Modelling land use change

Nederlandse Geografische Studies / Netherlands Geographical Studies

ISSN 0169-4839

Netherlands Geographical Studies 386

Modelling land use change

Improving the prediction of future land use patterns

Ton de Nijs

Utrecht 2009

Koninklijk Nederlands Aardrijkskundig Genootschap
Faculteit Geowetenschappen Universiteit Utrecht

Promotoren:
Prof.dr. S.M. de Jong
Prof.dr. E.J. Pebesma

Examen Commissie:
Prof.dr.ir L.O. Fresco, Universiteit van Amsterdam
Prof.ir. N.D. van Egmond, Universiteit Utrecht
Prof.dr. A. van Veen, Universiteit Twente
Prof.dr. F. Canters, Vrije Universiteit Brussel
Dr.Ir. P. H. Verburg, Wageningen Universiteit

ISBN 978-90-6809-429-9

Graphic design, cartography and figures:
GeoMedia (Faculty of Geosciences, Utrecht University)

Printed in the Netherlands by A-D Druk b.v. – Zeist

Contents

Figures		**9**
Tables		**11**
1	**Introduction**	**13**
1.1	Societal Relevance	13
1.2	Scientific Relevance	14
1.3	A framework for quality assurance	16
1.3.1	Improving the Conceptual Model	17
1.3.2	Improving the Model	18
1.3.3	Improving the Application	18
1.4	Objectives and approach	19
1.4.1	Improving the Conceptual Model	20
1.4.2	Improving the Model	21
1.4.3	Improving the Application	21
1.5	Thesis outline	22
	Note	24
2	**The Environment Explorer: spatial support system for integrated assessment of socio-economic and environmental policies in the Netherlands**	**25**
	Abstract	25
2.1	If only we knew…	26
2.2	An integrated spatial model of the Netherlands	27
2.2.1	Autonomous developments, intended and actual policies	27
2.2.2	Models coupled at 3 geographical levels	28
2.3	Using the Environment Explorer for integrated planning	32
2.3.1	State variables and Indicators	32
2.3.2	Scenarios, Strategies and Projects: interventions in the spatial structure	33
2.4	An extensive analysis	36
2.5	An evaluation	36
3	**Constructing land use maps of the Netherlands in 2030**	**39**
	Abstract	39
3.1	Introduction	39
3.2	The Individual World scenario	40
3.3	The Environment Explorer	41
3.4	Model input and assumptions	43
3.5	Future land use in an individual world	46

3.6	Discussion	46
3.7	Conclusions	48
4	**A method to analyse neighbourhood characteristics of land use patterns**	**49**
	Abstract	49
4.1	Introduction	49
4.2	Methods	52
4.2.1	Characterisation of neighbourhood characteristics	52
4.2.2	Logistic regression	54
4.2.3	Data	54
4.3	Results	56
4.3.1	Neighbourhood characteristics of land use configuration in 1989	56
4.3.2	Region-dependent variability in neighbourhood characteristics	59
4.3.3	Neighbourhood characteristics of changes in land use between 1989 and 1996	61
4.3.4	Scale dependence of the results	62
4.3.5	Logistic regression: conversion probabilities based on neighbourhood characteristics	63
4.4	Discussion and conclusions	65
5	**Determinants of land use change patterns in the Netherlands**	**69**
	Abstract	69
5.1	Introduction	69
5.2	Determinants of land use change	70
5.2.1	Biophysical constraints and potentials	70
5.2.2	Economic factors	71
5.2.3	Social factors, irreversibility and uncertainty	71
5.2.4	Spatial interaction and neighbourhood characteristics	72
5.2.5	Spatial policies	72
5.3	Methods and data	73
5.3.1	Methods	73
5.3.2	Data	74
5.4	Long-term land use changes	80
5.4.1	Model definition	80
5.4.2	Results	81
5.5	Recent land use changes: 1989-1996	84
5.5.1	Location characteristics	84
5.5.2	Location and neighbourhood characteristics	90
5.6	Discussion and conclusion	92
6	**The Map Comparison Kit**	**95**
	Abstract	95
6.1	Introduction	95
6.2	The Map Comparison Kit	98
6.2.1	Software	98
6.2.2	Handling different measurement scales	99
6.3	Map comparison methods	102

6.3.1	Visual versus automated map comparison	102
6.3.2	Cell-by-cell map comparison	103
6.3.3	Fuzzy-based map comparison	103
6.3.4	Map similarity measures other than Kappa	105
6.4	Comparing land use maps	106
6.4.1	Simulating land use	106
6.4.2	Map comparison	106
6.5	Final remarks	109
7	**Measuring the performance of land use models. An evaluation framework for the calibration and validation of integrated land use models featuring cellular automata**	**111**
	Abstract	111
7.1	Introduction	111
7.2	Methodology	112
7.2.1	Calibration procedure	112
7.2.2	Validation procedure	115
7.3	Results	116
7.3.1	The pilot case: Environment Explorer for the Netherlands.	116
7.3.2	Further experience with Estonia and northern Italy	118
7.4	Conclusions	120
8	**Spatial uncertainty in land use models. An alternative method to estimate uncertainty in logistic regression methods**	**123**
	Abstract	123
8.1	Introduction	123
8.2	Model	124
8.3	Uncertainty	124
8.4	Results	127
8.5	Discussion and Conclusions	129
9	**Future land use in the Netherlands: evaluation of the National Spatial Strategy**	**131**
	Abstract	131
9.1	Introduction	131
9.2	Environment Explorer	134
9.3	The Trend and Sprawl scenario	135
9.4	Results	138
9.4.1	Urbanisation probabilities in 2030	138
9.4.2	Pressures on land in the restrictive areas	139
9.4.3	Urban development in the National Landscapes	139
9.4.4	Space in the concentration areas	140
9.5	Conclusions and future directions	141

10	**Estimating the influence of the neighbourhood in the development of residential areas in the Netherlands**	**145**
	Abstract	145
10.1	Introduction	145
10.2	Data	146
10.2.1	Land use	146
10.2.2	Composition of the neighbourhood	147
10.2.3	Policy and suitability maps	148
10.2.4	Study area and data sampling	149
10.3	Methods	149
10.3.1	Single-parameter impact estimation.	151
10.4	Model validation	153
10.5	Model results	154
10.5.1	Policy and suitability	154
10.5.2	Initial land use	155
10.5.3	Influence of the composition of the neighbourhood	155
10.5.4	Influence of the cluster size	162
10.5.5	Probability map	163
10.5.6	Spatial autocorrelation	163
10.6	Discussion and conclusions	164
10.6.1	Methodological aspects	164
10.6.2	Determinants of residential growth	164
11	**Synthesis**	**167**
11.1	Introduction	167
11.2	Improving the Conceptual Model	168
11.3	Improving the Model	170
11.4	Improving the Application	171
11.5	Main findings and developments	172
Summary		**175**
Samenvatting		**179**
Dankwoord		**185**
Curriculum Vitae		**187**
Publications		**189**
References		**191**

Figures

1.1	Definition of the steps involved in the development of a simulation model.	17
1.2	Structure of the thesis in relation to the stages in model development.	23
2.1	National, regional and local level in the Environment Explorer.	27
2.2	Four elements determining the dynamics at the local level.	31
2.3	Spatial fragmentation, one of the indicators in Environment Explorer.	34
2.4	Graphical user interface of the Environment Explorer.	34
2.5	Probability map for the land use 'industry' in the Environment Explorer.	35
3.1	Four scenarios in Nature Outlook 2 and their major characteristics.	41
3.2	National, regional and local level in the Environment Explorer.	42
3.3	Schematic representation of neighbourhood in the Environment Explorer.	42
3.4	Land use in 2030 for the Individual World scenario.	46
4.1	Configuration of neighbourhoods used in this study.	52
4.2	Land use in the Netherlands in 1989 and 1996.	55
4.3	The logarithm of the enrichment factor as a function of distance.	57
4.4	Neighbourhood characteristics for grassland in the Netherlands.	59
4.5	Region-specific neighbourhood characteristics for residential area.	60
4.6	Neighbourhood characteristics for locations converted into industrial/ commercial, recreational and residential area.	62
4.7	Neighbourhood characteristics for residential area using high resolution data.	63
4.8	Locations with high probability for conversion into residential area and actual locations with new developments.	64
4.9	Procedure for defining transition rules within a cellular automata land use change model.	66
5.1	Location of forest/nature areas and probability for the occurrence of forest/ nature based on the logit model.	82
5.2	Some example maps of the variables used in this study.	83
5.3	Observed land use changes and probabilities based on a logistic regression model without and with neighbourhood factors.	88
5.4	Neighbourhood characteristics for locations converted into industrial/ commercial, recreational and residential area.	90
6.1	The Hobbits: spot the 15 differences in the drawings.	97
6.2	Impression of the user interface of the Map Comparison Kit.	98
6.3	Window showing all ten comparison methods of the Map Comparison Kit.	99
6.4	Advanced settings window for the Fuzzy Set algorithm.	101
6.5	Three distance-decay functions available in the Fuzzy Set algorithm.	101
6.6	Two pairs of maps showing differences in similarity, but identical results for Percentage-of-Agreement and the Kappa statistic.	104

6.7	The Fuzzy Set Map Comparison of the two pairs of maps reflecting the level of similarity.	104
6.8	Actual land use for 1995 and simulated land use for 2020.	107
6.9	Map differences for two single land use categories using 'Per Category' comparison method.	108
6.10	Equal-unequal map and corresponding Kappa statistics, using 'Cell by cell', comparison.	108
6.11	The Fuzzy similarity map and corresponding Fuzzy Kappa statistic.	109
7.1	Flowchart overview of the full calibration procedure.	113
7.2	Distribution of activity growth error over the regions.	117
7.3	Spatial distribution of disagreement in 2000.	117
7.4	Cluster size distribution, for clusters of urban cells.	118
7.5	Land use maps used as input to the cluster analysis.	119
7.6	Moving window based-structure comparisons to validate the global behaviour of the Estonia model.	121
8.1	Relationship between the logit value and working residuals.	125
8.2	Percentiles of the working residuals and number of cases per binned logit value.	125
8.3	Logit, standard error, 2.5 and 97.5 percentile of the working residuals.	127
8.4	Probability, probability at lower and upper prediction limit and the prediction range for the residential land use model.	128
9.1	The Environment Explorer simulates land use developments as a function of suitability, policy, accessibility, current land use and the influence of the neighbourhood.	132
9.2	Size distributions of urban clusters according to Zipf's Law.	133
9.3	Urbanisation probabilities in the Randstad conurbation in 2030 under the Trend scenario.	138
9.4	Probability of housing development in the restrictive areas.	139
9.5	Difference in urbanisation probabilities between the Trend and Sprawl scenarios	140
10.1	Comparison of the estimated regression coefficients for the neighbourhood variables with regression coefficients of the basic model.	152
10.2	Agreement between the simulated and observed land use maps.	153
10.3	Contribution of the neighbourhood of various land use categories in the logit model for the region of Amsterdam, the Netherlands.	157
10.4	Neighbourhood rules of the neighbourhood model showing the influence of various land uses on residential developments with increasing distance.	158
10.5	Neighbourhood rules of the neighbourhood model showing the influence of various land uses on residential developments with increasing distance.	159
10.6	Probability of residential developments according the neighbourhood model near Amsterdam, the Netherlands.	160
10.7	Variogram of Pearson residuals of a null model with only an intercept, the basic model and the neighbourhood model.	162
10.8	Spatial autocorrelation decreases with increasing distance from existing urban locations.	163
11.1	Definition of the steps involved in the development of a simulation model.	167

Tables

1.1	Overview of the main modelling approaches applied to simulate land use developments.	15
3.1	Results of the calibration and validation of the Environment Explorer.	43
3.2	Detailed specification of the Individual World scenario in the Environment Explorer.	44
4.1	Land use classification used in this analysis.	55
4.2	Enrichment factors of the land use pattern in 1989 for a 3x3 Moore neighbourhood.	58
4.3	Land use before conversion for locations converted into respectively industrial/ commercial area, residential area or recreational area.	61
4.4	Coefficients of a logistic regression model explaining the spatial distribution of new residential area by the enrichment of the neighbourhood.	64
5.1	Land use classification used in this analysis.	75
5.2	Description of spatial explicit allocation factors used in this study.	78
5.3	Binary logit estimates for land use patterns in 1989.	81
5.4	Binary logit estimates for the spatial pattern of arable land in 1989.	84
5.6	Binary logit estimates of land use changes between 1989 and 1996.	87
5.7	Residential area: estimates for the coefficient of a logit model of changes between 1989 and 1996 in different regions.	89
5.8	Industrial/commercial area: estimates for the coefficient of a logit model of changes between 1989 and 1996 in different regions.	89
5.9	Estimation results of logit models including neighbourhood characteristics.	91
7.1	Goodness of fit measures for the calibration and validation period for the model and the naïve predictor.	116
7.2	Contingency table summarising changes from 1989 to 1996 in the Netherlands and overall changes.	116
7.3	Contingency table summarising changes in Estonia and overall changes.	120
7.4	Contingency table summarising changes in Northern Italy and overall changes.	120
9.1	Results of the calibration and validation of the Environment Explorer based on the Fuzzy Kappa.	134
9.2	National targets per land use category for the period 2000-2030.	136
9.3	Definition of restrictive areas per land use.	137
9.4	Available, required and developed land in the concentration areas under the Trend scenario.	141
10.1	Description of land use categories.	147
10.2	Radius and lower and upper limit of the ring-shaped neighbourhoods.	147
10.3	Description of policy and suitability maps.	149

10.4	Number of cells and weights in the logistic regression for cells changing to residential area and cells not changing to residential area.	149
10.5	Neighbourhood variables included in the basic model.	151
10.6	Estimated regression coefficients and significance of the policy and suitability maps in the neighbourhood model.	155
10.7	Estimated regression coefficients and significance of initial land uses in the neighbourhood model.	155
10.8	Estimated coefficients and significance of the aggregated neighbourhood variables.	156
10.9	Size of small, median and large clusters and the median partial logit.	161

1 Introduction

1.1 Societal Relevance

Man has been altering his environment since prehistoric times. First, our predecessors used fire to flush out game. With the arrival of agriculture, forests made way for crops and pastures; this was followed by the industrial and demographic revolution and the subsequent growth of cities. The world's population has expanded to 6.7 billion and the effect of human activities on the land and the environment has increased enormously. The impact of urban sprawl on the environment has been acknowledged for several decades (Gottmann and Harper, 1967, Soule, 2006). Nowadays land use change is seen as a fundamental part of global environmental change and sustainability research (Turner et al., 2007). Land use and climate change are inseparably related through the global carbon cycle (Zaehle et al., 2007) and their impacts on the ecosystems (Schröter et al., 2005).

A wide range of land use models has been – and will continue to be – developed in order to study the changes in land use that will occur in the future. To explore the long-term dynamics of global change climate change models are integrated with global land use models, such as the IMAGE-model (Bouwman et al., 2006). To assess the impacts of the Common Agricultural Policy (CAP) and Bio Fuel Directive in the European Union, future land use developments have been simulated in the Eururalis project (Verburg et al., 2008). Most land use models, however, have been developed for the regional and metropolitan scale, with the aim to address planning problems or to evaluate future spatial developments(Pontius Jr et al., 2008).

The Netherlands is faced with it own diverse planning problems. With no fewer than 16.4 million people residing on some 41,500 km^2 of land that is at or below sea level, every square meter of land in used in the Netherlands. Moreover, approximately 92.9 million chickens, 11.3 million pigs and 3.8 million cows can also be found in this country (CBS, 2008). There are well-developed agricultural, industrial, transport and services sector. The economic heart of the Netherlands, the Randstad conurbation consisting of Amsterdam Rotterdam, Utrecht and The Hague, includes the largest harbour in the world and Amsterdam Airport Schiphol. All of these activities are located in space, are connected by infrastructure and have an impact on the environment. Economic growth and increasing prosperity lead to the growth of urban areas, infrastructure connections, roads and, in most cases, environmental impacts. Planners try to steer spatial developments through a wide range of interventions that either constrain certain developments (e.g., restrictive greenbelt policies) or favour them (e.g., designation of economic development zones or ecological corridors), thereby reducing conflicting interests and environmental impacts.

The first spatial policies plans in the Netherlands date from 1958 and these have been updated regularly since then. New plans have been drawn up for the creation of about 2,500 km^2 of forests and nature reserves by 2018 (LNV, 2002, MNP, 2002c). In addition, roughly 1,500 km^2 of land will be needed for new residential and industrial areas by 2030 (ABF, 2002).

According to current policy plans and trends, 10% of the land use will change over the next 30 years. How will these developments affect the future land use patterns in the Netherlands? How should we cope with potential conflicting spatial developments, the increasing needs for housing and residential developments while preserving our natural landscapes? Land use simulation provides a tool by which to depict possible future spatial developments, providing policymakers with crucial information to deal with the conflicting interests between urbanisation, nature conservation, industrialisation, water management and agriculture.

1.2 Scientific Relevance

Hand-in-hand with the development of computer technology, scientists started to build models that would simulate the processes taking place in the real world. The first 'large-scale urban models' or LSUMs as they were called, emerged in the 1960s as part of an effort to modernise planning and make the field more scientific. One should, however, bear in mind that these 'large-scale' models were small enough to run on a computer with 5K of random access memory (RAM). As a result, these models were not capable of simulating the urbanisation processes that were taking place in a realistic manner and of solving urban planning problems. At that time Lee (1973) wrote his famous 'Requiem for large-scale models', finishing off most of the remaining real world applications. From then on land use modelling became more or less '*a university matter*', as is reflected in the comprehensive overview given by Batty (1994).

Due to this attitude towards modelling, the number of real world applications with a high spatial resolution did not begin to increase until the mid-1990s. In 1994, Landis (1994) presented his Californian Urban Future (CUF) model, a metropolitan planning model designed to help planners, policymakers and citizens "*to create and compare alternative futures.*" In the second-generation CUF model introduced in 1998, the uncalibrated 'developer-driven' approach was substituted by a multi-nomial logit approach to urban land use change (Landis and Zhang, 1998a, Landis and Zhang, 1998b). Engelen et al. (1995) published their first application of a cellular automata model simulation of the effects of climate change on small islands; this subsequently formed the basis for an integrated dynamic regional modelling system (White and Engelen, 1997b, White et al., 1997). Clarke et al. (1997) applied a self-modifying cellular automaton simulating model (SLEUTH) to estimate the impact of urbanisation on the San Francisco Bay area. This model was calibrated on historical maps using observed data for the years 1940, 1954, 1962, 1974 and 1990. Veldkamp and Fresco (1996) developed CLUE-CR, a multi-scale land use model to simulate land use scenarios for Costa Rica. The linear regression model used in the early version of CLUE was subsequently replaced by binary logistic regression models used to estimate the various land use classes. Kitamura et al. (1997) and Morita et al. (1997) describe the development and application of a multinomial logit model to simulate land use developments in a case study area in Japan. By the end of the 1990s, researchers have re-focused on land use models and the number of real world applications has started to grow.

The first reviews of land use models, which started to appear by the end of the 1990s, generally discussed the wide variation in land use models or focused on specific characteristics of specific models. Agarwal et al. (2000) review nineteen land use change models using a classification scheme based on three components – space, time and human decision-making- and use this classification scheme to assess future directions for incorporating social drivers in land use models. In the Web Book of Regional Science Briassoulis (2000) reviews a wide range of land use change models, distinguishing four main categories: statistical and econometric models, spatial interaction models, optimisation models and integrated models. Each of these modelling traditions is discussed by Briassoulis following a set of features: underlying theory, purpose, aggregation (spatial, temporal, sectoral, land use), model specification, land use determinants, data used and real world applications.

Verburg et al. (2004c) discuss the current practices and research priorities relevant to land use change modelling and review the large variety of land use models currently in use on the basis of six important concepts of land use modelling: (1) level of analysis; (2) cross-scale dynamics; (3) driving forces; (4) spatial interaction and neighbourhood effects; (5) temporal dynamics; (6) level of integration. The main modelling approaches applied to simulate land use developments reported in this review are presented in Table 1.1. In all of these approaches, land use developments are allocated to a land use map given the characteristics of a location and its neighbourhood – i.e., the determinants of land use change.

Two different perspectives to modelling are distinguished in Table 1.1: the macro- and the micro-level perspective. At the macro-level perspective, macro-economic theory or a systems approach (Odum, 1983) is used to simulate changes in land use. Most of these models use statistical techniques, logistic regression or neural networks to relate the characteristics of a location and its neighbourhood to historical land use changes. Cellular automata models often rely on expert judgment to define the interaction between the land use at a certain location and the characteristics of the location and its neighbourhood.

At the micro-level, changes in land use are simulated based on the behaviour and the decisions of individual land owners. In micro-economic models these decisions are driven by the objective to maximise expected returns or utility from the land, using economic theory. Irwin and Geoghegan (2001) review the advances in the development of spatial land use change models focusing on the economic process associated with land use change. One of the drawbacks

Table 1.1 Overview of the main modelling approaches applied to simulate land use developments based on Verburg et al. (2004c).

Method	Examples
Macro-level perspective	
logistic regression	(Landis, 1994, Landis and Zhang, 1998a, Landis and Zhang, 1998b, Verburg et al., 1999, Waddell, 2000, Schneider and Pontius, 2001, Serneels and Lambin, 2001, Verburg et al., 2001, Overmars et al., 2003, Aspinall, 2004, Verburg et al., 2004b)
neural networks	(Li and Yeh, 2002, Pijanowski et al., 2002, Yeh and Li, 2003, Dai et al., 2005, Pijanowski et al., 2005, Nemmour and Chibani, 2006, Vafeidis et al., 2008)
cellular automata	(White and Engelen, 1993, Batty and Xie, 1994, White and Engelen, 1994, Clarke et al., 1997, White and Engelen, 1997a, Clarke and Gaydos, 1998, Wu, 1998, Li and Yeh, 2000, Wu and Webster, 2000, de Almeida et al., 2003)
Micro-level perspective	
micro-economic	(Bockstael, 1996, Chomitz and Gray, 1996, Geoghegan et al., 1997, Nelson and Hellerstein, 1997, Pfaff, 1999, Geoghegan et al., 2001, Irwin and Geoghegan, 2001, Nelson and Geoghegan, 2002)
multi-agent	(Ligtenberg et al., 2001, Otter et al., 2001, Parker et al., 2003, Deadman et al., 2004, Huigen, 2004, Ligtenberg et al., 2004, Brown et al., 2005, Manson, 2005, Huigen et al., 2006, Liu et al., 2006)

of these models is the assumption that the individuals are perfectly rational optimisers with free access to information, foresight and infinite analytical ability (Parker et al., 2003). Multi-agent models also focus on human actions and decisions but instead of having the goal of maximising returns, the agents make inductive, discrete and evolving choices to achieve their goals. Parker et al. (2003) present an overview of multi-agent land use models, combining a cellular landscape model with agent-based representations of decision-making versus alternative land use modelling techniques.

The development of geographical information systems, modelling software and ever increasing computer capacities brings the development of spatially detailed land use models within the reach of many potential users (Karssenberg et al., 2001). Current models are being applied for spatial explicit explorations and predictions. Decision makers are placing increasing more reliance on model results when there is a need to evaluate spatial policy. As such, these stake-holders should have some idea about the validity and uncertainty of the models. Moreover, specific requirements have been defined for the use of models in policy applications. According to the handbook of Good Modelling Practice (van Waveren et al., 1999), models applied in policy applications should be subjected to various analytical tests: calibration, validation, verification, global behaviour analysis, sensitivity analysis and robustness tests. Furthermore, since the public discussion in the press in the Netherlands (NRC, 1999) on the uncertainty of model predictions, a Guidance on Uncertainty Assessment and Communication (van der Sluijs et al., 2003) has been developed which describes how to cope with the uncertainties of model predictions within the framework of policy evaluation. In essence, this guideline requires that the uncertainty of model predictions be defined and described along with the conclusions on policy.

1.3 A framework for quality assurance

With the aim of establishing quality criteria for simulation models, Refsgaard and Henriksen (2004) propose a framework for quality assurance. They distinguish four steps in the development of a simulation model: Analysis, Programming, Model set-up and Simulation (Figure 1.1). The credibility of these steps is described in the outer circle: Model confirmation, Code verification, Model calibration and Model validation. The framework proposed by Refsgaard and Henriksen (2004) is applied in this thesis to determine the relevant scientific research for improving the quality of land use simulation models.

The primary focus of this thesis is not the Programming and Code verification steps but rather the improvement of:

- the conceptual model, including Analysis and Model confirmation;
- the model, including Model set-up and Calibration;
- the application[1] of the model, including Simulation and Validation.

The following sections describe these three stages in model development. The definitions of terms according to Refsgaard and Henriksen (2004) are presented at the beginning of each section, followed by a discussion of the research priorities.

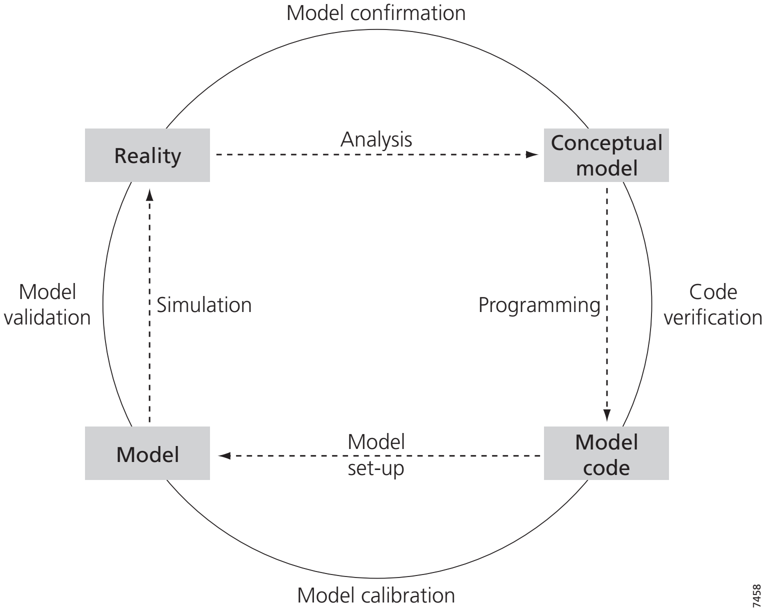


Figure 1.1 Definition and terminology of the four steps involved in the development of a simulation model and the necessary quality assurance as proposed by Refsgaard and Henriksen (2004).

1.3.1 Improving the Conceptual Model

The Conceptual Model is defined by Refsgaard and Henriksen (2004) as "a description of reality in terms of verbal descriptions, equations, governing relationships or 'natural laws' that purport to describe reality." The term analysis is not defined by Refsgaard and Henriksen (2004) but instead is taken to encompass various ways to extract the relevant processes in the (conceptual) model from reality. In terms of confirmation of the conceptual model, they state that it "should follow the standard procedures for confirmation of scientific theories. This implies that conceptual models should be confronted with actual field data and be subject to critical peer reviews. Furthermore, the feedback from the calibration and validation process may also serve as a means by which one or a number of alternative conceptual model(s) may be either confirmed or falsified."

In land use change research, land use data and observations are used to analyse the main determinants of land use change. Lambin et al. (2001a) describe the wide range of determinants used in the land use models from various disciplinary perspectives. In practice, all processes influencing land use change interact, resulting in complex spatio-temporal relations depending on local cultural, socio-economic and biophysical context (Verburg et al., 2004c). Given this geographical variation in the determinants of land use change, detailed empirical analysis of these determinants are necessary in order to provide a land use model with a solid foundation. Specific attention should be paid to the influence of the neighbourhood factors. As stated by Verburg et al. (2004c) "the theoretical basis of the quantification of the neighbourhood functions for the cellular automata is poor. Quantification of this type of relations is now mostly based upon expert knowledge and not on observations. It is recommended that a more sophisticated and reproducible way is developed to define these neighbourhood effects."

1.3.2 Improving the Model

A 'Model' is defined by Refsgaard and Henriksen (2004) as: "A site-specific model established for a particular study area, including input data and parameter values." In the Model set-up, the necessary site-specific information is added to the model code. Model calibration is defined as: "The procedure of adjustment of parameter values of a model to reproduce the response of reality within the range of accuracy specified in the performance criteria." The performance criteria are defined as: "Level of acceptable agreement between model and reality. These performance criteria apply both for model calibration and model validation."

Calibration plays an important role in the selection and quantification of the determinants of land use change. For land use models based on stepwise logistic regression or neural networks, the calibration takes place while fitting the model on the data. These statistical models are limited in terms of the number of neighbourhood variables that can be included. Most neighbourhood variables that express the influence of one land use on another land use at various distances are highly correlated, making the regression coefficients highly variable and poorly determined. If the influence of the neighbourhood is important in determining the locations of land use change, new methodologies will be needed to cope with the multicollinearity of the neighbourhood variables in the statistical methods.

Standard calibration routines are applied in cellular automata-based land use models. Several publications address the calibration of these land use models and present techniques aimed at quantification; one of these is the neighbourhood effects (Clarke and Gaydos, 1998, Wu and Webster, 1998, Li and Yeh, 2001, Al-Kheder et al., 2007, Dietzel and Clarke, 2007). However, the calibration and validation of land use models are hindered by the lack of performance measures. These performance measures should measure the agreement between the model and reality by comparing the simulated and observed land use maps. However, more often than not, the validity of a land use model is assessed by a visual inspection of the simulated land use map, irregardless of whether this is a realistic approach.

First attempts to compare maps were on a cell-by-cell approach. Monserud and Leemans (1992) started to use the Kappa statistic to compare global vegetation maps. The Kappa statistic is a more robust measure than the simple percentage agreement calculation since it accounts for the agreement occurring by chance. This method was subsequently improved by making a distinction between quantification and location errors (Pontius, 2000). Pontius and Schneider (2001) validated their land use model using the 'Receiver Operating Characteristic' (ROC) method. Power et al. (2001) and Hagen (2003) were the initiators of using Fuzzy set theory to compare land use maps. It is therefore evident that good performance measures should be defined and developed in order to be able to compare the results of the various land use models.

1.3.3 Improving the Application

The framework proposed by Refsgaard and Henrikson does not define the final product of the simulation nor the object of the validation. As it certainly is not 'Reality', the term Application has been suggested as a replacement. Refsgaard and Henriksen (2004) define 'Simulation' as the "Use of a validated model to gain insight into reality and obtain predictions that can be used by managers" and 'Model validation' as "Substantiation that a model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model." Assessments of the model predictions are very important in the context of this uncertainty.

