Chapter 11
Comparisons of Validated Agent-Based Model and Calibrated Statistical Model

Il-Chul Moon and Jang Won Bae

11.1 Introduction

Modeling and simulation are an abstracted generation of a part of the real world, and their validity on the generation is the holy grail of the modeling and simulation study. In spite that this is a key factor in the credibility of simulation, still there is a large unchartered area in the validation of modeling and simulation. Prof. Tuncer Ören acknowledged this challenge in his article (Sheng et al. 1993; Yilmaz and Ören 2009). He claims that there are three major reasons which make the validation a challenge: philosophical, definitional, and theoretical reasons. He points out the philosophical problem by quoting Thomas Kuhn and Karl Popper, and he contrasts two arguments: "…theories are confirmed or refuted on the basis of critical experiments designed to verify the consequences of theories" from Kuhn and "… as scientists we can never validate a hypothesis, only invalidate it" from Popper. If we take Kuhn's positivism argument, it is imperative to validate a simulation model by the consequence of the model execution, yet Popper might agree that a simulation model will be only invalidated when a real-world case opposing to the simulation result is found.

Prof. Tuncer Ören also points out the various definitions of model validation. For instance, a model validation can be measured by the standard of either correct representation or acceptable representation. This definition difference inherits the Kuhn's view and the Popperian view, respectively. When we narrow down the scope to the computer simulation, the Society of Computer Simulation (SCS) defines the validation as "substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model." This definition seems to be a mixture of

I.-C. Moon (✉) • J.W. Bae
Department of Industrial and Systems Engineering, KAIST, Daejeon, South Korea
e-mail: icmoon@kaist.ac.kr; repute82@kaist.ac.kr

L. Yilmaz (ed.), *Concepts and Methodologies for Modeling and Simulation*,
Simulation Foundations, Methods and Applications,
DOI 10.1007/978-3-319-15096-3_11

arguments from correct representation and acceptable representation. *The satisfactory range of accuracy* suggests that the correct representation with a certain level of quantitative confidence is required, yet *the intended application of the model* asserts the importance of the acceptance of the model.

Prof. Tuncer Ören argues that one distinction of the validation study in the modeling and simulation is the relative validity concept from the experiments of simulation models. Since a simulation is a limited representation of the real world, and because the limited representation is determined by the modeler, the modeler is able to adjust the scope to make the model more valid. The simulations often used as a part of virtual experiments which is a part of the typical scientific method. The virtual experiments consist of (1) a simulation model, (2) generated/collected/observed data, and (3) the experiment framework. The scope of the model depends on the model features as well as the used dataset in the experiment design.

Though time has passed since Prof. Tuncer Ören's works on validation, this community still struggles to have better theoretical frameworks as well as case studies on validation. Furthermore, when the users of simulation models call for the predictive results, simulation models are considered to be weaker solutions than statistical models only designed for predictions. Such perception is in the line of the continuing argument of *correctness vs. acceptance* of models. However, some simulation models might be better at the prediction as well as the replication to a certain scope at a certain problem set. This chapter introduces such case to see whether there is a chance for a simulation model to be good at the two aspects simultaneously.

11.2 Background

Agent-based model consists of multiple agents, environments, and interactions among the agents and environments (Yilmaz and Ören 2009; Bonabeau 2002). Such model structure reflects a holistic view where the aggregate of a system is different from the sum of its components. This view has been recognized as an efficient method to analyze complex systems that contain a large number of components and interactions among them. From this reason, agent-based model has been applied to understanding, replicating, and resolving problems in various domains, such as sociology (Schelling 1971), economy (Tesfatsion 2003), biology (Grimm et al. 2005), etc. Agent-based modes are generative models so that they can provide unrecorded, yet important information to model users. For example, in the disaster event (Lee et al. 2014), how the disaster response organizations efficiently reacted is hard to be evaluated, because their efficiency is rarely recorded in such immediate events. However, agent-based model, as a generative model, can help us to estimate such efficiency using several partial data which seem not to be related to the efficiency.

Although agent-based model provides such invaluable information, outsider as well as insider of the agent-based model community often considers about its accuracy for the prediction (Moss 2008; Brown et al. 2005; Windrum 2007).

