International Environmental Modelling and Software Society (iEMSs)
2012 International Congress on Environmental Modelling and Software
Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig, Germany
R. Seppelt, A.A. Voinov, S. Lange, D. Bankamp (Eds.)
http://www.iemss.org/society/index.php/iemss-2012-proceedings

Human decision making for empirical agent-based models: construction and validation

Grace B. Villamor[ab], **Meine van Noordwijk**[b], **Klaus G. Troitzsch**[c], **Paul L.G. Vlek**[a],
[a]Center for Development Research (ZEF), University of Bonn, Germany
[b]World Agroforestry Centre (ICRAF), Southeast Asia Regional Office, Indonesia
[c]Institute for Economic and Computer Science,
University of Koblenz-Landau, Germany
(*gracev@uni-bonn.de)*

Abstract: Results from agent-based or multi-agent simulation (AB/MAS) modelling can provide relevant information for policy makers, scientists and stakeholders about the boundary conditions of rural development and the uncertainties involved in land-use/cover change (LUCC). However, the process of model validation that can build trust in the outcomes for new parameter conditions and in future scenarios is not a trivial problem. Apparently, no common measure of the degree of confounding between parameterization and validation data sets exists. The current lack of success and the effort necessary for validating the models can be traced to the weak theoretical representation of human decision making in current models. Thus, this paper reviews various ways to represent land-use decision making using AB/MAS models. It briefly describes process-based decision making as an alternative approach to address the problem of weak theoretical representation of human decision making, and presents a case study of an agent decision-making model applying an empirical validation technique.

***Keywords**:* land-use decision making, agent-based/multi-agent simulation models, empirical validation

1 INTRODUCTION

In the field of land-use science, the integration of human and environment dynamics is possible through the use of agent-based or multi-agent simulation (AB/MAS) models. One of the main strengths of these models is their ability to simulate the implications of human decision-making processes explicitly (Parker et al. 2003; Matthews et al. 2007; Smajgl et al. 2011; Villamor et al. 2012). These models could thus provide valuable prognoses about future land uses and also about the likely consequences and tradeoffs of land-use change and conservation policies. In spite of this advancement of AB/MAS models, various studies identify model weaknesses and limitations. Among these are: (i) difficulty in validation and verification, (ii) shortage of effective architectures and protocols to represent agents and their interactions (i.e. stylized ways), and (iii) poor representation of learning processes of real-world decision making. These limitations boil down to the weak representation of real-world human decision making, which is due to the difficulty in collecting empirical data on a system level (i.e. parameterization) and identifying its underlying causes (Heckbert et al. 2010). In this paper, we will try to briefly describe the common and currently approaches for modelling decision-making and then, concisely assess the way the validation is conducted. In the last part of this paper, a specific decision-making model construction is presented with an empirical validation technique application using the indirect calibration method (Windrum et al. 2007).

2 METHODOLOGY

Data collection: A literature review was conducted using the Web of Science to search the target topics. A combination of the following keywords was searched: "agent-based model" or "agent based simulation", "multi-agent model" or "multi agent simulation", "land use change" or "land cover change", and "human decision making". Only models that explicitly describe the decision making for empirical investigation of land-use/cover change were selected. In each selected model, the objective of the decision-making model/architecture applied and the validation techniques were the main aspects used to assess the model.

For the case study, a household survey was conducted with 95 households (out of 551 households) to elicit the agents' characteristics and behavioral responses. The survey was conducted between February and March 2010 in three villages of Bungo District, Jambi Province, Indonesia. In the survey questionnaires, two main conditions are explored, namely 1) the current condition of the agent, the household profile, and the farm-holding characteristics from which the current land-use choice was generated, and 2) under certain conditions or situations in which the agent will likely perceive and behave as if the condition existed (i.e. if supported by financial investment in the next 5 to 10 years, under payments for ecosystem services or PES through conservation agreement scheme). We also asked the reasons for choosing the land use in order to understand the actual motivations and preferences behind the decision.

