Validation of agent-based land use model by Markovian model

Application to forest-agriculture transitions in Madagascar

D. Hervé[*] , J.-P. Müller[**] , V. Ratiarson[***] , B. Ramamonjisoa[***]

[*] *IRD, UMR 220*
Montpellier, France
dominique.herve@ird.fr

[**] *CIRAD, UPR GREEN*
Montpellier, France
jean-pierre.muller@cirad.fr

[***] *Université de Fianarantsoa, ENI*
Fianarantsoa, Madagascar
r_venot@yahoo.fr,bertinram@yahoo.fr

Abstract. *The determination of transition rules that farmers adopt to manage crop-fallow after forest clearing, is essential for deciding a sustainable strategy for forest conservation. The effect of the type of farms with respect to these transition rules in forest border may mitigate incentive measures planned by forest conservation policy. Agent-base modeling (ABM) of land use is a relevant approach to manage the dynamics of heterogeneous mosaic landscapes such as the border of the Malagasy Eastern rainforest. Transition rules between six land uses (forest, fallow, crop, grass, plantation and paddy field) are formalized at a plot level. A historical database containing transitions between the first four land use states was used to calibrate transition models for the ecological and farmer land use dynamics. Three land-use models have been built: (1) a Markov chain (stochastic), (2) a timed automaton (deterministic), (3) and an agent-based model, which introduces the farmers. The land use ABM allows to test scenarios of deforestation with both varying initial population and farm spatial organization, size or strategy. The land use ABM is first calibrated via a timed automaton, fitting time delay parameters, the duration of each land use state (fallow, crop, grass), and the number of*

V2CS special issue of Studia Informatica Universalis, pages 33 à 61

cropping cycles since the first forest clearing. It is then validated with the help of a Markovian model, comparing two transition matrices with χ^2 metrics. The two transition matrices were respectively created with historical data of plot land use, and with simulated data produced by the land use ABM. We finish with a general discussion on the validation of such a complex system with a simple mathematical model.

Keywords: *rainforest, transition, farmer, land use dynamics, agent-based model, timed automaton.*

---

1. Introduction

Land use dynamics have been studied with different kinds of tools:

– differential equations [DBB97],

– stochastic models: Markov chains to model landscape spatial dynamics [Tur87, LDL85, MM94, Lui02] or vegetation dynamics [Bal00, Del94], stochastic cellular automata identifying transition rules for spatial cells [LF02, LDM$^+$05],

– and rule-based models as timed automaton [LC00, CL03], transition rules model [CRP01] or agent-based model [CBND02].

This range of models, from the more aggregated to the more explicit one, reveals that they no treat the same question, and that hypotheses, objectives and scales varied. To this variety of models corresponds a variety of validation methods. In a way, the ABM is more descriptive and explicative, but the way it is validated need to be formalized.

Deforestation of the rainforest has been studied in Eastern Madagascar mainly by spatial analysis of past aerial and satellite images [GS90]. In the case of these rainforests, it has been demonstrated a return to a forest state if no crop, no fire, and no pasture are applied to a plot after crop abandonment, during a long time (30 to 50 years). But few researches were dedicated to post-forest land use dynamics and forest regeneration [RSC07]. Prospective models are lacking to predict the impacts of incentive conservation measures.

We are tackling the following questions: what is the applied problem, the global context of malagasy rainforest? How to influence farmer practices to obtain a required landscape? How to involve farmers in forest conservation with clear incentives measures? The more general question is then to identify the effect of farm diversity on land use decision and forest conservation. The main objective of our modeling process is to link rainforest conservation objectives with farmer's strategies and practices to be able to reduce the deforestation trends via farm incentives. Deforestation must be studied into the forest and on the borders, since post forest land uses in slopes increase after paddy field saturation is achieved in marshlands [RRRH10]. However, the transfer of forest management from public Water and Forest administration to local communities, via local contracts (Gelose in 1996, effective since 2000), limited the access to forest resources. Since 2000, farmers have then adapted their production system to this reduced access to forest land [Toi09].

Farmers took two kinds of decision: to slash and burn a new plot into the forest, and to let a cultivated plot into fallow (plot abandonment), when the plot does not produce enough. Post-forest land use is described, at the plot level, both as crop-fallow successions after the first forest clearing and as vegetation successions after each plot abandonment. Studying land use dynamics at the plot level (1) enables to link the resulting landscape to farmer practices, (2) supposes plot distribution into farms and number and size of these plots (that means farm diversity) have influence in land use decision and forest conservation. In the case of forest conservation, it helps to design economic incentives devoted to the farmers to reduce the incidence or frequency of slash and burn practices.

An agent-based model (ABM) was designed to model deforestation in Madagascar rainforest and to test explicit scenarios pertinent to account for farmers' decision making. The use of ABM deals with the Malagasy deforestation problem via the post-forest transitions operated by farmers. However, the calibration and validation of such a model raise particular challenges. The large number of parameters in an ABM makes its calibration difficult [GM07]. In order to tackle this problem, we propose to calibrate the agent behaviors through an average behavioral model using a timed automaton model (TAM), which is easier to calibrate on a mean behavior.

The ABM validation relies on the adequate production of a global phenomenon from local interactions. In our case, the local interactions are land use transitions at a plot scale, based on historical records of crop-fallow successions after the first forest clearing. The global phenomenon to produce is the evolution of a landscape in terms of spatialized land use successions. From plots to landscape, as it is impossible to reproduce exactly what happened, we propose to statistically summarize the global phenomenon with a Markov chain model (MCM). This technique enables us to compare both observed field data and data issued from ABM simulations.

In the following sections, we first present a state of the art on model validation and more specifically on ABM validation. Then, we describe the adopted methodology of validation before showing and discussing our results. We finish with a general discussion on the validation of such a complex system with a simple mathematical model.