To compare with the model for accurate prediction, such as statistical model, agent-based model might have less accuracy in the prediction. Statistical model describes how a set of variables are related to the other set of variables mathematically. From such mathematical bases, statistical models often represent the accurate predictions. This accuracy depends on the proper data for the prediction, yet finding such data in the real world would be another problem.

Hence, this chapter intends to investigate the differences of the model for regeneration of the real world, i.e., agent-based model, and the model for accurate predictions of the real world, i.e., statistical model. The investigations include quantitative comparisons of the two types of models. As a case study, an agent-based model for city commerce was compared to the corresponding statistical model. This particular comparison, again, casts light on the trade-off of different contributions from different models.

11.3 Validation of Agent-Based Models

The validation of simulation models can be completed at the various levels and through diverse methodologies (Moss 2008). These different types of validations range from the qualitative assessments of models, i.e., quality assurance of models, to the quantitative validations, i.e., correlation between the simulated world and the real world. Some models in a certain virtual experiment can be verified by the quantitative analyses, while other models are only able to be qualitatively validated.

Let's imagine that there is a traffic simulator designed for urban areas. This traffic simulator might be validated for daily traffic estimation which can be validated by historic records of the simulated area. Since it is daily estimation, the validation data would be sufficient, and the details of the model would be easily expected by the modelers. On the other hand, this traffic simulator might be used for the traffic estimation on city-scale evacuation. This case would be very difficult to be numerically validated because it would be impossible to obtain sufficient datasets to be compared with. Moreover, the modeled details of the evacuation simulator should be different from the models for the daily traffic estimations. These hypothetical scenarios reveal that some models might be quantitatively validated by the support of the experimental framework (see Fig. 11.1), and other models are only subject to qualitative quality assurance of the model.

When the validation of a simulation model is perceived as quality assurance, there are two different approaches to view the simulation model. The first perspective is treating the simulation model as software artifact, which is true because the simulation model to be validated needs to be implemented as a software executable (Yilmaz and Ören 2009). Then, we can apply diverse software quality assurance techniques, for instance, CMMI (Chrissis et al. 2003). This is a comprehensive quality checking over the modelers, the simulation model before implementation, and the simulation software after implementation. The other perspective is adopting a qualitative simulation validation approach. Such qualitative assessment often

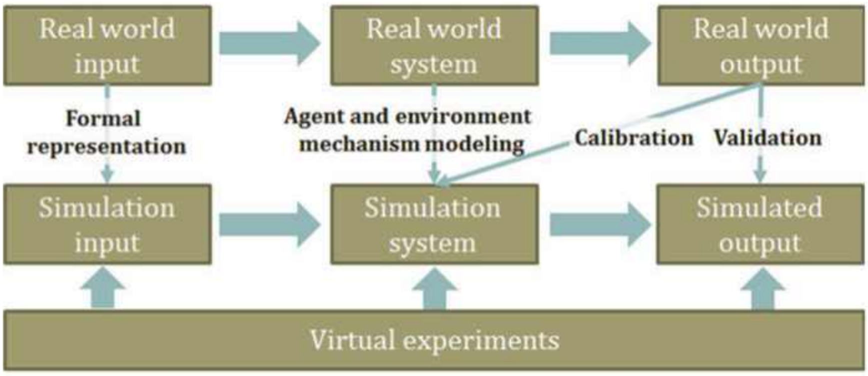

Fig. 11.1 Simulation-based experiment flow and the interaction between the real-world data and the simulated data

relies on the knowledge of subject-matter experts (Sargent 2005). The experts receive the simulation results in various formats, i.e., visualization of simulation progresses, response surface analysis (Inman et al. 1981), etc. With the provided materials from simulations, the experts decide to accredit the simulation model for its validity.