Study area: The case study was conducted in the villages of Lubuk Beringin, Laman Panjang, and Desa Buat, which cover a total area of 157 km^2. The area is dominated by rubber agroforests that support both high lowland biodiversity and the livelihoods of the people there. Except for Desa Buat, these villages are considered poor and have poor access to market roads and electricity infrastructure due to their distance from the district center.

Data analysis: Statistical analysis was performed using SPSS and Stata version 12 to obtain the decision rules and stylized facts. First, the household categorization was done using principal component analysis (PCA) and cluster analysis (*K*-means cluster or KCA). Then, the behavior of household types regarding land-use choice and preferred land use under certain conditions was estimated using multinomial and binary logistic regression. Results were evaluated and validated using role-playing game (Villamor and van Noordwijk 2011), historical land-use change assessment (Villamor 2012), literature review, participants observation, and expert knowledge. The detailed results are not presented in this paper.

3 RESULTS

3.1 Human decision making

Out of 160 articles identified in the Web of Science, only 8 articles specifically targeted the modelling of decision making and its impact on land use/cover (Table 1). Agent decision making is represented in many ways. However, the architectures can be generalized in two broad approaches – heuristic and optimization. The reason for choosing an approach is not clear. Nonetheless, many used heuristics or combination of optimization, probably due to the bounded view of rationality that better describes the way human process information to output decisions (Kahneman et al. 1982; Chion et al. 2011). Nevertheless, each approach has its own strengths and weaknesses (Table 2). For example, in the MARIA model (Multi-agent reasoning in Amazonia) Cabrera et al. (2010) explicitly contrasted linear programming against decision tree using the same objectives, household demographics, wealth and information. They concluded that decision-making methods significantly affect the land-use trajectories of household agents. Accordingly, optimizing agent actions are far too extreme, and readily respond to changing factors (e.g. price changes), while heuristic agents show a realistic response to small changes.

Table 1: Empirical AB/MAS models for land-use science

Decision-making model approach	ABM/MAS model example
a) Heuristic behavior (e.g. decision tree, satisficing)	SMASH – Spatialized multi-agent system of landscape colonization by ash (Gibon et al. 2010); ABM/LUCC model (Valbuena et al. 2010); SOME – SLUCE's original model for exploration; DEED –Dynamic ecological exurban development (Brown et al. 2007); LUDAS – Land use DynAmic Simulator (Le et al. 2008); and SYPRIA – Southern Yucatan Peninsular Region Integrated Assessment (Manson 2005)
b) Optimizing behavior (e.g. genetic programming, mathematical programming, and neural networks)	LUCIM - Land use change in the Midwest (Hoffmann et al. 2003); MP-MAS (Schreinemachers and Berger 2011); and MARIA – Multi-agent reasoning in Amazonia (Cabrera et al. 2010)*

Note: * combination

Table 2: Strength and weaknesses of two decision-making model approaches (updated from Schreinemachers and Berger 2006)

Decision architecture	Strengths	Weaknesses
a) Heuristic	• Uses simple rules that guide human decision making • Recognizes the limited cognitive capabilities of humans in decision making (i.e. bounded rationality) • Allows decision processes • Allows participatory process to validate results (i.e. companion modeling) • Calibration is quick and easy	• Difficult to construct correct sequence of decisions (decision tree) • Difficult to identify appropriate conditions or saturation levels of the set variables or parameters • Prone to model artifacts (i.e. in using single set of heuristics) • Economic tradeoffs could not be captured (method relies on pre-determined conditions that are sequentially and independently evaluated)
b) Optimization	• Able to select the best or optimum decision from a range of feasible alternatives • Able to re-allocate resources to attain a higher level of goal satisfaction in which inefficiencies are eliminated • Able to incorporate risks and uncertainties • Can accommodate large number of conditions and actions of agents • Allows assessing the structural sources of inefficiencies	• Agents are always assumed rational with full access to information • Few real empirical applications • Calibration is time consuming • Recognized as a 'black box', since model outcomes are better investigated through sensitivity analysis • Inability to model cognitive processes