2. State of the arts: model validation

2.1. *Model validation*

[ARB06] defines "verification" as checking the conformity of an implemented software against specifications, "calibration" as identifying where, in the space of parameters, the model has the expected properties, and "validation" as meeting the user needs with the delivered software.

The verification and the validation of a model are relative to the model objective, either predictive, or for understanding how does the system operate. [Ryk96] suggests, before any kind of validation, to define not only the objective but also the criteria of acceptance and the context of operation of the model.

Many validation tests are conducted all over the modeling and simulation process, during the initial stages of data collection and analysis, when the conceptual model is built, as well as during the production of results by the operational model. [LF05] suggested that model validity must be assessed relatively to the objective at each step of the modeling and simulation process because each step brings its own source of errors. Many authors [CH97, ARB06] agree that validation is not universal, nor unique. Validation techniques depend on the model objective and its domain of application.

Other kinds of validation have been proposed, either model-centered or user-centered. Model-centered validation, called functional validation, corresponds to the design of an experimental plan to analyze the influence of several combined factors on a global result of the model. User-centered validation is defined by the problems that are sought to be solved by the use of the model. Coherency (the model is not internally contradictory) or veracity (the model does not contradict the reality), evaluated by experts of the same scientific domain, users or actors, increase trust in the results of the model [Bom97]

Role-playing games or visual simulations are considered as forms of validation [ARB06, Guy06]. Visual interactive simulation allows modifying the parameter values during simulation. With graphic animations, the user may gain understanding of the model behavior, gain implication if he changes the parameter values, and link these results to his knowledge of the real system.

Validation may be a condition of confidence that the model would be used in the future. [Sar84] defined the validation as a proof that the model has a reasonable margin of confidence in its domain of application. The model is considered valid when the simulated data are very similar to the observed data. We will keep this last definition in the following.

2.2. *ABM validation*

How much confidence can we give to an agent-based model? Many users prefer ABM to mathematical equations because those models are more intuitive. However [Def05] asks: "do they really increase our knowledge on the studied dynamics?" Due to their complexity and high number of parameters, ABM are generally more difficult to calibrate and validate than the other modeling approaches[BP07]. Several tests exist in the object-oriented approach when validation, as a measure of similarity with observed data, is not sufficient. [Axe97] claimed for such a model comparison along with sensitive analysis in the case of ABM applied to social sciences: "computational modeling would have never provided the clear sense of domain of validity that typically can be obtained for mathematical theories" [Axe97] .

[Bou95] proposes to validate ABM at a more global level of organization. The artificial universe of ABM has been built with forms, typologies, distributions, and qualitative models, from which a global behavior emerges. An ABM may be validated when compared with the formalization of this knowledge at a superior level of organization.

Moreover, to assure robust conclusions, many simulations have to give the same results. Temporal dynamics at the local scale allow simulations on hundreds of years, without possible data to confirm these dynamics. Spatial distribution at a global scale may be compared to actual maps and, as a consequence, may be more easily used in validation protocols. [CBND02] adopted such a spatial validation with an ABM at the village level. This ABM has been validated by simulation: (1) testing hypothesis on the relative impact of some factors (access, land property rules, agriculture-livestock interactions) on agrarian dynamics, and (2) evaluating the gap between simulated and actual maps of land use. Simulations are initialized by the map of the simulated village, by a geographic information system with information on land use in 1990, including access, land availability for cropping and socio-economic data on the villages. Different ABM validation processes have been proposed, but without total satisfaction. The consequences are three: ABM is difficult to compare to other models, ABM prospective use is criticized and more confidence is needed for ABM.

Lobry proposed to build the mathematical theory of ABM [TDM+07]. Checking the coherency of a process may be currently done by mathematics, but only if the process is homogenous and the number of agents is small [TDZ08]. In the case of a complex system unsolved with mathematical equations, simulation is necessary. But computer sciences may not solve every kind of problem. This explains why multi-agent systems are normally more criticized than others.

[Gin03] proposes a statistical approach of ABM, rethinking the building, understanding and use of complex systems in the basis of repeated simulations with their respective statistical treatment. Calculating the variance of the response of the model, parameter per parameter, on the total range of variation of each parameter, allows to test how much each parameter contributes to the response variability of the model. Sensitivity analysis can be used on continuous and discrete parameters, initial conditions, stochastic components, and numeric resolution. However, any systematic exploration would need heavy statistics tools because the number of variables is high [GM07]. The ABM produces too much information, and the problem is how to control this complexity. Two options are available:

1) To start with simple mechanisms and to progressively complexify them.

2) To differentiate part of the system, express all its complexity, and then add progressively other mechanisms.

We selected the first option: to accept to have a first look on a simplified image of the reality, namely the changes and sequences of four land use states using Markov chains. Our aim is to use this mathematical model to validate an ABM. We illustrate this option in the case of deforestation in Madagascar, with an ABM based on post forest dynamic of land use system. The objective, hypothesis, method and initial state of this model are defined in [RHRM11].

3. Materials and methods

3.1. *Available databases*

Both the historical land-use database and the farm typology database were elaborated on the western border of the forest corridor of Fianarantsoa, which links the national parks of Ranomafana and Andringitra. Generally speaking, ecological and socio-economic data are rarely collected together.

The land-use database results from direct observations of annual plot use between 2003 and 2006 and inquiries about plot use since first forest clearing (1973) with owners, neighbors and local guides. Since the data is historical from an initial forest cover to an actual agricultural use, it appears necessary to cross different information sources, especially direct observation with historical surveys.