Similar to the levels of quality assurance, the quantitative validation has diverse levels of validation. The fundamental of the quantitative validation is the comparison between the simulated data and the real-world data with statistical approaches. However, the comparison points would be different from the intended acceptance level, which we and Prof. Tuncer Oren discussed in the Background section. A certain virtual experiment might be only interested in the change point of the target system, which is known as tipping point (Davis et al. 2007). Also, another experiment setting would concentrate on the performance and the state dynamics at the end of the target period or at the convergence of the system states. These quantitative validations aim at the point validation, which limits the validity of the simulation to a certain point of the simulated period. The point validation can be simply performed by checking the difference between the simulated point and the real-world point with a null hypothesis. Often, confidence intervals, T-tests, F-tests, etc. are the statistical tools to be used. The experiments with point validation have lower intended acceptance level than the validation on the complete timeline of the simulated period because the point validation does not ensure the generation process of the simulated world to be close to the real world. From this perspective, the point validation is quite analogical to the statistical models for predictions. The statistical models might not produce the generation process of the target system, and the statistical models are only *trained* to be close to the prediction point of the system.

When modelers need to validate the generative process of the simulations, it is necessary to validate the matching between the simulated data and the real-world data, and this process is often called as trend validation (Kleijnen 1995; Barlas 1996). This is different from the point validation which requires a correlation on a single time-step. Because the trend validation requires match the trend, not a point, this validation needs a further calibration than the before. The calibration now includes adjusting the parameters controlling the temporal flow of the simulation. The trend validation can be performed by multiple statistical methods. Firstly, each point over the simulated period might be tested with the point validation technique, such as *T*-test. Secondly, the overtime data from the real world is fitted to an auto-regression model, either linear or nonlinear; then we can see the fitness of the simulated data to the auto-regression model. Finally, there are techniques for temporal data comparisons, such as dynamic time warping. The dynamic time warping technique selects the most similar temporal flow between two overtime data, so this supplements the temporal flow discrepancy between a simulation and the real world.

11.4 Prediction with Agent-Based Model

In this section, we presented city commerce models for a quantitative comparison of validated agent-based model and calibrated statistical model.

11.4.1 Case Study: City Commerce Models

The agent-based model was developed in our previous paper (Lee et al. 2013) to estimate the impact on the city commerce by relocation policy. Relocation policy, which moves city functions from the centralized city to newly developed city, has been implemented in several countries such as the UK, Ireland, and Germany (Marshall 2007). In Korea, the government recently executed a relocation policy to resolve problems from overpopulation in Seoul, which is the capital of Korea, and achieve the balanced regional development. By the relocation policy, some of government branches from Gwacheon city, near Seoul, moved to a newly built city, located about 100 km south from Seoul.

While the government cares about the positive effect of the relocation policy, on the other side, several researchers concern about potential negative impacts associated with relocation, such as depressed economy in the region which people left. For investigating such negative impacts from the relocation policy from Korea, we developed an agent-based model for Gwacheon city.

To evaluate the changes, we established an assumption: the change of city commerce might have a positive correlation with the mobility of the population in Gwacheon city. So, the agent-based model describes a daily movement of the citizen in Gwacheon city. More specifically, the daily movement indicates a traffic flow in a day with respect to the job types of peoples in the city and daily time schedules associated with the jobs. To represent such characteristics in the agent-based model, we applied three kinds of real data: micropopulation data, time-use data, and GIS data.

Micropopulation data are detailed statistics about population in a certain region. In the United States, the micropopulation data are called the Public Use Microdata Sample (PUMS); in Korea, the data are named the Micro Data Service System (MDSS). MDSS contains a collection of attributes, such as the individual's address, occupation, family composition, education level, and so on. Because there is a concern about privacy violation, the dataset is usually provided as an anonymized sample. We could obtain a 5 % sample of population data, which contains 1,189 population data out of 23,780 populations in Gwacheon city, from MDSS, and use them to identify jobs of each agent (see the top of Fig. 11.2).