3.2 Validation

“How robust and reliable are the decision-making models?” and “How adequately do they represent the human system being modelled?” are questions often difficult to answer. Validation is a critical issue especially when using AB/MAS models to model land-use decisions. From the review (Table 1), the most common approach is through expert opinions (e.g. Valbuena et al. 2010; Brown et al. 2007). However, Heckbert et al. (2010) criticized this approach, since different stakeholders have different subjective understandings of the systems; the model might be an accurate representation of some views but inaccurate representation of others (Moss 2008). In other studies, model outcomes are validated by comparing simulated results to survey data or to literature (e.g. Le et al. 2008; Schreinemachers and Berger 2011) or by statistical validation (i.e. goodness of fit). Matching a model’s component structures and processes to structures and processes in the system being modelled is performed more in a conceptual way than by testing a one-to-one accuracy. In empirical AB/MAS LUCC models, agents are real human decision makers and in modelling them, we can’t ignore the domain of social sciences, where internal (i.e. causality), construct (i.e. confounded) and structural validity (i.e. reflecting the theory behind the simulation model; Troitzsch 2004) are fundamental part of the scientific discipline. However, these essential aspects are not included in the descriptions of current AB/MAS models.

3.3 Case study: application of LUDAS framework and indirect calibration

Various sources in literature suggest the use of process-based decision making and extending process-based modelling to the socio-economic components (Pahl-Wostl 2002, 2007; Barthel et al. 2008; An 2011). This is because humans make decisions in response to changing natural environments that will in turn change the context for future decisions. Accordingly, process-based decision models are those capturing the triggers, options, and temporal and spatial aspects of an actor’s reaction in a relatively direct, transparent and realistic way. For example, in dealing with the uncertainty of assumptions in models and data, an accepted way of reducing uncertainty or showing the influence of uncertainty processes on model results is by modelling the actual processes (Barthel et al. 2008). Thus, substantial efforts should be invested in process-based decision-making mechanisms or models to better understand the socio-ecological systems (An 2011). In the case study presented below, a process-based decision making model is constructed based on the possible behaviour of the household agents in the study area under certain conditions or situations as part of the decision process. The decision process includes a time element that is pertinent for establishing causal relationships (van Belle 2008) derived from a cross-sectional survey. First, we will briefly describe the agent decision-making model of the LUDAS framework as an empirically based decision model. Then, we will embed causal structures with the intention of applying the indirect calibration method (Windrum et al. 2007) for internal and construct validity.

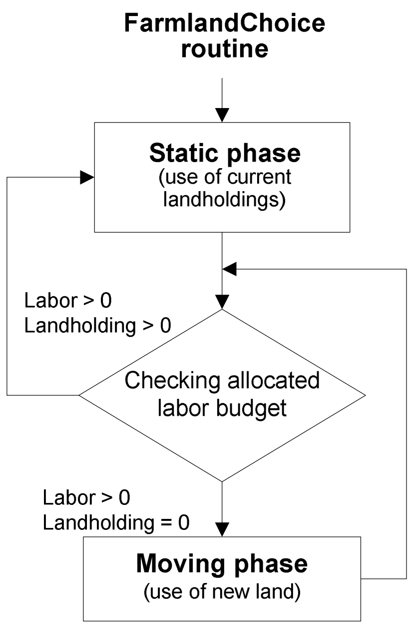

Figure 1. Decision algorithm of LUDAS model (Le 2005)

In the LUDAS model (Le et al. 2008), the household agent’s decision making and actions with respect to land use is defined by the *FarmlandChoice* procedure. This procedure consists of two separate phases: 1) static phase (use of old landholding), and 2) moving phase (use of new land) (Figure 1). A utility function of land-use options is applied in both phases using a standard regression-based approach.