The typology database is issued from a PhD thesis in agronomy and geography realized during the period 2004-2009 in the Western and Eastern sides of the forest corridor of Fianarantsoa [Toi09]. This farm typology was built with an emphasis on farm spatial organization and on farm adaptation to forest conservation. It resulted in 5 types of farms. To simplify the use of this data in an ABM, these 5 types of farms were brought together into 3 types of farms: savannah, mixed and forest types. In fact, the landscape bordering the Fianarantsoa corridor is structured in three bands, defined by their distance to the forest: savannah, savannah-forest, and forest.

3.2. *Land use transitions*

Mixed forest and farm landscape dynamic is summarized into six land use states: forest, crop, fallow, grass, paddy field and plantation, and the possible transitions between these six states (Figure 1). Paddy field and plantation are states of long duration, resulting from heavy transitions. They remain the same for either a long time (plantation) or even definitively (paddy fields). Therefore, the active transitions we may observe through historical data concern only four states: forest, crop, fallow and grass. We call "fallow" what occurs to the plot after harvest, with a shrub regeneration, and "grass" the herbaceous cover that appears only after a determined number of crop-fallow cycles. In order to analyze these data, we first consider stochastic chains of land use transitions.

In order to analyze the behavior of the individual farmers, they are considered to follow the same decision rules for the duration of each of the three temporary land uses. The number of crop-fallow cycles defines the cultivation intensity. This additional parameter is necessary because a high number of cycles is an indication of the occurrence of grass after crop.

In the following, each of the possible land use is coded as a letter as follows: forest$\rightarrow$ F, crop$\rightarrow$ C, fallow$\rightarrow$ J, grass$\rightarrow$ G, plantation$\rightarrow$

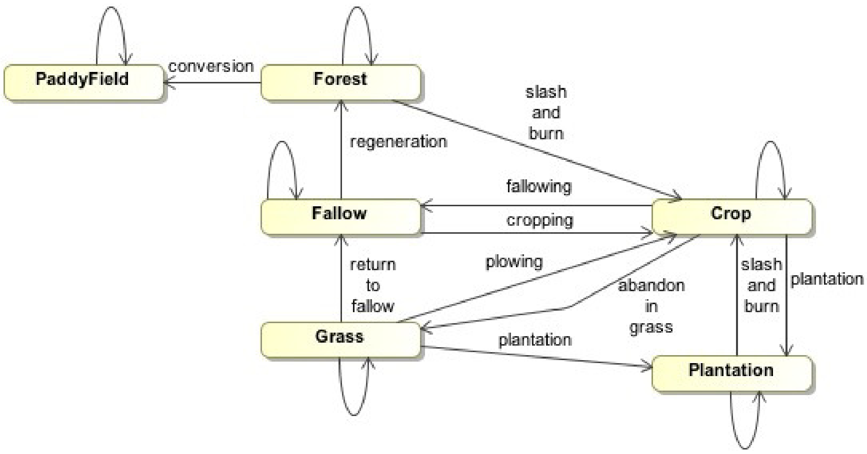

Figure 1: **The six land use states and their transitions**

P, paddy field$\rightarrow R$. Only the four first land use states are considered with their duration. As the collected data include annual plot use, the durations can be calculated for each land state.

To choose the plot to cultivate, the farmer has the choice among various initial land cover. The farmers ordinate the states previous to cropping in an order which respects two principles: to maintain soil fertility and to reduce hand work. The farmers invest their work in plots where they expect higher yields. Manually ploughing of grassland is considered less work than forest slash and burn, but more work than cultivating a previously harvested plot (continuous cropping). The ordering based on increasing invested work is: $C \rightarrow C, G \rightarrow C, F \rightarrow C$. Globally, fallow age is a good indicator of soil fertility. The ordering based on increasing soil fertility is then: $C \rightarrow C, G \rightarrow C, J \rightarrow C, F \rightarrow C$.

But this order of preference for crop precedence: Forest (F) > Fallow (J) > Plantation (P) > Grass (G) > Crop (C) cannot be respected in a context of forest conservation. Forest being protected, farmers may choose between two strategies when they adapt their current ordering:

1) $J \rightarrow C$, before $G \rightarrow C$, before $F \rightarrow C$, before $C \rightarrow C$, when farmers accept to preserve forest from slash and burn, but at last, they

will choose a new plot in forest before accepting to cultivate one year more a plot which does not produce enough. In this case, they exploit mature forest soil fertility to preserve the soil fertility into their own plots.

2) $J \rightarrow C$, before $G \rightarrow C$, before $C \rightarrow C$, before $F \rightarrow C$, when farmers accept to preserve forest from slash and burn, choosing other land covers before crop, including last crop. This ultimate strategy supposes that continuous cropping is possible, either using new cropping techniques available to restore soil fertility, or with a systematic weed control.

The TAM and the ABM have been designed with the order (1).

3.3. *The models*

3.3.1. *The MCM*

The temporal sequences of plot land use are complex data. They include first forest clearing date, crop successions, crop and fallow sequences, post-cropping vegetal successions. Many works have used Markov models to represent the dynamics of vegetation or land cover [LDL85, Tur87, Bal00, Lui02, RRRH10]. In our work, Markovian matrices are used to summarize the complex information on changes and sequences of land use with four states, on a stochastic basis. Formally, given a set of possible states for each plot of the landscape, in our case forest, crop, fallow and grass, a Markov state is a vector representing the probability distribution over this set of states. In our case, the probability distribution represents the probability for a plot to be in a given state in the landscape at a given time. A transition matrix encodes the evolution of this landscape at each step (here the step is one year). The multiplication of the Markov state vector by the transition matrix provides the new probability distribution at the next time step. Therefore, given sequences of landscape states, the calibration consists in building a Markov matrix. Given the initial distribution of land cover for the landscape, the MCM is able to reproduce the sequence of distribution of plots on the landscape.

