

# Calibrating and validating the Land Use Scanner algorithms

15TH European Colloquium on Theoretical and Quantitative Geography  
September 7-11, 2007, Montreux, Switzerland

W.Loonen<sup>1,2</sup>, E.Koomen<sup>2</sup>, M. Kuijpers-Linde<sup>1,2</sup>

<sup>1</sup> Netherlands Environmental Assessment Agency (MNP)

<sup>2</sup> Vrije Universiteit Amsterdam, Department of Spatial Economics/SPINlab

## ABSTRACT

The Land Use Scanner is a spatial model that simulates future land use. Since its development in 1997 it has been applied in many policy-related land use projects. In 2005 a completely revised version became available. In this version land use can be modelled at a more detailed 100 meter resolution as opposed to the original 500 meter resolution. Furthermore the new version offers the possibility to model homogenous cells that offer a discrete description of land use, in stead of the original continuous description that listed the fraction claimed by different types of land use in each cell.

In this paper we describe the calibration and validation of the two modelling approaches available in the Land Use Scanner model. We used multinomial logistic regression methods to obtain the suitability values for the different land-use types. The resulting simulations are then compared to the observed land use in the base year (calibration) and a future year (validation).

## KEYWORDS

Land-use modelling, calibration, validation, multinomial logistic regression, spatial optimization

## INTRODUCTION

The Land Use Scanner is a spatial model that simulates future land use. Since the development of its first version in 1997 it has been applied in many policy-related studies of land-use change. Applications include, amongst others: the simulation of future land use following different scenarios [1;2], the evaluation of alternatives for a new national airport [3], the preparation of the Fifth National Physical Planning Report [4] and an outlook for the prospects of agricultural land use in the Netherlands [5]. Apart from these Dutch applications the model has also been applied in several European countries [6;7] A full account of the original model is provided in [8]. For an extensive overview of all Land Use Scanner related publications the reader is referred to [www.lumos.info](http://www.lumos.info).

A seriously revised version (4.7) of the model became available in 2005. This new version offers the possibility to use a 100x100 metre grid, covering the terrestrial Netherlands in about 3.3 million cells. This resolution comes close to the size of actual building blocks and allows for the use of homogenous cells that only describe the dominant land use. The previous version of the model had a 500 metre resolution with heterogeneous cells, each describing the relative proportion of all present land-use types. Together with the introduction of homogenous cells using a dominant land use a new algorithm has been developed that finds the optimal allocation of land use given the specified demand and suitability definition. This new approach is referred to as the discrete model, as it uses a discrete description of land use per cell: each cell is assigned only one type of land use from the total range of possible land-use types. The original

model is in this paper referred to as the continuous model, since it uses a continuous description of land use per cell.

The Land Use Scanner now has a flexible layout that allows for the selection of five different resolutions, ranging from 100 to 10,000 metre, and the choice of the discrete or continuous model, thus providing a total of 10 basic model types. In this paper we focus on the calibration and validation of the two available 100 metre models. This choice is motivated by the fact that we are especially interested in 1) assessing the potential of the new fine resolution and 2) comparing the performance of the two available algorithms. We furthermore want to pinpoint the most important location factors in the suitability map definition and assess the relative weighing of the suitability maps for the different types of land use.

## METHODOLOGY

The analysis is divided in separate *calibration* and *validation* phases. The *calibration* aims at reproducing land use of the base year (1993) on which the model is tuned. The second, *validation* phase simulates future (2000) land-use patterns based on the relations established in the calibration phase and compares these with observed land-use patterns of the same year. For this analysis we use a simplified model configuration that uses nine major types of land use.

Three different methods are used to obtain the suitability maps for the calibration and validation attempts. The first and most simple method only takes the actual (1993) land use as the suitability specification and then uses the allocation algorithm to reconstruct these land use patterns. This very basic test is included to show that the model is capable of exactly reproducing current land use. The two other methods apply multinomial logit analysis to specify the suitability values and use a random set of observations consisting of about 1.600.000 cells. This calibration set represents about half of the total number of observations. The other half of the cells is used as an independent set that allows us to test validity of the observed relations on an independent dataset. The multinomial logit analysis provides an autologistic and a full factor specification. In the autologistic specification the suitability values only depend on the land use in the surrounding cells. For these surroundings we distinguish the 8 immediately surrounding neighbours (ring 1) and the following 40 (rings 2 and 3). The probability that a certain land use occurs is estimated here as a function of the total number of cells that belong to each of the nine distinguished land-use types. Implementation of the suitability values that are derived in this way causes the model to perform in neighbourhood-oriented manner that is similar to classical Cellular Automata. In the full factor specification we add a number of location characteristics to the variables related to the surrounding land use. These extra independent variables relate to a number of driving forces that are considered important in land use development: accessibility, spatial policies and the presence of main underground infrastructure.

To compare the maps of the different model runs, we use simple cell-to-cell comparisons of simulated and observed land-use in either 1993 (calibration) or 2000 (validation). These comparisons are made for all simulated land-use types separately, to individually assess their performance. We have not calculated an overall degree of correspondence for the whole simulation as this would be largely equivalent to the value for the prevailing land-use type (agriculture). The selected, straightforward comparison approach is easy to comprehend and very informative. Other, more complex comparison methods that deliver, for example, (Fuzzy) Kappa statistics or the log likelihood values are more difficult to interpret and have therefore not been applied.