The indirect calibration (IC) approach as one of the empirical validation approaches described by Windrum et al. (2007) was applied, which focuses on the parameters (drawing from stylized facts and empirical datasets) that are consistent with the output validation. It consists of a four-step approach to empirical validation:

- **Step 1**: Identify a set of stylized facts that the modeller is interested in reproducing;
- **Step 2**: Develop a model based on the empirical evidence regarding the agents and rules;
- **Step 3**: Use the empirical evidence regarding stylized facts to restrict the space of parameters and test the statistical regularities; and
- **Step 4**: Identify the causal mechanisms that underlie the stylized facts.

Following the above IC steps in the construction of the decision-making procedure of the LUDAS framework, the stylized facts (i.e. current land-use choice probabilities) derived from the empirical data (**step 1**; Table 3) will constitute the land-use choice model. This (current) land-use choice is mathematically stated in a multi-nominal logistic form integrated in both the static and moving phases of the *FarmlandChoice* routine (**step 2**; Figure 1). With this decision-making model, although empirically based, predicting the possible response of the agents to changing natural and political environments is limited. The model is thus only appropriate for describing the baseline scenario. This is because when the agent is ready to open new land due to an increase in economic and/or demographic factors while the environment is dynamically changing (moving phase) the preference coefficients used in calculating land-use probabilities are still fixed or constant (Le 2008).

Table 3: Summary of land-use choice probabilities and willingness to adopt PES (2010)

Land-use type	Probability (%)			
	Current/ baseline	Future*	Adopt PES	
			Yes	No
Household agent type 1			81	19
Rubber agroforest	33	87		
Monoculture (rubber or oil palm)	1	13		
Rice paddy	66	0		
Household agent type 2			92	8
Rubber agroforest	99	47		
Monoculture (rubber or oil palm)	1	19		
Rice paddy	0	0		

Note: * Scenario: If supported by financial investments in the next 5 to 10 years.
For detailed logistic regression results see Villamor 2012.

To realistically simulate future and other scenarios (i.e. under the PES scheme), we need to make the socio-economic component of the agents process based, (**step 3**) and at the same time embed the causal structure to achieve **step 4** of the IC approach. Thus, we need to include the time element (i.e. if supported by financial investments in the next 5 to 10 years) and decision process (i.e. willingness to adopt PES) to integrate the causal mechanism explicitly. We therefore reconstruct new decision algorithms that incorporate new the stylized facts derived from Table 3 (i.e. future land-use choice and willingness to adopt PES probabilities).

For instance, the difference between Figure 2a and Figure 1 is that the stylized facts derived under future conditions (Table 3) were used to construct the new land-use choice model and replaced the current land-use choice model in the moving phase (*highlighted in green)*. In this way, the household agents have new preference coefficients according to the new conditions of the scenario being explored while incorporating the temporal aspect (i.e. the next 5 to 10 years) and proposed option (i.e. with financial investments). On the other hand, stylized facts derived from the decision process of adopting PES (Table 3) were used to construct the decision-making model for the PES scenario (Figure 2b). The PES-adoption sub-model (*highlighted in yellow*) is nested within the current land-use choice model in the static phase. If based on the preference coefficients of PES-adoption variables (Villamor et al. 2011), the household agent's probability is 1, then do rubber agroforest, if otherwise look for other land-use types for current landholdings. For the preliminary results of these new decision-making sub-models and **steps 3** and **4**, see Villamor (2012) and a companion paper for this conference [Villamor et al. 2012].

