Jul 1st, 12:00 AM

Participatory Land Use Modeling with Bayesian Networks: a Focus on Subjective Validation

Enrico Celio
Sibyl H. Brunner
Adrienne Grêt-Regamey

Follow this and additional works at: https://scholarsarchive.byu.edu/iemssconference

Celio, Enrico; Brunner, Sibyl H.; and Grêt-Regamey, Adrienne, "Participatory Land Use Modeling with Bayesian Networks: a Focus on Subjective Validation" (2012). International Congress on Environmental Modelling and Software. 141.
https://scholarsarchive.byu.edu/iemssconference/2012/Stream-B/141

This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
Participatory Land Use Modeling with Bayesian Networks: a Focus on Subjective Validation

Enrico Celio*, Sibyl H. Brunner*, Adrienne Grêt-Regamey*
*ETH Zurich, Institute for Landscape and Spatial Development, Planning of Landscape and Urban Systems (PLUS), Wolfgang-Pauli-Strasse 15, 8093 Zürich, Switzerland. Emails: ecelio@ethz.ch, brunner@nsl.ethz.ch, gret@nsl.ethz.ch

Abstract: Taking into account land use models in spatial planning processes would be a valuable support for facilitating the development and the acceptance of innovative land use management strategies. But, the utilization of such models in spatial planning processes is rarely done. One condition for implementation is trust and credibility in the model, which can be increased through validation. We argue that subjective validation can be a major contribution to efforts advancing the implementation of models. The presented land use model applies Bayesian networks (BN) and incorporates secondary data as well as local actors’ characteristics. Stakeholders from the study area and experts were involved in designing and updating the network in workshops and questionnaires. The subjective validation procedure is based on the concepts of conceptual and operational validation and these are applied in a workshop format. Results of the subjective validation process show that experts have understood mechanisms of the BN and their inputs can be effectively used for model improvement. The parameterization of the BN was adapted and the understanding of the processes improved. Comparing the results of the subjective validation with the outcome of an objective validation approach reveals surplus concerns concerning credibility of the model and mutual learning processes.

Keywords: land use modeling; Bayesian networks; decision model; validation; subjective validation

1 INTRODUCTION

Land use change is dynamic [Schneeberger et al. 2007] and is one of the most important and intense drivers altering the provision of landscape goods and services [MA 2005]. Past and current land use changes (e.g. urban sprawl) have had negative effects on the amount and quality of provided goods and services. Addressing these impacts is a key challenge for spatial planning, but such efforts are often not effective and the land use changes are not stopped. Integrating land use models to support management and decision making in spatial planning might support such processes, but decision-makers are often skeptic towards their utilization and only a few decision support systems are used in practice. Skepticism arises mainly because the individual case is never tapped, problem framing together with the stakeholders is avoided, and the model is not transparent to the target group [cf. van Delden et al. 2011]. Learning with the help of models can also be a fruitful process as new perspectives are gained and the discussion is rationalized. But, requirements for a successful application of land use models have not fully been defined yet. Following McIntosh et al. [2008], a model needs to be developed collaboratively, purpose of modeling
must be clear, modelers need to understand the users’ needs, and credibility must be established. We argue that, as a preliminary, the developed model must be understood by all involved persons (e.g. experts involved in set up process) as well as the target group (e.g. public official that applies the model). Hence, a model needs to match - to a certain extent – the perceptions of the target group and be based on local actors’ knowledge to be better accepted and implemented by the target group in spatial planning.

Spatially explicit land use models using Bayesian Networks (BN) have two advantages compared to other land use modeling approaches: firstly, uncertainty inherent to decision-making and secondly, stakeholder knowledge can be accounted for in a spatially explicit manner. The variability of stakeholder knowledge over time is however a major challenge for the validation of these models. It is likely that local actors’ decisions are different in different time periods because they face other basic conditions. This cannot be validated objectively. Proper validation is then no longer a question of data availability, but also of good collaboration with the target group. We argue that validation itself should have a quantitative or objective part [cf. Kok et al. 2001, Pontius and Millones 2011] as well as a subjective part (involving stakeholders [cf. Sargent 2010]) to take into account possible changes in stakeholders’ decisions. Despite the acknowledged importance of validation, a majority of models “lack a validity check” [Kok et al. 2001]. Especially subjective validation is less described in literature, even though it is a mandatory mean for securing credibility in participatory modeling [Prell et al. 2007, Voinov and Gaddis 2008] and when complexity of integrated models increases [Nguyen et al. 2007]. Only few examples describe subjective validation processes. Nguyen et al. [2007] use a scenario building-process to validate model outcomes and Millington et al. [2011] conduct interviews with local stakeholders for the evaluation of agent-based land use models.