According to these definitions, a validated model should be used in those applications for which the uncertainty in the model predictions has been assessed. The portability of the model from one situation or site to another situation or site should be valid. However, this is not the case for most applications of land use models. As already stated, land use models are still limited by the definition of good performance measures to assess the accuracy of the calibration and validation. Once performance measures are available, the model can be validated using data on historical developments.

However, most land use simulation models are applied in long-term scenario analysis to facilitate the answering of questions such as: what spatial developments will take place if the economy grows with 4% a year, the population starts to shrink and oil prices rise by 300%? The basic question implicit in these applications is: how to simulate the future land use developments given the available input information and scenario elements? Validation of these applications are – according to the definition of Refsgaard and Henriksen (2004) – not possible as no data will be available reflecting these societal changes.

The results of the model applications strongly depend on the accuracy of the model predictions. For example, any misclassification in the observed land use maps has the potential to seriously impact on simulated land use. Moreover, it is also difficult to draw policy-relevant conclusions when the model predictions are highly uncertain. Given the intrinsic uncertainty in model predictions, research can be directed towards improving the model on these points. Methods to address uncertainty in land use change models are starting to appear in the literature (Eckhardt et al., 2003, Ward et al., 2003, Pontius and Spencer, 2005).

The uncertainties can be derived from the model validation using data on historical developments. For most simulation models, the uncertainty can be determined using Monte Carlo simulations in which the input parameters are varied according to their statistical distribution. For the case of a logistic regression model, it should be possible to use the standard error and residual variance to estimate the uncertainty in the model predictions. For scenario applications, it may be possible to extend the uncertainties derived from the historical model validation to the predictions of the future land use maps.

The major challenge, however, remains the development of quality assurance tests – i.e., checking the credibility of future land use patterns using the characteristics of the spatial distribution of land use categories in the land use map. Such characteristics include fractal dimensions (Batty and Longley, 1994, White et al., 1997), Zipf's Law (Zipf, 1949) or other power law distributions.

1.4 Objectives and approach

The overall objective of this thesis is: *to improve the quality of land use simulation models and the prediction of future land use patterns.*

The main land use model selected to simulate future land use developments in this thesis is the Environment Explorer (Engelen et al., 2003). This model was developed at the National Institute of Public Health and the Environment (RIVM) in the Netherlands in collaboration with the Research Institute of Knowledge Systems (RIKS) to evaluate the impacts of future spatial developments. The Environment Explorer consists of a geographically detailed, dynamic land use model developed by White and Engelen (1997b). This model was adapted to the Dutch situation and extended with a regional and national module to account for developments on both the national and regional scale and with spatially detailed indicators describing the potential impacts on nature and landscape (de Nijs et al., 2001b).

In this thesis the quality of land use simulation models and the prediction of future land use patterns have been enhanced according the aforementioned stages in model development as proposed by Refsgaard and Henrikson.

1.4.1 Improving the Conceptual Model

The 'conceptual model' has been improved through analysis of changes in land use maps. The questions addressed are:

1. How to analyse the neighbourhood characteristics of land use patterns?

The interactions between the various types of land use in the neighbourhood forms one of the main driving factors in the Environment Explorer. In this thesis a method for the analysis of neighbourhood interactions is introduced given the observed land use changes. The method describes the neighbourhood characteristic of a land use map, revealing the features of those locations more prone to changes in land use and helps improve our knowledge of neighbourhood interactions. Moreover, the method may be used to quantify neighbourhood interactions in land use models in general and in the Environment Explorer specifically.

2. What are the determinants of land use change in the Netherlands?

A large number of determinants, in addition to neighbourhood characteristics, may affect land use change. In this thesis, an empirical study is presented that reveals the main determinants of land use change for a wide range of land use types. The study is not limited to single disciplinary approach but, rather, analyses the impact of different determinants derived from various land use allocation theories. The study also addresses the historical land use developments that have shaped the current land use pattern as well as the recent land use changes that will shape future land use developments. The main objective of the study is to increase our understanding of the land use change processes and to present a basis for the specification of land use change models.

3. What influence has the neighbourhood on land use developments? Is it possible to build a hybrid model incorporating neighbourhood interactions from a cellular automata-based model into a logistic regression model?

Neighbourhood interactions, which form the essence of cellular automata-based models, are being increasingly used in statistical and spatial econometric land use models. In most cases, simple variables are included in the set of explanatory variables. Once the impact of the neighbourhood on the land use developments are defined, it should be possible to build a hybrid model integrating the neighbourhood interactions from a cellular automata-based model into a logistic regression-based land use model. With the aim of including all neighbourhood interactions in a logistic regression model, a new regression methodology has been developed in this thesis that resolves the issues of spatial autocorrelation and multicollinearity in the neighbourhood variables. It will be shown that the influence of the neighbourhood on residential developments in the Netherlands depends on the land use types, the distance and the cluster size of the land use categories.

1.4.2 Improving the Model

The model has been improved through the development of calibration tools and performance measures. The questions addressed in this stage are:

4. How to compare maps for the calibration and validation of spatial simulation models?
Map comparison techniques are important in many explicitly spatial studies. Specifically, the calibration and validation of land use models require the application of map comparison techniques. How well do these models predict historical land use developments? Objective measures for map (dis)similarity are needed in order to calibrate and validate land use change models such as the Environment Explorer. In the Map Comparison Kit (MCK) various ordinal and nominal map comparison techniques have been combined in one software package, including recent developments in Fuzzy set map comparison.

5. How to calibrate integrated land use models featuring cellular automata?
The Environment Explorer is an integrated land use model consisting of a spatial interaction (gravity) model at the regional level and a cellular automata-based land use model. The regional model defines the amount of land use change in the land use model. The land use model, in turn, provides the regional model with information on the mean suitability and availability of land. Both models include a large set of highly uncertain parameters. A major challenge in the application of the model is the calibration and validation. This paper describes some steps towards a fully automated calibration of the Environment Explorer.

1.4.3 Improving the Application

The model applications have been improved by the development of a technique to define the model uncertainty and simulating future land use developments given the available input information and scenario elements. Questions addressed are:

6. How to define the spatial uncertainty of models of land use change?
During the last decade, many land use models have been developed to assess future land use developments. Decision makers who rely on these models should have concept of the scale of uncertainty of these models. Moreover, given the uncertainty in model predictions in general, research can be directed towards improving the model on these points. Under most conditions, uncertainty in the predictions of logistic regression models can be derived from the standard error of the model and the residual variance. However, in our case the uncertainty in a model of residential development in the Netherlands turned out to be very large due to the large variance in the working residuals. An alternative, novel methodology is proposed here that employs the relationship between the predicted logit values and working residuals.

7. How to create spatially detailed land use maps of the Netherlands in 2030 given the scenario elements, story lines and data on national, regional and local land use developments.
This application of the Environment Explorer was applied in the National Nature Outlook 2 (RIVM, 2002) to evaluate the effects of alternative socio-economic developments on nature

and the landscape. In Outlook 2, four scenarios were translated into spatially detailed land use maps of the Netherlands in 2030. The study describes the construction of these maps using one scenario as an example. Hereto, national economic and demographic developments have been extended with assumptions on the spatial behaviour of the different actors, activities and land use functions.

8. How will the spatial strategy affect future land use patterns in the Netherlands?
In the Evaluation of the National Spatial Strategy (RIVM, 2004), the Environment Explorer was applied to evaluate the effects of the future spatial developments, addressing specific questions such as: What effect will the spatial strategy have on the national landscapes? Will the restricted areas remain restricted? Is there enough space in the urban concentration areas to accommodate all spatial developments?

In this application the parameter simulating the stochastic perturbation was calibrated to best fit the distribution cluster sizes according Zipf's Law. This was followed by a Monte Carlo simulation to define the probability of urbanisation in the future land use maps.

1.5 Thesis outline

The papers in this thesis are presented in chronological order, thereby reflecting an increase in the knowledge and ambition of the author over time. All papers have been published in an international scientific journal or reader, with the exception of chapter 10, which has been accepted for publication. The papers are, in a slightly modified form, included in this thesis as separate chapters:

Chapter 2 is titled 'The Environment Explorer: spatial support system for integrated assessment of socio-economic and environmental policies in the Netherlands' and is published in its original form in *Integrated Assessment* (Engelen et al., 2003);

Chapter 3 is titled 'Constructing land use maps of the Netherlands in 2030' and is published in its original form in the *Journal of Environmental Management* (de Nijs et al., 2004);

Chapter 4 is titled 'A method to analyse neighbourhood characteristics of land use patterns' and is published in its original form in *Computers, Environment and Urban Systems* (Verburg et al., 2004a);

Chapter 5 is titled 'Determinants of land use change patterns in the Netherlands' and is published in its original form in *Environment and Planning B Planning and Design* (Verburg et al., 2004d)

Chapter 6 is titled 'The Map Comparison Kit' and is published in its original form in *Environmental Modelling and Software* (Visser and De Nijs, 2006). This paper received an award from the Editorial Board at the Fourth Biennial meeting of the International Environmental Modelling and Software Society held in Barcelona 2007 for its quality and relevance;

Chapter 7 is titled 'Measuring performance of land use models: an evaluation framework for the calibration and validation of integrated land use models featuring cellular automata' and has been presented at 14th European Colloquium on Theoretical and Quantitative Geography, Tomar, Portugal (Hagen-Zanker et al., 2005b);

Chapter 8 is titled 'Spatial uncertainty in land use models. An alternative method to estimate uncertainty in logistic regression methods' and has been presented at 15th European

Colloquium on Theoretical and Quantitative Geography. Montreux, Switzerland (de Nijs and Pebesma, 2007);

Chapter 9 is titled 'Future land use in the Netherlands: evaluation of the National Spatial Strategy' and is published in its original form in *Planning Support Systems: Best Practice and New Methods*, edited by Stan Geertman and John Stillwell (de Nijs, 2009);

Chapter 10 is titled 'Estimating the influence of the neighbourhood in the development of residential areas in the Netherlands' and has been accepted for publication in *Environment and Planning B Planning and Design* (de Nijs and Pebesma, accepted)

Chapter 2 gives a general introduction to the objective, structure and application of the Environment Explorer. The structure of this thesis in relation to the stages in model developments is given in Figure 1.2.The work presented in chapters 4, 5 and 10 was undertaken to improve the Conceptual Model of the Environment Explorer. Chapters 6 and 7 focus on the improvement of the Model and describe methods and tools that have been developed due to the need to compare maps in order to calibrate and validate the land use model. Chapters 3, 8 and 9 focus on the improvement of the application of land use models. An alternative methodology for quantifying the spatial uncertainty in land use models is presented in chapter 8. Chapters 3 and 9 present two applications of the Environment Explorer. The synthesis and main conclusions of this thesis are presented in chapter 11.

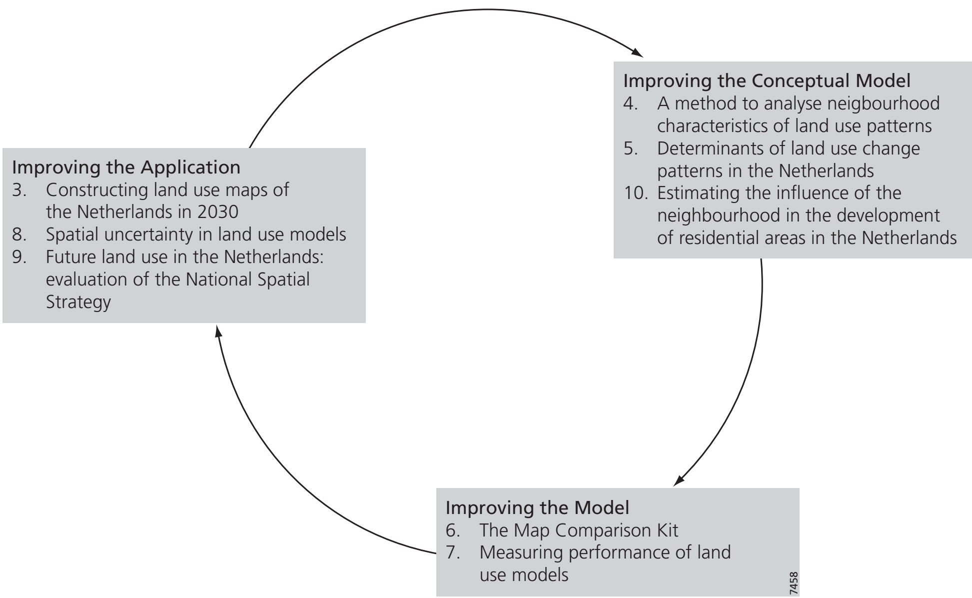


Figure 1.2 Structure of this thesis in relation to the stages in model development according Refsgaard and Henriksen (2004).

Note

1. The Refsgaard & Henriksen framework does not define the final product of the simulation nor the object of the validation. As it certainly is not Reality, it is suggested here that the term Application be used

2 The Environment Explorer: spatial support system for integrated assessment of socio-economic and environmental policies in the Netherlands

Reprinted from *Integrated Assessment*, 4. Engelen, G., White, R. & de Nijs, T. C. M. (2003) The Environment Explorer: Spatial Support System for Integrated Assessment of Socio-economic and Environmental Policies in the Netherlands. Pages 97-105.

Abstract

Environment Explorer is a system developed to support spatial scientists, planners and decision makers at the regional and national levels in the Netherlands, to help them analyse a wide range of social, economic and environmental policies and their associated temporal and spatial dynamics. The core of this system consists of linked dynamic spatial models operating at both the macro- and the micro-geographical scales. At the macro scale, the modelling framework integrates several component submodels, representing the natural, social and economic sub-systems. At the micro level, cellular automata based models determine the fate of individual parcels of land, based on institutional, physical and environmental factors as well as on the type of activities in their immediate neighbourhoods. The approach chosen enables the straightforward integration of detailed physical, environmental and institutional characteristics as well as the particulars of the transportation infrastructure and permits a very detailed representation of the evolving spatial system. As part of the policy support system, the models are supplemented with dedicated tools for interactive design, analysis and evaluation of the policy interventions and scenarios to be tried out. The system covers the entire territory of the Netherlands and represents processes at the national, the regional (40 economic regions) and the cellular (25 ha cells) levels. It runs on top of detailed GIS information and generates future land use and land cover for the period 2000 till 2030. The quality of the policies tried out is expressed in some 40 economic, social and environmental indicators available in the model as dynamic maps. The application has been developed over the past 5 years. It has been used at the national and the provincial level for the preparation of spatial policy documents. Some conclusions relative to the development and the use of the system are presented.

2.1 If only we knew…

With a population of some 16 million living on 41,500 km^2 of land, the Netherlands is a small, but densely populated country. It is primarily a large Delta shaped by the Rhine, Meuse and Scheldt rivers. No less than 5 million people live on some 8,500 km^2 of land that is at or below sea level. Certainly, they live behind dikes, protected from the sea and the rivers, nonetheless; they live in a country that is troubled by on the one hand subsiding land due to tectonic movements and continued oxidation of the peat substrate and on the other hand increasing amounts of water running down the main rivers. At the same time, the Netherlands is one of the Western European countries that still witnesses a growing population and consequently is faced with an ever-increasing demand for space to house its population and their economic activities. In some scenarios a growth of over 1.5 million people in the next 20 years is expected.

In the given circumstances strict policies and control on how the limited available land is used seconded by an alert and elaborate monitoring system and a set of state of the art instruments for exploring strategies and policies developed to tackle threats detected, are the best available options to keep the country afloat. With the increasing amount of stress exerted on the system and the complexities of the problems posed, the needs for appropriate planning instruments change rapidly too. In particular the need for instruments supporting a truly integrated approach to spatial planning problems is very urgent. Four aspects of planning problems and the systems in which they are set are of particular importance. First and most importantly, these systems are *integrated wholes*. Thus, while a planner or policymaker may intervene directly in only a limited part of the system, linkages will transmit consequences of the policy to many other parts of the system. Conversely, the problems the planner is dealing with may have had their origins in actions that were taken in other parts of the system in an attempt to resolve other problems. Second, human systems and the natural systems in which they are imbedded, are *dynamic and evolving*; they are never in equilibrium. Policymakers thus intervene in a changing system and at certain critical points; the consequences of even a small intervention may be of major importance yet may be entirely unanticipated. Third, these systems are *inherently spatial*. In space, human and natural processes occur in more or less defined clusters of high and low concentration and typified by periods of high and low activity. The consequences of planning policies depend on the spatial context within which they are implemented, as well as on the way they alter that context. Fourth, the world is one of *uncertainty* and while increased knowledge and improved modelling tools may lessen that uncertainty, they cannot eliminate it. Policies therefore need to be designed to incorporate uncertainty, rather than assume that it does not exist.

This paper presents the Environment Explorer an integrated spatial planning instrument developed to accommodate the four characteristics of planning problems just described and built to design, analyse and evaluate long-term policies relative to the physical environment in the Netherlands in an economic, social and ecological context. The development of Environment Explorer started in 1997 (de Nijs et al., 2001b). Since then, it has evolved into a powerful Policy Support System for integrated spatial planning, supported by the National Institute for Public health and the Environment (RIVM), as well as the National Institute for Marine and Coastal Management (RIKZ), the National Institute for Inland Water Management and Waste Water Treatment (RIZA) and the Transport Research Centre (AVV) of the Ministry of Transport, Public Works and Water Management. It has become a highly interactive, transparent instrument currently used at the National and the Provincial levels. For example the province of Utrecht used Environment Explorer in its search for new residential and industrial locations as part of its new provincial master plan (van Delden et al., 2003).

2.2 An integrated spatial model of the Netherlands

2.2.1 Autonomous developments, intended and actual policies

The primary goal of Environment Explorer is to explore the effects of (alternative) policy options on the quality of the socio-economic and physical environment and, with this information at hand, to stimulate and facilitate awareness building, learning and discussion prior to the decision-making proper. To this end, the system combines autonomous developments with policy-induced changes to form integral pictures of possible futures for the Netherlands and evaluates their relative value on the basis of social, economic and ecological criteria. It does not seek to optimise the separate economic, ecological and social dimensions, rather to maximise the whole. Although this means losing some detail, the benefit of the approach is the strong integrative and interactive nature of the resulting system, in which highly dynamic, autonomous processes play a key role.

The motor driving the spatial changes in Environment Explorer is fuelled by economic and demographic developments: supply and demand in both qualitative and quantitative terms. These processes operate at different spatial scales and are thus represented in the models. In

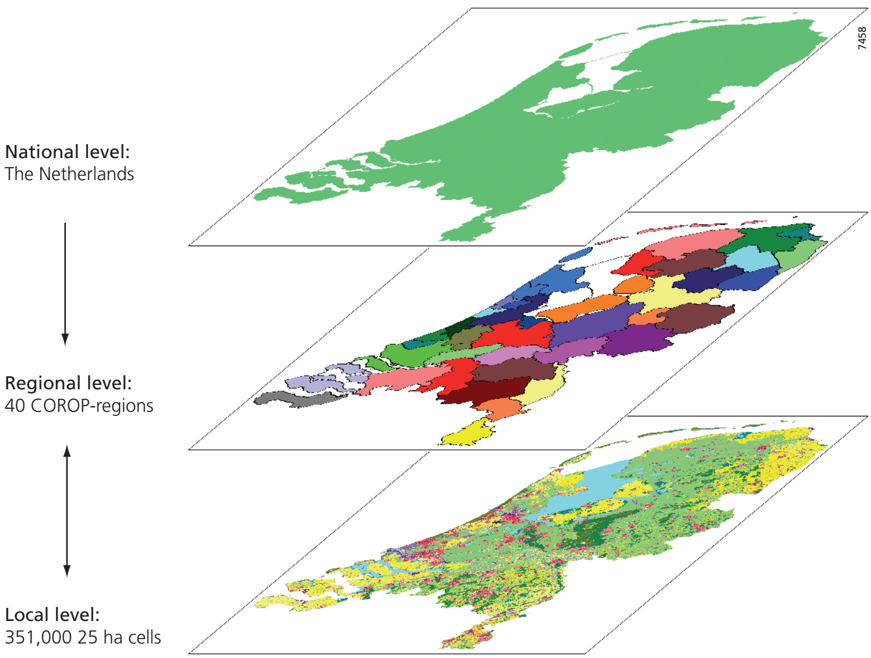


Figure 2.1 The Environment Explorer model represents processes at three spatial levels: national, regional and local.

fact, these dynamics are very much represented as the competitive reality of 'survival of the fittest': it is the function that is most powerful and that generates the highest added value per unit of area that will be most successful in claiming parcels of land. Similarly, it is the region that offers the most attractive alternative for economic and residential activity that will attract most businesses and residents. Government, represented by the analyst using the system, has the task of safeguarding collective interests, including protection of the economically weak, the social values, the open space and the natural assets in general. The actual and the intended policy actions to counter the initiatives of the 'free market' players can be entered into the Environment Explorer system by means of zoning maps and control parameters acting as constraints upon the autonomous dynamics of the system.

2.2.2 Models coupled at 3 geographical levels

The core component of Environment Explorer is an explicitly dynamic land use-transportation model applied to the full territory of the Netherlands. In order to represent the processes that make and change the spatial configuration of the country, it features a layered model representing processes operating at three geographical levels: the national (1 region), the regional (40 economic regions) and the local (351000 cellular units of 25 ha each) (Figure 2.1) (For a detailed discussion of the model, including its mathematical description, the user is referred to (de Nijs et al., 2001b).

At the national level, the scale of the entire country, the model integrates national figures taken from economic, demographic and environmental growth scenarios considering developments in the Netherlands in the context of Europe and the world beyond and prepared by the Dutch planning agencies or IPCC (de Nijs et al., 2004). From these, growth figures for the national population, the activity per economic sector and the expansion of particular natural land uses are derived and entered in the model as trend lines.

The economic activities are condensed into eight main sectors: crop farming, dairy farming, greenhouse farming, other farming, industry, services, socio-cultural activities and recreation. The population is assigned to two residential categories: high and low-density residential. The natural land use categories are: wetlands, forests and extensive grasslands. The choice of the categories is based on their distinct spatial requirements and specific spatial behaviour, the quality, match and availability of data at the three geographical levels of the model and lasts but not least the end-use requirements of the model: the particulars of spatial policies and spatial planning criteria.

The national growth figures entail changes in land use through changes in economic output and land required for carrying out the activities, through changes in land required for housing and through the expansion of natural land uses. But they do not say where land use changes will occur, or whether the changes are possible given constraints on the amount and quality of land available. These locational aspects are modelled at the regional level and at the local level. In the first instance, the national growth figures are a constraint for the models at the regional level.

At the regional level consisting of 40 large economic regions (called COROP regions), a dynamic spatial interaction based model (Wilson, 1974, White, 1977, White, 1978, Allen and Sanglier, 1979b, Allen and Sanglier, 1979a) caters for the fact that the national growth will not evenly spread over the country, rather that regional inequalities will influence the location and relocation of new residents and new economic activity and thus drive regional development. The model arranges the allocation of national growth as well as the inter-regional migration of activities and residents based on the relative attractiveness of the regions. One of three principles is applied:

1. For the allocation and relocation of people, industrial and socio- cultural activities, a standard potential based model is applied: each region competes with all the other regions for new residents and new activities in each of the eight economic sectors on the basis of its geographical position relative to the other regions, its employment level, the size of its population, the type and quantity of activity already present and its location relative to the public and the private transportation systems. In addition to these and novel in the context of interaction based models, summarised cellular measures obtained from the model at the local level characterising the space within the regions are factors determining the relative regional attractiveness. The latter are: the abundance of good quality land, the zoning status of that land and its accessibility relative to the waterways, roads and public transportation.
2. For other activities, such as services, for which economic considerations are more important, a relative profit criterion is applied. In this case, the relative costs of producing and shipping the goods as well as the buying power of the customers is an additional criterion in the determination of the attractiveness of a region.
3. Finally, for the natural and agricultural categories, only the abundance and the quality of the land for sustaining the particular ecosystem or activity are criteria in the determination of the attractiveness.

In the dynamic context of the model, the attractiveness of regions evolves with the changes in each component on which it is based. Conversely, changes in regional attractiveness result in changing numbers of people and levels of activity as well as altered pressure on the transportation infrastructure. Four submodels can be distinguished:

1. A regional economic module calculates the amount of production and employment for each of the economic sectors, its allocation and reallocation among the regions.
2. A regional demographic module deals with the growth of the regional population: its allocation and reallocation among the regions and the demand for housing.
3. A transportation module deals with the changes in the characteristics of the transportation infrastructure, the flows of people and goods travelling over it, the congestion of the networks and its consequences on interregional distances and accessibility. The interregional distances are expressed in generalised transportation costs calculated on the basis of the costs per kilometre travelled, the costs per hour travelled and the parking costs. The transportation module is in fact a four stage transportation model fully linked to the other modules at the regional, but also at the local level and solved at each simulation time step. The transportation system consists of: the railways and railway stations, the navigable waterways and the road network (the LMS-road network used by the Ministry of Transport, Public Works and Water Management) consisting of the motorways, main national and regional roads. People choose to travel via the road or railway system, while goods are assumed to be transported over the roads only. The geographical layout of the networks, as well as the quality and capacity of the links determine the flows of goods and people. On the basis of different motives for travelling, the production of traffic is calculated. The total volume of traffic between regions is split over the public and private transportation system in function of the generalised costs. Next, the volumes are allocated to the shortest routes linking the centres of the regions and the links constituting them. Thus the congestion can be calculated as well as its impact on the interregional distances. These affect the long-term migration of

activities and residents as well as the shortterm movement of goods and people between the different regions.

4. A land claim module translates the regional growth numbers into spatial claims. The latter are passed on to the model at the local (cellular) level for a detailed allocation. Two principles are applied at this level: (1) a claim for land is fixed and passed on as a hard constraint. This principle reflects the fact that for particular activities policies determine the amount and location of land that is to be created or to be preserved in a region. It applies mostly for the natural land categories and recreational land. Or, (2) the principle of supply and demand is applied to regulate the densification of land use as well as its spatial allocation. This principle applies in particular to housing and most of the economic activities.

At the local level the detailed allocation of economic activities and people is modelled by means of a cellular automata based land use model (Couclelis, 1985, White and Engelen, 1993, Batty and Xie, 1994, Engelen et al., 1995, White and Engelen, 1997a). To that effect, the Netherlands is represented as a mosaic of 351000 grid cells of 25 ha each (500 m on the side). Each cell is modelled dynamically and together the cells constitute the changing land use pattern of the country. Land use is classified in 17 categories, 10 of which are so called land use 'functions' and modelled dynamically. The land use function categories are chosen with a view to guarantee a one-to-one relation with the economic and residential categories at the regional level. In principle, it is the relative attractiveness of a cell as viewed by a particular spatial agent, as well as the local constraints and opportunities that cause cells to change from one type of land use to another. This model is driven by the demands for land per region generated at the regional level. In fact there are 40 identical cellular models running in parallel: one for each COROP region. Four elements determine whether a piece of land (each 25 ha cell) is taken in by a particular land use function or not (Figure 2.2):

1. The physical suitability. Suitability is represented in the model by one map per land use function modelled. The term suitability is used here to describe the degree to which a cell is fit to support a particular land use function and the associated economic or residential activity for a particular activity (Wright, 1990)). It is a composite measure, prepared in a Geographical Information System (GIS), on the basis of some 15 factor maps determining the physical, ecological and environmental appropriateness of cells. Factors used are among others: elevation, soil quality and stability, agricultural capacity, air quality, noise pollution, etc.
2. The zoning or institutional suitability. Zoning too is characterised by one map per land use function. It is another composite measure based on master plans and planning documents available from the national or regional planning authorities including among others ecologically valuable and protected areas, protected culturescapes, buffer areas, etc. For three planning periods, to be determined by the user (example: 2000-2005, 2005-2015 and 2015-2030), the map specifies which cells can and cannot be taken in by the particular land use. For the analysis of policy and planning alternatives, it is of paramount importance that suitability and zoning can be handled separately. Zoning is a man made instrument for imposing constraints or stimulating particular trends, while suitability is most often a fact of life and an intrinsic quality of the area. Changing suitability requires usually an engineering action in the physical environment, such as altering slopes, filling in land, building infrastructures, etc., while changing zoning requires first of all an intervention in the legal and institutional environment.