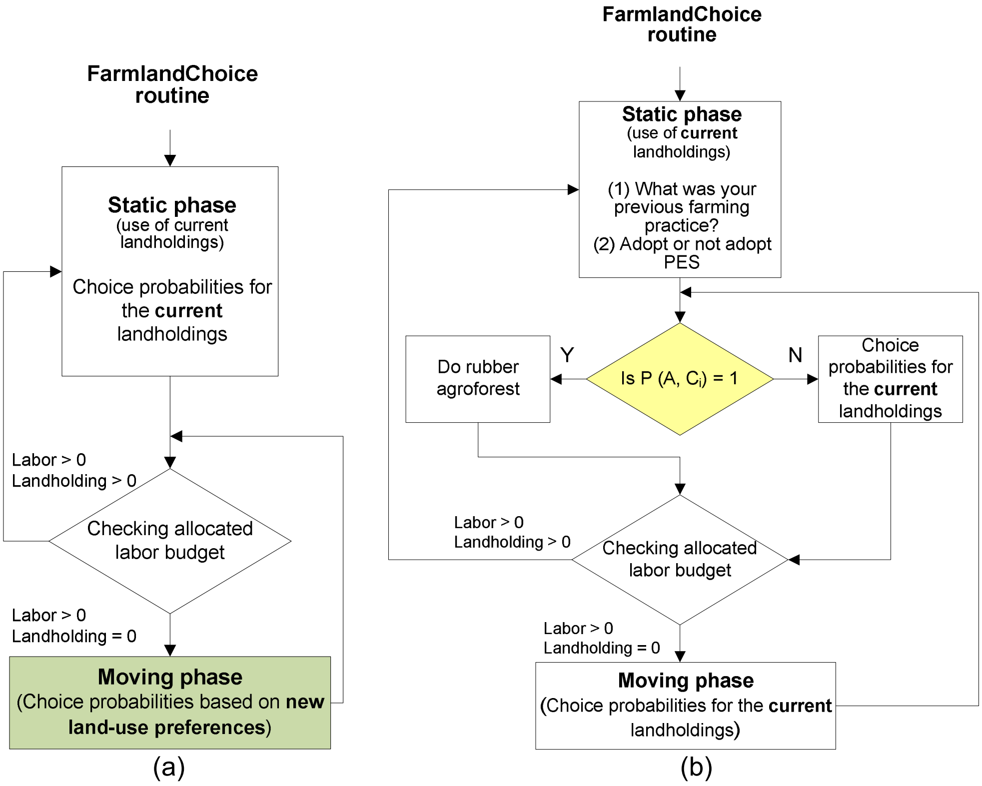

Figure 2: Flow chart of process-based decision making for condition with new preferences (a) and PES willingness (b). *P (A, C_i)* refers to probabilities (*P*) for agent (*A*) to choose the choices (C_i).

4 DISCUSSION

With the process-based decision making integrated in the AB/MAS model framework, we could address the following fundamental questions pertinent to empirical validation (i.e. **step 4** of IC approach):

Internal validity (i.e. what causes what?): The temporal element, in this case *'if supported by financial investment or subsidies for the next 5 to 10 years*', is one step to embed causality (van Belle 2008). To show this, for example, the scenario of the current land-use choices suggests that 99% of the type 2 (relatively poor) households will choose rubber agroforests while 1% are in monoculture rubber (Table 3). However, if offered financial support through investments in the future, the type 2 agents will behave differently. The probability of choosing rubber agroforest decreases by 50%, while for monoculture rubber it increases by almost 50%. This suggests a risk-taking behavior of the type 2 households due to the offered financial support; at the same time we could determine the significant factors affecting this behavior (Villamor 2012). Thus, we can better explain the possible changes (i.e. micro-structure) through a given parameter combination and initial condition, which is in accordance with step 4 of the IC approach.

Construct validity (i.e. am I really measuring the construct that I want to study? How adequately do I represent the human system being modeled?): In order to estimate the central decision process more prospectively, one could estimate some parameters of the decision process directly. Here, the decision process of 'adopting PES' and 'if supported by financial investments' are modeled directly in the decision making of human agents. However, the challenge lies in collecting and identifying different causal factors, which is important in explaining the processes and outcomes observed. Thus, expertise on subject matter is very useful to rule out competing explanations (Freedman 2010) while a variety of parameterization techniques are needed (Schreinemachers and Berger 2011).