3.3.2. *The TAM*

In the TAM, the dynamics is represented by a timed automaton. To make this process spatially explicit, this model has been implemented as a cellular automaton where each cell represents a plot driven by a timed automaton. A global (landscape level) algorithm selects a new plot to be cultivated after a plot has been abandoned, to maintain the same number of cultivated plots and to stabilize the food supply of the families. The priority to select a new plot is defined by the transition rules (3.2). The timed automaton has four parameters: the three time delays for crop, fallow and grass, and the number of crop-fallow cycles, since the first forest clearing. The crop time delay means the number of annual successive crops. The crop-fallow cycle is determined by the respective crop and fallow delay times; it begins with crop and ends with fallow, before another cropping. Time delays are calculated from the database for each of the land use state: crop, fallow and grass. These time delays can also be obtained by calculating statistics on the time series: average, standard deviation, and median.

However, using the timed automaton presents some limits. On one hand, population scenarios with timed automata are rough and indirect, for example, changing the one-to-one ratio to a one-to-two ratio. With ABM, these scenarios can be more detailed and explicit with, for example, changes in initial populations or demographic rates. On the other hand, farmer population is not homogeneous, therefore, an average behavior is not realistic. The richest farmers are those who have enough resources to contract labour to slash and burn forest and, therefore, to exploit forest soil fertility. The adaptation capacity to a limited access to forest plots depends on the spatial distribution of plots within the farm [TSHL11].

3.3.3. *The ABM*

These arguments made us choose an ABM in which:

1) The spatial constraints are supposed to influence farmers' adaptation to a limited access to forest; they are summarized by distance from forest: plots into the forest, near the forest, out of the forest. We divided the landscape into three bands parallel to the forest linear corridor, from the East to the West: forest, forest border, savannah.

2) The demographic process is a key factor in explaining land use transitions. The ABM allows to explicit demographic parameters as initial population and demographic growth rate.

3) Land use decision may vary around average decision rules, according to farm strategies. To maintain the food supply of the family is a survival strategy that can be qualified as "simple reproduction". To open more plots than actually needed by the family is a growing land access strategy to secure future access to land for the next generation; it can be qualified as "enlarged reproduction".

The resulting ABM is illustrated by an UML conceptual model (Figure 2). Following [TSHL11], we introduce a typology of only three types of farms (forest, mixed and savannah farms), depending on the distance to the forest. The village territory is composed of three zones, forest, forest border and savannah. The population of plots managed by a farm is part of the village territory. Land use states differ between rain fed slopes and plane marshlands. Finally the time delays and the number of cycles issued from the timed automaton along with the farm typology are used to generate the ABM.

4. The validation process

Figure 3 illustrates the whole calibration/validation process. First the observed data, i.e. sequences of land cover successions, are used to calibrate a Markov chain, producing a transition matrix (1). This transition matrix is considered as the reference summary of the land use dynamics. The timed automaton model, parametrized by the time delays for each state, and the number of crop-fallow cycles, then run to simulate the land use dynamics. The generated data is used to calibrate a new Markov chain, producing a new transition matrix (2). The comparison between the observed transition matrix (1) and the matrix issued from timed automaton simulation (2) is used to adjust the time delays and the number of cycles until both matrices are close enough. Notice that the calibration process produces an adjustment of the model parameters while the validation process produces a yes/no answer (is the model valid or not?).