3. The accessibility. The accessibility for each land use function is calculated in the model relative to the transportation system consisting of the railways and railway stations, the navigable waterways and the road network. It is an expression of the ease with which an activity can fulfil its needs for transportation and mobility in a particular cell. It accounts for the distance of the cell to the nearest link or node on each of the infrastructure elements, the importance and quality of that link or node and the needs for transportation of the particular activity or land use function.
4. Dynamics at the local level. While the above three elements are introduced in the model to determine the non-homogeneous nature of the physical space within which the land use dynamics unfold, there is a fourth and important aspect, namely the dynamic impact of land uses in the area immediately surrounding a location. This is no longer the domain of abstract planning, but rather that of the reality on the ground representing the fact that the presence of complementary or competing activities and desirable or repellent land uses is of great significance for the quality of that location and thus for its appeal to particular activities. For each location, each cell that is, the model assesses the quality of its neighbourhood: a circular area with a radius of 4 km containing the 196 nearest cells. For each land use function, a set

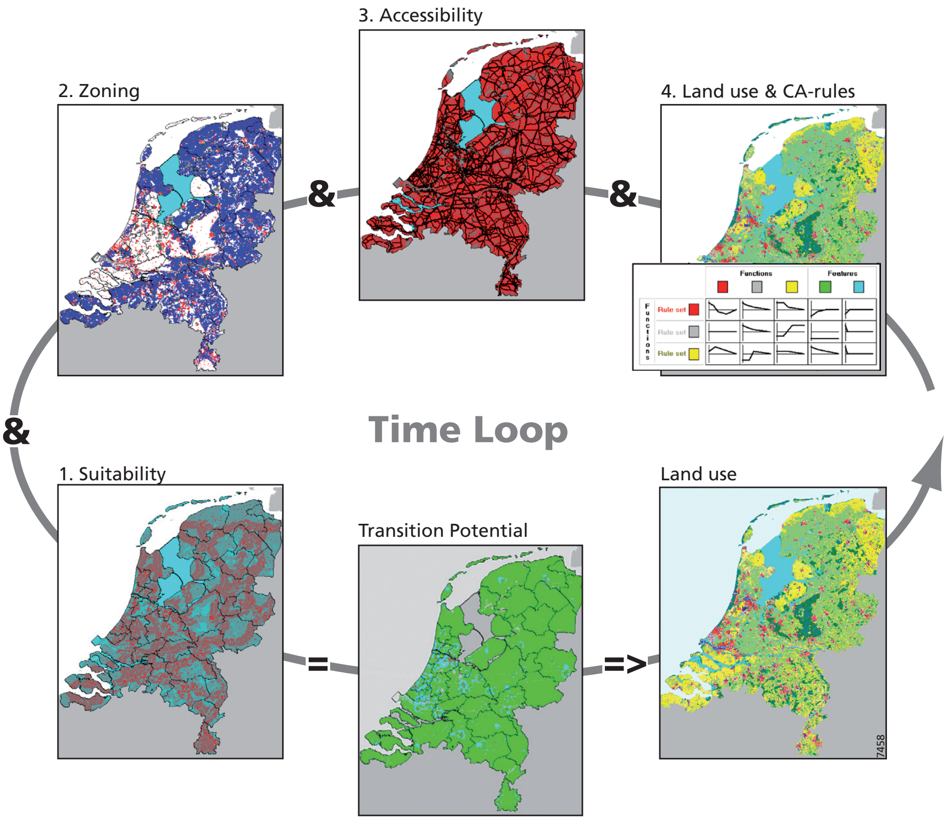


Figure 2.2 Four elements determine the dynamics at the local level.

of rules determines the degree to which it is attracted to, or repelled by, the other functions present in the neighbourhood. The strength of the interactions as a function of the distance separating the different functions within the neighbourhood is articulated in the rules. If the attractiveness is high enough, the function will try to occupy the location, if not, it will look for more attractive places. New activities and land uses invading a neighbourhood over time will thus change its attractiveness for activities already present and others searching for space. This process explains the decay of a residential neighbourhood due to the invasion by industrial or commercial activities, as well as the revival of decayed neighbourhoods initiated by the arrival of few high quality functions like parks, exclusive office buildings, high-end condominiums, etc. The rules determining the interactions between the different functions, the inertia, the push and pull forces and economies of scale, are defined as part of the calibration of this cellular automata model.

On the basis of these four elements, the model calculates for every simulation step the transition potential for each cell and function. In the course of time and until regional demands are satisfied, cells will change to the land use function for which they have the highest potential. Consequently, the transition potentials reflect the pressures exerted on the land and thus constitute important information for those responsible for the design of sound spatial planning policies.

The linkage between the models at the national, regional and local levels is bi-directional and very intense: the national growth figures are imposed as constraints on the regional models, the regional models distribute and allocate the national growth to the 40 regions and impose the regional growth numbers on the cellular models. Finally, the cellular models determine at the highest level of detail where the growth is likely to take place. In this process, the cellular models return to the regional models information on the quality and the availability of space for further expansion of each type of economic or residential activity. This information is an input into the spatial interaction calculations at the regional level and it influences strongly the relative attractiveness of the individual regions. As regions in the course of time are gradually running out of space for one or the other activity, they will lose part of their competitive edge and will exert less attraction. Growth is consequently diverted to other, more attractive regions.

2.3 Using the Environment Explorer for integrated planning

2.3.1 State variables and Indicators

From the model description, it is clear that Environment Explorer generates output at the regional and the local levels. Typically the model is run for the 30-year period between 2000 and 2030, but other time intervals are possible too. Results are calculated and visualised on a yearly basis. At the regional level, the population, as well as the employment and production figures in each economic sector are calculated. At the local level the resulting new land use map is generated and presented for every simulated year. In addition to these and based on the regional and local state variables, the model calculates some 40 spatial indicators expressing changes in the economic, social, or environmental status of the country, the regional and cellular entities. Together they constitute important information relative to the merits of one or the other project, policy or strategy tried out with Environment Explorer. Each indicator is in itself a more or less elaborate sub-model that may require specific additional information. Indicators include among others: access to employment (economic), cost of land (economic), built-up area (social), open space (social), recreational space per inhabitant (social), noise pollution and emissions due to traffic (environmental), congestion on the road system (economic), flooding risk (social), residential density (social), spatial fragmentation (environmental, Figure 2.3), etc. Like the other state variables, indicators are calculated on a yearly basis and are available in Environment Explorer in the form of dynamic maps, time charts and numeric output.

2.3.2 Scenarios, Strategies and Projects: interventions in the spatial structure

Environment Explorer has been designed for use as an analytical tool: it offers analytic capabilities to policymakers in government departments ranging from the municipal to the national levels. But, it has an important role as a tool for communication too: it stimulates collaboration, discussion and consultation among the different planning institutions and departments. That is why the instrument has been equipped with a state of the art graphical user interface (Figure 2.4), providing access to all the variables, parameters and maps used at every level of detail. It offers the users the ability to create and enter policy variants interactively by defining and adjusting values within a preset range of values and context. Over and above economic and demographic scenarios, which can be stated by means of dedicated dialogues, tables and graphs, the model can take into account spatial scenarios, such as those proposed by various government departments and entered in the form of alternative sets of numbers and maps. The suitability and zoning maps as well as the transportation networks are available in the model with the appropriate editors enabling interactive changes by the user so that infrastructure projects or particular policy decisions on opening or closing areas for development can be tried out experimentally. Thus the visions and strategies as options for future spatial, economic and social policies, proposed by the individual governmental institutions are pooled and presented as detailed spatial blueprints of the Netherlands and translated into relevant societal terms. Their impact and effectiveness in the short and long term future of the country is calculated, visualised and available for further analysis and discussion.

In order to construct, amend and evaluate integrated spatial strategies and scenarios, a number of additional tools are available. In particular, for preparing input maps, Environment Explorer is equipped with an overlay tool, an instrument geared at the creation of the suitability and zoning maps used at the local level of the model. It takes (factor) maps from a GIS as an input and combines them into a single composite map by weighing the relative importance of the information represented on the maps. The weights are set interactively through the manipulation of sliders on the screen. The composite map changes accordingly and instantly. The output can be exported straightforwardly into Environment Explorer or back into the GIS if desired. But, the overlay tool is equally used for carrying out multi-criteria analysis on the spatial outputs of Environment Explorer and in particular on sets of indicator maps. The result of this analysis is one composite indicator map, or a series of such maps: one per year simulated, reflecting the concerns expressed in the weights and criteria selected. It thus is instrumental in evaluating the particular merits of policy options tried out in terms of their multi-faceted impacts, the spatial patterns generated and the precise timing of events and developments.

The analyse tool is an instrument enabling the pairwise comparison of the many maps generated in typical runs of Environment Explorer, containing categorical data or data on a ratio or ordinal scale. It is an essential instrument for comparing and analysing the spatial effects of the alternatives generated. To that effect it is equipped with fuzzy set map comparison

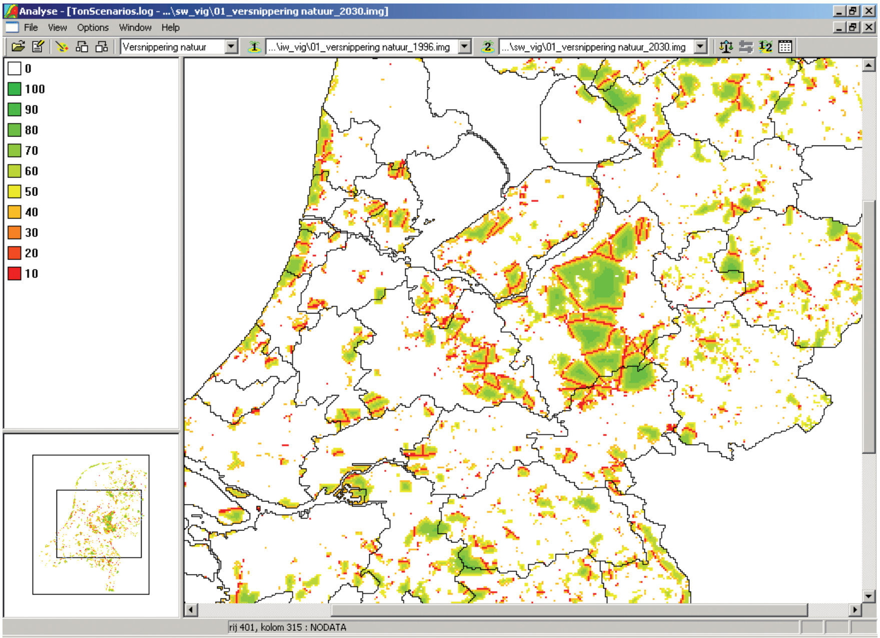


Figure 2.3 The degree of spatial fragmentation of natural areas calculated as an indicator in Environment Explorer.

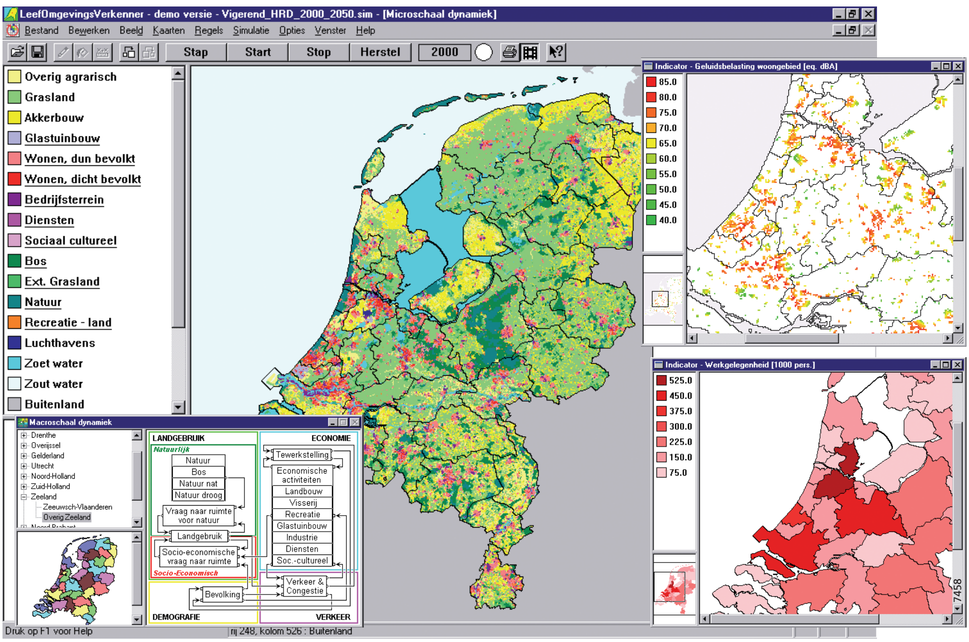


Figure 2.4 The Environment Explorer is equipped with a state of the art graphical user interface.

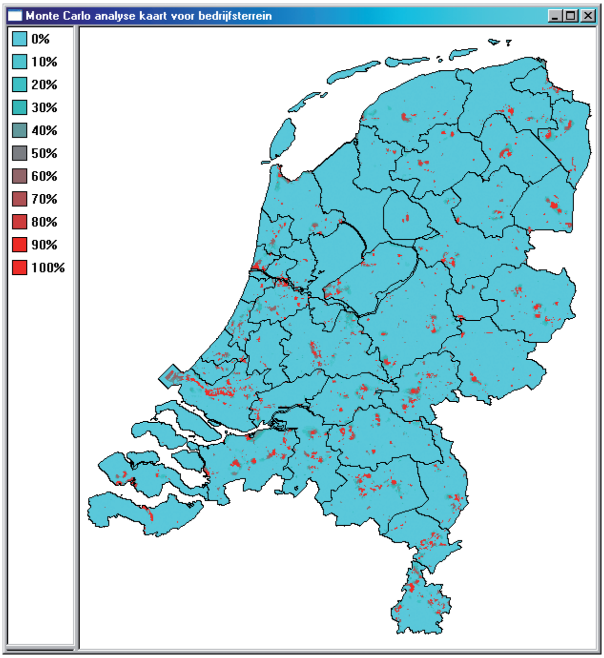


Figure 2.5 Map representing the probability that a cell is taken in by the land use 'industry' as the result of the uncertainty in a population growth parameter.

techniques capable of detecting qualitative similarities between maps (Power et al., 2001, Hagen, 2003), as well as other comparison tools and statistics such as the Kappa statistic (Monserud and Leemans, 1992).

The overlay tool and the analyse tool are of particular use in interactive sessions with stakeholders and representatives of different planning departments as they focus the discussion on the importance of particular factors in the determination of the physical or institutional appropriateness of areas for one or the other use. Similarly, they enable a straightforward evaluation of alternatives designed and tried out in collaborative working sessions.

The Monte Carlo tool is geared at using the model in a stochastic mode. In particular it supports the execution of multiple runs with the model in which particular combinations of parameters are varied stochastically within predefined ranges representative of their inherent uncertainty. As a result, for each land use function modelled, a land use probability map is made available, summarising the multiple run and computed as the proportion of runs in which the cells were taken in by the land use (Figure 2.5). The particular value of such probability map is that it demonstrates to the planner the existence of spatial bifurcations in the spatial system, appearing suddenly as one or more parameters pass a critical value (White and Engelen, 1997a). Most important to the planner is the knowledge about bifurcations that cause important qualitative changes in the system due to variations in parameter values that are well within the uncertainty ranges. Another practical value of the land use probability maps is that they show the planner where in space certain risks may appear in more or less consistent manners and independent of the uncertainty in one or more parameters (Engelen et al., 1996). If activities appear consistently in areas where they cause environmental, social or economic stress, then the planner needs to work out additional restrictive measures. Similarly, if activities systematically avoid expanding into areas designated for growth, then new stimulating measures need to be tried out.

2.4 An extensive analysis

The prime purpose of Environment Explorer is to represent at a high level of abstraction the autonomous dynamics that change the face of the Netherlands. Actual and intended policies and plans are introduced. They constrain and steer this autonomous growth. Confronted with adverse trends and growth of the system in economic, social or environmental terms, the planner can intervene and change existing policies or define new ones in an effort to bring the system onto a more favourable path of development. The many features of Environment Explorer are specifically intended for this purpose and support the user in his search for a better or more acceptable evolution of the system. Its integrated nature enables analysing direct but also indirect consequences of interventions, its dynamic nature enables exploring immediate effects, but also those that become visible at later stages, its spatial nature enables evaluating impacts at the national, the regional and the local scale and its stochastic mode enables studying the effectiveness of planning options under conditions of uncertainty.

2.5 An evaluation

Now that the system is in its fifth year of existence, a number of conclusions can be drawn relative to its development and use.

Overall, the appropriateness of Environment Explorer as an analytical instrument for the design and evaluation of spatial plans as well as a tool for communication about such plans is granted by most users at the provincial and the national levels. All agree that it provides insight in the interconnected nature of different functions, processes and cause-effect relations. It makes the consequences of policy interventions explicit in the domain of the specialist user but also in that of colleagues and counterparts working in the other domains. Its availability enables the calculation of more alternatives than would normally be possible and hence permits more alternatives to be considered and it enables an objective evaluation of the results generated. Environment Explorer stimulates the creativity of technicians and policymakers. With the tool at hand, they generally work more systematically and intensively on the definition of the set of evaluation criteria for the alternatives prior to their elaboration and develop more alternatives before they home in on ‘politically acceptable’, ‘likely’, ‘no-regret’ or ‘easy to implement policies’. However, Environment Explorer represents a complex and complicated system. Consequently it is difficult to keep the tool itself from being complicated and despite the fact that major effort has gone into rendering it as transparent and user-friendly as possible, it only meets the expectations of part of its intended end-users in this respect. Technicians, familiar with GIS and models having worked their way trough the technical documentation, are happy with the system as it is. For occasional and non-technical users this is much less the case. They tend to get lost in the open nature of Environment Explorer and have suggested a number of additional tools to streamline policy analyses. Among others, the design and evaluation of alternative policies using terms and a language better understood by policymakers and planners has been advocated by many. That is also why hands-on training workshops have been organised and sessions, involving groups of specialists working as a team on national or regional planning documents, have been facilitated or assisted by those already familiar with the system.

Instruments like Environment Explorer depend extensively on good quality data. In particular high-resolution land use maps are required for the base year of the model, but also for years in the past in order to determine trends, calibrate and validate the model. This is a problem because good land use maps are rare. Most often the detail, the number and the set of land use categories mapped are insufficient to be useful. Moreover, the definitions and categories change over time. Thus, making the land use maps consistent between years becomes a laborious prerequisite. Next, the match between on the one hand the land use categories of the land use map (at the local level) and on the other hand the activity classes (at the regional and national levels) represented in economic and census tables, is not necessarily one to one. Very often it will be necessary to go back to the most fine-grained representation of the census data in order to establish a workable match.

The calibration and validation of this type of model is far from easy or fast. This is partly due to the limited availability of good quality data, but also to the integrated nature of the model: all linked processes need to be calibrated in isolation and in combination in order to generate reliable results. Moreover, it has been the experience of Environment Explorer that the model needs to be recalibrated regularly when additional processes are built-in or when better data become available. An automated or semi-automated calibration procedure is much wanted to take care of this problem. Such procedure is currently under development (Straatman et al., 2004).

Finally, Environment Explorer is a data intensive system. This is true for the base data required to set-up the system but also for the definition of the scenarios, strategies and alternatives tried out. The system facilitates handling these with great ease, yet it is important to have a good documentation method to keep track of the data used in the exercises carried out.

3 Constructing land use maps of the Netherlands in 2030

Reprinted from *Journal of Environmental Management*, 72. De Nijs, A. C. M., De Niet, R. & Crommentuijn, L. Constructing land use maps of the Netherlands in 2030. Pages 35-42. Copyright (2003), with permission from Elsevier

Abstract

The National Environmental Assessment Agency of the RIVM in the Netherlands is obliged to report on future trends in the environment and nature every 4 years. The last report, Nature Outlook 2, evaluated the effects of four alternative socio-economic and demographic scenarios on nature and the landscape. Spatially detailed land use maps are needed to assess effects on nature and landscape.

The objective of the study presented here was how to create spatially detailed land use maps of the Netherlands in 2030 using the Environment Explorer, a cellular automata based land use model to construct land use maps from four scenarios. One of these is discussed in great detail to show how the maps were constructed from the various scenario elements, story lines and additional data and assumptions on national, regional and local land use developments.

It was the first time in the history of our outlooks that consistent, spatially detailed land use maps of the Netherlands for 2030 were constructed from national economic and demographic scenarios. Each map represents a direct reflection of model input and assumptions. The maps do not show the most probable developments in the Netherlands but describe the possible change in land use if Dutch society were to develop according to one of the four scenarios. The large (societal) uncertainties are reflected in the total set of future land use maps. The application of a land use model such as the Environment Explorer ensures that all relevant aspects of a scenario, i.e. economic and demographic developments, zoning policies and urban growth, are integrated systematically into one consistent framework.

3.1 Introduction

The Netherlands is a small but densely populated country. About 16 million people live in an area of 41,500 km^2. The economic heart of the country is the so-called Randstad. The Randstad comprises the major cities of Amsterdam, Rotterdam, Utrecht and The Hague and incorporates more than 5 million inhabitants. During the last century almost all the land development was for agricultural use, so that forest and nature reserves now cover less than 5,000 km^2. To increase the total land for nature areas in the Netherlands, about 2,500 km^2 of agricultural land is to be transformed to new forests and nature reserves by 2018 (LNV, 2002, MNP, 2002a). Furthermore, roughly 1,500 km^2 land will be needed for the development of residential and industrial areas by 2030 (ABF, 2002). According to current policy plans and trends, 10% of the land use will change in the next 30 years.

The National Environmental Assessment Agency (MNP) is obliged to report on future trends in the environment and nature every four years in a so-called outlook. Nature Outlook 2 evaluated the effects of alternative socio-economic developments on nature and the landscape (MNP, 2002b) within a project consisting of three steps: formulating scenarios, constructing land use maps and estimating effects on nature and landscape. This article will focus on the second step, constructing land use maps. This step was aimed at how to create spatially detailed land use maps of the Netherlands in 2030 from the various scenario elements, story lines and additional data on national, regional and local land use developments. These land use maps should reflect the main assumptions and basic principles of the scenarios for all relevant land uses in order to estimate the potential effects on nature.

In total, four scenarios were translated into spatially detailed land use maps of the Netherlands in 2030. Only one scenario, Individual World, will be described here to show in detail how the maps were constructed. The following sections will go on to discuss the characteristics of the Individual World scenario, the land use model, model input and assumptions and future land use maps. Scenario implementation is also discussed and conclusions are drawn on how to construct future land use maps from scenarios.

3.2 The Individual World scenario

Scenario analysis has evolved into a standard methodology in environmental sciences for analysing the effects of different driving forces and assessing the associated uncertainties. The scenarios developed for Nature Outlook 2 are related to the IPCC scenarios (IPCC, 2000), which are constructed on two axes. The first axis describes the degree of 'globalisation' versus 'regionalisation', while the second axis covers the 'individual orientation on material values' versus a more 'cooperative orientation respecting social and ecological values'.

These two basic future trends were adapted for Nature Outlook 2 in terms of five, quite different, entities: living, working, agriculture, nature and government (Luttik et al., 2002), which, in turn, were analysed for possible developments in the Netherlands, including their spatial implications. The different elements from these thematic analyses were combined into four integrated spatial scenarios (Figure 3.1).

The Individual World scenario is characterised by a global orientation. Government withdraws from many areas. Market principles dominate in spatial planning and restrictive zoning policies are being phased out, giving way to urban sprawl and the scattered development of villages in green areas. Only the river forelands of the Rhine, Meuse and IJssel are restricted to urban developments for optimal inland navigation.

Besides large suburbs near the access and exit points of motorways, small 'green villages', consisting of about 100 houses with an average per house of 2500 m^2, develop in regions with low noise levels and located near nature, forest or extensive agricultural areas. The 'Waterman' Plan, consisting of the offshore development of a residential island near The Hague has been carried out.

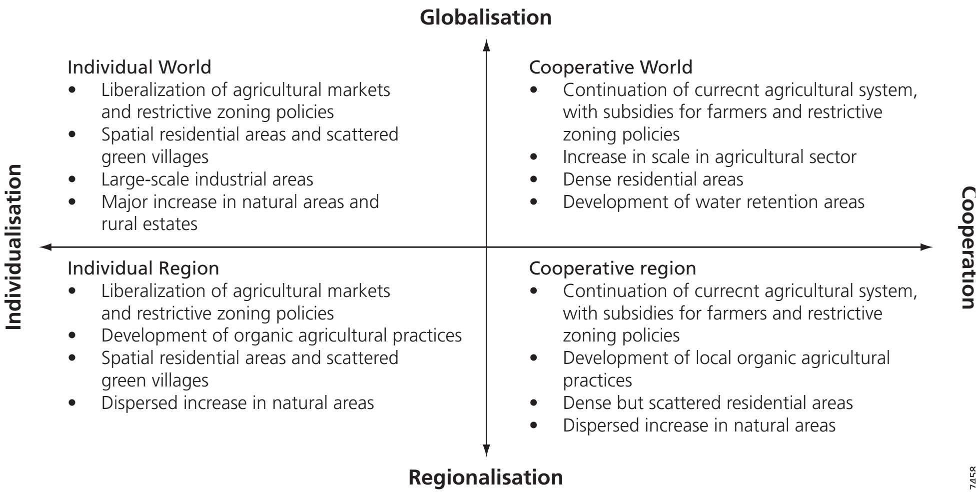


Figure 3.1 The four scenarios in Nature Outlook 2 and their major characteristics.

The greatest demand for industrial and commercial areas is near the urban areas close to access roads to motorways. Accessibility by public transport is less important. Huge industrial and commercial complexes develop at the cost of smaller or more poorly located sites. Furthermore, the harbour area of Rotterdam has been extended into the sea with the development of the so-called Second (12 km^2) and Third (25 km^2) 'Maasvlakte'. Project developers have created three new lakes in the Randstad, offering all kinds of water recreation. A border lake around the reclaimed Noordoostpolder has been created as stated in the Fifth Spatial Policy Plan (VROM, 2001).

Agricultural subsidies are diminished, leading to a sharp decrease in agricultural area. Less profitable agricultural land will be left fallow. Large pieces of agricultural land will be bought up by nature conservation organisations for the development of nature reserves and by private individuals for building their own rural estates.

3.3 The Environment Explorer

The Environment Explorer, used in this study to construct the future land use maps (Engelen et al., 2003), is a so-called cellular automata (CA)-based land use model of the Netherlands distinguishing 17 land use types. The model simulates the development of ten 'active' functions: high and low-density residential areas, industrial areas, commercial services, public services, recreation, greenhouses and three types of forest and nature. The model also distinguishes three 'passive' functions: grassland, arable land and other agriculture and four 'static' features: airports, fresh water, salt water and foreign land. The passive functions are not simulated but can switch to one of the active functions, while the static features cannot change in land use type but will influence the allocation of the active functions.

The national or regional developments of the active land uses are allocated on a 500 m grid-based land use map of the Netherlands (Figure 3.2). The model calculates the transition potential for each grid cell for each active land use. The land use developments in each region are

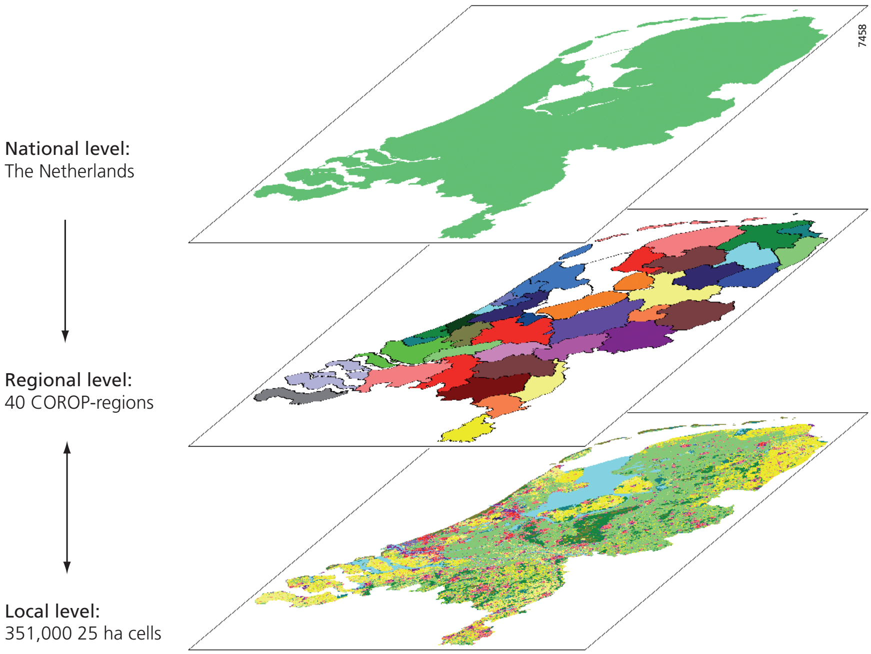


Figure 3.2 The three spatial levels, national, regional and local, represented in the Environment Explorer.

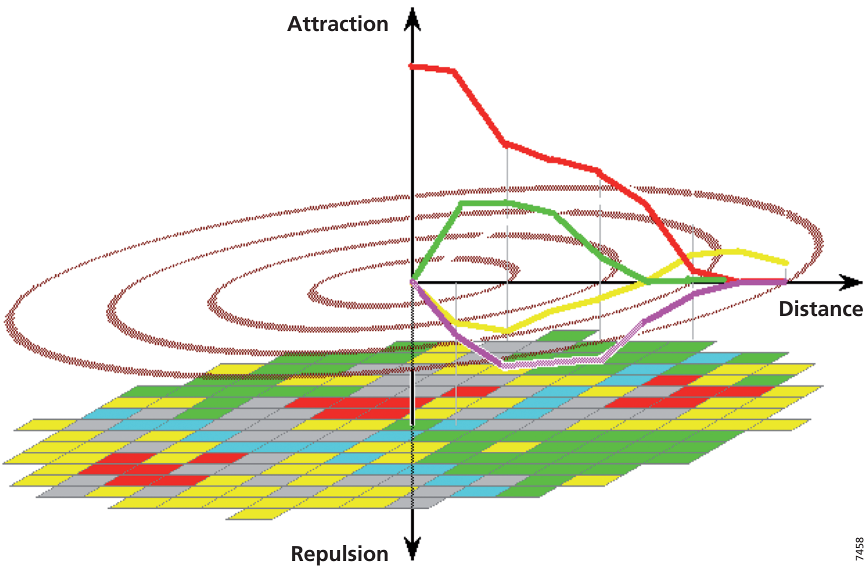


Figure 3.3 Schematic representation of the 196 cell neighbourhood in the Environment Explorer. One set of the CA rules is shown superimposed on the neighbourhood map.

Table 3.1 Results of the calibration of the Environment Explorer for 1989-1993, where simulated land use in 1993 and observed land use in 1989 are compared with observed land use in 1993 for Kappa, K-location, K-histo and Fuzzy Kappa.

Land use map 1	**Simulated 1993**	**Observed 1989**
Land use map 2	**Observed 1993**	**Observed 1993**
Kappa	0.933	0.948
K-location	0.949	0.958
K-histo	0.983	0.989
Fuzzy Kappa	0.925	0.942

allocated to the grid cell with the highest transition potential. This transition potential (*TP*) is the weighted sum of the neighbourhood potential (*NP*), the policy or zoning map (*ZM*) and the suitability map (*SM*):

$$TP_{k,i} = NP + c_1 ZM_{k,i} + c_2 SM_{k,i} \qquad \text{(eq. 3.1)}$$

in which c_1 and c_2 are two weighting factors, *k* the land use and *i* the cell indices. The neighbourhood potential is calculated with a CA model.

The behaviour of the CA model is defined by a set of rules describing the relative influence of all 17 land uses on the 10 active land uses within a radius of eight cells (Figure 3.3). In this way, the relative attraction or repulsion of the active land uses by all other land uses is described. The zoning map indicates where land use may and may not be developed according to policy plans. A suitability map is defined for each land use, indicating the relative suitability of a grid cell. For a more detailed discussion, including the mathematical description, the reader is referred to the Technical Documentation (de Nijs et al., 2001b).