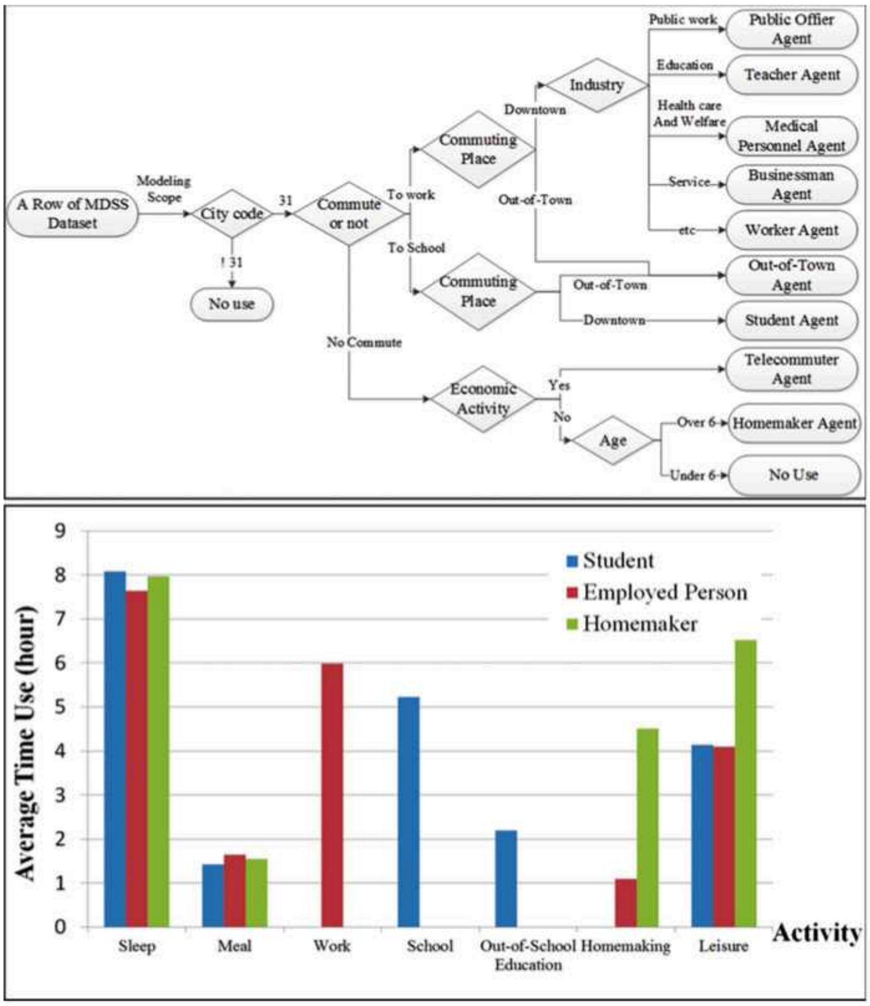

Fig. 11.2 (*Top*) Flowchart of agent-type categorization with MDSS dataset and (*bottom*) an example of time-use statistics of daily activity by student, employed persons, and homemaker

In many countries, time-use data is used to gauge productivity, daily life, and infrastructure efficiency of a population. In Korea, the Korean National Statistical Office provides time-use data that specify how much time an individual with a certain job spends performing a certain activity (see the bottom of Fig. 11.2). This data provides three types of modeling information. First, from the time-use data, we enumerate activity states of agent types and their transitions. Second, the transition time for states of an agent is also specified by the dataset. Lastly, we can develop commuting time of agents, which is important in analyzing and simulating traffic patterns, shopping-in behavior, and regional characteristics.

In the agent-based model, agents are corresponding to vehicles, and they move through the road network in Gwacheon city. Thus, the structure of the road network affects to the simulation objective which is to see the traffic flow in the target city. To reflect geospatial information to the agent-based model, we utilized GIS data of the target region. The data were downloaded from OpenStreetMap, and the data include the information of roads and buildings, such as coordinates, type, and identification. Figure 11.3 shows GIS data in Gwacheon city and replicated road network in the agent-based model using the GIS data. In particular, we selected the commercial buildings of interests to see the city commerce in the agent-based model (buildings with numbering in the right of Fig. 11.3).

On the other hand, some researchers raised a doubt about why agent-based model is applied to investigate the change of the city commerce. They stated that for the city commerce model, statistical model would be a better choice than agent-based model from the perspective of the accuracy in the model prediction.