Structural validity (i.e. given parameter combinations and initial condition, do the emerging macrostructures sufficiently resemble observable macrostructures?): The agent decision is validated *ex post*. In this type of validity, the validation of models can be interpreted as validation of theories, i.e. finding out whether the intended application of a theory or observations to which the theory refers exists (Troitzsch 2004). A companion paper for this conference (Villamor et al. 2012) describes some of the results that shed light on the concept of PES and its conditionality.

5 CONCLUSIONS

Incorporating the decision-making processes intends to better represent the preferences and perceptions of household agents in order to clarify the scenarios built to explore the possibilities, i.e. the opportunities and dangers of an uncertain future. With these new decision-making models, time-related questions (i.e. in the next 5 to 10 years) and possible behavior of the agents (i.e. if PES scheme will be adopted based on the real pilot PES projects in the study area) form a new basis of more adequate decisions of the household agents. Also, the use of process-based decision making together with the IC approach in empirical validation of AB/MAS models strengthens the causal mechanism of the model. This could be possibly done both in heuristic and optimatization decision-making routines/approaches. However, Windrum et al. (2007) stated a number of unresolved issues for this technique (e.g. alternative strategies for constructing models, over-parameterization, etc.). Thus, the recommendation is to keep the model as simple and descriptive as possible.

ACKNOWLEDGEMENT

We thank Guido Lüchters and anonymous reviewers for the comments and suggestions on the earlier draft.

REFERENCES

An L., Modeling human decisions in coupled human and natural systems: review of agent-based models, *Ecological Modeling*, doi:101016/j.ecolmodel.2011.07.010, 2011.

Barthel R., S. Janisch, N. Schwarz, A. Trifkovic, D. Nickel, C. Schulz and W. Mauser, An integrated modelling framework for simulating regional-scale actor responses to global change in the water domain, *Environmental Modelling & Software,* 23, 1095-1121, 2008.

Brown, D.G., D. Robinson, L. An, J.I. Nassauer, M. Zellner, W. Rand, R. Riolo, S.E. Page, B. Low and Z. Wang, Exurbia from the bottom-up: confronting empirical challenges to characterizing a complex system, *Geoforum*, doi:10.1016/j.geoforum.2007.02.010, 2007.

Cabrera, A.R., Deadman, P., Brondizio, E. and Pinedo-Vasquez, *Exploring the choice of decision making method in an agent-based model of land use change*, In: D. Swayne, W. Yang, A. Voinov, A. Rizzoli and T. Filatova (Editor), 2010 International Environmental Modelling and Software Society Ottawa, Canada, 2010.

Chion, C., P. Lamontagne, S.Turgeon, L. Parrot, J.-A. Landry, D.J. Marceau, C.C.A. Martins, R. Michaud, N. Menard, G. Cantin and S. Dionne, Eliciting cognitive

