. Validation of models: statistical techniques and data availability. In proceddings of the 1999 wintersimulation conference.

Law, A. M., and Kelton, W. D., 1991. Simulation modelling and analysis. McGraw-Hill, New York, second edition.

Niazi, M., Siddique, Q., Hussain, A., Kolberg, M., 2010. Verification & Validation of an Agent-Based Forest Fire Simulation Model. Proceedings of the Agent Directed Simulation Symposium 2010, as part of the ACM SCS Spring Simulation Multiconference, pp142-149.

Parker, D. C., Manson, S. M., Janssen, M. A., Hoffmann, M. J., & Deadman, P., 2002. Multiagent systems for the simulation of land-use and land-cover change: A review. Annals of the Association of American Geographers, 93, 314–337.

Polhill., J. G., Izquierdo, L. R., and Gotts, N. M., 2005. The ghost in the model (and other effects of floating point arithmetic). Journal of Artificial Societies and Social Simulation, vol. 8, no. 1.

Sargent, R. G., 2004. Validation and verification of simulation models. Proceedings of the Simulation Conference, pp. 17-28.

Windrum, P. F., Fagiolo, G., & Moneta, A., 2007. Empirical validation of agent-based models: Alternatives and prospects. Journal of Artifi cial Societies and Social Simulation, 10(2), 8.

Xiang, X., Kennedy, R., and Madey, G., 2005. Verification and Validation of Agent-based Scientific Simulation Models. *Agent-Directed Simulation Conference*, San Diego, CA, April, pp. 47-55.