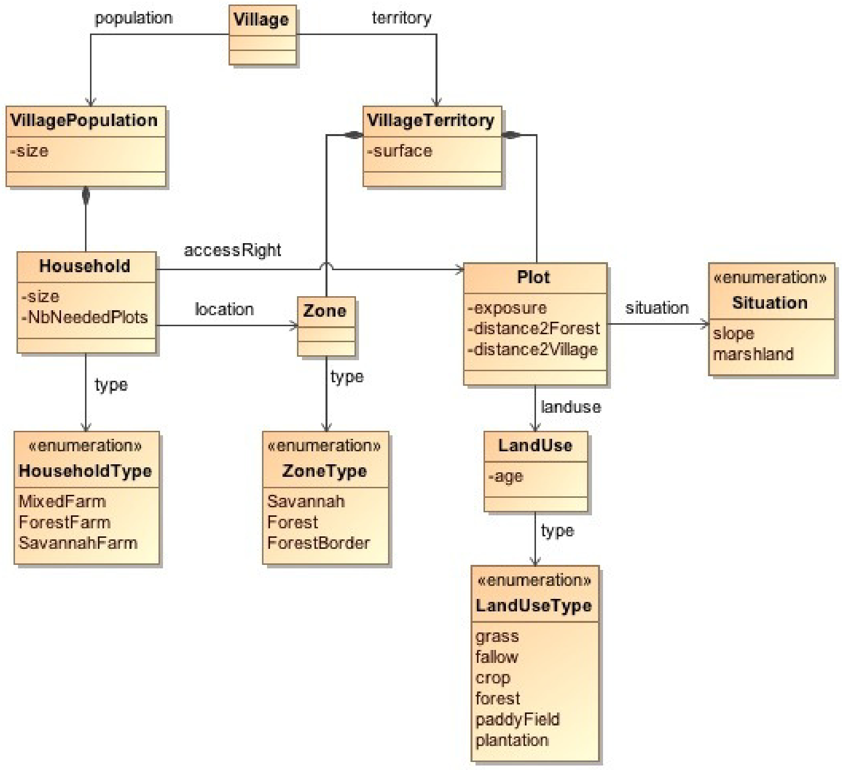

Figure 2: **The ABM conceptual model**

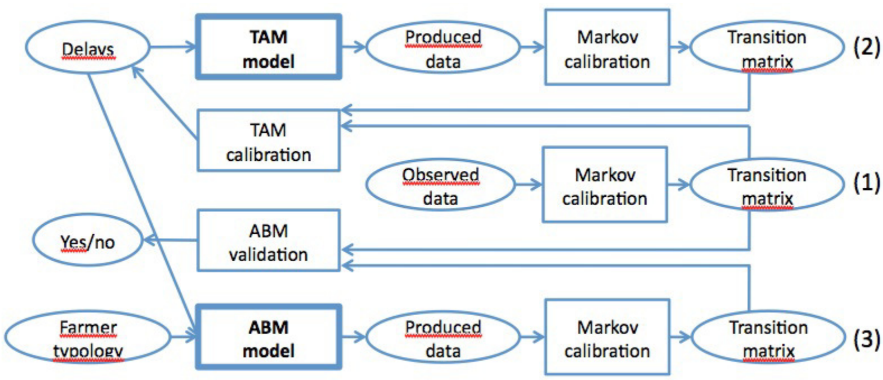

Figure 3: **The whole calibration/validation process**

In our case, we take care of validating the model with another set of data than the data used to calibrate, both on another period and on another site. The model is considered validated when the simulated data are very similar to the observed data. Two land-use data subset were then separated on a geographical basis: 104 plots dispersed in the forest border and 66 plots grouped in two small watersheds. The first dataset, covering a period of 34 years (1973-2006), was used to calibrate the timed automaton. The second dataset, which covers a period of 22 years (1985-2006), was used to validate the model.

The time delays and number of cycles, calibrated above, are used as parameters in the ABM. The ABM simulation produces new data on which a new Markov chain can be calibrated (3). The resulting matrix is used to compare the observed data and the ABM simulated data. The comparison between real data and simulated data is tested by a χ^2 comparison of two transition matrices. These two matrices are the summary of crop and fallow successions issued respectively from the simulation of the calibrated ABM (3) and the observed dataset (1).

If we make the hypothesis that both processes are the same, the simulated transition matrix should be identical to the observed transition matrix. For $i, j = 1 \ldots m$, statistic of test 1 is χ^2 with $m(m-1)$ degrees of freedom, where m denotes the number of possible states [AG57]. Probability p_{ij} is empirically estimated by the ratio $p_{ij} = n_{ij}/n_i$ where n_{ij} is the number of transitions from state i to state j and n_i is the number of landscape units in the i state [Ber88].

We use the same method of comparison for validation as for calibration: Markovian matrices with χ^2 test. The way we use calibration and validation for each kind of model is summarized in table 1.

5. Results

As described in the methodology, we have first calibrated and validated an MCM in order to represent in a synthetic way, the dynamics of the land use transitions. The obtained matrix will be used to calibrate and validate the other two models: the TAM and the ABM.

Table 1: Strategies of calibration and validation for the different model types used in this study.

Model type	Calibration	Validation
Markov model	Mathematical formalization easy to calibrate with a transition matrix if this matrix is time-homogenous.	Aggregated model easy to validate with a transition matrix calculated on another set of data than the set of data used for calibration.
Timed automaton	Calibration by simulation with the hypothesis that transition rules are unique and stable in time. Comparing real and simulated Markovian matrices allows to adjust the four parameters of Timed automaton.	Validation by simulation on another set of data with the parameters adjusted by calibration.
ABM	Global calibration difficult because parameters are too many. Partial calibration centered on time delays, using the adjusted parameters defined by timed automaton.	Validation by simulation on another set of data with the adjusted parameters defined by timed automaton.

5.1. *TAM calibration*

The timed automaton is calibrated and validated with the help of a Markov chain according to the χ^2 metric [AG57], from two samples of plots initially covered with forest and considered independent and not spatially localized. The timed automaton is calibrated by simulating during a 34 years with the first sample of historical transitions with 4 land use states and then validated during a 22 years with the second sample since forest clearing as to minimize the distance between two Markov transition matrices, which summarize the simulated and observed dynamics [RHRM11].

The calibration of the timed automaton is the adjustment of the values of the four parameters of the transition rules, the time delays for crop, fallow and grass states and the number of crop/fallow cycles. Reiterative simulation is managed for each combination of four parameters in their intervals of possible discrete values (time delays). The adjusted parameters are those that minimize the distance between the two Markovian transition matrices, the observed and the simulated. The acceptance of fitted parameters is based on the results of the χ^2 test and our knowledge about farmer practices. The calibration results are shown in the figure 2; the resulting fitted parameters are given in the table 3.

In the table 3, we also mentioned the statistics directly made on the observed time series. The histograms reveal that the probability distribution follows a decreasing exponential law of parameter λ. Therefore the mean $\overline{\sigma}$ is an estimation of λ ($\lambda = 1/\overline{\sigma}$). The results were obtained by removing the last homogeneous sequence from each plot because it very often corresponds to a definitive abandonment into fallow or grass. By just calculating the median, we find almost directly the fitted parameters. Further work is necessary to better understand the small differences. The advantage of our method is to be completely independent of the kind of model we are calibrating and validating. The corresponding model is really used as a black box.

Table 2: Comparison of Markovian matrices with observed and timed automaton simulated transitions for calibration

	Observed matrix				Automaton simulated matrix			
States	Forest	Fallow	Crop	Herb	Forest	Fallow	Crop	Herb
Forest (F)	0	0	1	0	0	0	1	0
Fallow (J)	0	0.59	0.41	0	0	0.63	0.37	0
Crop (C)	0	0.26	0.72	0.02	0	0.25	0.71	0.04
Herb (H)	0	0	0.19	0.81	0	0	0.18	0.82
χ^2 test, risk α=5%, freedom degree=12 Statistic test: 5.52 ; critical threshold: 31.03								

Table 3: Observed and calibrated values of times delays and number of crop-fallow cycles

	Parameters values			
Land use states	Crop (years)	Fallow (years)	Grass (years)	Nb crop-fallow cycles
Observed parameters values	1 to 9	1 to 12	2 to 7	2 to 4
Mean, standard deviation, median	2.59, 2.82, 2	2.89, 4.40, 2	4.28, 3.24, 5	2.88, 0.61, 3
Parameters with the minimum distance between both Markovian matrices	2	2	3	4

5.2. *ABM setting and validation*

Calibration based of farm typology and decision rules for each agent would introduce many parameters, making very difficult any kind of calibration. It is the reason why we used an existing farm typology and the parameters fitted for the TAM to set the initial state of the ABM.