For the comparison of accuracy from the two models, we develop several statistical models using a linear regression method. It is difficult to consider the above real data, such as MDSS, time-use data, and GIS data, in the development of statistical models, so we applied another data to the model development. In the development of statistical models, data that directly reflect the traffic flow would be

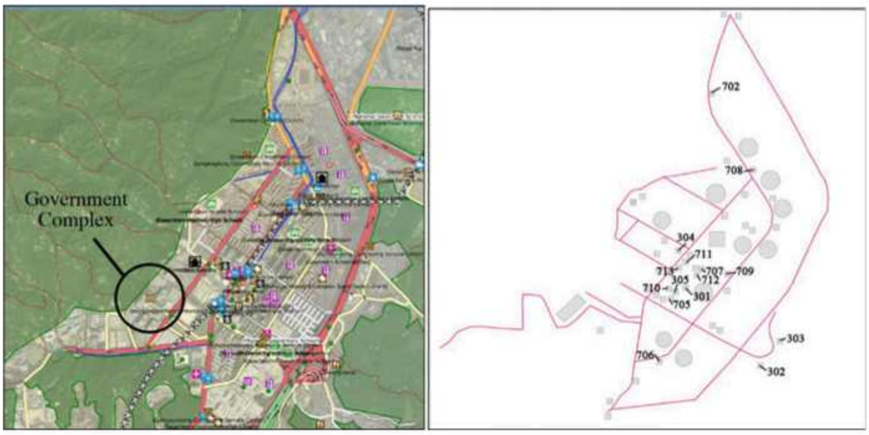

Fig. 11.3 (*Left*) GIS data of Gwacheon city from OpenStreetMap and (*right*) replicated road network and commercial buildings of ID in Gwacheon city using GIS data

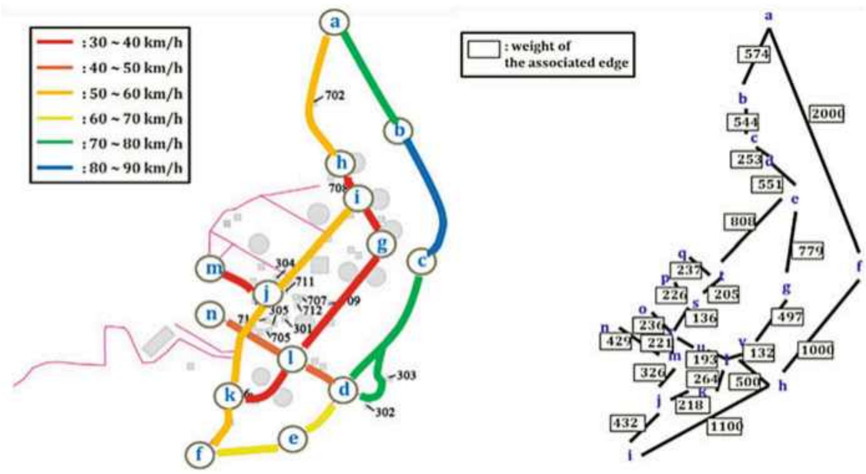

Fig. 11.4 (*Left*) Average traffic velocity of the roads and (*right*) road network with edge weight of road distance

more appropriate, so we collected two kinds of such data: traffic velocity and centrality of the road network.

The traffic velocity indicates average velocity of the roads in Gwacheon city, and it can be a direct barometer for the traffic flow. Traffic velocity of the roads in Gwacheon is opened in Intelligent Transport Systems (ITS) in Gwacheon. Based on the traffic velocity from ITS Gwacheon, we calculated the average velocity of the roads at each hour for a week during 6 months (Mar 2012–Jun 2012 and Mar 2013–Jun 2013). The left in Fig. 11.4 shows the average velocity of the roads in Gwacheon city.

The other is the road network structure derived from GIS data in Gwacheon city. GIS data provide the information of junctions and roads so that the road network can be developed by its vertices as the junctions and its edges as the roads. The right of Fig. 11.4 illustrated the road network of which edge weights represent the distances between two junctions. Using this network structure, we can calculate measures for the network centralities for the traffic flow: degree and betweenness centrality.

Degree centrality is defined as the number of edges that are incident upon a vertex (Freeman 1979). The degree centrality can be interpreted as the immediate risk of a node for catching whatever is flowing through the network, which means that higher degree centrality shows more central node in the network. In the road network, degree centrality (C_{D}) of a junction (j_i) is expressed by Eq. (11.1), where deg (j_i) means the degree centrality of junction j_i and $\sum_{i\gamma_{ij}}$ indicates the number of roads connected to junction j_i

$$C_{\mathrm{D}}(j_i) = \deg(j_i) = \sum\nolimits_{i\,\in\,\text{all junctions}} \gamma_{ij} \tag{11.1}$$

Because the degree centrality is purely a local measure, we applied betweenness centrality, which is a useful measure for both the load and the importance of a node.