In this contribution, we use a BN linked to spatially explicit data to model land use changes in a case study in Switzerland. The final aim of the modeling process is to display most probable land use changes under different socio-economic conditions, which help rationalize current decisions in the spatial planning of alpine and pre-alpine regions and support the identification of key influence factors on land use decisions. The BN approach allows including secondary data (geodata, statistical data) and local actors’ characteristics (primary data, e.g. data gained from a questionnaire), thus, integrating local actors’ intention to alter their land use. To engage in the efforts of advancing participatory modeling, we shortly explain the model and the underlying modeling process as well as the quantitative validation reclining on Pontius (2011) and focus mainly on subjective validation to improve model credibility and accuracy. We test the model output as well as the behavior of the BN in a workshop and propose improvements of the model setup. The results show that experts can be fully integrated in the validation procedure and how subjective validation can contribute to improve the modeling process.

2 METHODS

The land use model was set up in three modules: (1) Elaboration of Bayesian networks (BN) to model land use decisions in an experts’ process, (2) updating of the Bayesian networks by surveys, and (3) Java scripting to integrate spatially explicit data with the other two modules. The BN elaboration process was conducted in an iterative manner by alternating group and single face-to-face meetings. The second module improves the BN by integrating spatially explicit characteristics of local actors in agriculture, forestry and in the settlement. For the technical implementation of the model, a Java console is used to integrate spatial data and BN as well as to calculate time steps.

Figure 1 shows a simplified BN for decisions made on agriculture land. Two separate BN have been elaborated for forestry and the settlement. The variables influencing land use decisions are represented by nodes characterized by different states and connected through arcs showing causal relationships among these
variables. The causal connections are quantified through conditional probability tables (CPT) [cf. Kjaerulf and Madsen 2008], thus, taking into account the inherent variable uncertainty and their relationships. Central to all elaborated BN are nodes that reflect intention of the decision maker and his or her actual behavior (this distinction can be found in Ajzen [1991]).

![Diagram of Bayesian network](image)

Figure 1: Simplified Bayesian network depicting nodes and states for modeling agriculture land use change

The experts had to spend approximately ten hours each for the set up process (Module 1). Due to high requirements of system knowledge the number of experts was left to six persons (two per modeled land use sector: agriculture, forestry, settlement). Experts hold mainly higher functions in Cantonal administration, agricultural or spatial planning consultation and forestry. The co-construction process was led by the modeler: Single face-to-face meetings were alternated by group meetings. To ensure maximum transparency the experts were informed about all the discussions held in single-meetings. Their output (e.g. weightings, appraisals of probabilities) was constantly given back to the others for finding consensuses. This procedure enabled recording uncertainties in experts’ judgment as well as ensuring the voicing of every participant [cf. Voinov and Bousquet 2010]. Additionally, local actors are surveyed to update the BN according to spatially explicit characteristics of the actors.

The model was quantitatively validated just before the validation workshop by comparing the modeled land use maps and maps of observed land use data [Pontius and Millones 2011]. Pontius and Millones [2011] recommend using allocation and quantity disagreement which indicate whether a modeled outcome meets the amount of pixels in a category (quantity disagreement) and whether the pixels are at a location depicting the category in question:

Overall quantity disagreement:

$$QD = \frac{\sum_{i=1}^{J} \left| \left( \sum_{i=1}^{J} p_{ij} \right) - \left( \sum_{i=1}^{J} p_{ij} \right) \right|^2}{2}$$  \hspace{1cm} (1)

Overall allocation disagreement:

$$AD = \frac{\sum_{i=1}^{J} \left( \min \left( \sum_{i=1}^{J} p_{ij}, \sum_{i=1}^{J} p_{ij} \right) \right)^2}{2}$$  \hspace{1cm} (2)

where $p$ is the portion of a sample map compared to the total area, $J$ is the total number of land use types, $i$ and $g$ are the different land use types. When a reference map is available for all pixels, then the portions of the sample map can be expressed as number of pixels in each category. The applicability of the approach is dependent on the availability of historical data. For our case study, we
have generated forest data for 1986 and settlement data for 2006 from analog archive data as well as agricultural data for 2006 from a Cantonal database.

For the qualitative (subjective) validation, a workshop with all the experts was held as an intermediate breakpoint of the entire model development process. The workshop aimed at ensuring relevance of the chosen variables (e.g. land use categories) as well as to test the model outputs for accuracy. All six experts took part in the workshop. The scientist group consisted of three persons to guide and assist in the different exercises. The outcome of the exercises was recorded in audio and in writing. The meeting lasted approximately two hours and was held at a meeting room of the Cantonal administration.