Therefore, to set the ABM parameters, we used the time delays already fitted by timed automaton calibration, i.e. the time delays of three land use states, crop, fallow and grass and the number of cropping cycles. In the ABM, each agent is considered as a rule-based model, in conformity with its main objective and the respective times of crop-fallow successions. The ABM uses also i.e. the parameters introduced by the farm typology. In reference to farm typology, we need 12 other variables to describe farm relation to forest corridor. These twelve variables are (Table 4):

1) The three types of farm,

2) The corridor zones as defined by the distance from forest (three zones: savannah, forest border, forest),

3) Three land uses: paddy-field, crop and fallow. Note that grass land use state is not specified. Note that these land uses are the agricultural land uses, which include the three land use states of timed automaton model (land uses with transitions) and the paddy-field land use without transitions,

4) the number of plots by farm type and by corridor zone (3 types and 3 zones, generating 9 variables). These complementary variables are used to define the initial state of the model.

Table 4 shows the 12 additional variables to add to the 4 already defined time delays and crop intensity parameters. The calibration of a model with 16 variables is heavy and the protocol of calibration is not defined here. Obviously, it would need more detailed data. As the transition rules are the same for all farmers, the differences between farms are due to the number of plots (size) and the spatial plot distribution, and not to the time delays of the four selected parameters.

Table 4: Farm typology according to the spatial distribution of plots

FARM TYPE	ZONE			Total plots	No paddy fields plots	Farm house locali- -sation
	Savannah	Forest border	Forest			
Savannah farms (12%) [1 Farm]	- 2 paddy - 1 crop	- 2 crops - 2 fallows		7	5	Savannah
Mixed farms (58 %) [6 Farms]	- 2 paddy - 1 crop	- 1 paddy - 1 crop	- 1 paddy - 1 crop - 1 fallow	8	4	Savannah
Forest farms (30%) [3 Farms]		- 1 crop	- 3 paddy - 1 crop - 1 fallow	6	3	Forest

The validation was achieved on the global result of the ABM, the land use landscape that results from the farm distribution and the farm strategies. The results are shown in the figure 5.

Table 5: Comparison of Markovian matrices with observed and ABM simulated transitions for validation

	Observed matrix				ABM simulated matrix			
States	Forest	Fallow	Crop	Herb	Forest	Fallow	Crop	Herb
Forest (F)	0	0	1	0	0	0	1	0
Fallow (J)	0	0.65	0.35	0	0	0.69	0.31	0
Crop (C)	0	0.32	0.65	0.03	0	0.29	0.69	0.02
Herb (H)	0	0	0.17	0.83	0	0	0.17	0.83
χ^2 test, risk α=5%, freedom degree=12 H_0: observed matrix is identical to simulated matrix Statistic test: 4.99; critical threshold: 21.03 H_0 approved at a 5% risk: no significative difference between both matrices								

6. Discussion

In Madagascar, forest conservation is viewed as a spatial protection of a specific forest territory. Our hypothesis is that farmer adaptation is mainly a spatial adaptation. A spatial typology has been built and tested by ABM on forest-agriculture transitions. This spatial typology does not validate the time delays. The parameters of this spatial typology are issued from field studies. These parameters have not been used in the ABM calibration but they are used to initialize this model. Thus the ABM is calibrated with the four fitted parameters issued from the TAM.

Additionally, a Markov model has been used for calibration of the TAM and for both validations.

The three kinds of models used in this modeling process are part of the ABM validation. Each one is dedicated to a particular kind of use. In order to make it explicit, we propose a set of descriptors to qualify these kind of models: the structure of the initial state, and the possible thematic objects (table 6); the mathematical, computational and thematic objectives or questions, and their respective answers, relatively to hypothesis (mathematical and computational) or assumptions (thematic) (table 7); finally the calibration and validation process, and the used scenarios (table 8).

7. Conclusion

Model validation is a complex process. It needs to define a protocol and intermediate steps. Progressive validation could be another but unexplored solution. Relatively to ABM validation, our proposal is to rely on one of the property of multi-agent systems, i.e. to produce a global phenomenon from local interactions. Therefore, its success relies on its ability to accurately produce this global phenomenon. Consequently, we proposed to describe the global dynamics using a Markov model in order to summarize it in a simple, stochastic, way. Comparing the global dynamics produced by the ABM simulation with the observed global dynamics allows to validate the model as being able to reproduce the targeted behavior. But it does not imply that the underlying local interactions are the right ones. It only implies that the hypotheses made at the local level do not contradict the globally observed behavior. However these local interactions themselves need to be calibrated and validated. Being, in principle, simpler than the complete ABM system, we used a third formalism, the timed automaton, in order to calibrate and validate the local behaviors (here, the transition rules). The general lesson learned from our proposed methodology is the usefulness of using a set of carefully chosen formalisms, and not only a single one, to deal with the various aspects of both the ABM and the targeted research questions. Having used three formalisms; the Markov chain, the timed automaton and the ABM, we proposed a synthetic analysis of the con-

Table 6: Comparative synthesis of the 3 models (1)

Model	Object		Initial state
	Formal	Thematic	
Markov (stochastic)	Vector state and transition matrix (mathematical)	Temporal sequences of vegetal cover in each plot: crop and fallow successions	Any kind of land use. In our case, all forest.
Timed automaton (deterministic)	Finite state automaton with timed transition rules (computational)	Time delays and transition rules per plot, extracted from owners inquiries and direct observations. Transition rules are considered as the same mean rules for all peasants.	Number of plots initially forested and number of plots initially cultivated, on the basis of one plot per person.
Agent-based (ABM)	Multi-agent system with its agents and environment	- Decision rules - Farmers - Landscape	Farm typology defined by: 3 zones by their distance from the forest, Number of farms per type of farms Number of plots per farm and per zone.