The load of a node describes a global effect in a network, whereas the importance of a node shows a local effect of the node. In the road network, betweenness centrality (C_B) of a junction (j_i) is expressed by Eq. (11.2). In Eq. (11.2), σ_{st} is the total number of shortest paths from node s to node t, and σ_{st} (j_i) is the number of those paths that pass through j_i:

$$C_B = \sum\nolimits_{s \neq t \neq j_i} \frac{\sigma_{st}(j_i)}{\sigma_{st}}, \text{ where } s, t \in \text{all junctions} \quad (11.2)$$

11.4.2 Comparison of Predictions from Agent-Based and Statistical Models

To evaluate our assumption, which is that the city commerce would have positive correlations with the mobility of the citizen, we developed agent-based model describing the daily movement in the target city and statistical model using real data that is directly related to the assumption. Now, we intend to compare the accuracy of their predictions by calculating their correlations to a value in the real world.

We collect an indicator of the real city commerce in the target region, which are the rent rates of the commercial buildings in the target region. The direct measure of commercial status of a city might be a sales amount of shops and malls in the target region, but such information is difficult to collect in the city-wide area. Therefore, we chose and collected the indirect measure, and this measure can be collected by personal visits to real estate agencies in the target region.

Before the comparison, we needed a value for calculating the correlations from the agent-based model related with the interested buildings. To do that, we counted the number of the passing-by agents which go through front roads around the interested areas. Similarly, for the statistical models, we mapped (1) traffic velocities to the buildings by calculating the average velocity of the roads around the building and (2) network centralities to the buildings by calculating the average centrality for the roads and the junctions around the building.

We started correlation analyses of the two from drawing scattering plots between the rent rates and the data from the agent-based model and the statistical model (see Fig. 11.5). Although it is difficult to see strong correlations between the two, they represented different trends: a scattering plot for the number of passing-by agents illustrates a positive correlation to the rent rates, yet the plot for traffic velocity shows little correlation, and the plot for network centralities shows, even, negative correlations to the rent rates. Also, we can find that several outliers in each scatter plot and some of them are included in all the scatter plots, such as building 711, 706, and 302.

Since we cannot confirm the correlations from Fig. 11.5 and say that the assumption is surely true, we calculated the correlations between the rent rate and the four data. In particular, we calculated three different correlations in each case: Pearson's value correlation, Spearman's rank correlation, and Kendall's tau rank