The workshop had two parts and three exercises were elaborated [cf. van Delden et al. 2011]: Part A on conceptual model validation concentrated on the question ‘are the foundations (theories and assumptions) of the model credible?’ [cf. Sargent 2010] and part B on operational model validation asked ‘is the simulation credible? Can the tool be used as a decision support tool for securing multifunctional cultural landscape?’ [cf. Sargent 2010].

Exercise “A.1.” consisted of visualizing scenarios in the BN software (in a node-graph-structure) and discussing the related land use change implications calculated by the model. The scenarios were intuitively formulated by the experts [cf. Nguyen et al. 2007] and entered into a BN interface on computers (using Netica 4.16). Different settings of the BNs were tested following a guided questionnaire by comparing land use changes modeled by the BN to those expected by the experts. In exercise “A.2.”, the probability distribution of future land use was traced over time: The probability distribution of the land use after different time steps was presented to the experts and their comments on the land use developments were recorded to improve the model parameterization. “B.1.” consisted of a sequence of land use development over time [cf. face validity, scenario in Sargent 2010]: A land use scenario which has been regionalized on the base of another research project in which scenarios were designed for the alpine- and pre-alpine area (cf. http://www.cces.ethz.ch/projects/sulu/MOUNTLAND) was presented to experts. They were asked to draw on a map where land use changes should be different from the modeled ones and to give reasons for their disagreements. The following checks were not part of the described validation procedure: data validity; computerized model verification, specification verification, implementation validation [cf. Sargent 2010].

Maps displayed in the result section are a cut-out of 2.5 km² of the total study area (catchment of the river Kleine Emme, mostly Canton of Lucerne, Switzerland, approx. 481 km²).

3 VALIDATION RESULTS

The modeled maps show mainly two changing components: the settlement development and the changing intensity in agriculture (Figure 2). Focusing first on objective validation, allocation disagreement and quantity disagreement of the validation sample are shown in Table 1. It is obvious that on this stage of development, modeled output is closer to the reference map of 2006 than to the observed land use in 2011. Modeled outputs for 2011 are however much better than a random land use change simulation. Comparing the two reference maps the disagreement indicators reflect the actual land use change.

| Table 1: allocation and quantity disagreement for the validation sample |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| allocation disagreement     | 0.00                            | 0.02                            | 9.58                          | 0.81                          |
| quantity disagreement       | 0.74                            | 7.22                            | 17.10                         | 6.71                          |
Comparing the modeled map 2011 with the reference of 2011 two main inconsistencies are detected: (1) In the settlement area, thirty cells should have been modeled as undeveloped in 2011, but are depicted in the modeled map as developed. In this case the reference map truly depicts the situation at some point in 2010 when the Cantonal administration had conducted the analysis of undeveloped parcels. Between 2010 and 2011, the undeveloped parcels were mainly developed. Therefore, the respective parameters in the CPT of the BN were not changed to fit the reference map state of 2011. Thus, the model better reflect the current state. (2) In the north-eastern agricultural area the reference depicts extensive use whereas the land uses modeled remain intensive. If the land use model is run to 2016, approximately half of the cells that are extensive in the reference map 2011 (but are not in the model output 2011) turn to extensive agricultural land use also in the modeled map. If this half is changed to extensive the allocation and quantity disagreement change to 0.02 and 3.64, respectively.

Table 2 shows the outcome of the workshop on the subjective validation of the model, divided into conceptual and operational validation. Results of the workshop are suggestions for improvements of the model as perceived by the experts.

Table 2: Results of validation workshop

<table>
<thead>
<tr>
<th>A. Conceptual model validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.1. Visualization of scenarios in Bayesian networks</td>
</tr>
<tr>
<td>Experts identified inexact network nodes and quantified necessary changes:</td>
</tr>
<tr>
<td>• In the agricultural BN: Node ‘agricultural environment’ needs to have a bigger</td>
</tr>
</tbody>
</table>

Figure 2: Modeled land use and comparison with reference maps. Data: GIS Kanton Luzern, swisstopo.
influence on node ‘intention’. ‘Intention’ has a too small influence on target node ‘land use’.*

• In the forestry BN: Node ‘environmental awareness’ needs to have a stronger influence on node ‘intention’. Changing the type of forest owners, corporations should be closer to a private sector actor than to a state actor influencing node ‘intention’.

• In the settlement BN: In node ‘attitude towards a possible development’, a third state (called “indifference”) should be added in node ‘attitude’. The big influence of node ‘time distance’ to agglomeration should be reflected and network connections changed to achieve more subtle distinctions in target node ‘land use’.

A.2. Probability distribution traced over time

• Probabilities need to be displayed in the maps.
• Very small probabilities of states which are impossible should be zero.
• Constraints (slope, skiing slopes etc.) should be taken into account in the modeling.