Table 7: Comparative synthesis of the 3 models (2)

Model	Question		Answer	
	Formal	Thematic	Formal	Thematic
Markov (stochastic)	Is there a stationary state ?	Do we obtain a stable mosaic landscape ?	Homogeneity of the transition matrix Ergodicity Vector made of state probabilities	Partial answer Historic of a sample of spatialised plots.
Timed automaton (deterministic)	How to represent time delays by a minimal automaton ?	What is the medium behaviour of peasants to cover aliment needs of the family?	Time delays in the transition graph and motor for abandoned plot replacement.	Average time delays in each of the three states and number of crop fallow cycles (obtained by calibration) Motor replacement of abandoned plot.
Agent-based (ABM)	What is the global phenomenon that emerges from the individual behaviours? What kind of interaction between agents and between the agents and the environment generates a global given dynamics?	What landscape is build from an initial distribution of farms in the landscape ?	History of land use state successions. The fitted time delays correctly generate the sequences we want to reproduce, the real observed sequences.	Spatial typology does not answer to the question of time delays in each land use state.

Table 8: Comparative synthesis of the 3 models (3)

Model	Calibration	Validation	Scenarii
Markov (stochastic)	Building the transition matrix from data.	Comparing matrices calibrated with two different sets of data	By modifying the initial state.
Timed automaton (deterministic)	Fitting the parameters of the transition rules.	Validation by simulating with another set of data, using the values of fitted parameters by calibration.	With fixed time delays, deforestation scenarii varying the rate of abandoned plot remplacement. Formally, these scenarii are cellular automaton scenarii, not timed automaton.
Agent-based (ABM)	Calibration meets difficulties if too much parameters have to be estimated by simulation. Some parameters of agent behavior may be qualitative.	Validation by simulation is difficult if too much parameters	Complex scenarii Forest cover evolution varying initial pressure on land, demographic growth rate and land use intensification strategies.

tribution of each of them for various questions to be answered around a complex model for both calibration and validation.

The definition and realization of a global evaluation of an ABM simulation still leave a conceptual unsolved problem. Even if the desired global phenomenon is produced, the calibration of a large number of parameters at the local level remains to be done, as well as their validation. A method could be to validate individual strategies using the farm typology, and then the effect of the interactions (in our case the plots exchanges). However, it remains problematic with complex systems in general, where the global behavior is more than the sum of the individual ones.

References

[AG57] T. W. Anderson and L. A. Goodman. Statistical inference about Markov chains. *Annals of Mathematical Statistics*, 28:89–109, 1957.

[ARB06] Fréderic Amblard, Juliette Rouchier, and Pierre Bommel. *Evaluation et validation de modèles multi-agents*, chapter 2, pages 87–122. Hermès Science, 2006.

[Axe97] R. Axelrod. *The complexity of Cooperation : Agent-Based Models of Competition and Collaboration*. Princeton University Press, 1997.

[Bal00] H. Baltzer. Markov chain models for vegetation dynamics. *Ecological Modelling*, 126:139–154, 2000.

[Ber88] A. Berchtold. *Chaînes de Markov et modèles de transition. Applications aux sciences sociales*. Hermès Science, 1988.

[Bom97] Pierre Bommel. Recherche d'une méthodologie pour la validation des modèles multi-agents. application à la modélisation de la pêche artisanale au Sénégal (modèle MOPA). Master's thesis, Université de Montpellier II, 1997.

[Bou95] François Bousquet. Représentation de la complexité des exploitations halieutiques. In Francis Laloë,

H. Rey, and J.-L. Durand, editors, *Questions sur la dynamique de l'exploitation halieutique*, pages 141–166. ORSTOM/IFREMER, 1995.

[BP07] Müller Jean Pierre Bommel Pierre. *Agent-based modelling and simulation in the social and human sciences*, chapter An introduction to UML for modelling in the human and social sciences, pages 273–294. Bardwell Press, 2007.

[CBND02] Jean-Christophe Castella, Stanislas Boisseau, T. Tran Ngoc, and D. Q. Dang. Samba, un système multi-agents pour la compréhension des dynamiques agraires: cas des zones de montagne du bassin du fleuve rouge (Vietnam). In *Collection Colloques et Séminaires*, page 20. IRD Editions, 2002.

[CH97] P. Coquillard and David Hill. *Modélisation et simulation d'écosystèmes. Des modèles déterministes aux simulations à évènements discrets*. Recherche en Ecologie. Masson, 1997.

[CL03] Marie-Odile Cordier and C. Largouët. Interprétation d'une séquence d'images à l'aide d'un modèle d'évaluation du système observé. application à la reconnaissance de l'occupation du sol. In *Colloque Automatique et Agronomie*, Montpellier, 2003. CNRS-INRA-ENSAM.

[CRP01] F. Colin, D. Raclot, and C. Puech. Traitement d'image en mode parcellaire couplé à un système de règles de décision. application à la classification de l'occupation du sol en zone de petit parcellaire agricole. In Thérèse Libourel, editor, *Géomatique et espace rural*. CIRAD, 2001.