In a previous study, the Environment Explorer, including the set of rules for the CA model, was calibrated for the development of the land use over the 1989-1993 period. The final model result was compared with observed land use data (CBS, 1993, CBS, 1997). The Kappa statistic, K-histo and K-location, defined according to Pontius (2002) and the Fuzzy Kappa statistic according to Hagen (2003) are all shown in Table 3.1. Here, simulated land use in 1993 and observed land use in 1989 are compared with observed land use in 1993. The results of the calibration are discussed in Section 6.

3.4 Model input and assumptions

The Environment Explorer was used to situate the growing demands on land according to the various trends in the scenarios (de Nijs et al., 2002). The original set of land uses in the model was adjusted for Nature Outlook 2 to reflect the main scenario assumptions. The total number of land uses in the model is limited to 17. In order to accommodate two new land uses, industrial area and commercial and public services were aggregated into one land use 'industrial/commercial' area. The two vacant land uses were defined as green villages and as an extra natural land use. In total, four natural land uses were defined given the type of nature management and use: almost natural, half natural, recreational and extensive agricultural.

Table 3.2 Detailed specification of the Individual World scenario in the Environment Explorer

Scenario elements	**Implementation**
Residential areas	
High density 122 km^2	Regional trends
Low density 489 km^2	Regional trends
Green villages 486 km^2	Regional trends
Regional distribution of residential developments conforming to previously developed Trend scenario (MNP, 2001)	Regional trends
Restriction of residential developments in the riverbed of the Rhine, Meuse and IJssel	Zoning map
Location of low and high-density residential areas near access roads to motorways	Suitability map
Location of green villages in nature reserves, forests and extensive grassland	Suitability map and CA rules
Location of green villages in regions with low noise levels (< 50 dB(A), MNP, 2001)	Suitability map
Development of small green villages (< 1 km^2)	CA rules
Offshore development of the so-called Waterman Plan near The Hague	ArcGIS
Industrial/commercial areas	
Industr./Comm. Areas 683 km^2	Regional trends
Regional distribution of developments conforming to previously developed Trend scenario (MNP, 2001)	Regional trends
Restriction of developments in the riverbed of the Rhine, Meuse and IJssel	Zoning map
Location of industrial/commercial areas near motorways and access roads	Suitability map
Development of large-scale industrial/commercial areas, industrial complexes	CA rules
Extension of the Rotterdam harbour into the sea, with the so-called Second Maasvlakte (12 km^2) and Third Maasvlakte (25 km^2)	Zoning and Suitability map
Recreation	
Recreation 300 km^2	Regional trends
Regionally distributed according to the relative number of inhabitants per region in 2030	Regional trends
Nature reserves	
Almost natural 4550 km^2	Regional trends
Half-natural 1625 km^2	Regional trends
Recreational 325 km^2	Regional trends
Ext. grasslands 0 km^2	Regional trends
Regionally distributed according to relative share of sandy and peat soils	Regional trends
Location of nature reserves on less profitable agricultural soils	Suitability map
Extension near existing nature reserves	CA rules
Water	
Fresh water 100 km^2	
Development of new recreational lakes in the Krimpenerwaard and the Alblasserwaard	ArcGIS
Development of a border lake around the reclaimed Noord-Oost Polder as suggested in the Fifth Spatial Policy Plan (VROM, 2001)	ArcGIS

Once the set of the land uses had been defined, an initial land use map was constructed on the basis of the 1996 Soil Statistic map of the Netherlands (CBS, 2000). Agricultural land use in this map was reclassified with the Land use map of the Netherlands (SC-DLO, 1999) and residential area was reclassified according the Residential Typology (ABF, 1998).

The development of each land use in the Environment Explorer is defined by four inputs: a policy map, a suitability map, CA rules and regional development. The specific scenario elements are used in the definition of these four elements. Table 3.2 gives a detailed overview of the model specification.

For each land use, zoning and suitability maps were defined according the characteristics of the scenario. The zoning maps include the planned developments for all land uses (VROM, 1999), while the forelands of the Rhine, Meuse and IJssel are restricted to urban land use development. The suitability maps for the urban land uses include the access roads to motorways. Suitability of green villages was determined from the combination of a noise map (Dolmans et al., 2000) and a soil map. The noise map is used to identify the relatively quiet areas in the Netherlands, while the soil map identifies the less profitable soils, which are assumed to lay fallow.

The influence of the neighbourhood on the spatial dynamics is defined by the CA rules set. The calibrated set of CA rules of the Environment Explorer was applied for all original land uses that did not change. Rules that correspond best to the new land use definitions were chosen from the calibrated set of CA rules for the new land uses. In this way, the new land use, i.e. industrial/commercial areas, is based on the original rules for industrial areas. The majority of the original aggregated land uses consists of industrial area. The positive attraction of industrial cells to other industrial cells was increased at larger the distances within the neighbourhood to realise larger industrial areas. The rules for green villages are based on the low density residential areas. However, to create small green villages occupying only one cell of 25 ha, the rules were modified. The original attraction of this land use to cells with the same land use has been changed to repulsion so as to create small villages occupying only one cell.

Land use developments on a national scale comprise 12,200 ha high-density residential area, 48,900 ha low density residential area, 48,600 ha green villages, 30,000 ha recreational areas, including parks, sports and camping grounds and 68,300 ha industrial and office area. The regional development of residential and industrial/commercial areas are based on the 'Trend' scenario, developed in the previous assessments of the Fifth Spatial Policy Plan (MNP, 2001, VROM, 2001). The national development of recreational area is distributed over the regions according to the number of inhabitants per region in 2030 given in the Trend scenario.

Total development of natural areas in the Individual World scenario is assumed to be 650,000 ha. A major part of these developed areas (400,000 ha) consists of rural estates. The national development is allocated to the region given its relative share of the national amount of land with agriculturally less profitable soils. The less profitable soils are the sandy soils with a clay content of less than 10% and subsiding lands due to the oxidation of the peat substrate. Finally, agricultural land use diminishes by 860,000 ha in this scenario, almost 20% of the total area of the Netherlands.

Because of practical reasons, the new fresh water lakes, the border lake around the reclaimed Noordoostpolder and the 'Waterman' Plan have been added to the simulated land use map in ArcGIS.

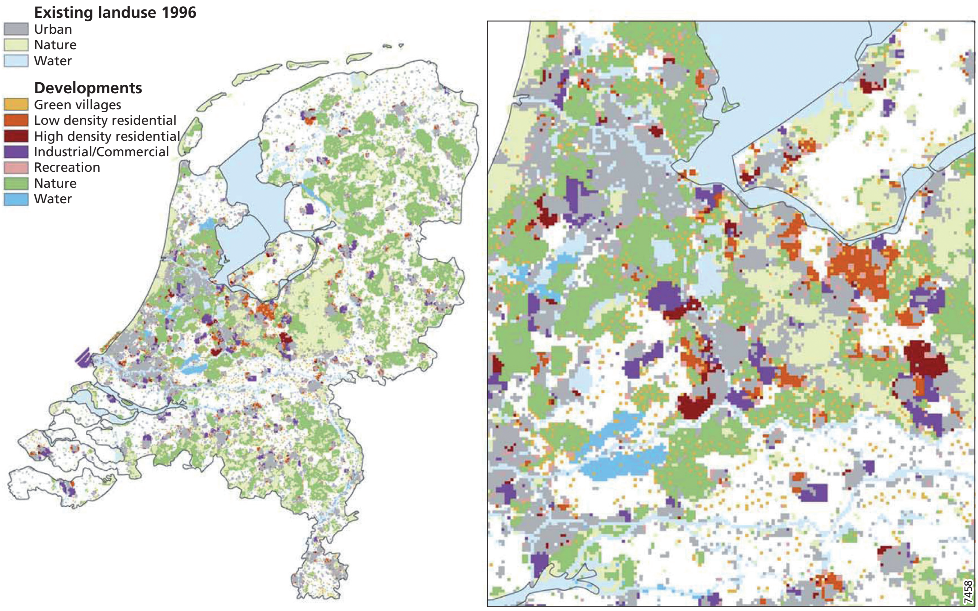


Figure 3.4 Land use in 2030 for the Individual World scenario on national and regional scales.

3.5 Future land use in an individual world

The future land use in 2030 is shown in Figure 3.4. This map directly reflects the assumptions concerning urban and rural developments. The Randstad region urbanises to a high degree, especially around Utrecht in the centre of the Netherlands. One of the new high density residential areas is located more to the east, in the Gelderse Vallei between Veenendaal and Ede. Large residential suburbs develop in green rural areas, where former spatial policy restrictions to protect the cultural or ecological value of the area were lifted. Green villages develop scattered throughout the open space between the Rhine and Meuse and in nature reserves, forests and extensive grasslands. The large increase in residential areas and rural estates causes a reduction in the available land for new developments, resulting in relocation of development to the surrounding regions. Huge, well accessible industrial/commercial complexes develop near the access roads to motorways.

Agriculture disappears from the less profitable sandy and subsiding peat soils, introducing opportunities for development of nature reserves and rural estates. The agricultural area decreases the most in this scenario, by 36% relative to the situation in 1996.

3.6 Discussion

In this study future land use maps were constructed from different scenarios. National development of urban and natural areas was translated into detailed spatial maps of the land use in 2030. One should be very careful in further application of these results. In the case projected here they have been used to show the potential effects on nature and the landscape and to see whether policy objectives can be realised in the context of various societal developments.

Application of a land use model such as the Environment Explorer will ensure that future land use maps are linked to current spatial developments and also to the way the spatial socio-economic system functions (Helling, 1998). This analysis adheres to historical trends as far as possible. For this reason the model was calibrated over the period, 1989-1993. The results of the calibration in Table 3.1 showed the observed land use in 1989 to be a better predictor of the situation in 1993 than the simulated map of 1993. This seems to be the case for many spatial allocation models (Pontius et al., 2003). In itself, this may not be so peculiar. It is difficult to get a historically consistent set of land use maps to calibrate the land use model. Observed land use maps for 1989 and 1996 in the Netherlands are not consistent in time and differ to a large extent because of small changes in the definitions of the land use categories. These small changes have a large effect on the dominant land use at a 500-m resolution. For instance, 15% of the recreational area in the observed land use map of 1989 is not present at the same location in the observed land use map of 1993. The inconsistency will most probably decrease with smaller grid sizes. The inconsistency in land use maps does not directly influence the Kappa statistics, but indirectly influences the behaviour of the model as defined in the CA rules. Furthermore, the spatial allocation of developments is, to some degree, random (White and Engelen, 1993), which makes it hard to simulate the exact location of land use developments. The initial land use map will be a better predictor for relative short simulation periods when spatial developments are relatively small. Over a longer period, spatial developments will increase and the simulated map will, most likely, perform better, compared to the initial map because the total amount of changed land use increases.

The development of low and high residential areas adheres more strictly to the calibrated model than new defined functions such as green villages. Low and high residential developments represent an extrapolation of current trends, while development of green villages, as assumed in the scenario, should be viewed as a transition in the spatial behaviour of the residential actors (Black et al., 1998, Aspinall, 2004). The addition of a new form of land use such as green villages is a way to incorporate these assumed transitions. The behaviour of such a new land use to explore the effect of a transition can, by definition, not be calibrated on trends in historical data. In the long term, these transitions in society form major uncertainties. Scenario analysis is a way to cope with these uncertainties. Four spatially detailed land use maps were constructed for evaluating the effects of scenarios on environment and nature in the Nature Outlook 2. Each of these maps reflects model input and assumptions. The maps do not show the most probable developments in the Netherlands but describe the possible change in land use if Dutch society were to develop according to one of the four scenarios. The total set of future land use maps reflects the large (societal) uncertainties that are characteristic for such long-term analyses on a spatially detailed scale.

The construction of future land use maps can be improved by:

- making causal, consistent and plausible scenarios relating past and present developments to future ones, including potential and probable transitions (CPB, 1996, Myers and Kitsuse, 2000);
- utilising the five driving forces in the development of the scenarios: international political developments, economy, demography, technology and socio-cultural;
- including the most sensitive and uncertain variables of the spatial allocation model in the definition of the scenarios.

The reliability of the spatial allocation models should be improved by means of calibration and validation of the land use patterns they describe (Straatman et al., 2004). The advances in satellite based land use maps and new methodologies to compare observed and simulated maps offer good possibilities to calibrate and validate the models on an independent data set and to perform uncertainty and sensitivity analyses.

Additional analysis of observed spatial developments are needed to support the models (Verburg et al., 2004d). Only a relatively small number of analyses for intercomparison have been reported, especially for urban developments. Results of these studies are important to set model parameters within realistic ranges. More comparable studies for regions with different spatial developments may shed some light on the variability of location processes in the different land uses.

In the case of CA based land use models, a methodology should be developed to distil the essential set of rules from observed changes in land use maps. The set of rules driving multi-state CA models yields a huge set of parameters and this hampers model calibration. Analyses of models for different areas e.g. water and soil quality, ecological effects and traffic and transport, show that in most cases only a limited number of variables, 10-15, actually determine the results of the model.

3.7 Conclusions

For the first time in the history of our outlooks, consistent, spatially detailed land use maps of the Netherlands in 2030 have been constructed from national economic and demographic scenarios. Construction of these maps does require extra information on spatial distribution at regional and local level. National economic and demographic scenarios have been extended with assumptions on the spatial behaviour of the different actors, activities and land use functions.

The application of a land use model such as the Environment Explorer ensures that all relevant aspects of a scenario, i.e. economic and demographic developments, zoning policies and urban growth, will be integrated systematically into one consistent framework. Each of these elements can be defined by the user on the basis of historical trends, scenario assumptions or urban designs, in which prescriptive design methods and descriptive long-term projection methods are linked.

The assumptions in the scenarios and their implementation in the allocation model define the resulting land use maps for 2030. The maps do not show the most probable developments in the Netherlands but describe the possible change in land use if Dutch society is to develop according to one of the four scenarios. The large (societal) uncertainties are reflected in the total set of future land use maps. The set of land use maps is important for evaluating potential effects on nature and landscape according to the trends in society. And the final issue is what the effect will be if the Netherlands does indeed develop according to one of these scenarios?

4 A method to analyse neighbourhood characteristics of land use patterns

Reprinted from *Computers, Environment and Urban Systems*, 28., Verburg, P. H., de Nijs, T. C. M., Ritsema van Eck, J., Visser, H. & Jong, K. d. A method to analyse neighbourhood characteristics of land use patterns. Pages 667-690. Copyright (2004), with permission from Elsevier

Abstract

Neighbourhood interactions between land use types are one of the factors often included in spatially explicit analyses of land use change. They are considered to be particularly relevant in the context of urban growth and are often addressed both in theories of urban development and in dynamic models of (urban) land use change. Neighbourhood interactions are also included in many land use change models of which a large group uses cellular automata (CA) to model the neighbourhood interactions.

This paper introduces a method for analysing the neighbourhood characteristics of land use. For each location in a rectangular grid, the enrichment of the neighbourhood according to specific land use types is studied. The application of this method to the situation in the Netherlands indicates that different land use types have clearly distinct neighbourhood characteristics. Land use conversions can be explained, for a large part, by the occurrence of land uses in the neighbourhood.

The neighbourhood characterisation introduced in this paper can help to further unravel the processes of land use change allocation and assist in the definition of transition rules for CA models.

4.1 Introduction

Changes in land cover and land use are among the most important human-induced changes that have a major impact on the functioning of the Earth system (Turner II et al., 1990, Lambin et al., 2001a). Apart from impacts on biodiversity, climate change and global warming (Tyson et al., 2001), land cover and land use change can also, indirectly, influence the vulnerability of places and people to climatic, economic or socio-political perturbations (Kasperson et al., 1995, Kasperson and Kasperson, 2001). In terms of the consequences at local and regional levels, therefore, the spatial patterns of land use change are as relevant as the aggregate volume of change.

Researchers from different scientific disciplines have recently addressed land use change issues with the aim of acquiring a better understanding of the causes and consequences of land use change and of exploring the extent and location of future land use changes. A unifying hypothesis that links researchers from different disciplines is that humans respond to cues from both the physical and socio-cultural environments. Land use change is therefore often seen as a function of socio-economic and biophysical factors that are referred to as the 'driving factors' of land use change (Turner II et al., 1993). Driving factors that influence the magnitude and extent of land use change are often related to the functioning of local and national markets, policy and demographic conditions.

One of the factors often included in spatially explicit analyses of land use change is the interaction between neighbouring land use types. This interaction is considered to be particularly relevant in the context of urban growth as neighbourhood interactions are often addressed based on the notion that urban development can be conceived as a self-organising system in which natural constraints and institutional controls (land use policies) temper the way in which local decision-making processes produce macroscopic urban form. Different processes can explain the importance of neighbourhood interactions. Simple mechanisms for economic interaction between locations are provided by the central place theory (Christaller, 1933), which describes the uniform pattern of towns and cities in space as a function of the distance that consumers in the surrounding region travel to the nearest facilities. Spatial interaction between the location of facilities, residential areas and industries has been given more attention in the work of Krugman (Fujita et al., 1999, Krugman, 1999). Here, spatial interactions are explained by a number of factors that cause the concentration of urban functions (centripetal forces: economies of scale, localised knowledge spill-overs, dense labour markets) and others that lead to a spatial spread of urban functions (centrifugal forces: congestion, land rents, factor immobility etc.).

The objective of this paper is to introduce a method for the analysis of neighbourhood interactions based upon an empirical analysis of changes in land use pattern. This method should facilitate the testing and validation of hypotheses on neighbourhood interactions. An additional aim of the method is to assist modellers in the implementation and quantification of neighbourhood interactions in land use models.

Neighbourhood interactions are an important component of many land use models. The most common method currently used to implement neighbourhood interactions in land use change models is cellular automata (CA). Cellular automata were originally conceived by Ulam and Von Neumann in the 1940s to provide a formal framework for investigating the behaviour of complex, extended systems(Von Neumann, 1966). In land use models CA typically model the transition of a cell from one land use to another depending on the land use of cells within the neighbourhood of the cell. Cellular automata are used in almost all land use change models for urban environments (White and Engelen, 1997a, Wu, 1998, Candau, 2000, Ward et al., 2000, Jenerette and Wu, 2001, Sui and Zeng, 2001, Torrens and O'Sullivan, 2001), and their use has been expanded to simulate other processes of land use change; for example, (Messina & Walsh, 2001) use CA-based models to study land use and land cover dynamics in the Ecuadorian Amazon, an area where tropical forest is converted into agricultural land. Applications of CA for land use change modelling in which both urban and rural land uses are considered are provided by Engelen et al. (1995) and White and Engelen (2000).

The aim of a CA model is to realistically simulate land use and land cover change; as such, the proper definition of the transition rules of the CA model is both critical and essential. Land use change is the result of a complicated decision-making process; however, the transition rules of CA models are often defined on an ad hoc basis. In most articles describing CA models, the definition of the transition rules is given little attention, and what is lacking is a method that describes how the transition rule set can be derived. In a recent editorial on research priorities for CA and urban simulation, Torrens and O'Sullivan (2001) argue that rather than developing theories through the exploration of hypothetical concepts of urban dynamics, urban CA models are now mostly technology driven.

Recently, different approaches have evolved to better match the transition rule set with reality. Sui and Zeng (2001) use historical conversions of land use to derive empirical evidence for the importance of the different factors and use multiple regression techniques to quantify the weights of the different factors within the transition rules. Other authors use advanced calibration methods for the model as a whole to fine-tune the coefficients of the transition rules based on a number of pattern and quantity measures (Clarke et al., 1996, Messina and Walsh, 2001, Silva and Clarke, 2002, Straatman et al., 2004).

The calibration of CA transition rules is complex due to the many interacting coefficients that do not necessarily yield unique solutions: different processes (rule sets) may lead to identical patterns. Calibration, therefore, does not always lead to new understandings of the relative importance of the different coefficients and is inappropriate for testing hypothesis on the underlying factors of urban development. The same argument holds for other methods that calibrate the transition rule set without explicating the relations used. Li and Yeh (2001, 2002) propose a method that overcomes the definition problem of the transition rules of a CA model by training artificial neural networks. However, neural networks do not provide insight into the relations actually used in modelling, leaving the user uninformed about the possible lack of causality in the relations that are used in the model. The method of Yang and Billings (2000a, 2000b), which solves this inverse problem of CA based on genetic algorithms, also has a number of drawbacks. This method is, at present, only operational for simple, binary patterns. Land use patterns with multiple different land use types are much more difficult to unravel.

The importance of neighbourhood interactions in land use modelling and the drawbacks of calibration and automatic procedures for deriving transition rules for neighbourhood relations in land use change models call for new approaches to explore neighbourhood relations in land use change. Contrary to the calibration methods mentioned above, the aim of the method reported in this paper is to explore neighbourhood relations rather than to calibrate the CA transition rules directly. Such an exploration of neighbourhood relations is the first step towards narrowing down the solution space for implementing neighbourhood relations in land use change models.

In order to provide a generic method, we do not limit ourselves to specific land use transitions or specific sizes of the neighbourhood. In many studies, the size of the neighbourhood is chosen arbitrarily, and only the direct neighbourhood of a location is taken into account (e.g. Von Neumann or Moore neighbourhoods). Other authors have argued that human activities are influenced by a wider space, making a flexible definition of the neighbourhood essential (White and Engelen, 2000). We will follow this latter approach and study the interactions between different land use types in the Netherlands. In the last section of this paper we will discuss the possible use of the method for land use change modellers.

4.2 Methods

4.2.1 Characterisation of neighbourhood characteristics

Neighbourhood characteristics are calculated when a situation requires the analysis of relationships between locations, rather than an interpretation of the characteristics at individual locations. In raster-based geographic analysis, neighbourhood operations are used to compute a new value for every location as a function of its neighbourhood. A neighbourhood is any set of one or more locations that bear a specified distance and/or directional relationship to a particular location, the neighbourhood focus (Tomlin, 1990). The operations that are used to calculate neighbourhood characteristics are called convolution, spatial filtering, or focal functions (Bonham-Carter, 1994, Burrough and McDonell, 1998). Various statistics can be used to characterise the neighbourhood of a location. In our study, to characterise the neighbourhood of a location in a land use map, we have defined a measure that is based on the over- or under-representation of different land use types in the neighbourhood of a location. This measure, the enrichment factor (F), is defined by the occurrence of a land use type in the neighbourhood of a location relative to the occurrence of this land use type in the study area as a whole, as follows:

$$F_{i,k,d} = \frac{n_{k,d,i} / n_{d,i}}{N_k / N} \qquad \text{(eq. 4.1)}$$

where $F_{i,k,d}$ characterises the enrichment of neighbourhood d of location i with land use type k. The shape of the neighbourhood and the distance of the neighbourhood from the central grid-cell i is identified by d. Figure 4.1 shows the shape of the neighbourhoods used in this study. $n_{k,d,i}$ is the number of cells of land use type k in the neighbourhood d of cell i, $n_{d,i}$ is the total number of cells in the neighbourhood, N_k is the number of cells with land use type k in the whole raster

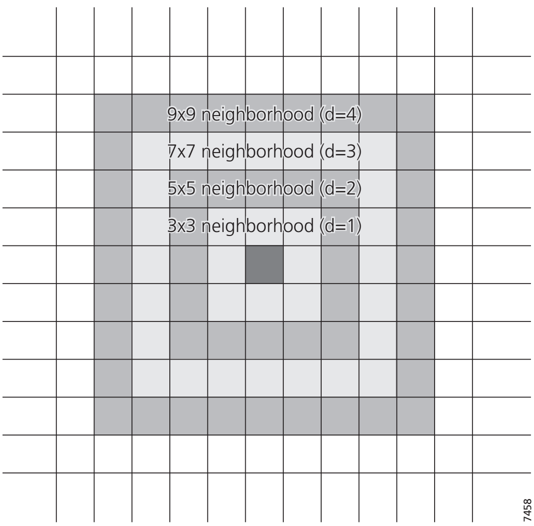


Figure 4.1 Configuration of neighbourhoods used in tis study.

and N is all cells in the raster. Consequently, if the neighbourhood of a certain grid cell contains 50% grass, whereas the proportion of grass in the country as a whole is 25%, we can characterise the neighbourhood by an enrichment factor of 2 for grassland. When the proportion of a land use type in the neighbourhood equals the national average, the neighbourhood is characterised by a factor of 1 for that land use type. An under-representation of a certain land use type in the neighbourhood results in an enrichment factor of between 0 and 1.

For each grid cell i, this neighbourhood characteristic results in a series of enrichment factors for the different land use types (k). The procedure is repeated for different neighbourhoods located at different distances (d) from the grid cell in order to study the influence of distance on the relation between land use types. In this study, we have used square rings at a distance d from the central cell as neighbourhoods (based on the Moore neighbourhood, see Figure 4.1). A C++ program was used to perform all calculations.

The average neighbourhood characteristic for a particular land use type l ($\overline{F}_{l,k,d}$) is calculated by taking the average of the enrichment factors for all grid cells belonging to a certain land use type l, as follows:

$$\overline{F}_{l,k,d} = \frac{1}{N_l} \sum_{i \in L} F_{i,k,d} \qquad \text{(eq. 4.2)}$$

where L is the set of all locations with land use type l, and N_l is the total number of grid-cells belonging to this set.

The enrichment factor defined above is comparable to the location quotient that is often used in economic geography (Smith, 1975). The average factors are symmetrical for land use in the central cell (l) and land use in the neighbourhood (k), i.e., the value of the average enrichment factor is equal for the enrichment of the neighbourhood of land use l with land use k and the enrichment of the neighbourhood of land use k with land use l. Small deviations can occur through edge effects.

Confidence intervals for testing the significance of this characteristic cannot be calculated following standard procedures. Spatial autocorrelation causes all statistical tests to be biased and is therefore inappropriate for determining the significance of the calculated characteristics (Anselin, 1988). Therefore, we have characterised the variability by the standard deviation and studied regional variability in the characteristics by comparing aggregated results with region-specific results. The standard deviation is defined by:

$$s_{l,k,d} = \sqrt{\frac{1}{(N_l - 1)} \sum_{i \in L} (F_{i,k,d} - \overline{F}_{l,k,d})^2} \qquad \text{(eq. 4.3)}$$

Regional variability was determined by comparing the average characteristics for the country as a whole with the characteristics specifically derived for a number of biophysical and administrative regions.

In this study, the neighbourhood characteristics for land use were determined for the land use pattern of the Netherlands in 1989. The characteristics of the locations where land use changed between 1989 and 1996 were specifically selected, and the average neighbourhood characteristics of these locations were determined separately.

4.2.2 Logistic regression

An analysis of the (average) neighbourhood characteristics alone does not indicate to what extent the spatial pattern of land use can be explained by the neighbourhood characteristics. An analysis of the explanatory capacity of the enrichment factors is helpful for determining the relevance of neighbourhood interactions for the particular case study. To assess the explanatory capacity of the enrichment factors, we have used these in a logistic regression model to relate the location of changed land use with the calculated enrichment factors. Other probabilistic methods, such as those based on Bayes' theory and the related 'weights of evidence' approach, could also have been used for this purpose. However, logistic regression is more suitable because the 'enrichment factor' has continuous values instead of the binary values used as independent variables in the 'weights of evidence' approach (Bonham-Carter, 1994). In logistic regression, the probability of the conversion of a grid cell (P_i) is described as a function of a set of enrichment factors, as follows:

$$Log\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \beta_{grass,d=1}F_{i,grass,d=1} + \beta_{forests,d=2}F_{i,forest,d=2} + \ldots + \beta_{k,d}F_{i,k,d} \qquad \text{(eq. 4.4)}$$

where the independent variables ($F_{i,k,d}$) are the enrichment factors of the individual grid-cells i of the neighbourhood d with land use k, and $\beta_{k,d}$ are the coefficients to be estimated with a maximum likelihood estimation. The number of independent variables included in the equation can be very large when many interactions at different distances are included in the specification. The selection of interactions included depends on the theoretical considerations of the researcher and the complexity envisaged. The resulting probabilities can be compared with the locations that actually changed to determine the goodness of fit of the regression model. The goodness of fit is a measure of the amount of spatial variation of land use change that can be explained by the neighbourhood characteristics. When a large variability is prevalent in the neighbourhood characteristics, the level of explanation of the regression equation will be low. High levels of explanation indicate that neighbourhood interactions should be taken into account in subsequent analytical or modelling efforts.

In this study we use the ROC (Relative Operating Characteristic) to evaluate the goodness of fit of the logistic regression model (Swets, 1986, Pontius and Schneider, 2001). The ROC is based on a curve relating the true-positive proportion and the false-positive proportion of the complete range of cut-off values in classifying the probability. The ROC statistic measures the area beneath the curve and varies between 0.5 (completely random) and 1 (perfect discrimination).

4.2.3 Data

We used the 1989 and 1996 maps of land use produced by the Central Bureau for Statistics (CBS Bodemstatistiek). Since 1989 these maps are based on aerial photographs (scale 1:18,000) and 33 different land use types are distinguished. These maps were checked for inconsistencies by the National Institute of Public Health and the Environment, rasterised to 25×25 m (Raziei and Evers, 2001) and combined with the LGN database ('Landelijk Grondgebruiksbestand Nederland'), another series of high-resolution land use maps produced by Wageningen University and the Research Center (Alterra Green World Research) based on the interpretation of Landsat_TM remote sensing images (de Wit et al., 1999). The integration of both maps made it possible to sub-divide the agricultural land use type (de Nijs et al., 2001a). For this

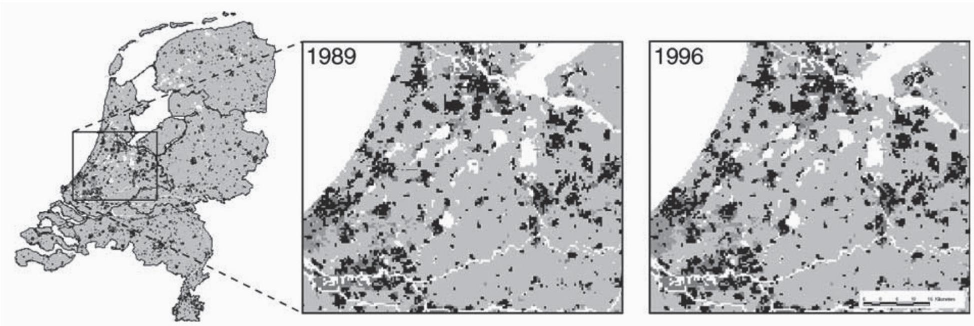


Figure 4.2 Land use in the Netherlands in 1989 and 1996 at a resolution of 500·× 500 m; dark grey shades indicate greenhouses, recreational area and industry/commercial area; black indicates residential area and airports.

study, the data set was aggregated to a 500×500-m resolution. This aggregation was based on the majority-rule. As such a procedure could yield a bias, which would especially affect land use types with a relative small coverage, causing them to disappear (Milne and Johnson, 1993, Moody and Woodcock, 1994, He et al., 2002), we constrained the aggregation by requiring that there was a correspondence between the total areas of each land use type in both the high- and low-resolution maps.