B. Operational validation

B.1. Animation of land use development

• In the agricultural sector: The modeling output for the scenario is credible, but should be compared to simulations in valley regions. Not in all the study area an extensification takes place.
• In the forestry sector: The time horizon of modeling needs to be extended to simulate changes in the forest.
• In the settlement sector: Undeveloped building sites inside building zones need to be visible (a patchy impression should be visible). A comparison with developments in other villages of the study area needs to be considered.

* Used to demonstrate consistency of model in the next paragraphs.
educated, had an innovative agricultural environment and worked as a part-time farmer. This shows the model reactions on local actors’ characteristics, but also how to use BN as a mean to communicate and learn from the model.

4 DISCUSSION

This contribution shows that experts can be fully integrated into land use model validation if they understand the underlying processes. A full validation procedure consists of two parts: objective and subjective validation. Being able to back up subjective validation with quantitative material from objective validation completes the picture for experts and the modeler. Furthermore, quantitative indicators serve as criteria to cross-check the implemented changes from subjective validation. At the same time, indicators of an objective validation are often difficult to communicate and iterative learning from model results is hindered. In our case, where availability of historical data is a critical issue, subjective validation becomes key. Especially in research targeting local actors, subjective validation allows a mutual learning process and becomes a plus for researchers and actors: the feedback induces model improvements and the co-production of knowledge can be implemented in every ones working fields. Both benefits became evident during the validation workshop: On the one hand, the feedbacks from the workshop exercises were valuable for the model improvements. On the other hand, the validation workshop showed that the modeling efforts can reveal new ideas and these are being understood by the other participants. The workshop helped to bridge ‘thought styles’ [cf. Pohl et al. 2010] as the thinking about the causal relations and its probabilities catalyzed rational arguments. BN proved to support such a subjective validation process and to improve the understanding of involved experts for land use changes as the process is not a black box modeling technique. Results from the conducted exercises can help setting up further validation workshops. We could identify three valuable elements of such a process: (1) Discussing the behavior of BN on a computer is a valuable mean for validation. The limited capability of human beings to estimate a probability is thus actively addressed as effects of changing probabilities are instantly recognized [cf. Kynn 2008, O’Hagan et al. 2006]. (2) The tracing of a single cell through time enhances understanding of the model and its underlying functions. (3) Through presenting land use scenarios, discussion was directed to a regional perspective of integrated spatial planning. Presenting these exercises rather at the beginning of the workshop contradicts the recommendation of Voinov and Gaddis [2008], but allows prime attention from the experts coming directly from their daily business, thus diminishing biases introduced by discussions on the modeling process, and fosters a focused discussion on the matter directly following the model demonstration. One may criticize that the same experts have set up and validated the model. This procedure checks whether the experts’ knowledge is appropriately implemented in the computerized model (whether experts recognize their work and this was confirmed). We argue that the bias of the experts’ group was very small as there was a time gap of about four months between the last interaction and the validation workshop. A further validation with completely external experts or local actors with an improved model is planned. In conclusion, besides enhancing the understanding of land use decision processes, such BN-based land use decision models, can provide important information for facilitating the development and the acceptance of innovative land use management strategies. In a modeling process with target groups outside academia, subjective validation is one major point to gain a high degree of confidence in the model.

ACKNOWLEDGEMENTS

This work was funded by the Swiss National Science Foundation (SNF) through National Research Program 61 (“Sustainable Water Management”). The authors
much appreciate the review of earlier versions of this paper by members of the

REFERENCES

Ajzen, I. The Theory of Planned Behavior. Organizational Behavior and Human


Kok, K., Farrow, A., Veldkamp, A. and Verburg, P. H. A method and application

Kynn, M. The 'heuristics and biases' bias in expert elicitation. Journal of the Royal


McIntosh, B. S., Giupponi, C., Voinov, A. A., Smith, C., Matthews, K. B., Monticino,

Millington, J. D. A., Demeritt, D. and Romero-Calcerrada, R. Participatory

Nguyen, T. G., Kok, J. L. d. and Titus, M. J. A new approach to testing an

O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H.,

Pohl, C., Rist, S., Zimmermann, A., Fry, P., Gurung, G. S., Schneider, F.,

Pontius, G. R. and Millones, M. Death to Kappa: birth of quantity disagreement and

Prell, C., Hubacek, K., Reed, M., Quinn, C., Jin, N., Holden, J., Burt, T., Kirby, M.


Schneeberger, N., Bürgi, M. and Kienast, P. D. F. Rates of landscape change at the

van Delden, H., Seppelt, R., White, R. and Jakeman, A. J. A methodology for the
design and development of integrated models for policy support. Environmental Modelling & Software. 26, 3, 266-279, 2011.

Voinov, A. and Gaddis, E. J. B. Lessons for successful participatory watershed