[DBB97] A. P. Dobson, A. D. Bradshow, and A. J. M. Baker. Hopes for the future: Restoration ecology and conservation biology. *Science*, 277:512–515, 1997.

[Def05] Guillaume Deffuant. Can we learn something from agent-based models? some lessons from the european project images. In *Modèles et simulations multi-agents de systèmes complexes pour les sciences de l'homme et de la société : principes et méthodes de conception et d'usage*. Ecole de Porquerolles - CNRS, September 2005.

[Del94] Philippe Delcros. *Ecologie du paysage et dynamique végétale post-culturale en zone de montagne*. PhD thesis, Université Joseph Fournier, Grenoble, 1994.

[Gin03] Vincent Ginot. Exploration numérique des propriétés d'un modèle. In Jean-Pierre Müller, editor, *10èmes journées de Rochebrune, Rencontres interdisciplinaires sur les systèmes complexes naturels et artificiels: Le statut épistémologique de la simulation*. ENST, 2003.

[GM07] Vincent Ginot and Hervé Monod. *Exploring Models by Simulation: Application to Sensitivity Analysis*, chapter 3, pages 63–92. The Bardwell Press, 2007.

[GS90] G.M. Green and R. W. Sussman. Deforestation history of the eastern rain forests of Madagascar from satellite images. *Science*, 248:212–215, 1990.

[Guy06] Paul Guyot. *Simulations multiagents participatives*. PhD thesis, Université Paris VI, 2006.

[LC00] C. Largouët and Marie-Odile Cordier. Modélisation par automate temporisé pour aider à l'identification de l'occupation du sol. *Reconnaissance des Formes et Intelligence Artificielle*, 2:285–294, 2000.

[LDL85] E. Lippe, J. T. Desmidt, and D. C. Glena Lewin. Markov models and succession : a test for heathland in the Netherlands. *Journal of Ecology*, 73:775–791, 1985.

[LDM+05] S. Ladet, M. Deconchat, Claude Monteil, J.-P. Lacombe, and G. Balent. Les chaînes de Markov spatialisées comme outil de simulation. *Revue de géomatique*, 15:159–173, 2005.

[LF02] A. T. S. Landzer and V. D. Fillar. Probabilistic cellular automata model and application to vegetation dynamics. *Community Ecology*, 3(2):159–167, 2002.

[LF05] Jean Le Fur. Approches de la validation des modèles de système complexe: exemples issus de simulations multiagents de secteurs halieutiques. Atelier Modélisation Environnement, NSS-D et IRD, Montpellier, January 2005.

[Lui02] J.C. Luijten. A systematic method for generating land use patterns using stochastic rules and basic landscape characteristics : results for a colombian hillside watershed. *Agriculture, Ecosystems and Environment*, 95:427–441, 2002.

[MM94] M. R. Muller and J. Middelton. A markov model of land-use change dynamics in the Niagara region, Ontario, Canada. *Landscape Ecology*, 9(2):151–157, 1994.

[RHRM11] Venot Ratiarson, Dominique Hervé, Cyprien Rakotoasimbahoaka, and Jean-Pierre Müller. Calibration et validation d'un modèle de dynamique d'occupation du sol post-forestière à base d'automate temporisé à l'aide d'un modèle markovien. application à la transition forêt-agriculture à Madagascar. *Cahiers de l'Agriculture*, 20:274–279, 2011.

[RRRH10] Cyprien Rakotoasimbahoaka, Venot Ratiarson, Bertin Ramamonjisoa, and Dominique Hervé. Modélisation de la dynamique d'aménagement des bas-fonds rizicoles en forêt. In A. Sbihi, E. Badouel, and I. Lokpo, editors, *Actes X CARI*, volume Modélisation des Systèmes Complexes, pages 293–300, 2010.

[RSC07] J. R. Randriamalala, G. Serpantié, and S. M. Carrière. Influence des partiques culturales et du milieu sur la diversité des jachères d'origine forestière (Hautes-Terres, Madagascar). *Revue d'écologie (Terre Vie)*, 62:169–189, 2007.

[Ryk96] E. J. Rykiel. Testing ecological models: the meaning of validation. *Ecological Modelling*, 90(2):229–244, 1996.

[Sar84] R.G. Sargent. A tutorial on verification and validation of simulation models. In S. Sheppard, U. Pooch, and D. Pegden, editors, *Proceedings of the 1984 Winter Simulation Conference*, pages 115–122. IEEE, 1984.

[TDM+07] Jean-Pierre Treuil, Alexis Drogoul, Claude Miller, Jean-Pierre Müller, and Dominique Hervé. De l'équation mathématique au programme informatique, dialogue entre mathématiciens et informaticiens. *Nature Sciences Sociétés*, 15:33–43, 2007.

[TDZ08] Jean-Pierre Treuil, Alexis Drogoul, and Jean-Daniel Zucker. *Modélisation et simulation à base d'agents.*

Exemples commentés, outils informatiques et questions théoriques. Dunod, 2008.

[Toi09] Aurélie Toillier. *Capacités d'adaptation des agriculteurs à la conservation des forêts dans le corridor Ranomafana-Andringitra (Madagascar). Perspectives pour un aménagement intégré des territoires*. PhD thesis, AgroParisTech, 2009.

[TSHL11] Aurélie Toillier, G. Serpantié, Dominique Hervé, and Sylvie Lardon. Livelihood strategies and land use changes in response to conservation : Pitfalls of community-based forest management in Madagascar. *Journal of Sustainable Forestry*, 30(1):20–56, 2011.

[Tur87] M. G. Turner. Spatial simulation of landscape changes in Georgia: a comparison of 3 transition models. *Landscape Ecology*, 1(1):29–36, 1987.