Based on this procedure, consistent land use maps at 500×500 m were produced for 1989 and 1996 and reclassified into ten land use types relevant to the analysis presented in this paper (Figure 4.2). Table 4.1 provides a description of the land use types.

Table 4.1 Land use classification used in this analysis

Land use type	Description
Other agriculture	Agricultural land not belonging to greenhouses, grassland or arable land, includes horticulture, orchards etc.
Grassland	Grasslands, incl. semi-natural grasslands
Arable land	All arable lands
Greenhouses	Greenhouses
Residential areas	Residential areas and social-cultural facilities, incl. houses, roads within residential areas, schools, hospitals, churches etc.
Industrial/commercial	Mining areas, industries, harbours, shopping malls, prisons and all service industries
Forest/nature	Forests and natural areas, incl. peat areas, swamps, heather etc.
Recreation	Parks, sport fields, kitchen garden complexes, camp sites etc.
Airports	Airports
Water	Water

4.3 Results

4.3.1 Neighbourhood characteristics of land use configuration in 1989

The characteristics of the neighbourhood of the different land use types in 1989 for the smallest neighbourhood (d=1; 3×3 neighbourhood) are listed in Table 4.2. An enrichment of the neighbourhood is indicated when the value for a land use type exceeds 1; a less than average occurrence of the land use type is present in the neighbourhood when the value falls between 0 and 1. The large values at the diagonal of the table represent the positive spatial autocorrelation of all land use types. Based on these results, it is possible to make a broad classification of the spatial clustering of land use types. Residential, industrial and recreational areas occur in clusters, whereas forest/nature and recreational areas also show a slightly positive neighbourhood relation. Arable land is often in the neighbourhood of other agricultural land, and grassland has a positive neighbourhood relation only with other grassland. Greenhouses are found in the neighbourhood of the residential areas, especially in the densely populated western part of the Netherlands. However, the high standard deviations of the characteristics for greenhouses indicate the high variability in the neighbourhood characteristics of this land use type.

Results for larger neighbourhoods are given for a number of land use types in Figure 4.3. These graphs present the average enrichment factor ($\overline{F}_{l,k,d}$) as a function of the distance (d; see Figure 4.1). The enrichment factor is presented at a logarithmic scale to obtain an equal scale for land use types that occur more frequently than average in the neighbourhood (enrichment factor > 1) and land use types that occur less often than average in the neighbourhood (enrichment factor < 1). For all land use types considered, the neighbourhood tends to become less specific with increasing distance from the central cell. When values are close to 0, the composition of the neighbourhood is similar to the land use composition of a random selection of grid-cells. At a distance of approximately 15 cells (d=15; the neighbourhood is located at between 7.5 and 10.5 km from the central cell), large deviations from the average composition no longer occur. The most specific neighbourhood characteristics are found for neighbourhood sizes smaller than 5 (between 2.5 and 3.5 km from the location). Over the whole range of neighbourhoods studied and for most land use types, the neighbourhoods are enriched the most by the land use type itself. The large extent of a number of grassland and forest/nature patches (Veluwe area in central Netherlands for forest/nature and the polder landscape of the western Netherlands for grassland) causes a relatively large, positive autocorrelation over a large distance. Recreation has a high positive spatial autocorrelation for a small neighbourhood, but industrial/commercial areas are, at a certain distance, more dominant in the neighbourhood, indicating the location of recreational as well as industrial/commercial areas at the edge of towns and cities. The positive enrichment of the residential area with industrial/commercial land use suggests that these land use types are preferably allocated close to residential areas where both an adequate labour force and a market for products are available. However, residential areas and industrial compounds are preferably separated by some distance as a consequence of the potential negative impact of industrial compounds on the livelihood. This is represented by the somewhat lower enrichment of the neighbourhood of residential area by industrial area at a short distance compared to the enrichment at a larger distance from the residential area. Another striking result, which is characteristic of the Dutch landscape pattern, is the negative relation between residential area and forest/nature at a short distance that turns into a slightly positive relation at larger distances. Historically, arable land and grassland surrounded settlements in the central and eastern part of the Netherlands, while, at an increasing distance, wild lands were used for fuel wood and extensive grazing (heather fields). These patterns are still clearly visible in large parts of the Netherlands and explain the neighbourhood characteristics. This pattern may be reinforced by planning policies in the last decades that have aimed to protect open areas that were seen as

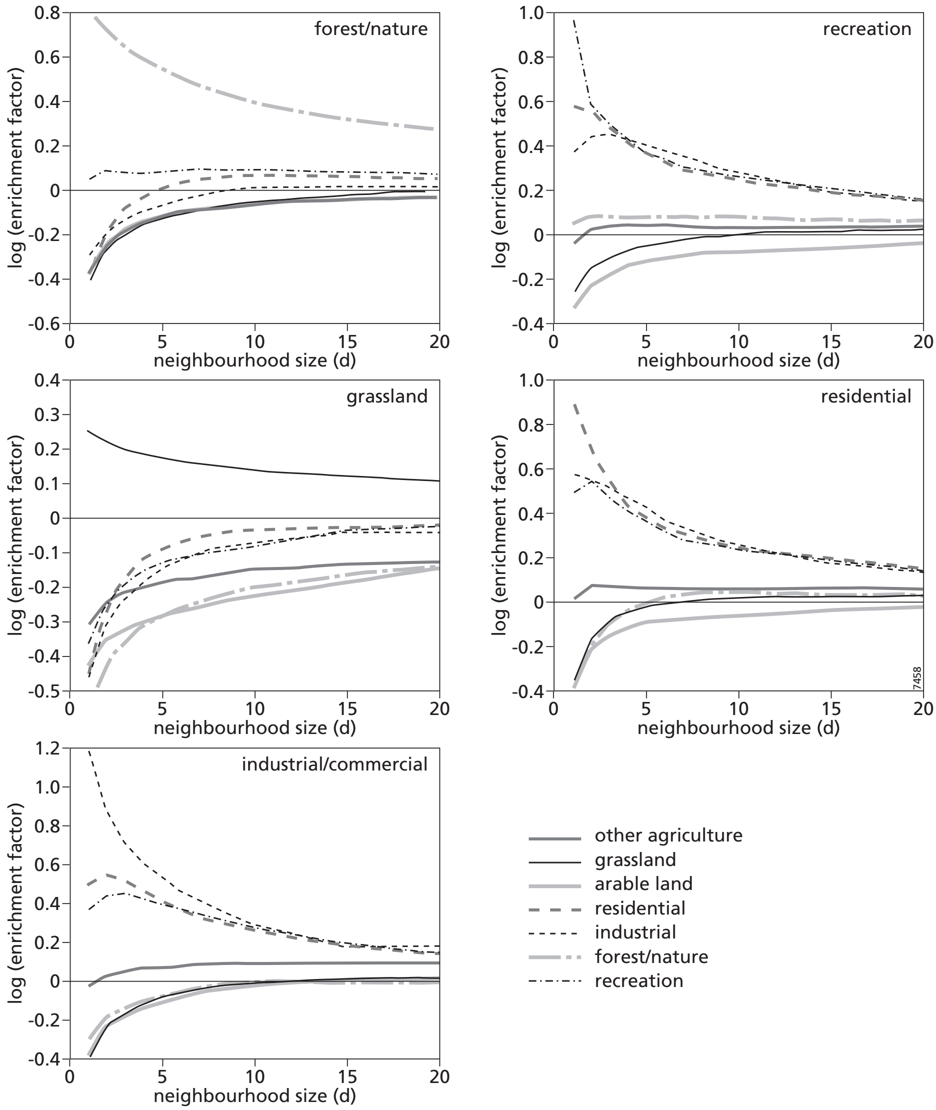


Figure 4.3 The logarithm of the enrichment factor (log ($\overline{F}_{l,k,d}$))as a function of the distance of the neighbourhood from the central cell (*d*; see figure 4.1). Each of the five graphs indicates the neighbourhood characteristics for a specific land use type.

Table 4.2 Enrichment factors ($\overline{F}_{i,k,d=1}$) of the land use pattern in 1989 for a 3x3 Moore neighbourhood (d=1); standard deviations between brackets.

Land use type in central cell (*l*)	**Land use in neighbourhood (*k*)**									
	Other agriculture	**Grassland**	**Arable land**	**Greenhouses**	**Residential**	**Industrial**	**Forest/nature**	**Recreation**	**Airports**	**Water**
Oth. agriculture	3.7 (2.4)	0.6 (1.3)	1.4 (1.8)	1.7 (1.9)	1.0 (1.5)	0.9 (1.4)	0.4 (1.0)	0.9 (1.4)	2.0 (2.1)	0.2 (0.6)
Grassland	0.6 (0.7)	2.0 (0.8)	0.7 (0.8)	0.4 (0.6)	0.4 (0.6)	0.4 (0.6)	0.4 (0.6)	0.6 (0.7)	0.3 (0.5)	0.1 (0.4)
Arable land	1.4 (1.5)	0.7 (1.1)	3.3 (1.8)	0.4 (0.8)	0.4 (0.9)	0.4 (0.9)	0.4 (0.8)	0.5 (0.9)	0.8 (1.2)	0.1 (0.4)
Greenhouses	1.7 (12.2)	0.4 (5.4)	0.4 (4.7)	126.5 (88)	1.6 (12.8)	1.4 (12.9)	0.2 (3.8)	1.8 (14.0)	0.5 (5.9)	0.1 (2.8)
Residential	1.0 (2.1)	0.5 (1.3)	0.4 (1.2)	1.6 (2.4)	7.9 (4.4)	3.1 (3.9)	0.4 (1.4)	3.8 (4.1)	0.3 (1.2)	0.2 (0.9)
Industrial	0.9 (3.2)	0.4 (2.0)	0.4 (2.1)	1.4 (3.2)	3.1 (5.5)	15.3 (12)	0.5 (2.3)	2.3 (5.1)	1.2 (3.8)	0.5 (2.8)
Forest/nature	0.4 (1.1)	0.4 (1.1)	0.4 (1.1)	0.2 (0.7)	0.4 (1.2)	0.5 (1.4)	6.5 (3.0)	1.1 (2.0)	0.7 (1.7)	0.2 (0.9)
Recreation	0.9 (2.8)	0.6 (2.2)	0.5 (2.0)	1.8 (4.1)	3.8 (5.4)	2.3 (4.4)	1.1 (3.7)	9.2 (8.7)	0.5 (2.4)	0.3 (2.0)
Airports	2.0 (23.1)	0.3 (8.7)	0.8 (15.3)	0.5 (8.0)	0.3 (7.4)	1.2 (20.5)	0.7 (12.6)	0.5 (9.6)	433 (363)	0.1 (3.7)
Water	0.2 (0.5)	0.1 (0.5)	0.1 (0.4)	0.1 (0.5)	0.2 (0.5)	0.5 (0.9)	0.2 (0.7)	0.3 (0.8)	0.1 (0.3)	4.7 (1.3)

characteristic of Dutch landscapes by redirecting residential development to areas where it is not visible from a large distance.

4.3.2 Region-dependent variability in neighbourhood characteristics

Different biophysical conditions and a different settlement history can cause differences in the spatial pattern of land use in different parts of the study area. This could lead to different characteristics of neighbourhood composition. These differences are represented by the standard deviations listed in Table 4.2. However, these differences in neighbourhood composition are partly caused by differences in the distribution of land use types over the country. Most forest/nature areas in the Netherlands are found in the central and eastern part of the country, whereas the western part is more heavily urbanised. When regional enrichment factors are calculated to compare the neighbourhood characteristics, this effect is compensated for because the enrichment factor is based on the regional frequency distribution of land use types (N_k) in Equation 4.1. The variations in the enrichment factor between regions therefore represent differences in relative neighbourhood composition.

We studied regional variations in neighbourhood characteristics for two cases. One case consisted of studying the effect of differences in biophysical conditions based on dividing the

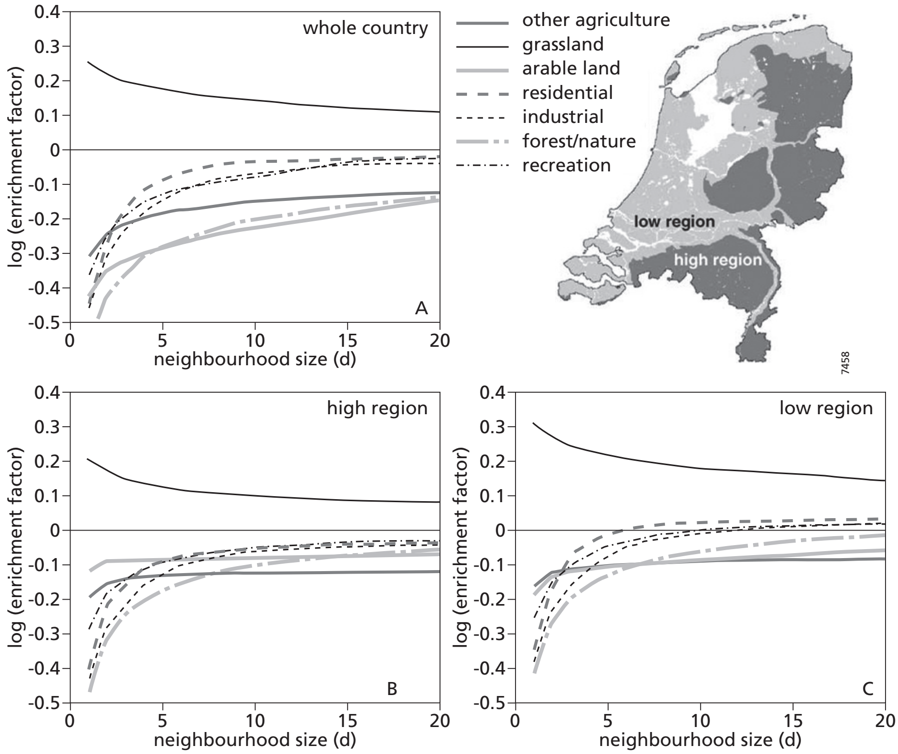


Figure 4.4 Neighbourhood characteristics (log enrichment factor) for grassland for the Netherlands as a whole (A), for the higher region (B) and the lower region (C).

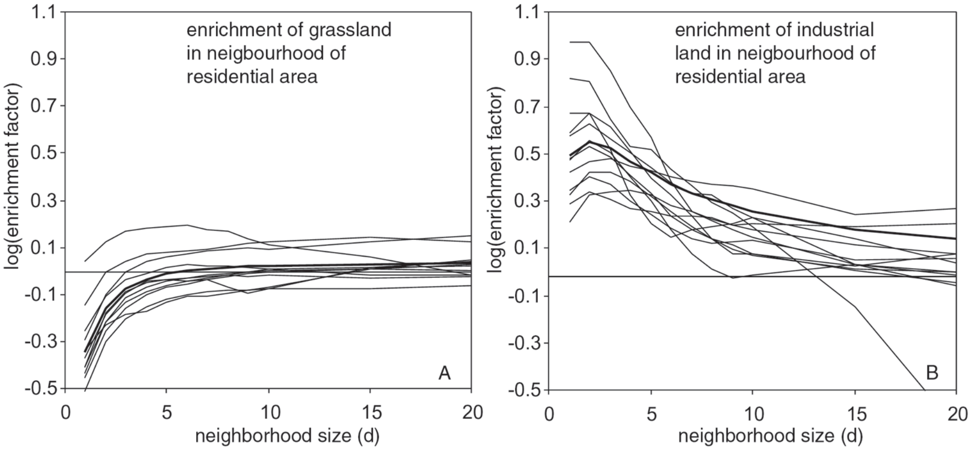


Figure 4.5 Region-specific neighbourhood characteristics for residential area for all provinces of the Netherlands based on the land use pattern in 1989; (A) enrichment with grassland and (B) enrichment with industry/commercial.

Netherlands into two regions according to its geomorphology. It was expected that the landscape pattern and, therefore, the neighbourhood characteristics would be very different for the high and low parts of the Netherlands. The high region of the country is dominated by sandy soils and the low region by mostly clayey (fluvial and marine) and peaty soils. The neighbourhood characteristics for these regions were calculated independently and compared to characteristics for the country as a whole (Figure 4.4). The general composition of the neighbourhood of grassland is similar in both landscapes: a large positive enrichment with grassland in the neighbourhood and, mostly, negative enrichments for all other land use types. However, a more detailed analysis of the results also shows large differences in neighbourhood characteristics between the two landscapes. In the low lying part of the Netherlands, the landscape is dominated by grassland areas in which villages and cities are developed at the somewhat higher locations (levees) along rivers and creeks. Residential area and grassland are direct neighbours. Arable land is situated in separate patches in the polders that have clayey soils suitable for arable land. In the southern and eastern part of the Netherlands, the pattern of villages, grassland and arable land is more mixed. This results in a lower enrichment of the neighbourhood with grassland itself and a somewhat higher contribution of arable land in the neighbourhood. Maize cultivation and grassland occur on the same fields, often as a cropping rotation.

In the second case, the variability of neighbourhood characteristics was addressed by calculating the neighbourhood characteristics for the 12 provinces of the Netherlands separately. Figure 4.5 gives the results for the different provinces for the enrichment with grassland and industrial land, respectively, in the neighbourhood of residential area. The shape of the curves for the different provinces is very similar, indicating that, despite large differences in settlement history, biophysical and socio-economic conditions, similar neighbourhood relations exist. The strength of the relation differs by province, indicating differences in the distance over which the relation affects land use. A more detailed analysis of the processes underlying these differences, however, is beyond the scope of this study. One of the provinces included in the analysis only has a limited industrial/commercial area, mainly concentrated in one compound, causing a negative value for the relation with residential area at larger distances.

4.3.3 Neighbourhood characteristics of changes in land use between 1989 and 1996

The results presented above all apply to the neighbourhood of the land uses as they existed in 1989. However, for a better understanding of land use change, it is of interest to investigate the neighbourhood characteristics of the locations where land use actually changed. The major land use changes in the Netherlands during the period 1989 to 1996 are the increase in industrial/commercial area, residential area and recreational area. We have characterised the neighbourhoods of the new locations of these three land use types.

Table 4.3 lists the existing land uses in 1989 at locations that have been developed between 1989 and 1996. It shows that a large part of the newly developed locations were formerly used for a land use type classified as 'other agriculture'. Almost 4% of the recreational area in 1989 has been converted into residential area in 1996, while new recreational areas have been developed on agricultural lands. The use of 'other agricultural' land and grasslands for urban developments is related to the occurrence of this land use type in the western part of the Netherlands where many new residential areas have been developed. The 'other agriculture' land use type includes horticultural land use, which historically develops in areas surrounding the cities due labour availability, transportation time and costs, but also because of the location of the cities on the somewhat higher, more suitable grounds, on the levees of the rivers and behind the coastal dunes where the soil is especially suitable for the culture of flower-bulbs. Careful analysis of the data indicates that building lots are often, unjustly, classified as 'other agriculture'. This is another reason for the high percentage of 'other agricultural' soils that have been converted into residential and/or industrial/commercial area.

Figure 4.6 presents the neighbourhood characteristics of the newly developed locations. The characteristics are determined by calculating the average characteristics of the neighbourhood in 1989 of all locations that have changed into a particular land use type between 1989 and 1996. For all of the land use types considered, we find that the new developments are close to already existing occurrences of the same land use type. New residential area, however, is also located in the neighbourhood of existing recreational and industrial areas, indicating the outward expansion of existing towns and cities. Urban growth is clearly much more important

Table 4.3 Land use before conversion for locations converted into respectively industrial/ commercial area, residential area or recreational area between 1989 and 1996.

Land use in 1989	Distribution (%) of converted land over land use types in 1989			% of the total area of a land use type in 1989 converted between 1989 and 1996		
	Industrial/ commercial	Residential	Recreational	Industrial/ commercial	Residential	Recreational
Oth. agriculture	57.7	60.4	38.3	2.4	3.7	1.6
Grassland	21.7	15.0	34.6	0.3	0.3	0.4
Arable land	7.0	7.0	9.3	0.2	0.3	0.3
Greenhouses	0.8	1.1	1.0	1.1	2.1	1.2
Forest/nature	7.3	3.8	16.8	0.3	0.2	0.7
Recreational	5.5	12.7	-	1.1	3.8	-

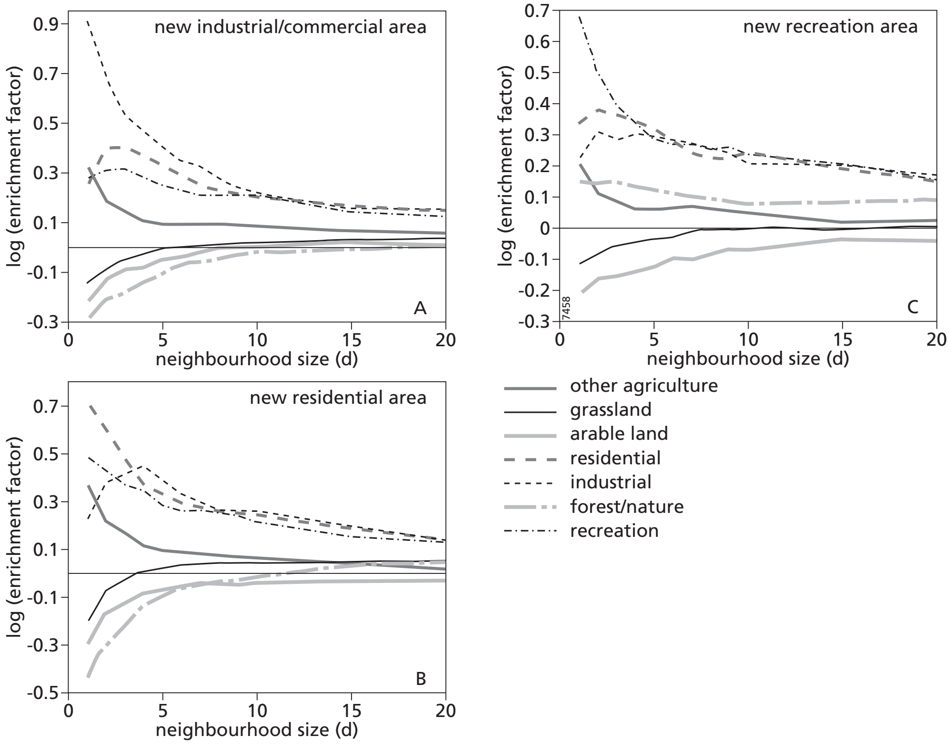


Figure 4.6 Neighbourhood characteristics (log enrichment factor) as a function of the neighbourhood size for locations converted into industrial/commercial land (A), recreational area (B) and residential area (C)

than new residential developments far from existing residential areas. For residential areas, the same observation holds as for the 1989 land use pattern: although a positively influence of industrial/commercial compounds is found at all distances, the new residential areas are preferably not located in the direct neighbourhood of these industrial/commercial compounds. The maximum enrichment of the neighbourhood with industrial/commercial area is located at a distance of 4 units (2-3 km from the new residential location). The maximum enrichment of the neighbourhood with industrial/commercial area for new recreational areas is found at a smaller distance, indicating that the area between residential and industrial land uses is often used for recreational purposes. For all three presented land use types, a large similarity between neighbourhood characteristics for the newly developed locations and the existing locations in 1989 is found (Figures 4.3 and 4.6).

4.3.4 Scale dependence of the results

The influence of data resolution on the neighbourhood characteristics was tested by conducting a neighbourhood characterisation similar to that presented in the preceding sections, but on a data set with a different resolution. To this end, we used the original data set of land use in the Netherlands described in the methods section with the original 25×25-m resolution. Since

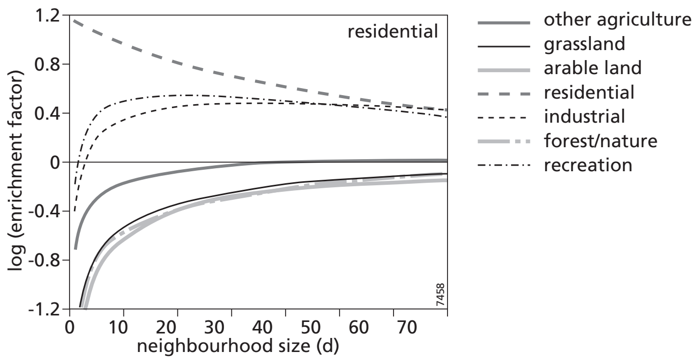


Figure 4.7 Neighbourhood characteristics (log enrichment factor) as a function of the neighbourhood size (cell size: 25 m) for residential area in 1989 using high resolution data.

this is the same data set that the aggregated 500×500-m data were based, the results for both data sets can be compared. Figure 4.7 presents the resulting neighbourhood characteristics for residential area in 1989. The neighbourhood characteristics are different from the 500×500-m resolution results (Figure 4.3) in that at short distances, residential area is only positively related to other neighbouring residential land use, but at a somewhat larger distance, recreational area and industrial area appear frequently in the neighbourhood of residential area. For a neighbourhood size of 20 units (approx. 500 m) the characteristics of the high-resolution neighbourhood are comparable with the results for the smallest neighbourhood at the coarse resolution, corresponding with the same distance from the central cell. The maximal occurrence of recreational area in the neighbourhood is found for approximately 20 units, while the maximal occurrence for industrial/commercial area is found at 40 units. This latter observation is also found for the coarse resolution data between 1989 and 1996.

4.3.5 Logistic regression: conversion probabilities based on neighbourhood characteristics

In order to assess the explanatory power of the enrichment factors, we fit a logistic regression model explaining the spatial distribution of newly developed residential areas between 1989 and 1996 using the enrichment factors of the individual cells as explanatory variables. Locations classified as water, residential or industrial/commercial areas in 1989 were excluded from the analysis. Based on the neighbourhood characteristic of the newly developed residential areas (Figure 4.6) and our understandings of the processes leading to locational decisions, we decided to include only the enrichment of the neighbourhood with residential, industrial/commercial and forest/nature as explanatory variables. Despite the large neighbourhood effect observed for recreational area, we did not include the enrichment with recreation as an explanatory factor because the development of recreational areas mostly follows after new residential areas are established. We included the enrichment for the nearest neighbours (d=1) and neighbours located at between 1500 and 2100 m from the location (d=3) to represent the influence of neighbouring land use types. The neighbourhood relation is most pronounced for the immediate neighbours. For the relation between residential and industrial/commercial area, the enrichment factor was highest for locations between 1500 and 2100 m from the central cell (d=3). These neighbourhood sizes turned out to give the highest level of explanation in a number of trials.

Table 4.4 Coefficients of a logistic regression model explaining the spatial distribution of new residential area by the enrichment of the neighbourhood (n=166,079).

Variable	Estimated coefficient	Odds ratio
Constant	-5.537*	-
$F_{residential,i,1}$	0.364*	1.438
$F_{industrial,i,1}$	0.023*	1.023
$F_{forest/nature,i,1}$	-0.201*	0.818
$F_{residential,i,3}$	0.090*	1.094
$F_{industrial,i,3}$	0.028*	1.029
-2 Log likelihood	9478	
Chi-square	2594*	
ROC	0.91	

*significant at $p < 0.01$

The enrichment factor for forest/nature in the larger neighbourhood (*d*=3) did not significantly contribute to the regression model and was therefore excluded. The resulting coefficients of the regression equation are given in Table 4.4.

The model has a good explanatory power as can be seen from the ROC value of 0.91. The ROC value can vary between 0.5 (completely random) and 1 (perfect fit); consequently, the value of 0.91 for the spatial distribution of new residential area indicates that it is possible to predict new residential area locations reasonably well based on neighbourhood characteristics. Based on the regression model, conversion probabilities can be calculated for every location that can potentially be converted into residential area. Locations with high probabilities are shown in

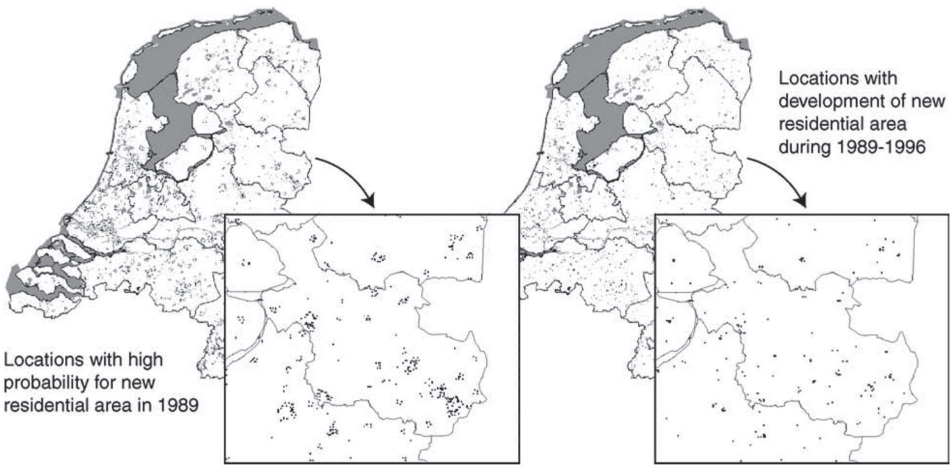


Figure 4.8 Locations with high probability for new conversion into residential area in 1989 based on logistic regression using neighbourhood characteristics as the independent variable (left) and actual locations with new development of residential area between 1989 and 1996 (right).

Figure 4.8 and can be visually compared to locations that were actually converted into urban area during 1989-1996. This visualisation confirms the high ROC value: most conversions did indeed take place at locations with a high conversion probability. However, the figure reveals that many other locations in the neighbourhood of the actual conversions also have a high probability, indicating that other factors also influence the allocation decisions, such as policy considerations, ecological values, soil suitability, tenure status among others.