processes underlying patterns of human-wildlife interactions for agent-based modelling, *Ecological Modelling*, doi:10.1016/j.ecolmodel.2011.02.014, 2011.
Freedman D.A., *Statistical models and causal inference*, Cambridge University Press, Cambridge, 2010.
Gibon, A., D. Sheeren, C. Monteil, S. Ladet and G. Balent, Modelling and simulating change in reforesting mountain landscape using a social-ecological framework, *Landscape Ecology*, 25:267-285, 2010.
Grimm, V., E. Revilla, U. Berger, F. Jeltsch, W.M. Mooij, S.F. Railsback, H.H. Thulke, J. Weiner, T. Wiegand and D.L. DeAngelis, Pattern-oriented modeling of agent-based complex systems: lessons from ecology, *Science*, 310:987-991, 2005.
Heckbert, S., T. Baynes and A. Reeson, Agent-based modeling in ecological economics, *Annals of the New York Academy of Sciences*, 1185:39-53, 2010.
Hoffman, M., H. Kelley, and T. Evans, *Simulating land-cover change in South-Central Indiana: an agent-based model of deforestation and afforestation,* In Complexity and Ecosystem Management: The Theroy and Practice of Multi-Agent Approaches.M. Janssen (ed), Northhampton, Massachusetts: Edward Elgar Publishers, pp. 218-247, 2003.
Kahneman, D., P. Slovic and A. Tversky, *Judgement under uncertainty: heuristics and biases,* Cambridge University Press, Cambridge, MA, 1982.
Le, Q.B., S.J. Park, P.L.G. Vlek and A.B. Cremers, Land-use dynamic simulator (LUDAS): a multi-agent system model for simulating spatio-temporal dynamics of coupled human-landscape system. I. Structure and theoretical specification, *Ecological Informatics*, 2:135-153, 2008.
Manson, S.M., Agent-based modeling and genetic programming for modeling land change in the Southern Yucatan Peninsular Region of Mexico, *Agricultural Ecosystem Environment*, 111:47-62, 2005.
Matthews, R., N. Gilbert, A. Roach, J. Polhill, and N. Gotts, Agent-based land-use models: a review of applications, *Landscape Ecology,* 22:1447-1459, 2007.
Moss, S., Alternative approaches to empirical validation of agent-based models, *Journal of Artificial Societies and Social Simulations*, 11, http://jasss.soc.surrey.ac.uk/11/11/15.html. 2008.
Pahl-Wostl, C., Transition towards adaptive management of water facing climate and global change, *Water Resources Management* 21(1):49-62, 2007
Pahl-Wostl, C., Towards sustainability in the water sector- the importance of human actors and processes of social learning, *Aquatic Sciences* 64:394-411, 2002.
Parker, D.C., S.M. Manson, M.A. Janssen, M.J. Hoffmann, and P.J. Deadman, Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review, *Annals of the Association of American Geographers*, 93:314-337, 2003
Schreinemachers, P. and T. Berger, An agent-based simulation model of human-environment interactions in agricultural systems, *Environmental Modelling & Software*, 26:845-859, 2011.
Schreinemachers, P. and T. Berger, Land-use decisions in developing countries and their representation in multi-agent systems, Journal of Land Use Science, 1:29-44, 2006.
Smajgl, A., D.G. Brown, D. Valbuena, and M. Huigen, Empirical characterisation of agent behaviours in socio-ecological systems, *Environmental Modelling & Software*, 26:837-844, 2011.
Valbuena, D., P.H. Verburg, A. Veldkamp, A.K. Bregt, and A. Ligtenberg, Effects of farmers' decision on the landscape structure of a Dutch rural region: an agent-based approach, *Landscape and Urban Planning*, 97:98-110, 2010.
van Belle G, *Statistical rules of thumb*, Wiley, New York, 2008.
Villamor, G., M. van Noordwijk, K.G. Troitzsch and P.L.G. Vlek, Human decision making in empriical agent-based models: pitfalls and caveats for land-use change policies, 26th European Conference on Modelling and Simulation, Koblenz, Germany, submitted.
Villamor, G., *Flexibility of multi-agent system models for rubber agroforest landscapes and social response to emerging reward mechanisms for ecosystem services in Sumatra, Indonesia,* Unpublished manuscript, University of Bonn, Bonn, 2012.
Villamor, G. and M. van Noordwijk, Social role play games versus individual perceptions of conservation and PES agreements for maintaining rubber agroforests in Jambi (Sumatra), Indonesia, *Ecology and Society*, 6(3):27, 2011.
Villamor, G.B., M. van Noordwijk, Q.B. Le, B. Lusiana, R. Matthews, and P.L.G. Vlek, Diversity deficits in modelled landscape mosaics, *Ecological Informatics*, 6:73-82, 2011.
Troitzsch, K.G, *Validating simulation models*, Proceedings of the 18th European Simualtion Multi-Conforence, pp.98-106, 2004.
Windrum, P., G. Fagiolo and A. Moneta, Empirical validation of agent-based models: alternatives and prospects, *Journal of Articifial Societies and Social Simulations*, 10 (2):8, 2007.