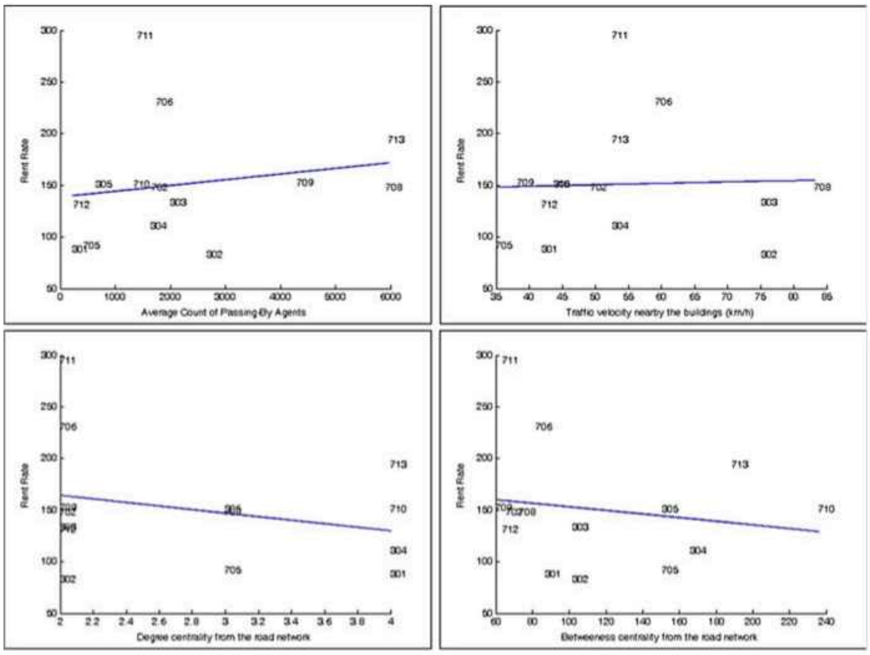

Fig. 11.5 Scattering plots and its linear fitted lines for the rent rates from the real world: (*left-top*) number of passing-by agents, (*right-top*) average of traffic velocity, (*left-bottom*) degree centrality, and (*right-bottom*) betweenness centrality

correlation. Table 11.1 shows the summary of this correlation analysis. When we include every building of interests, the number of passing-by agents and betweenness centrality is recorded the highest positive and negative correlation values, respectively. However, when we exclude the common outliers in Fig. 11.5, the correlation value of the number of passing-by agents becomes 68.3 and 68.5 % in value and rank correlations. This would be a good quality of validation considering the difficulty of the validation of social simulations. However, the correlation values of other data are much lower than the one of the number of passing-by agents.

Also, we built linear regression models for the rent rate as an independent variable and the four factors as the dependent variables (see Table 11.2). Model M1 including all the four variables shows the highest value of R-squared (0.680), which suggests that the prediction from this model is more reliable than any other models with confidence. However, some of the dependent variables showed p-value over 0.05. To find a significant regression model, we performed the backward elimination to M1 and, eventually, developed M6. M6 holds one dependent variable, as the number of the passing-by agents with p-value <0.05 and shows higher R-square value. On the other hand, other factors, such as traffic velocity, degree

Table 11.1 Correlation analysis between the rent rate and passing-by agent counts (agent), average traffic velocity (velocity), degree centrality (degree), and betweenness centrality (betweenness)

Correlation	14 buildings				11 buildings (excluding outliers: building 711, 706, and 302)			
	Agent	Velocity	Degree	Betweenness	Agent	Velocity	Degree	Betweenness
Pearson	0.184	0.036	−0.213	−0.266	0.683	0.239	0.115	−0.078
Spearman	0.337	0.110	0.062	−0.162	0.685	0.211	0.175	−0.068
Kendall	0.256	0.091	0.069	−0.122	0.537	0.133	0.133	−0.043

Table 11.2 Linear regression models for the rent rate with varying variables

Model	Variables	Std. coefficients	p-value	R^2
M1	Agent	0.878	0.019	0.680
	Velocity	−0.238	0.424	
	Degree	−0.481	0.204	
	Betweenness	0.519	0.177	
M2	Velocity	0.234	0.534	0.137
	Degree	−0.356	0.508	
	Betweenness	0.471	0.392	
M3	Agent	0.844	0.017	0.544
	Velocity	−0.266	0.373	
M4	Agent	0.706	0.023	0.499
	Degree	−0.072	0.780	
M5	Agent	0.713	0.018	0.536
	Betweenness	0.206	0.418	
M6	Agent	0.703	0.016	0.494
M7	Velocity	0.184	0.589	0.034
M8	Degree	−0.038	0.911	0.001
M9	Betweenness	0.170	0.616	0.029

Agent the number of passing-by agents, *Velocity* average of the traffic velocity, *Degree* degree centrality, and *Betweenness* betweenness centrality

centrality, and betweenness centrality, are not significant variables in all the regression models.

The above correlation analyses and linear regression models show that the number of passing-by agents from the agent-based model has strong relationship with the rent rates from the real world. These results indicate that the agent-based model is well validated and expected to generate reliable predictions in the assumption, yet traffic velocity and network centralities, which are directly related to the assumption, did not provide significant predictions. It is because both traffic velocity and network centralities are certainly related to the daily movements of the citizens, but each is not sufficient to express them. On the other hand, the agent-based model is capable of representing such dynamic behaviors, that's why the number of passing-by agents from the agent-based model shows the highest correlation.

11.5 Conclusion

Agent-based model has been applied to various problems, yet the method has been questioned for its prediction accuracy. It is because that contrary to statistical models, the prediction of the agent-based model is difficult to be evaluated, even if it is validated. However, as a generative model, agent-based model provides invaluable results which cannot be generated by statistical models. In certain cases, these invaluable results provide more accurate predictions than statistical models.