4.4 Discussion and conclusions

This paper has introduced a simple method for exploring and quantifying the neighbourhood characteristics of land use. The results can be used to verify hypotheses on the interaction of land use types and to suggest possible explanations for unaccountable neighbourhood relations between land use types. Neighbourhood effects alone are, however, not the only relevant locational factor describing the spatial pattern of land use. Other factors, such as accessibility, environmental suitability, spatial policies among others also influence the pattern of land use. The neighbourhood characteristics presented in this paper are, therefore, not suitable for explaining large-scale patterns in land use, which are determined by other factors. In the Netherlands, remaining nature areas are very much related to the location of poor, sandy soils that were unsuitable for agricultural use. Many cities are located in the western part of the Netherlands because of the historical advantage of these locations in terms of industrialisation and urbanisation. At a more detailed level, neighbourhood interactions do explain spatial differences in land use pattern and can indicate which locations are suitable for future conversions based on the neighbouring land use types. In the case study reported here for the Netherlands, the predictability of land use change based on neighbourhood characteristics alone turned out to be relatively high, stressing the importance of including neighbourhood interactions in studies of land use change. A more detailed analysis of the different factors underlying the allocation is presented by Verburg et al. (2004d).

The method described here for quantifying the neighbourhood characteristics of land use is of special interest for land use change modellers. The method can be used to assist in establishing the definition of transition rules related to neighbourhood interactions in these models. These neighbourhood interactions can be incorporated in land use change models in different ways. Models based on transition probabilities, such as Markov chain models (Li and Reynolds, 1997, Thornton and Jones, 1998), statistical models (Chomitz and Gray, 1996, Veldkamp and Fresco, 1997, Schneider and Pontius, 2001, Serneels and Lambin, 2001), or those based on other probabilistic methods, such as the 'weights of evidence' approach (de Almeida et al., 2003), can incorporate the enrichment factor as an explanatory variable. Many models that explicitly address neighbourhood interactions are based on CA, and the 'enrichment factor' can assist the modeller in formulating the transition rules for CA. It is, however, not possible to directly translate the neighbourhood characteristics into transition rules for CA. Not all neighbourhood relations identified with the enrichment factor have a causal explanation, and some of the observed neighbourhood effects are the indirect result of other interactions. Furthermore, when used for calculating land use transitions, CA models work with smaller time-steps (usually 1 year) than is usually possible in an empirical analysis of neighbourhood characteristics. These short time-steps result in the emergence of complex patterns that cannot be deduced analytically back into transition rules (Torrens and O'Sullivan, 2001). Therefore, the results of this analysis can only be applied to directly derive CA transition rules when land use data with a high temporal resolution, similar to the time-steps of the CA model, are available. The increasing availability of remote sensing images makes this a realistic option. When land use data are not available at a high temporal resolution, the results of land use pattern explorations are still of value to CA modellers. The empirical results identify the most important neighbourhood relations and their spatial extent for a particular case-study. After the empirical results have been interpreted, a selection of interactions can be made that is used as a starting point for defining the transition rules in the CA model. Such a selection should be based on the researcher's careful analysis of the theoretical considerations that give rise to spatial interactions between land use types and the empirical results obtained by the method presented in this paper. Calibration methods for CA models (Clarke et al., 1996, Messina and Walsh, 2001, Straatman et al., 2004) can now be used to modify the transition rules such that model results are closer to reality. The empirical analysis has (most likely) narrowed down the solution space for such a calibration considerably, which makes the calibration simpler and more realistic. This procedure is summarised in Figure 4.9. The use of empirically derived relations provides a much better starting point for calibration than transition rules based on expert judgement alone.

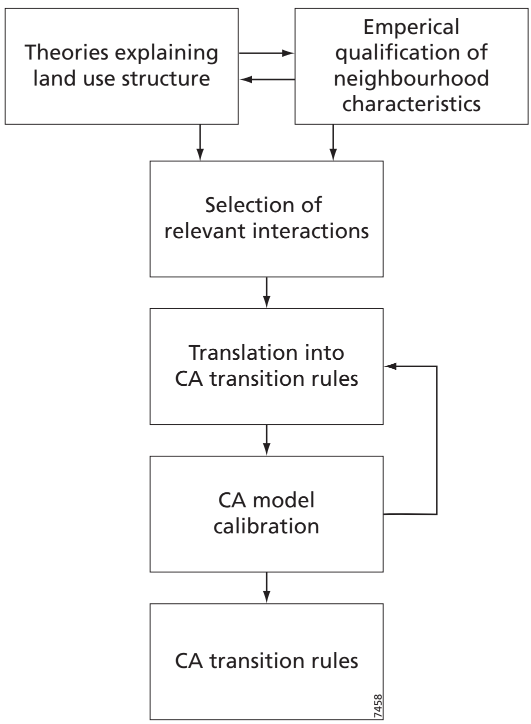


Figure 4.9 Procedure for defining transition rules within a cellular automata (CA) land use change model based on an empirical characterisation of neighbourhood characteristics.

The method is particularly suitable for CA models that include multiple land use types and a flexible definition of the neighbourhood, such as the models of White and Engelen (1997a, 2000). However, the 'enrichment factor' can also be used in situations where only one land use transition is considered (e.g. the conversion of non-urban to urban) or where a fixed size of the neighbourhood is assumed.

Although some of the results from our study for the Netherlands suggest straightforward distance-decay functions for the interaction between land uses suffice, this does not mean that it is easy to determine the neighbourhood interactions based on expert knowledge alone. The results also indicate that interactions between land uses differ by region, scale and time-period. The method introduced here allows an exploration of the implications of these differences for land use modelling. Most land use change models that include neighbourhood interactions (Wu, 1999, Ward et al., 2000, Li and Yeh, 2002, Soares-Filho et al., 2002) use uniform transition rules for the whole study area. Our analysis has shown that it is possible that neighbourhood interactions among land use types differ in different parts of a study area, due to different environmental or social-cultural conditions. If so, it would be advisable to use regional specific transition rules instead of uniform rules for the whole study area. The empirical analysis presented in this paper provides a good approach for exploring regional differences in neighbourhood interactions and can help the researcher decide upon the need for region-specific transition rules.

Another aspect of transition rules in land use models is the temporal stability of these rules. Changes in society, policy and land use pressure may lead to different interactions among land use types. The method presented here can be used to compare the neighbourhoods in different time periods in order to analyse temporal stability. However, high-resolution land use data are often difficult to obtain for different periods and, even if available, changes in data gathering methodology and the classification system still make it difficult to analyse time series. In the case study presented here, we were able to analyse the land use pattern in 1989 as well as the changes between 1989 and 1996. The land use pattern in 1989 reflects land use as it has changed since historical times. As such, it reflects how land use structure has evolved due to the interactions between land use and the social-economic and biophysical environment and among land use types itself.

One could argue that changes in the processes underlying the neighbourhood interactions would have caused differences between the neighbourhood characteristics of the land use pattern in 1989 and the neighbourhood characteristics of the changes between 1989 and 1996. However, for all studied land use types, the neighbourhood characteristics only showed subtle differences between the static and dynamic analysis, indicating a stability in neighbourhood interactions in time. This can be explained by a stability in the processes that underlie the neighbourhood interactions in land use. However, it should be noted that changes in urban and industrial/commercial land use have been most pronounced during the last few decades, thus heavily influencing the neighbourhood characteristics of the land use pattern in 1989. Similar observations of the temporal stability of functional relationships have been found in studies analysing the relation between land use and the underlying biophysical and demographic conditions for different periods of time (Hoshino, 1996, Veldkamp and Fresco, 1997).

In concluding, several remarks must be made concerning the use of the introduced methodology. The resulting neighbourhood characteristics are dependent on the resolution of the input data. Maps of different spatial resolution exhibit different spatial variability and structure.

The resulting neighbourhood characteristics are, therefore, scale-dependent. Land use types that tend to be neighbours at a certain resolution can be distant at another resolution. When the derived neighbourhood characteristics are used to assist the parameterisation of land use models with neighbourhood interaction (e.g. CA models), care must be taken that the neighbourhood characterisation is made at the same scale as that for which the transition rules are used.

The shape of the neighbourhood can also influence the results obtained. Here, we have used square Moore neighbourhoods because of computational advantages. However, especially for larger neighbourhoods, the square shape causes differences in the distance between the neighbourhood and the central cell that has no theoretical validity. It would, therefore, be better to use a more circular neighbourhood, such as the circular neighbourhoods used by (White and Engelen, 2000). Neighbourhoods based on the activity range of the agents of land use change, such as the network-based neighbourhoods used in graph-CA (O'Sullivan, 2001), are even more advanced options. Asymmetrical neighbourhoods may be relevant for other interactions; for example, for interactions between heavy industry and residential land use, the neighbourhood shape can be based on the prevalent wind direction. The next version of the software used to calculate the neighbourhood characteristics will enable a flexible definition of the shape of the neighbourhood.

The simple and straightforward method for analysing neighbourhood interactions in land use pattern introduced in this paper makes it worthwhile to include the exploration of neighbourhood interactions in assessments of land use change. The empirical characterisation can contribute to the identification of spatial relations between land uses underlying the spatial allocation of land use change. Future research should test the approach in a wider range of case-studies and evaluate the use of this method in defining well-informed transition rules in land use change models.

5 Determinants of land use change patterns in the Netherlands

Reprinted from *Environment and Planning B Planning and Design*, 31. Verburg, P. H., Van Eck, J. R. R., De Nijs, T. C. M., Dijst, M. J. & Schot, P. Determinants of land use change patterns in theNetherlands. Pages No. 125-150. Copyright (2004), with permission from Pion Limited, London

Abstract

Land use change patterns are the result of the complex interaction between the human and physical environment. Case-studies of the determinants of land use change can facilitate analyses aimed at determining which theory is appropriate in a particular region and stimulate the development of new theoretic understandings. We present here an empirical method for analysing the pattern of land use change that allows a wide range of factors, derived from different disciplines, to contribute to the explanation of land use change. The method is applied to the Netherlands based on an extensive database of land use change and its potential determinants. Historical as well as recent land use changes are studied. Historical land use change is mainly related to the variation in the biophysical environment. Levels of explanation are low due to the inability to address the temporal variation in location factors; in contrast, for the more recent changes in land use, high levels of explanation are obtained. The most important changes during this period are residential, industrial/commercial and recreational area expansion. The location of these changes can be explained by a combination of accessibility measures, spatial policies and neighbourhood interactions. Based on these results it is possible to define priority topics for in-depth analyses of land use change processes and suggest factors, relations and processes that need to be included in dynamic land use change models that support land use planning policies.

5.1 Introduction

The spatial configuration of land use is an important determinant of many ecological and socio-economical processes (Lambin et al., 2001b). As such, a better understanding of the determinants of the spatial configuration of land use is necessary to assess the impact of possible future developments on the environment, economy and society at large.

Theoretical understandings of the processes that lead to the spatial pattern of land use have been provided by researchers from different disciplines (Christaller, 1933, Alonso, 1964, Von Thünen, 1966, Fujita et al., 1999). Many empirical studies of land use change patterns have focused on tropical deforestation (Angelsen and Kaimowitz, 1999, Nelson et al., 2001, Walsh et al., 2001), while other researchers have focused on urbanisation from the perspective of economic geography (Arthur, 1994, Fujita et al., 1999, Krugman, 1999) or on the interaction of land uses through cellular automata (CA) models from a more technology-driven perspective (Wu, 1999, White and Engelen, 2000, Torrens and O'Sullivan, 2001). In practice, many processes that influence land use change interact and lead to complex patterns, depending on the local cultural, socio-economic and biophysical context at different spatial scales (Lambin et al., 2001b). For that reason, in this paper, we do not restrict ourselves to a single disciplinary perspective or typical land use conversion but will introduce a multi-disciplinary approach to analyse the impact of different determinants, derived from different theories explaining land use allocation, for a range of different land uses. For this analysis, we will use an extensive database from the Netherlands. Although urbanisation is the most important process of land use change in the Netherlands, we will also address the interaction of urban land uses with other land use types and address recent as well as historical land use change processes that have shaped the land use pattern. This analysis may contribute to our understanding of the interaction of land use change processes in the Netherlands and provide an empirical foundation for the specification of land use change models.

The following section discusses a number of theories of land use change relevant to the analysis of land use in the Netherlands. Based on these theories, we then specify an empirical model (section 3) that is subsequently applied to the long-term historical changes in land use (section 4) and the more recent changes that have occurred during the period between 1989 and 1996 (section 5). The findings are discussed and the applicability of the empirical method is evaluated in the last section.

5.2 Determinants of land use change

Land use change can be described as the complex interaction of behavioural and structural factors associated with the demand, technological capacity, social relations affecting demand and capacity and the nature of the environment in question. It is doubtful that a relatively simple explanation of why we transform the environment the way we do is forthcoming (Turner II et al., 1990). A theory of land use change needs to conceptualise the relations among the driving forces of land use change, their mitigating processes and activities and human behaviour and organisation. Although no one all-compassing theory of land use change exists, different disciplinary theories can provide some help for analysing aspects of land use change in concrete cases, with the synthesis of these theories being essential. In this paper, we use theoretical understandings from a range of disciplines that have focused on the analysis of the spatial aspects of land use change. Macro-processes that drive land use conversion, the proximate causes and underlying driving factors of land use change, such as population growth, migration and economic change, are not analysed if they are not directly related to the processes that determine the spatial pattern of land use change.

5.2.1 Biophysical constraints and potentials

The natural sciences, especially physical geography and agro-ecological disciplines, have provided insights into the constraints and possibilities of the natural environment as related to different land uses (Penning de Vries et al., 1992). Each location has specific soil characteristics and climatic conditions that determine the possibilities for natural and agricultural vegetation. Some location characteristics constrain the growth of particular crops, definitively excluding their presence at that location. Other conditions only result in lower potential yields, or additional inputs (fertiliser, land preparation, irrigation etc.) are required to make the location suitable for that particular land use type. Agro-ecologists use land evaluation techniques to determine the potential land uses for a location based on the biophysical characteristics of that location (Fischer et al., 2002). Biophysical conditions are also important factors for land use types other than agriculture; for example, the soil, geology and drainage conditions determine the suitability of a location for residential construction. It takes considerable investments to reclaim swampy land for residential construction; therefore, a preference will exist to direct these developments to locations better suited for construction. In most land use change analyses, the biophysical characteristics of the location have been incorporated as the potential gains that can be achieved by a particular land use type at that location.

5.2.2 Economic factors

A wide range of land use change studies is based on economic theory. Economists model the relation between the location factors and land use assuming that, in equilibrium, land is devoted to the use that generates the highest potential profitability (Chomitz and Gray, 1996, Irwin and Geoghegan, 2001, Nelson et al., 2001, Munroe et al., 2002). This assumption builds on the insights of Von Thünen (1966) and Ricardo (1817). In its original application, only the location relative to the market is of importance – all other landscape features are ignored. Land use structure was then the outcome of the competition for accessibility, leading to the well-known concentric circles of land use. In later years, this theoretical model was extended by many other factors that determine the rent at a location. The most well-known of these is Alonso's (1964) land market theory and model, which includes a number of other economic and environmental characteristics that determine land rent, such as infrastructure and soil quality. Despite developments in societies, accessibility and locational trade-offs are still viewed as central elements in the creation of land use structure in the twentieth-first century (Clark, 2000). The main assumption of this approach – that land use decisions are made by utility-maximising individuals – underlies all applications. Psychological research has suggested various modifications to this conception of human choice and has indicated the limitations of using this theory in understanding land use structure (Rabin, 1998).

5.2.3 Social factors, irreversibility and uncertainty

Social theories focus on factors which influence locational choices of households. Important determinants of these choices are individuals' cultural values, norms and preferences (lifestyles) and his/hers financial, temporal and transport means (Pahl, 1975, Kempen et al., 2000). In terms of spatial indicators, households are focused on the site and situation characteristics of locations. Site characteristics include housing prices, level of supply of consumer services, quality of landscapes and social composition of neighbourhoods (Geoghegan et al., 1997, Faust et al., 1999, Geoghegan et al., 2001, Mertens et al., 2002). The role of historical events and arbitrary decisions (uncertainty) is also stressed as being important. A small, arbitrary, decision can lead to new developments through the path dependency of some of the processes that determine land use allocation. The development of cities or commercial compounds is an example: the initial decision to locate a company at a certain location can be rather arbitrary. Economies of scale (Arthur, 1994) and other processes can reinforce such a development and cause the expansion of urban land uses at that location. Another cause of the reinforcement of certain land use change patterns is related to the irreversibility of some land use changes; for example, tropical forest cannot be (re-)created, and desertified land is very difficult to reclaim.

5.2.4 Spatial interaction and neighbourhood characteristics

Geographers have contributed to the analysis of land use pattern through the analysis of spatial interactions. Land use does not develop independently at each individual location; each development affects the conditions of neighbouring and distant locations. Simple mechanisms for the interaction between locations are provided by the central place theory (Christaller, 1933) that describes the uniform pattern of towns and cities in space as a function of the distance that consumers in the surrounding region travel to the nearest facilities. More complex spatial interactions are the subject of urban network theories (Hohenberg and Lees, 1985), with the location of land uses being dependent on interregional and international networks of economic, cultural and political relations (Castells, 2000). Spatial interaction between the location of facilities, residential areas and industries has been given more attention in the work of Krugman (Fujita et al., 1999, Krugman, 1999) where (Fujita et al 1999,Krugman 1999)spatial interactions are explained by a number of factors – those that cause the concentration of urban functions (centripetal forces: economies of scale, localised knowledge spill-overs, dense labour markets) and those that lead to a spatial spread of urban functions (centrifugal forces: congestion, land rents, factor immobility etc.). The direct interaction between neighbouring land uses is often assessed through CA that specifically address the push-pull between neighbouring land uses (Webster and Wu, 2001, Yeh and Li, 2002).

5.2.5 Spatial policies

Policies at national or sub-national level have a bearing on land use. This is particularly true for policies that have a spatial manifestation (e.g. the creation of conservation areas, land tenure changes or designated areas for subsidised developments) as these influence the spatial pattern of land use change. Spatial policies are specific for the study area and the time period studied. Therefore, any study of land use pattern change should be preceded by an inventory of the relevant spatial policies.

A hypothesis that unifies the different disciplinary theories is that actors respond to cues from both the physical environment and their socio-cultural and spatial environment, subsequently behaving to increase both their economic and socio-cultural well-being, with well-being being dependent on perception, scale and time. The empirical model used in this paper is based on this hypothesis; it allows the incorporation of a large number of factors important to land use change that are derived from the different disciplinary understandings of land use change described above.

5.3 Methods and data

5.3.1 Methods

Humans are expected to optimise 'well-being' by allocating land use conversions at locations with the highest 'preference' for the specific type of land use conversion at that particular moment in time. 'Preference' is an unobserved, dimension-less variable, defined by the economic returns, market competition, the socio-cultural context, arbitrary preferences and policy regulations. Preferences can differ for the different actors involved in land use decisions. In this study, 'preference' represents the outcome of the interaction between the different actors and the decision-making processes that have resulted in a spatial land use configuration. The 'preference' of a location is empirically estimated from a set of factors that are assumed to influence the 'preference' based on the different, disciplinary, understandings described in the previous section. This definition of 'preference' enables us to include a wide range of factors that can also include factors that explain the deviation from rational behaviour. The 'preference' is calculated as follows:

$$R_{i,k} = A_k G_i + B_k D_i + \ldots \qquad \text{(eq. 5.1)}$$

where R is the preference to allocate location i to land use type k, G_i is one of the biophysical characteristics of location i and A_k is the relative impact of this biophysical characteristic on the preference for land use k. D_i represents the distance of location i to the market, and B_k is the relative importance of being located close to the market for land use type k. Other factors can be added depending on the land use type considered and the region addressed. The exact specification of the model should be based on a thorough review of the processes important to the spatial allocation of land use in the studied region.

The statistical model developed is a binomial logit model of two choices: convert location i into land use type k – or not. Assuming that the preference $R_{i,k}$ is the underlying response of this choice, we can specify the following regression:

$$R_{i,k} = \beta_k x_i + u_i \qquad \text{(eq. 5.2)}$$

where the x_i's are the exogenous (explanatory) variables of the preference (R_{ik}), β_k are the coefficients to be estimated and the u_i's are random errors. Clearly, the preference for a location cannot be observed or measured directly, so $R_{i,k}$ is actually unobserved. Therefore, a dummy variable, y, is created:

$$y = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} \quad \begin{matrix} \text{if } R_{i,k} > R_{i,j} \text{ for all uses } j \neq k \\ \text{all others} \end{matrix} \qquad \text{(eq. 5.3)}$$

That is, location i is converted into land use k if the preference for land use k is higher than that for all other land use types; otherwise, no conversion into land use k is observed. If it is assumed that the u_i's have a logistic distribution, then this is a binomial logit of the form:

$$P(y=1\,(i.e.\,converted)) = \frac{\exp(\beta x)}{1+\exp(\beta x)} \qquad \mathrm{logit}(P) = \beta x \qquad \text{(eq. 5.4)}$$

In this study, we have estimated the coefficients in Equation 5.4 for land use changes at two different time scales. Firstly we have formulated a long-term model to describe the land use conversions that have led to the land use pattern of 1989 in the Netherlands. More recent developments have been studied in a model that describes the conversions in the main land use types between 1989 and 1996, a period in which the Netherlands experienced strong urbanisation.

As a consequence of the regional differences in the biophysical and socio-economic conditions between regions in the Netherlands it is possible that different relations between land use and location characteristics are prevalent in different regions. To verify this, we have conducted region-specific analysis for a number of land use types.

Instead of the binomial logit specification, we could have used a multinomial logit model addressing the probability of conversion of the different land use categories relative to a reference category. However, in this study we are mainly interested in the probability of a land use change at a certain location relative to all other options. Therefore, a dichotomous specification of the regression is chosen. Furthermore, this specification allows different explanatory factors for different types of land use change, which is essential for an appropriate specification of the model.

Stepwise logistic regression was used to estimate the coefficients of the defined model. The dependent variable is a binary presence or absence event, where 1 = presence (occurrence of a land use conversion) and 0 = other. Odds ratios are used to facilitate model interpretations. The odds ratio ($\exp(\beta)$) can be interpreted as the change in the odds for the considered event (i.e. land use change) with an increase of one unit in the corresponding factor.

All analyses were performed using pixels of 500×500 m as the unit of observation. Dependent and independent variables were therefore converted to a raster-based format. Spatial dependence can be expected in models in which relative location is important. Regressions based on dependent variables that exhibit spatial auto-correlation are influenced by the spatial structure (Anselin, 1988, Bell and Bockstael, 2000). In this study, we have minimised the influence of spatial auto-correlation on the regression results by using a sample (10% of all observations) of randomly distributed observations instead of the full data set. The inclusion of neighbourhood characteristics in the model specification explicitly addresses the issue of auto-correlation in land use patterns.

The goodness of fit of the logistic regression models is measured by the ROC (Relative Operating Characteristic; (Swets, 1986, Pontius and Schneider, 2001). The ROC is based on a curve relating the true-positive proportion and the false-positive proportion for a range of cut-off values used in classifying the probability. The ROC statistic measures the area beneath this curve and varies between 0.5 (completely random) and 1 (perfect discrimination).

5.3.2 Data

Land use

We used the 1989 and 1996 maps of land use produced by the Central Bureau for Statistics (CBS Bodemstatistiek). Since 1989 these maps are based on aerial photographs (scale 1:18,000)

Table 5.1 Land use classification used in this analysis

Land use type	Description
Grassland	Grasslands, incl. semi-natural grasslands
Arable land	All arable lands
Greenhouses	Greenhouses
Other agriculture	Agricultural land not belonging to greenhouses, grassland or arable land, includes horticulture, orchards etc.
Residential areas	Residential areas and social-cultural facilities, incl. houses, roads within residential areas, schools, hospitals, churches etc.
Industrial/commercial	Mining areas, industries, harbours, shopping malls, prisons and all service industries
Forest/nature	Forests and natural areas, incl. peat areas, swamps, heather etc.
Recreation	Parks, sport fields, kitchen garden complexes, camp sites etc.
Airports	Airports
Water	Water

and 33 different land use types are distinguished. These maps were checked for inconsistencies by the National Institute of Public Health and the Environment, rasterised to 25×25 m (Raziei and Evers, 2001) and combined with the LGN database ('Landelijk Grondgebruiksbestand Nederland'), another series of high-resolution land use maps produced by Wageningen University and the Research Center (Alterra Green World Research) based on the interpretation of Landsat_TM remote sensing images (de Wit et al., 1999). The integration of both maps made it possible to sub-divide the agricultural land use type (de Nijs et al., 2001a). For this study, the data set was aggregated to a 500×500-m resolution. This aggregation was based on the majority-rule. As such a procedure could yield a bias, which would especially affect land use types with a relative small coverage, causing them to disappear (Milne and Johnson, 1993, Moody and Woodcock, 1994, He et al., 2002), we constrained the aggregation by requiring that there was a correspondence between the total areas of each land use type in both the high- and low-resolution.

Based on this procedure consistent land use maps at 500×500 meter were produced for 1989 and 1996 and reclassified into ten land use types relevant to the analysis presented in this paper. Table 5.1 gives a description of the land use types.

Biophysical characteristics

Biophysical characteristics determine the suitability of a location for agricultural use, adding to the potential profit that can be gained by a certain type of land use at a certain location. In terms of agricultural use, soil conditions are very important. Instead of characterising the soil conditions by the soil unit classification, we have translated the soil units into properties relevant to agricultural and natural land uses, such as the alkalinity, texture and organic matter content of both top- and subsoil. Drainage is characterised by the depth of the groundwater and its fluctuations throughout the year. Another biophysical characteristic included in this study is elevation, which is directly related to agricultural suitability and historical settlement patterns when flooding was more frequent.

Because of the assumed collinearity between the soil variables, we have tested the coefficients of determination (R^2) of the multivariate relationships between one of the variables against all of the others. High correlations were found between the clay and loam contents of the soils and between the calcium contents of top- and sub-soils. In the analysis, the calcium content of the subsoil, the loam content of the topsoil and the clay content of the sub-soil are omitted. Consequently, all multivariate relationships are well below the critical value of 0.80 (Menard, 1995).

Socio-economic conditions

Many socio-economic conditions do not relate to the location itself but, rather, to the regional conditions, such as the presence of employment opportunities in the neighbourhood, the level of facilities among others. The influence of socio-economic conditions in the region can be best characterised by the access that a location has to these facilities. In many studies (Chomitz and Gray, 1996), therefore, the distance to the nearest location of employment or facility is calculated to represent these location factors. Distance by air is often used (Pijanowski et al., 2002, Thompson et al., 2002), without taking into account the layout of the infrastructure network and natural barriers (rivers, lakes etc.). This approach can result in unrealistic distance measures. Therefore, we generated advanced accessibility measures using FLOWMAP to better represent the accessibility as faced by the agents of land use change. FLOWMAP is a software package for spatial analysis with an emphasis on interaction modelling, network analysis and accessibility measures. It allows easy data exchange with GIS packages where the results can be further analysed and visualised using more sophisticated cartographic techniques (de Jong and Ritsema van Eck, 1996).

A common variable in economic models of land use change is the distance between the residential location and employment, as this is a proxy for the costs associated with travel. We have included this proxy by calculating the average distance of a location to the location of employment, identified by 100,000 and 500,000 jobs, respectively, to distinguish between regional and national centres of employment. Economic motives and travel time can also be important factors influencing a decision to live/work at a certain location. Therefore, all accessibility measures are expressed in travel times as well as in distance measures. Another important factor affecting decisions on land use allocation is the accessibility of facilities, including schools, shops, medical and social services among others. We have developed a proxy for this factor by calculating average distances and travel times to the nearest 100,000 and 500,000 inhabitants. Furthermore, we have included accessibility of forest, nature areas, lakes and the coast as a proxy for the ease of access to recreational facilities, which is assumed to positively influence the decision to live somewhere. Because all accessibility measures are calculated based on roads using the car as the vehicle, we have also included the travel time and distance to the nearest railway station as a factor indicating the ease of access to public transport for a location.

Location factors that are important for the allocation decisions of industry and commercial compounds are the ease of access to Amsterdam Airport Schiphol and Rotterdam harbour. Also, the location next to motorways is important for companies, not only because of easy access (as measured by the accessibility to motorway intersections), but also because of visibility.

Another factor that is assumed to influence land allocation decisions is noise. It is expected that the proximity of highways and airports has a negative influence on the 'preference' to use a location for new residential land use due to noise. We have included a map with all locations subject to high noise levels as calculated by the National Institute for Public Health and the Environment. This map is supplemented with a special map showing the noise and safety contours of air-traffic relative to the national airport.

Neighbourhood characteristics
The land use situation in the direct neighbourhood of a location can be an important determinant of land use change. The neighbourhood characteristics can be quantified by the 'enrichment factor' (Verburg et al., 2004a), which(Verburg et al 2003) is defined by the occurrence of a land use type in the neighbourhood of a location relative to the occurrence of this land use type in the study area as a whole, as follows:

$$F_{i,k,d} = \frac{n_{k,d,i} / n_{d,i}}{N_k / N} \qquad \text{(eq. 5.5)}$$

where $F_{i,k,d}$ characterises the enrichment of neighbourhood d of location i with land use type k. The shape of the neighbourhood and the distance of the neighbourhood from the central grid-cell i is identified by d. $n_{k,d,i}$ is the number of cells of land use type k in the neighbourhood d of cell i, $n_{d,i}$ is the total number of cells in the neighbourhood, N_k is the number of cells with land use type k in the whole raster and N is all cells in the raster. Therefore, if the neighbourhood of a certain grid cell contains 50% grass, whereas the proportion of grass in the country as a whole is 25%, we characterise the neighbourhood by an enrichment factor of 2 for grassland. When the proportion of a land use type in the neighbourhood equals the national average, the neighbourhood is characterised by a factor of 1 for that land use type. An under-representation of a certain land use type in the neighbourhood results in an enrichment factor of between 0 and 1.

For each grid cell i, this neighbourhood characteristic results in a series of enrichment factors for the different land use types (k). The procedure is repeated for different neighbourhoods located at different distances (d) from the grid cell in order to study the influence of distance on the interaction between land use types. In this study, we have used square rings at different distances (d) from the central cell as neighbourhood.