Both the calibration and validation phase only test the functioning of the allocation algorithms and the suitability definition. We do this by allocating the actual 1993 (calibration) and 2000 (validation) amounts of land used by the various functions, thus only focusing on the land-use patterns resulting from the simulation and not on the land-use quantities. A full account of the methodology and its results is provided

in [9].

## RESULTS

The overall impression of the allocation mechanisms applied in the Land Use Scanner is positive. Both allocation algorithms provide similar results in a relatively short time span that compare favourably with the observed 1993 and 2000 land use. The discrete algorithm seems to produce higher similarity, but we have to stress that different comparison methods have been used and the values can not be directly compared. Furthermore the regression analysis has been performed on the discrete dataset. This may also cause higher similarity in the discrete model.

In the present research current land-use patterns and especially the presence of a similar land-use type in the neighbouring cells proved to be an important component in describing the suitability for most land-use types. This points at the importance of spatial auto-correlation in (Dutch) land-use datasets. We should, however, note that the current study estimated these relations in a static (one year only) regression analysis and tested them on a short, seven-year period. Other factors may prove to become more important when land-use changes are analysed and tested for a longer period of, for example, several decades. In this respect it is important to note that the aim of the model is not to replicate the past, but to visualize the possible state of the future and, more specifically, the impact of anticipated spatial planning measures.

In future research we want to pay specific attention to changed land use as this is the most difficult part of the simulation. We are currently thinking of implementing transition costs into the model, to improve the capacities. These transition costs can also be estimated in the multinomial logit, but require the availability of a third intermediate time-step in the base date (e.g. 1996) to be able to calibrate the model on one period (year 1-year 2) and validate it on another subsequent period (year 2-year 3). For the validation of the Land Use Scanner we use the statistical relations derived from the 1993 land-use patterns to simulate land use in 2000. The simulated land use is then compared to the observed land use in the same year. As we want to focus our comparison on the impact of the suitability map definition, we are using the exact quantities of land use per function for 2000. The regional land-use claims thus correspond exactly to the quantities of observed land use in 2000. The validation results are described for the three calibration efforts presented in the previous chapter.

## REFERENCES

1. **Borsboom-van Beurden, J.A.M., Bouwman, A.A. and Boersma, W.T.** Spatial elaboration of long-term scenarios. Future land use in the Netherlands. Environment and Planning B: Planning and Design, accepted.
2. **Dekkers, J.E.C. and Koomen, E.** Land-use simulation for water management: application of the Land Use Scanner model in two large-scale scenario-studies. In: Koomen, E., Stillwell, J., Bakema, A. and Scholten, H.J. (eds.). Simulating land-use change. New York, Kluwer Academic Publishers, 2007.
3. **Scholten, H.J., Van de Velde, R.J., Rietveld, P. and Hilferink, M.** Spatial information infrastructure for scenario planning: the development of a Land Use Planner for Holland. In: Stillwell, J., Geertman, S. and Openshaw, S. (eds.). Geographical Information and Planning, pp. 112-134. Berlin/Heidelberg/New York, Springer-Verlag, 1999.
4. **Schotten, C.G.J., Goetgeluk, R.W., Hilferink, M., Rietveld, P. and Scholten, H.J.** Residential construction, land use and the environment. Simulations for The Netherlands using a GIS-based land use model. Environmental Modelling and Assessment 6, 2001.

5. **Koomen, E., Kuhlman, T., Groen, J. and Bouwman, A.A.** Simulating the future of agricultural land use in the Netherlands. *Tijdschrift voor Economische en Sociale Geografie (Journal of Economic and Social Geography)* 96(2): 218-224, 2005.
6. **Hartje, V., Klaphake, A., Grossman, M., Mutafoglu, K., Borgwardt, J., Blazejczak, J., Gornig, M., Ansmann, T., Koomen, E. and Dekkers, J.E.C.** Regional Projection of Water Demand and Nutrient Emissions - The GLOWA-Elbe approach. Poster presented at the Statuskonferenz GLOWA-Elbe II, Köln May 18-19, 2005.
7. **Wagtendonk, A.J., Julião, R.P. and Schotten, C.G.J.** A regional planning application of Euroscanner in Portugal. In: Stillwell, J.C.H. and H.J.Scholten (eds.). *Land Use Simulation for Europe*. Amsterdam, Kluwer Academic Publishers, 2001.
8. **Hilferink, M. and Rietveld, P.** Land Use Scanner: An integrated GIS based model for long term projections of land use in urban and rural areas. *Journal of Geographical Systems* 1(2), 1999.
9. **Loonen, W. and Koomen, E.** Calibration and validation of the Land Use Scanner, Milieu- en Natuurplanbureau, Bilthoven, 2007.

#### **AUTHORS INFORMATION**

Willem Loonen  
Willem.Loonen@MNP.nl  
Netherlands Environmental  
Assessment Agency

Eric Koomen  
Ekoomen@feweb.vu.nl  
Vrije Universiteit  
Amsterdam

Marianne Kuijpers-Linde  
Marianne.Kuijpers@MNP.nl  
Netherlands Environmental  
Assessment Agency