This chapter provides one example of the certain cases. To investigate relationships between the simulated and the real world data from the city commerce, we calculated the correlation between the two datasets, and the correlation was high. Although this example does not provide an answer for doubts about the accuracy of the agent-based model, it is sufficient to consider a trade-off for selecting a modeling method for different cases.

References

Barlas Y (1996) Formal aspects of model validity and validation in system dynamics. Syst Dyn Rev. http://www.ie.boun.edu.tr/labs/sesdyn/publications/articles/Barlas_1996.pdf

Bonabeau E (2002) Agent-based modeling: methods and techniques for simulating human systems. Proc Natl Acad Sci U S A 99(3):7280–7287

Brown DG, Page S, Riolo R (2005) Path dependence and the validation of agent-based spatial models of land use. Int J Geogr Inf Sci. http://www.tandfonline.com/doi/abs/10.1080/13658810410001713399

Chrissis MB, Konrad M, Shrum S (2003) CMMI guidlines for process integration and product improvement. http://dl.acm.org/citation.cfm?id=773274

Davis JP, Eisenhardt KM, Bingham CB (2007) Developing theory through simulation methods. Acad Manage Rev. http://amr.aom.org/content/32/2/480.short

Freeman LC (1979) Centrality in social networks conceptual clarification. Soc Netw 1(3):215–239

Grimm V, Revilla E, Berger U, Jeltsch F, Mooij WM, Railsback SF, Thulke H-H, Weiner J, Wiegand T, DeAngelis DL (2005) Pattern-oriented modeling of agent-based complex systems: lessons from ecology. Science (New York, NY) 310(5750):987–991. doi:10.1126/science.1116681. http://www.sciencemag.org/content/310/5750/987.short

Inman RL, Helson JC, Campbell JE (1981) An approach to sensitivity analysis of computer models: part II-ranking of input variables, response surface validation, distribution effect and technique synopsis. J Qual Technol. https://secure.asq.org/perl/msg.pl?prvurl=/data/subscriptions/jqt_open/1981/oct/jqtv13i4iman.pdf

Kleijnen JPC (1995) Verification and validation of simulation models. Eur J Oper Res. http://www.sciencedirect.com/science/article/pii/0377221794000166

Lee SH, Shin JS, Lee GH, Moon I-C (2013) Impact of population relocation to city commerce: micro-level estimation with agent-based model. In: Proceedings of the agent-directed simulation symposium, vol 11. Society for Computer Simulation International, Orlando

Lee G, Bae JW, Oh N, Hong JH, Moon I-C (2014) Simulation experiment of disaster response organizational structures with alternative optimization techniques. Soc Sci Comput Rev. doi:10.1177/0894439314544628

Marshall J (2007) Public sector relocation policies in the UK and Ireland. Eur Plan Stud 15(5):645–666

Moss S (2008) Alternative approaches to the empirical validation of agent-based models. J Artif Soc Soc Simul. http://jasss.soc.surrey.ac.uk/11/1/5.html

Sargent RG (2005) Verification and validation of simulation models. In: Proceedings of the 37th conference on Winter simulation. http://dl.acm.org/citation.cfm?id=1162736

Schelling TC (1971) Dynamic models of segregation. J Math Sociol 1(2):143–186. http://www.tandfonline.com/doi/abs/10.1080/0022250X.1971.9989794

Sheng G, Elzas MS, Ören TI Cronhjort BT (1993) Model validation: a systemic and systematic approach. Reliab Eng Syst Saf 42(2–3):247–259. doi:10.1016/0951-8320(93)90092-D. http://www.sciencedirect.com/science/article/pii/095183209390092D

Tesfatsion L (2003) Agent-based computational economics: modeling economies as complex adaptive systems. Inf Sci. http://www.sciencedirect.com/science/article/pii/S0020025502002803

Windrum P (2007) Empirical validation of agent-based models: alternatives and prospects. J Artif Soc Soc Simul. http://jasss.soc.surrey.ac.uk/10/2/8.html

Yilmaz L, Ören T (2009) Agent-directed simulation and systems engineering. Wiley, New York