Spatial policies
Policies that have spatial manifestations are included as a spatial explicit layer designating the specified areas. Very few new specific spatial policies have been implemented during the period for which land use changes were studied in detail (1989-1996). This period is characterised by a disengagement of the central government from many policies associated with the Dutch welfare state, leading to weak governmental control over housing. Emphasis was placed on promoting compact urbanisation by developing sites within and directly adjacent to cities (Dieleman et al., 1999). During the 1970s and 1980s, so-called growth centres and growth towns were designated to cater for a large portion of the new home construction in order to avoid suburbanisation and overcrowding of the main cities of the Netherlands and to protect the Green Heart as an open space in the Randstad [conurbation of the four largest Dutch cities Amsterdam, Rotterdam, The Hague and Utrecht and the surrounding areas]. This policy of controlled dispersal was dubbed 'concentrated deconcentration' (Faludi and Van der Valk, 1990) (Faludi and Van der Valk 1990) and involved 19 municipalities, two of which are new towns on land reclaimed from the sea. It is assumed that this policy still affected residential construction after 1989. Therefore, we have included the growth centre municipalities as an independent variable.

Table 5.2 Description of spatial explicit allocation factors used in this study

Variable	Description	Source
Biophysical variables		
ALTITUDE	Mean elevation calculated from a high-resolution Digital Elevation Model	Rijkswaterstaat (info on: http://www.minvenw.nl/rws/mdi/geoloket/)
GRWATER	Groundwater dynamics characterised by depth and fluctuation through the year ('grondwatertrappen'); higher categories are better drained soils	1:50.000 soil map of the Netherlands (Steur and Heijink, 1991)
Soil Characteristics	Properties of soils based on the 1:50,000 soil map	1:50.000 soil map of the Netherlands (Steur and Heijink, 1991) combined with physical/chemical characteristics of the difference soil types (Bodemkundig Informatie Systeem, Alterra, Wageningen University and Research Centre).
ORGMAT_TOP	Organic matter content of the top-soil (0-35 cm)	
ORGMAT_SUB	Organic matter content of the sub-soil (35-125 cm)	
CA_TOP	Calcium content of the top-soil (0-35 cm)	
CA_SUB	Calcium content of the subsoil (35-125 cm)	
LOAM_TOP	Loam content of the topsoil (0-35 cm)	
LOAM_SUB	Loam content of the subsoil (35-125 cm)	
CLAY_TOP	Clay content of the topsoil (0-35 cm)	
CLAY_SUB	Clay content of the subsoil (35-125 cm)	
pH_TOP	Acidity of the topsoil (0-35 cm)	
pH_SUB	Acidity of the subsoil (35-125 cm)	
Socioeconomic variables		
D_STATION	Road distance to nearest railway station	National Road Database (BASNET: Basisnetwerk Nederland (1990)); Adviesdienst Verkeer en Vervoer Rijkswaterstaat (AVV); Employment and population statistics: Central Bureau for Statistics
T_STATION	Travel time to nearest railway station	
D_AIRPORT	Road distance to Amsterdam Airport Schiphol	
T_AIRPORT	Travel time to Amsterdam Airport Schiphol	
D_HARBOUR	Road distance to Rotterdam harbour	
T_HARBOUR	Travel time to Rotterdam harbour	
D_WORK1	Average distance to nearest 100,000 jobs	
T_WORK1	Average travel time to nearest 100,000 jobs	
D_WORK2	Average distance to nearest 500,000 jobs	
T_WORK2	Average travel time to nearest 500,000 jobs	
D_HIGHWAY	Distance to nearest motorway intersection	
T_HIGHWAY	Travel time to nearest highway entrance	
D_CITY	Average distance to nearest 500,000 inhabitants	
T_CITY	Average travel time to nearest 500,000 inhabitants	

D_TOWN	Average distance to nearest 100,000 inhabitants	
T_TOWN	Average travel time to nearest 500,000 inhabitants	
D_HISTCITY	Historic travel time index for accessibility to nearest main city (>10,000 inhabitants in 1550)	Network structure derived from (Ministerie van Verkeer en Waterstaat 2001)
D_HISTTOWN	Historic travel time index for accessibility to nearest main town (5,000 – 10,0000 inhabitants in 1550)	
D_RECREATION	Distance to the nearest main forest/water area frequently used for recreation	National Institute for Public Health and the Environment National Institute for Public Health and the Environment
T_RECREATION	Travel time to the nearest main forest/water area frequently used for recreation	National Road Database (BASNET: Basisnetwerk Nederland (1990)); Adviesdienst Verkeer en Vervoer Rijkswaterstaat (AVV);
D_OPENWATER	Distance to the coast or main river (no network)	
NOISE	Areas with a cumulative noise exceeding 50dB(A)	
AIRP_NOISE	Areas with moderate and high noise (20 and 35 Ke zones) due to the Amsterdam Airport Schiphol	
HIGHWAY	Areas located within 750 meters from a motorway	
Neighbourhood interaction characteristics	Enrichment factors for neighbourhood at an average distance of 500 meter and 1800 meter	Calculations based on 1989 land use map
INFL_RES1	Enrichment with residential land use at 500 meter	
INFL_RES3	Enrichment with residential land use at 1800 meter	
INFL_INDUST1	Enrichment with industrial/commercial land use at 500 m	
INFL_INDUST3	Enrichment with industrial/commercial land use at 1800 m	
INFL_FOREST1	Enrichment with forest/nature at 500 m	
INFL_FOREST3	Enrichment with forest/nature at 1800 m	
Policy		
GRCENTRE	Designated municipalities to cater for a large part of new home construction to avoid suburbanisation and overcrowding of the main cities of the Netherlands	
GRHEART	Delineation of the central open space in the western part of the Netherlands ('The Green Heart')	

Another important item of Dutch spatial planning doctrine has always been the Green Heart policy (Dieleman et al., 2002). Spatial policies have been relatively successful in keeping the Green Heart as a central open space surrounded by urban development. We included the delineation of the Green Heart as an explanatory factor in the analysis.

A summary of all data included in the analysis is given in Table 5.2.

5.4 Long-term land use changes

In this section an analysis is made of the long-term changes in land use that have resulted in the land use pattern as it is found nowadays. We have used the landscape pattern of 1989 as a dependent variable. This pattern is the result of many historical developments that have converted the Dutch landscape from an uninhabited, natural landscape with large water-logged areas into a heavily urbanised landscape.

5.4.1 Model definition

Many factors that are commonly used to explain land use change patterns are endogenous to the processes studied at a long time scale. As such, measures indicating the current accessibility can be used when studying land use change for one or several decades as a factor explaining the allocation of new residential or commercial locations. At longer time scales, accessibility and urban expansion are closely linked, with complex cause-effect relations. The same type of reasoning also results in other socio-economic location factors, such as population pressure and labour availability, being unsuitable explanatory factors for the land use pattern. Instead, we assume that the historical development of the land use pattern is closely related to the bio-geophysical conditions of the land. Therefore, we included soil properties and elevation as explanatory factors. We have left out drainage characteristics since throughout almost the entire Netherlands the groundwater dynamics are artificially managed in accordance with the requirements for the specific land use. Such management weakens the connection between land use and the constraints set by the abiotic template (Bürgi and Turner, 2002). Therefore, groundwater data cannot be considered as independent variables in the analysis. Soil properties, especially organic matter content, can be influenced by land use and management as well (Sonneveld et al., 2002). For this part of the study, in which we have derived the soil conditions from the 1:50.000 soil map that is not directly linked to management units, we assume that this effect is negligible.

At the time scale considered, we could include only two variables representing the socio-economic opportunities for urban development. This first is the distance to open water (coast and main rivers), which is seen as a proxy for market access and trade, and which has, historically, always been an important factor for the economic development of the Netherlands. The second is a rough map providing an index of historical accessibility (valid for around 1600) to the important towns. It is assumed that the historical settlement places have triggered further development of the urban land uses due to the advances of larger economies (Arthur, 1994, Krugman, 1998).

Table 5.3 Binary logit estimates for land use patterns in 1989 (Exp(b) values).

	Forest	**Arable land**	**Grassland**	**Residential land**	**Industrial/ commercial land**
CONSTANT	0.14	0.10	0.62	0.70	0.06
ALTITUDE	1.03	*	0.98	*	1.01
SOIL					
ORGMAT_TOP	0.99	1.02	0.996	*	*
ORGMAT_SUB	1.04	0.98	*	*	*
CA_TOP	1.10	1.29	0.72	*	*
LOAM_SUB	0.93	1.02	1.01	1.01	*
CLAY_TOP	*	0.97	1.01	0.95	*
pH_TOP	0.64	*	1.17	*	*
pH_SUB	2.04	0.90	0.86	0.77	*
D_HISTTOWN	-	1.00004	-	0.9999	0.99998
D_HISTCITY	-	-	-	*	0.99999
D_OPENWATER	-	-	-	*	0.99
ROC-statistic	0.82	0.73	0.73	0.67	0.65

- not included in model specification
* not significant; excluded from model specification

The tested logit models for residential and industrial/commercial land are:

$$\log\mathrm{it}(p) = \alpha + \beta_1 SOIL + \beta_2 ALTITUDE + \beta_3 D\_HISTTOWN + \beta_4 D\_OPENWATER \qquad \text{(eq. 5.6)}$$

where *SOIL* includes organic matter, loam content, clay content, pH and calcium content (listed in Table 5.2). For grassland and forest, we have included only the biophysical factors,

$$\log\mathrm{it}(p) = \alpha + \beta_1 SOIL + \beta_2 ALTITUDE \qquad \text{(eq. 5.7)}$$

and for arable land, we also included the distance to historical towns because it is expected that food production would preferably be located close to the town:

$$\log\mathrm{it}(p) = \alpha + \beta_1 SOIL + \beta_2 ALTITUDE + \beta_3 D\_HISTTOWN \qquad \text{(eq.5.8)}$$

5.4.2 Results

Table 5.3 gives the estimated coefficients for the logit models describing the land use pattern for the main land use types in 1989. The presented Exp(*b*) values (odds ratio) indicate the change in the odds of the dependent variable (land use) with a change in the independent variable. Values between 0 and 1 indicate that the probability will decrease with increases in the value of the independent variable, while values above 1 indicate an increase in probability with increases in the value of the independent variable. The ROC values indicate that the spatial pattern of forest/nature can be reasonably explained by the independent variables. The model fit for arable land

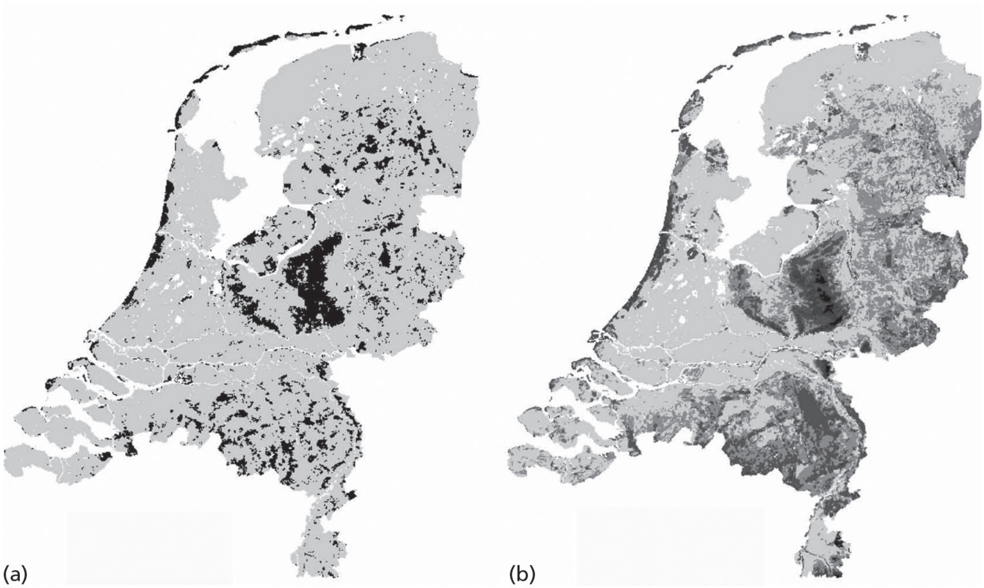


Figure 5.1 (a) Location of forest/nature areas in 1989 (black) and (b) probability for the occurrence of forest/nature based on the logit model (dark colour: high probability)

and grassland is not as good, and the independent variables only explain a small fraction of the spatial variability of residential and industrial/commercial land.

The model fit is illustrated in Figure 5.1 with a probability map for forest/nature based on the logit model in Table 5.3. A number of soil variables and the elevation contribute significantly to the explanation of the spatial pattern of forest/nature. In general, forest/nature is found on the somewhat higher locations, with soil properties that are clearly different from those of arable land, indicating that forest/nature remains in areas unsuitable for agriculture. The association with low pH of the topsoil relates to the podzolisation process that is dominant in most poor sandy soils in the Netherlands. Soil characteristics are important determinants of the distribution of arable land: the main arable areas are located on marine sediments in the northern and south-western part of the Netherlands and within the polder areas (positive contributions of loam and calcium content). An area with intensive potato cultivation is found in the Northern part of the country on reclaimed peat-bogs. After reclamation, the top-soil of these soils has a high organic matter content while the subsoil has a low organic matter content. The positive contribution of a high level of organic matter in the topsoil and negative contribution of low organic matter content in the subsoil make it possible to distinguish this area from the lowland peat areas in the western part of the Netherlands that are unsuitable as arable land.

We are less successful in explaining the spatial distribution of residential area. Few of the selected variables contribute significantly to the model. The positive association with loam content indicates the historical preference to use solid soils, mostly the levees of rivers, instead of the swampy areas in between the rivers. The distance to historical towns is an important factor, however, the low level of explanation indicates that many more processes have been important in the spatial allocation of new residential areas that are not included in the present model. The distance to historical cities and open water does not add to the explanation of residential land distribution but it does have an influence on industrial/commercial land allocation, indicating the importance of access to national and international markets for this land use type. It is possible that the location of residential and industrial/commercial concentrations can be better explained by self-reinforcement of industrial and residential locations through economies of scale and the emergence of new centres (Fujita et al., 1999, Krugman, 1999). Arthur (1994) illustrated that as soon as self-reinforcements or economies of scale are important, chance events become important determinants of the system of cities. Therefore, the observed pattern of residential and industrial locations cannot be explained by determinism alone without reference to chance events, coincidences and circumstances in the past. Chance events can never be adequately included in an empirical analysis of land use. However, the reinforcement of patterns and path-dependence can be studied at shorter time scales. In the next section we present an analysis of land use change between 1989 and 1996 that includes proxies (i.e. neighbourhood characteristics) for processes that lead to path-dependence in the model specification.

Another reason for the low explanatory power of the regression model for urban land use is the importance of time-dependency on factors relating to decisions on land use change. In the historical context, the distance between urban and agricultural areas may have been an important consideration when making land use decisions; however, modern transport facilities and storage capacity have resulted in it becoming an unimportant factor. Furthermore, land use policies as well as socio-economic conditions have seen large changes during the long period over which the land use pattern has developed. Therefore, it is difficult to explain the observed land use changes that have occurred over such a long period of time without explicitly addressing the changes in land use change processes over time.

Region-specific regression models for arable land were estimated for three different regions in the Netherlands (Figure 5.2; Table 5.4). The ROC value was higher for the northern and western regions than for the country as a whole, with different parameters found for the different

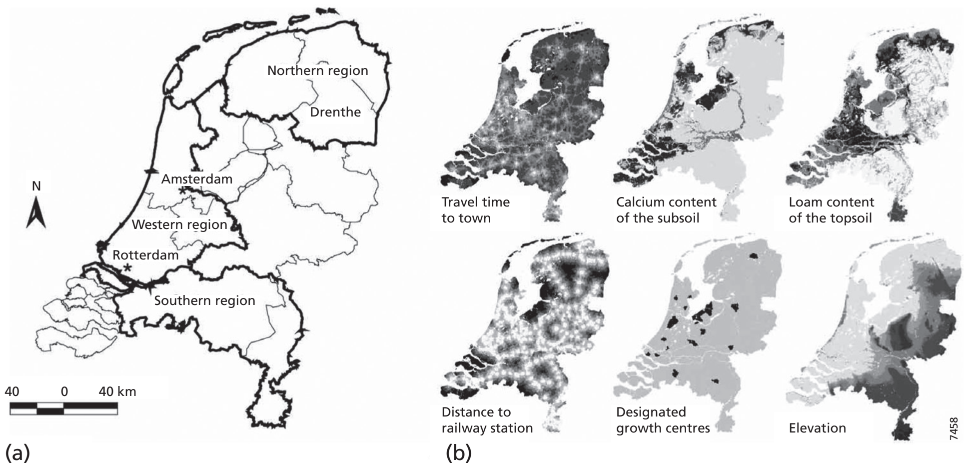


Figure 5.2 (a) Map of the Netherlands with provincial boundaries and regional sub-division and (b) examples of six variables used in this study.

Table 5.4 Binary logit estimates for the spatial pattern of arable land in 1989 (Exp(b) values).

	Western region	Northern region	Southern region
CONSTANT	0.02	0.07	0.11
ALTITUDE	0.78	0.99	*
SOIL			
ORGMAT_TOP	0.97	1.02	0.97
ORGMAT_SUB	0.96	0.97	0.99
CA_TOP	1.40	1.36	1.19
LOAM_SUB	1.02	1.02	1.01
CLAY_TOP	0.98	0.98	*
pH_TOP	*	*	0.95
pH_SUB	*	0.92	1.14
D_HISTTOWN	1.00005	1.00006	1.00002
ROC-statistic	0.89	0.76	0.63

* not significant; excluded from model specification

regions. The organic matter content of the top-soil has a positive relation with the occurrence of arable land only in the northern region (the afore-mentioned reclaimed peat bogs). The spatial pattern of arable land could be described particularly well in the western region because arable land is, in this area, only found on marine sediments. In the southern region, maize cultivation is the most important crop of the arable land use type. The lower dependence of maize cultivation on soil suitability leads to the lower level of explanation for the distribution of arable land in the Southern region.

The results of the regional analysis suggest that spatial disaggregation leads to more specific relations between the location characteristics and land use. Thematic disaggregation of the land use classes will certainly improve the results as well, such an analysis falls outside the scope of this study.

5.5 Recent land use changes: 1989-1996

In this section we analyse the determinants of the changes in land use pattern between 1989 and 1996. In the first part of this section we define and estimate a model that only includes location characteristics. In the second part we expand this model by including neighbourhood characteristics.

5.5.1 Location characteristics

Model definition

Models are defined to explain the major land use changes between 1989 and 1996. These changes include new residential developments (0.75% of the total area of the Netherlands), new areas allotted to industrial/commercial uses (0.49%) and new recreation areas (0.44%). Because a variety of different factors are assumed to determine the location of these land use changes, we estimated a different model for each of these three conversion types. The model for new residential developments is:

$$\operatorname{logit}(p) = \alpha + \beta_1 GRWATER + \beta_2 ALTITUDE + \beta_3 ACCESSIBILITY + \beta_4 NOISE + \beta_5\, POLICY \qquad \text{(eq. 5.9)}$$

ACCESSIBILITY includes accessibility measures that represent the distance and/or travel time to employment opportunities, urban facilities, infrastructure and recreation. We considered the following variables: D_STATION, T_STATION, D_WORK1, T_WORK1, D_WORK2, T_WORK2, D_HIGHWAY, T_HIGHWAY, D_CITY, T_CITY, D_TOWN, T_TOWN, D_RECREATION AND T_RECREATION (Table 5.2). POLICY variables include the growth centre policy (GRCENTRE) and the 'Green Heart' policy (GRHEART). Noise was expected to negatively influence the decisions to plan new residential areas. We included cumulative noise (NOISE) and noise and safety contours for the national airport (AIRP_ NOISE).

The estimated model for new industrial developments is:

$$\log\mathrm{it}(p) = \alpha + \beta_1 GRWATER + \beta_2 ALTITUDE + \beta_3 ACCESSIBILITY + \beta_4 POLICY + \beta_5 HIGHWAY \qquad \text{(eq. 5.10)}$$

where ACCESSIBILITY includes a somewhat different set of variables, including accessibility to the airport and Rotterdam harbour as potential factors determining the location of new industrial/commercial compounds. The full list of variables that represent the accessibility includes: D_STATION, T_STATION, D_WORK1, T_WORK1, D_WORK2, T_WORK2, D_HIGHWAY, T_HIGHWAY, D_CITY, T_CITY, D_TOWN, T_TOWN, D_AIRPORT, T_AIRPORT, D_HARBOUR and T_HARBOUR. In addition to the variables that describe the accessibility to infrastructure, we include a variable (HIGHWAY) that indicates the zone of 750 m along the main motorways that represents the preference to locate new industrial/ commercial compounds as much as possible near to the main infrastructure. For new recreation areas:

$$\log\mathrm{it}(p) = \alpha + \beta_1 ACCESSIBILITY + \beta_2\, POLICY \qquad \text{(eq. 5.11)}$$

where ACCESSIBILITY includes the distance/travel time to residential areas, existing recreation areas and infrastructure (D_HIGHWAY, T_HIGHWAY, D_CITY, T_CITY, D_TOWN, T_TOWN, D_RECREATION AND T_RECREATION). Bio-geophysical variables were supposed not to influence the allocation of new recreation areas.

Locations with land use classified as residential, industrial/commercial, airport or nature/ forest in 1989 were considered not to be potential locations for new residential or industrial/ commercial development and therefore excluded from the analysis. For the analysis of new recreational areas, the locations designated for recreational use in 1989 were also excluded from the analysis.

Table 5.5 Average characteristics of locations with and without new residential, industrial/commercial and recreational development, 1989–1996

Variable	Residentrial				Industrial/commercial				Recreation			
	without (n=104402)		with (n=1009)		without (n=104752)		with (n=559)		without (n=101279)		with (n=593)	
	average	SD	average	SD	average	SD	average	SD	average	SD	average	SD
D_WORK1	20.14	7.33	14.29	6.45	20.11	7.34	15.18	6.79	20.27	7.23	15.99	8.15
D_WORK2	41.9	15.6	32.45	13.47	41.86	15.6	32.9	13.98	42.16	15.46	33.23	16.63
T_WORK1	24.26	11.8	17.69	6.76	24.24	11.79	18.37	8.23	24.38	11.59	19.75	9.06
T_WORK2	42.25	17.37	32.98	11.76	42.22	17.36	33.4	12.63	42.46	17.17	34.67	15.59
D_TOWN	12.7	4.51	8.38	3.78	12.68	4.52	8.98	4.21	12.79	4.44	9.84	4.85
D_CITY	24.26	8.51	18.12	7.1	24.23	8.51	18.97	7.43	24.409	8.41	19.74	9.4
T_TOWN	16.58	8.88	11.62	3.95	16.56	8.87	12.1	5.2	16.65	8.64	13.3	5.4
T_CITY	27.93	13.1	21.13	7.44	27.9	13.09	21.83	8.6	28.05	12.91	23.32	10.41
D_HIGHW	11.81	8.48	7.1	6.67	11.8	8.48	7.61	7.75	11.91	8.46	8.4	7.55
T_HIGHW	15.26	13.65	9.36	7.87	15.24	13.64	9.69	9.17	15.36	13.52	11	9.68
D_STAT	10.98	6.99	6.68	5.95	10.96	6.99	7.58	6.56	11.06	6.96	8.49	6.91
T_STAT	16.75	12.91	11.02	7.87	16.72	12.9	12.11	9.03	16.82	12.78	13.42	9.39
GRCENTRE	2.95%		13.78%		3.00%		11.53%		2.87%		9.78%	
GRHEART	5.79%		2.68%		5.77%		3.79%		5.85%		3.04%	
GRWATER(1)	0.68%		0.69%		0.67%		1.06%		0.66%		1.01%	
GRWATER(2)	11.63%		10.01%		11.61%		11.53%		11.50%		11.90%	
GRWATER(3)	40.77%		34.79%		40.73%		37.33%		41.08%		32.30%	
GRWATER(4)	33.44%		31.02%		33.40%		35.05%		33.69%		32.00%	
GRWATER(5)	13.49%		23.49%		13.57%	65.59	15.02%	61.94	13.07%		22.60%	
AIRP_NOISE(1)	0.96%		2.57%									
AIRP_NOISE(2)	0.31%	8.67	0.29%	7.72								
D_RECR	16.25	9.5	13.91	8.13								
T_RECR	20.36		17.59									
NOISE	17.60%		44.20%									
D_HARB					128.6	45.64	101.19	41.88				
T_HARB					96.72	48.34	74.78	51.46				
D_AIRP					121.98	37.46	103.8	37.6				
T_AIRP					98.24		82.33					
HIGHWAY					15.33%		40.36%					

Results

The characteristics of locations chosen for development and the locations not chosen are calculated for each of the three types of land use change (Table 5.5). Existing urban locations and forest/nature areas are excluded from the analysis. Accessibility is important for all three land use changes studied: for all accessibility measures, differences are observed between the accessibility of developed and not developed locations. Development was higher with decreasing distance and shorter time to a city or town, job opportunities and road and rail infrastructure. It should be noted that while all accessibility measures used in this study are positively correlated, they still exhibit clear differences in spatial variability and pattern. While it was expected that new developments would avoid noisy locations, our results reveal that this is only true for the immediate neighbourhood of the national airport (AIRP_NOISE(2): 35 Ke zone); in other areas, new developments occurred at locations that have higher levels of noise than comparable locations elsewhere. This is a direct consequence of concentrations of population and the infrastructure needed to serve this population. Almost 14% of all new residential areas were constructed within the growth centres that were part of the 1980s housing policy. This result illustrates that even in the period considered in this study, there was still a tendency to concentrate housing in these growth centres. The growth centres are also important for the allocation of new commercial/industrial compounds, partly because of the attractive industrial location of some of the growth centres and the incentives provided by the municipalities where the growth centres are located (Faludi and Van der Valk, 1994). For commercial/industrial land, the influence of the proximity of a highway was a main determining factor. Not only the distance or travel time to a motorway intersection was important, also the location in the immediate neighbourhood of the motorway. Forty percent of all new industrial/commercial area was within 750 meter of a motorway while this was only 15% for locations not developed into industrial/ commercial land.

The results are summarised in logit models that were estimated independently for the three land use changes considered in this study. Table 5.6 reports the estimated coefficients and odds ratios for these models. Only a small number of variables was actually included due to multi-collinearity. Based on the supposed lack of causality in the relation between residential

Table 5.6 Binary logit estimates of land use changes between 1989 and 1996 (N~15750; thresholds for stepwise regression on 0.02/0.03)

	New residential area		**New industrial/commercial area**		**New recreation area**	
	Coefficient	**Exp(b)**	**Coefficient**	**Exp(b)**	**Coefficient**	**Exp(b)**
Constant	-1.34		-3.02		-2.58	
HIGHWAY	-	-	1.05**	2.84	-	-
D_TOWN	$-2.3\ 10^{-4}$**	0.9998	-2.0010^{-4}**	0.9998	$-2.40\ 10^{-4}$**	0.9998
D_STAT	$-9.8\ 10^{-5}$**	0.9999	-	-	-	-
GRCENTRE	0.69**	1.98	-	-	0.82*	2.27
GRHEART	-1.45**	0.23	-	-	-	-
ROC	0.81		0.80		0.75	

*significant at 0.05 level
**significant at 0.01 level

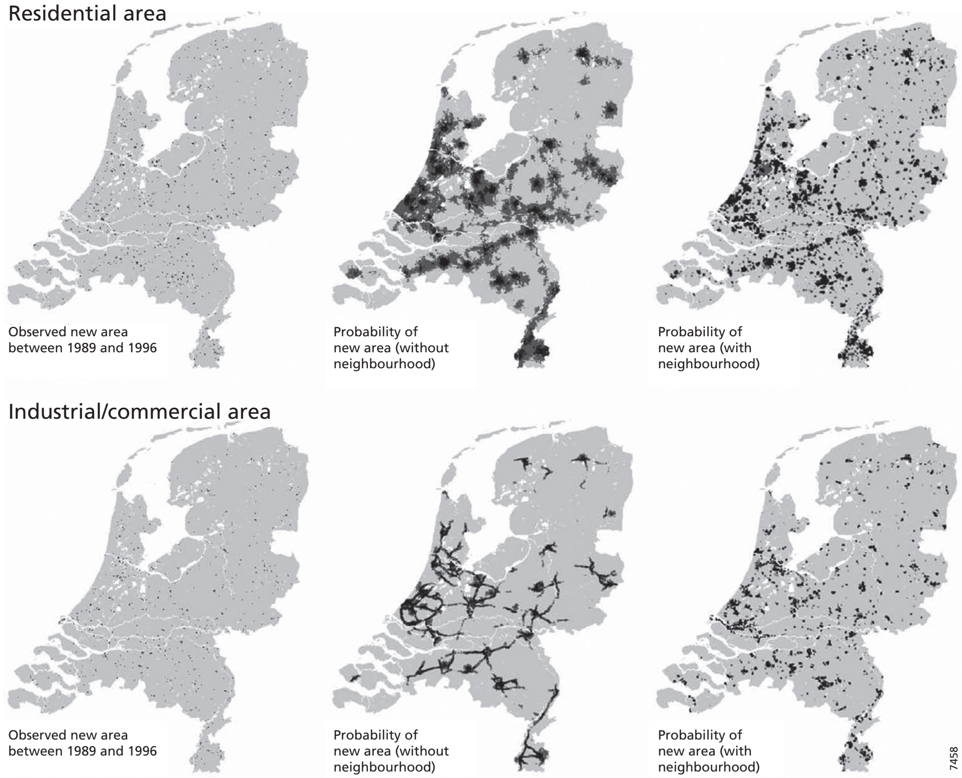


Figure 5.3 Observed changes in residential and industrial/commercial land use between 1989 and 1996 and probabilities based on the logistic regression model without and with neighbourhood factors.

land use and NOISE, we excluded this variable from the model. Variables that are significant in the model for new residential area relate to the biophysical conditions (dryer locations, lower reclamation costs), accessibility and policy. Only accessibility and location near highways were of significant importance for new industrial/commercial developments, while location close to concentrations of population and the growth centre policy were important for new recreational area. The probability maps for new residential area and new industrial/commercial area are shown in Figure 5.3. This figure shows that many more locations are indicated to have a high probability for conversion than have actually been converted. A considerable number of these locations have, however, already been converted to residential or industrial land uses and are therefore no longer candidates for conversion.

The logistic regressions mentioned above are based on an analysis of the Netherlands as a whole. Regional specific conditions are expected to influence the relations between location factors and land use change. Region-specific logit models for new residential and industrial/commercial area in the different regions are given in Tables 5.7 and 5.8. There are differences in the variables that prove most significant in explaining the pattern of land use change in the

Table 5.7 Residential area: estimates for the coefficient of a logit model of changes between 1989-1996 in different regions

	Whole country	Western region	Northern region	Southern region
Constant	-1.34	-1.39	-0.62	0.57
D_TOWN	$-2.3\ 10^{-4**}$	$-2.7\ 10^{-4**}$	-	$-4.4\ 10^{-4**}$
T_TOWN	-	-	-0.24^{**}	-
T_WORK1	-	-	-	-
D_WORK2	-	-	-	$-3.0\ 10^{-5**}$
D_STAT	-9.810^{-5**}	$-6.9\ 10^{-5**}$	-	-
T_STAT	-	-	-0.09^{**}	-
T_HIGHWAY	-	-	-	-
GRCENTRE	0.69^{**}	0.70^{**}	-	-
GRHEART	-1.45^{**}	-1.19^{**}	-	-
ROC	0.81	0.82	0.81	0.81

**significant at 0.01 level

Table 5.8 Industrial/commercial area: estimates for the coefficient of a logit model of changes between 1989-1996 in different regions

	Whole country	Western region	Northern region	Southern region
Constant	-3.02	-4.08	-4.29	-1.107
ALTITUDE	-	0.09^{**}	-	-
HIGHWAY	1.05^{**}	1.12^{**}	0.81^{**}	0.513^{**}
D_TOWN	$-2.3\ 10^{-4**}$	-	$-3.5\ 10^{-4**}$	-
T_TOWN	-	-0.10^{**}	-	-0.233^{**}
D_WORK1	-	-	$3.0\ 10^{-4**}$	-
D_WORK2	-	$-8.0\ 10^{-5**}$	-	-
T_CITY	-	0.11^{**}	-0.126^{**}	-
D_STAT	-	-	$-7.5\ 10^{-5**}$	-
D_AIRP	-	-	-	$-5.8\ 10^{-6**}$
GRCENTRE	-	0.89^{**}	3.317^{**}	-
ROC	0.80	0.75	0.82	0.75

**significant at 0.01 level

different regions. Groundwater level is a determinant of the location of new urban areas when the country is analysed as a whole; in contrast, in the individual regions, groundwater level does not significantly add to the explanation of the land use pattern. In the individual regions, different accessibility measures are highly capable of explaining the new residential locations; however, despite these differences, it is clear that new residential areas are preferably located with easy access to towns, cities and infrastructure. The growth centre and Green Heart policies only have a significant influence on the land use pattern in the Western region – where most growth centres are located. Similar observations hold for the regional variability in location factors determining the allocation of new industrial/commercial areas. In all regions, new industrial/

commercial areas are preferably located in the immediate neighbourhood of motorways, allowing easy access to consumers and labour forces. The growth centres in the western and northern regions attract industrial/commercial compounds. The variability explained by the regression models in all regions was found to be similar to the level of explanation for the country as a whole.

5.5.2 Location and neighbourhood characteristics

We analysed the neighbourhood interactions between land uses through the quantification of possible interactions between the land use types for locations with observed changes that occurred during 1989 and 1996. The enrichment factor was calculated for the neighbourhood of all locations with new residential, industrial/commercial and recreational areas based on the neighbouring land uses in 1989. Figure 5.4 presents the logarithm of the enrichment factor as a function of the distance from the location of land use change. Positive values indicate an enrichment of the neighbourhood of the location with the land use type, while negative values indicate that relatively few of the studied land use types are found in the neighbourhood. Distance-decay functions are found for all land use changes, indicating a decreasing deviation of the neighbourhood from the average land use composition. The interaction between residential

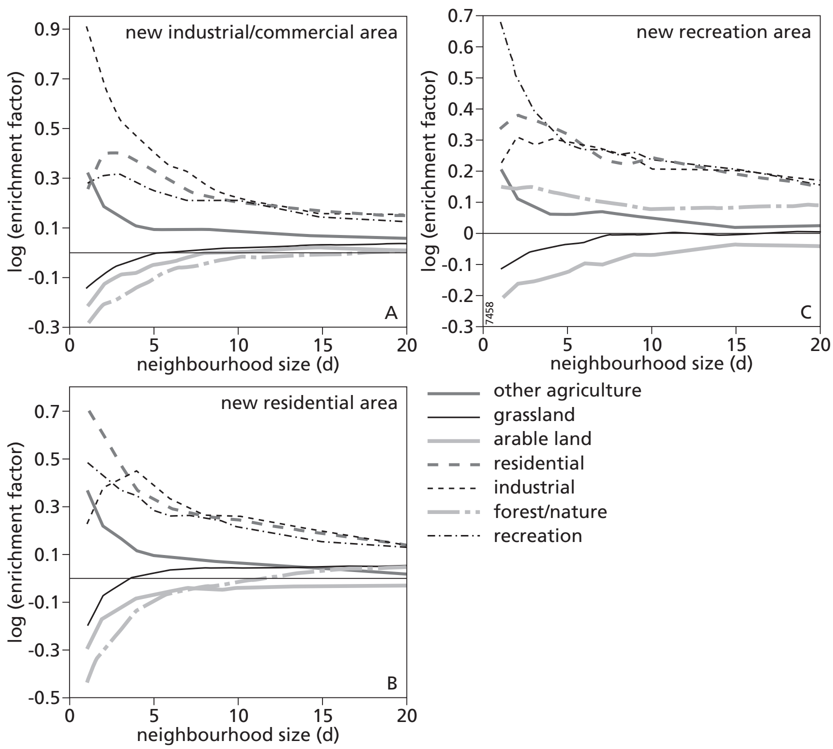


Figure 5.4 Neighbourhood characteristics (log enrichment factor) as a function of the neighbourhood size (d; number of grid cells of 500 m each) for locations converted into (a) industrial/commercial land, (b) recreational area and (c) residential area between 1989 and 1996.

area and industrial/commercial area does not follow a straightforward distance-decay function. Although a positive interaction between residential and industrial uses is found at all distances, the highest values are found at distances between 1 and 2 km (d=2 or 3) from the location itself. This indicates the tendency of industrial/commercial areas to locate in the neighbourhood of residential areas, but preferably in different compounds. The positive autocorrelation for residential and industrial uses indicates that urban growth (reinforcement of existing cities) is more important than new developments far from existing residential areas. The negative association of new residential and industrial developments with grassland, arable land and forest in the immediate neighbourhood confirms this observation. New recreation areas are positively associated with forest/nature in the neighbourhood, indicating that part of the new recreational facilities are created near to forest/nature areas that also have a recreational function. Other recreational areas are located in the immediate neighbourhood of residential areas, thereby providing facilities, such as sports fields, to the inhabitants.

Based on the results of this exploration, for each of the studied land use changes, we selected a number of interactions which were assumed to have a causal effect on the land use allocation. These factors were included in the logistic regression models of Table 5.6, which describe the location of land use change based on location characteristics. We included the influence of the immediate neighbourhood (d=1; 500-700 m) and the neighbourhood over which the interaction between industrial/commercial and residential area is maximal (d=3; 1500-2100 m). The influence of industrial and residential locations over larger distances is captured by the accessibility measures. For new residential land and for new industrial/commercial land, we included the neighbourhood enrichments with residential and industrial/commercial area;

Table 5.9 Estimation results of logit models including neighbourhood characteristics. Only the direction and significance of coefficients for the added neighbourhood variables are presented.

	New residential area	**New industrial/ commercial**	**New recreational**
Residential			
Close (INFLL41)	$+^{***}$	-	$+^{***}$
Distant (INFLL43)	-	$+^{*}$	+
Industrial/Commercial			
Close (INFLL51)	+	$+^{***}$	
Distant (INFLL53)	-	+	
Recreation			
Close (INFLL71)			$+^{***}$
Distant (INFLL73)			$+^{***}$
Forest/Nature			
Close (INFLL61)			$+^{**}$
Distant (INFLL63)			$+^{**}$
ROC (resp. without and with neighbourhood interaction)	0.81/0.90	0.80/0.91	0.75/0.85

*significant at 0.1 level
**significant at 0.05 level
***significant at 0.01 level

for new recreation area, we included the enrichment with residential, recreational and forest/ nature area. Table 5.9 summarises the results of adding these neighbourhood characteristics to the regression models. Not all neighbourhood factors significantly contribute to the model. However, in all cases, the variability explained by the model increased considerably, as indicated by the ROC values. Figure 5.3 shows that the inclusion of the neighbourhood characteristics also leads to better probability maps for the location of land use change.

5.6 Discussion and conclusion

The logit models for the location of land use change estimated in this paper indicate which processes are important to land use change in the Netherlands. Taken as a whole, the results show a close correspondence with our general understanding of the processes of land use change in the Netherlands. Historical land use pattern in Netherlands can, for a considerable part, be explained by the conditions of soil and landform. In particular, the location of agricultural land use and areas left to forest or nature can be explained by the suitability of the soil for agricultural purposes. Recent land use conversions, studied for the period between 1989 and 1996, are no longer related to the biophysical properties of a location; instead, accessibility, spatial policies and neighbourhood interactions are more important determinants of current land use changes. These factors are, to some extent, related to the historical developments in land use pattern. Recent developments in land use, therefore, follow a process of self-organisation in which centripetal forces lead to the growth of existing residential centres.

Residential land use patterns as observed in 1989 can not be explained very well by the spatial variation in location characteristics although it is possible that we missed some of the more important determinants of the suitability of the different locations for residential land use. It is also clear that location characteristics do not fully explain land use patterns: geography can engage in a process of self-organisation in which locations with seemingly identical potential end up playing very different roles (Krugman, 1999, Venables, 1999). However, the most important reason for the low level of explanation is likely related to the fact that a static analysis over a long time-span is not an appropriate method for representing the temporal dynamics leading to land use change. The method introduced here, therefore, is much more suitable for analysing the determinants of short-term changes in land use pattern. At this latter time-scale, the effect of self-organisation can be analysed by including variables that are directly related to the process of self-organisation (neighbourhood interactions, accessibility). In our analysis, these variables were identified to be significant contributors to the explanation of the geographical pattern of land use change with relatively high levels of explanation of the spatial variability. This result indicates that a land use change analysis based on neighbourhood characteristics and accessibility change is relevant and, as such, an important topic for further research. Therefore, CA may be suitable tools in models aimed at simulating these land use change processes.

The results of our study indicate that policies have an important influence on land use patterns in the Netherlands. Policy variables have not been the focus of attention in many studies because they are difficult to include in a quantitative assessment. However, our study indicates that omitting policy variables may cause an incomplete assessment. In the Netherlands, the growth centres have contributed to a large portion of the new residential developments. In our study period, 1989-1996, national policy changed in that the official growth centre policy came to an end, and most attention turned to reinforcing the role of the larger cities, using open spaces within the city or on the edge of the cities as preferential places for new residential and industrial use (Faludi and Van der Valk, 1990, 1994). This policy is still visible in the results of our analysis, as evidenced by the importance of parameters that represent the distance to the cities. However, it is impossible to determine, based on our analysis, if these relations are a direct effect of policy intervention or a consequence of other socio-economic processes that cause a concentration of residential and industrial/commercial areas.

Our analysis of the Netherlands can be summarised by the observation that multiple factors, derived from different theoretical frameworks, are needed to explain the pattern of land use change. One single disciplinary theory would have been insufficient to adequately address the complexity of the system.

The methods used in our study enable a simple, straightforward analysis of changes in land use pattern within the framework of geographic variability in location and neighbourhood characteristics. The results provide information that will both reveal and further our understanding of the processes that underlie decisions directly related to land use change. However, empirical relations do not straightforwardly translate into processes: different processes can lead to the same spatial pattern, and the same process can lead to different spatial patterns, such as by the amplification of chance events. Nevertheless, empirical analysis can suggest the mechanism and quantify the relation between land use and proximate factors. In-depth research, based on the results of empirical analysis, is necessary to reveal causality. In-depth studies on specific determinants of land use change for the Netherlands have been carried out by (Dieleman et al., 2002, van der Vlist et al., 2002), among others. Such research should also explicitly address the issue of spatial scale as it is incorrect to assume that results from an analysis based on area-level aggregate values apply at the individual level. Incorrect cross-level inferences, the so-called the 'ecological fallacy' (Robinson, 1950, Tranmer and Steel, 1998) are a common phenomenon in studies of complex systems.

The method described here can be of prime importance to modellers developing decision-support systems for land use planning and policy. Many land use models use a single set of factors, assumed to be of importance in the specific case-study, to forecast the location of future land use conversions. Most often, this selection of factors and processes assumed to be important is made without case-specific analysis and justification (Torrens and O'Sullivan, 2001). This lack of precision can result in a high level of uncertainty in land use projections due to the lack of causality in the underlying relations on which the extrapolations are based. When models are to be actively used in an impact analysis of future land use changes, planning and policy-making, theoretical understanding and empirical testing should underlie the specification of the simulation model. For specific case-studies, our method can be used to identify which processes and factors are of importance for the specific case-study and time period being analysed. Such information provides a suitable basis for specifying variables and identifying functional relations in a simulation model of land use change. In this context, our method can contribute to improved models of land use change and decision support.

6 The Map Comparison Kit

Reprinted from *Environmental Modelling and Software*, 21. Visser, H. & De Nijs, A. C. M. The Map Comparison Kit., pages 346-358.

Abstract

The comparison of maps is an important issue in environmental research. Maps are compared for many reasons: (i) to detect temporal/spatial changes or hot-spots (ii) to compare different models, methodologies or scenario's, (iii) to calibrate, validate land use models, (iv) to analyse model uncertainty and sensitivity and (v) to assess map accuracy. This paper addresses the quantification of map similarities and dissimilarities using the Map Comparison Kit (MCK) software. Software and documentation are publicly available on the RIVM website free of charge (www.rivm.nl/milieu/modellen). The main focus is on 'categorical' or 'nominal' maps. Four different nominal map-comparison techniques are integrated in the software. Maps on ordinal, ratio and interval scale can be dealt with as well. The software is unique in having two map comparison techniques based on *fuzzy-set calculation rules*. The rationale is that fuzzy-set map comparison is very close to human judgment. Both fuzziness in location and fuzziness in category definitions are dealt with in the software.

6.1 Introduction

The rapid growth of high-resolution spatial modelling, geographical information systems and remote sensing in recent decades has increased the need for map comparison methods. The importance of map comparison methods has been recognised and has stimulated growing interest among researchers (Monserud and Leemans, 1992, Metternicht, 1999, 2000, Winter, 2000, Pontius and Schneider, 2001, Power et al., 2001, Hagen, 2003, Pontius et al., 2004a).

In general, maps are compared for a number of reasons, four of which are listed here. First, we may want to compare maps generated by models under different scenarios and assumptions. In terms of a set of land use maps, we are interested in questions such as 'how similar are the maps?' or 'for which land use category are both maps most dissimilar?' Second, we may want to detect temporal changes. Since many maps have a temporal dimension, a spatio-temporal analysis may yield insight into economic and demographic developments. Third, we may want to calibrate/validate land use models. Land use models such as the Environment Explorer (Engelen et al., 2003) and the Land Use Scanner (Borsboom – van Beurden et al., 2002) generate land use maps starting with an observed land use map. It is important to know how well these models predict future developments and how we can optimise model output with the unknown parameters in the model. For such calibration problems we need an objective measure of map (dis)similarity. In fact, map comparison may be seen as finding a Goodness-of-fit measure. Finally, we may want to perform uncertainty and sensitivity analyses. There are many sources of errors in maps. By comparing model output to a reference map, we can detect and quantify such errors. The same holds for the accuracy of satellite-based maps. In fact, calibration is a form of uncertainty analysis.

In this article we review a software package that can be used for maps on different measurement scales, i.e. maps on nominal, ordinal, interval or ratio scales (definitions from Stevens, 1946). Nominal maps have the simplest and most elementary type of measurement scale, where objects/categories are only discriminated from another. However, the comparison of such maps is the most complicated of comparisons: for example, how can we compare the nominal object 'grassland' with 'residential', or 'glasshouses' with 'recreational'? Despite this difficulty, it is important to be able to have an exact quantification of differences between maps containing such categories. Ordinal maps are characterised by the property of order: we specify both the differences between objects/categories and the direction of those differences. Maps on the interval and ratio scale are based on continuous variables. They allow for an exact quantification of differences. The difference between an interval scale and a ratio scale is the presence of a fixed zero point in the latter scale (e.g., Fahrenheit temperatures are on an interval scale, while Kelvin temperatures are on a ratio scale).

The software introduced here incorporates four nominal map-comparison techniques, denoted as 'Per category', 'Cell by cell', 'Fuzzy Inference System algorithm' and 'Fuzzy Set algorithm', respectively, one ordinal method, denoted as the Fuzzy Set algorithm – advanced and six comparison methods for interval and ratio scale maps. The rationale behind this new software is the combination of standard cell-by-cell map comparison (the Kappa statistic and newly derived variants on Kappa) and recent developments in fuzzy set map comparison (the Kappa fuzzy). Up to now, there has been no software available for the latter techniques. A strong point of a fuzzy map comparison is that it resembles the way human observers compare maps.

A simple example is that of two checker boards, with one board turned around 90 degrees. If we regard the boards as maps with only two categories ('black' and 'white'), a cell-by-cell comparison will reveal that both maps are totally different. However, a human observer would judge both patterns as being highly similar. The same holds for fuzzy similarity. An in-depth experiment illustrating the similarity between fuzzy comparison and human judgement was reported by Hagen (2002a) for an Internet experiment showing ten map pairs (Kuhnert, 2002). A total of 127 participants assessed the map similarity (a figure between 0 and 1, with 0 = totally different and 1 = totally equal). The 'fuzzy Kappa', which we will describe in section 6.3, was strongly correlated to the human judgements ($R^2 = 0.88$).

The software described here is called the Map Comparison Kit or MCK for short. The MCK is a software package for 'state of the art' map comparison developed by the Research Institute for Knowledge Systems (RIKS) and the Netherlands Environmental Assessment Agency (MNP/RIVM). The MCK software was originally designed for the analysis of land use maps. However, the software is employable for many more GIS applications than these types of models (such as remote sensing, ecological models predicting the presence of plant or animal species, etc). In fact, the software yields general methods for pattern recognition.

A humoristic example of the latter ability is given in Figure 6.1. Here two 'maps' are analysed which look very similar but contain 15 major differences. Furthermore, the maps are not

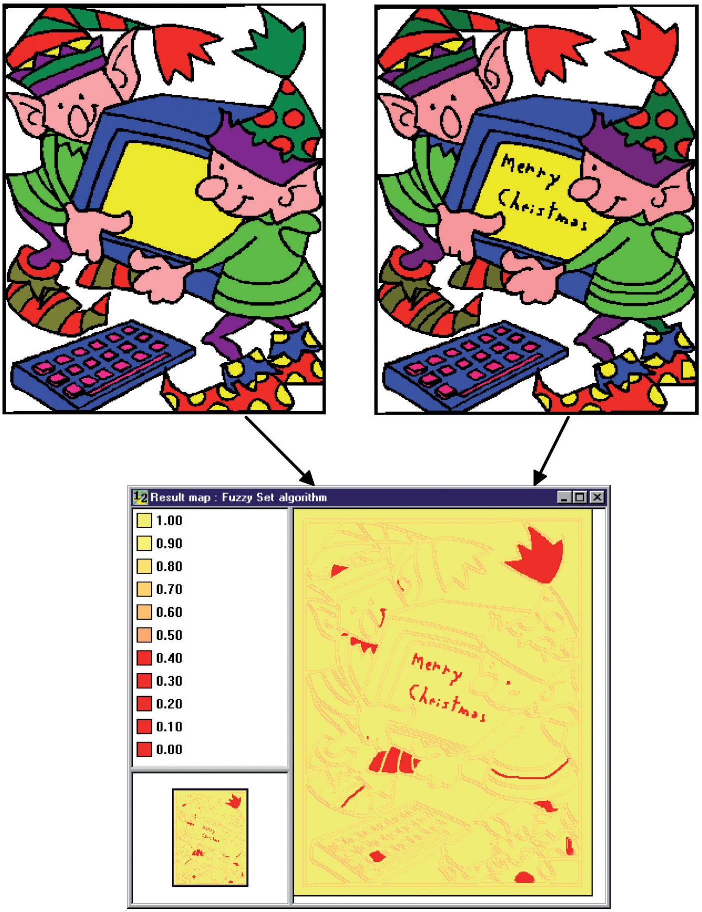


Figure 6.1 The upper panels show two bitmaps of a pair of hobbits The task is to spot 15 differences in the drawings ('maps'). Although not visible from these drawings, we shifted the original drawing, on the right, a few grid cells to the left. The lower panel is a fuzzy output map showing the differences between the two drawings. The orange contour lines are due to the shift of the right map. The panel easily discriminates the 15 differences we are looking for (clustered grids in the colour red).

co-registered because we shifted the complete original second map a few grids to the right (not visible to a human observer!). The fuzzy difference map corrects for the lack on co-registration and easily identifies the 15 differences.

This article provides an overview of the MCK software (section 6.2.1). We give examples of the software screens and on how to choose different scales and methods (section 6.2.2). The rationale underlying the different approaches present in the software is concisely dealt with in section 6.3. This is followed by an application of the software for land use maps (section 6.4). Finally, an evaluation of the software along with plans for the future is given in section 6.5. We draw the reader's attention to the fact that this article is not meant as a research article on map comparison techniques. The scientific background of MCK has been published by Hagen (2002b, 2002a, 2003) and Power (2001) and summarised by Visser (2004).

6.2 The Map Comparison Kit

6.2.1 Software

The first version of the MCK dates back to 1992 and was intended for analysing a series of land use maps. From 1992 onwards, the tool was steadily developed and refined as part of RIKS projects for various institutions. The latest extension of the MCK was developed in collaboration with and by order for the account of the Netherlands Environmental Assessment Agency (MNP/RIVM).

The MCK is designed for the analysis and comparison of raster maps, as illustrated in Figure 6.2. Maps may be on nominal, ordinal, interval or ratio scales. The software can also handle mixtures of nominal and ordinal scales.

In addition to its map comparison function, the MCK offers options to import/export comparative file formats (ArcInfo, Idrisi) and organise and visualise raster maps. The standard Windows copy functionality allows maps and statistics to be pasted in, for example, a MS-Office document in Word. A number of map datasets have been added to the MCK software

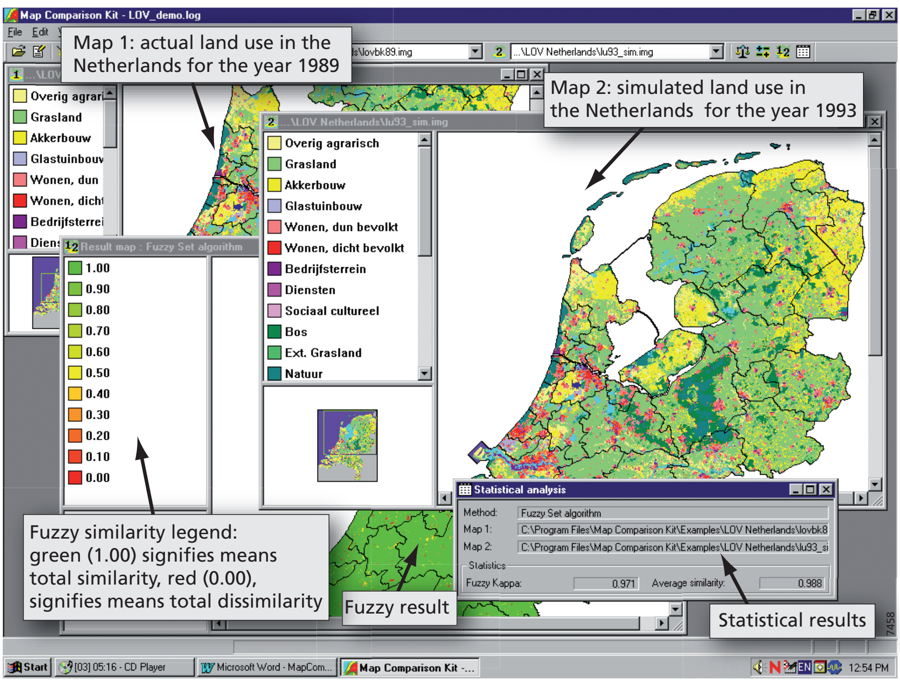


Figure 6.2 Impression of MCK user interface. Shown are two land use maps for the Netherlands, produced by the NEAA Environment Explorer. A fuzzy result map and corresponding statistical results are also shown.

to increase the user's understanding of the methods. Software and documentation can be downloaded from www.rivm.nl/milieu/modellen.

In addition to the stand-alone version, a Map Comparison Module, containing the 'Cell-by-cell', the 'Fuzzy Inference System' and the 'Fuzzy Set' algorithm is available (Uljee, 2003). The module can be incorporated into GIS packages such as ArcGis-8 and has the advantage that:

- it does not require the export and import maps,
- it can employ the most attractive features of GIS, such as zooming in on a small area and then carrying out the calculations within this sub-area,
- it facilitates an easy programming of a batch environment for MCK calculations where many map pairs occur.

6.2.2 Handling different measurement scales

The main window for selecting a comparison method is shown in Figure 6.3. There are ten comparison methods in total among the four measurement scales described in section 6.1. The upper box is for nominal maps and applies to the four methods, varying from 'Per category' and 'Cell-by-cell to 'Fuzzy Inference System' and 'Fuzzy Set algorithm'. The rationale behind these fuzzy methods is given in section 6.3.

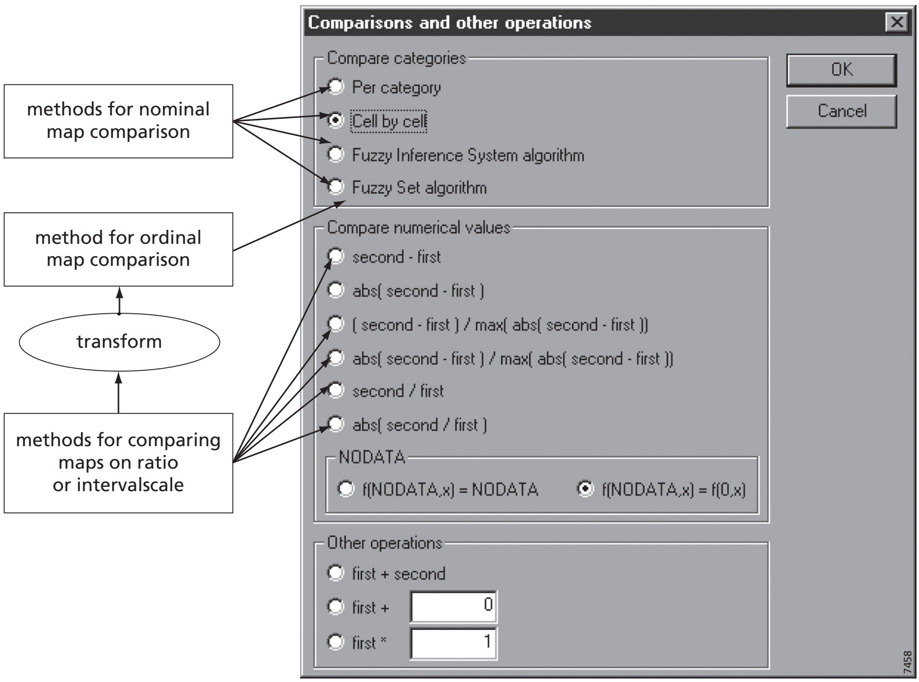


Figure 6.3 MCK algorithms window from which to choose one of ten map-comparison methods. The 'NODATA' and 'Other operations' lines define operations for composing a new map based on existing maps.

The 'Fuzzy Inference System' has an advanced screen on which default settings may be altered. This screen is shown in Figure 6.4. The upper part of the screen is for choosing fuzziness in location, and the lower part for introducing fuzziness in category definitions:

- The calculation of the fuzziness of location is based upon the notion that the fuzzy representation of a cell depends on the cell itself and, to a lesser extent, on the cells in its neighbourhood. The extent to which the neighbouring cells influence the fuzzy representation is expressed by a distance decay function. Three such functions are shown in Figure 6.5: a cone or 'linear decay' (defined by slope), an exponential decay (defined by halving distance) or a three-dimensional (3-D) Gauss curve (defined by variance) (see Bandemer and Gottwald, (1995). Which function is most appropriate and which radius of neighboorhood should be chosen depends on the nature of the uncertainty, vagueness of the data and the observer's tolerance for spatial error. From a theoretical point of view, there is no 'best alternative'. Hence, it is worthwhile to experiment with the size and form of the function.
- The Category Similarity Matrix can be applied to nominal maps where category definitions are (partly) unequal (i.e. maps with unequal legends) or vague. In fact, we may perform a semantic rather than a geometric map comparison by using vague category definitions. Vagueness can be introduced by setting off-diagonal elements in the Category Similarity Matrix to a number between 0 and 1. We note here that the approach above is, in principle, not meant for semantic map comparisons. The reader is referred to Rodriguez, Egenhofer and Rugg (1999) and Kokla and Kavouras (2001) for more details on semantic map comparisons.

Ordinal maps can be examined with the Fuzzy Set algorithm using the Category Similarity Matrix (Figure 6.4). We actually are defining a nominal set of maps where adjacent categories have vague transitions. The off-diagonal elements can be used to define the similarity between adjacent categories; for example, we may want to make the two categories, such as 'sparse residential' and 'dense residential', more or less similar by applying a value of, say, 0.5. A value of 1.0 would make both categories totally equal and would in fact define the new category, 'total residential'.

We note that the use of the Category Similarity Matrix allows us to handle pairs of maps which are partly nominal and partly ordinal. As in the example above: 'residential' categories have an ordinal character, while other categories are nominal.

Let us use another example. Suppose we have an ecological model predicting the presence ('1') or absence ('0') of a particular plant species by chance on a gridded map. Chance is expressed in the following categories: the chance of a plant being present in a particular grid lies between 0.0 and 0.25 (category 1), between 0.25 and 0.50 (category 2), between 0.5 and 0.75 (category 3) or between 0.75 and 1.0 (category 4). Clearly, adjacent categories are similar. We want to compare the chance map for 2004 with a predicted map for 2020. To employ the similarity between adjacent categories, we may choose the following Category Similarity Matrix:

$$\begin{pmatrix} 1.00 & 0.70 & 0.40 & 0.10 \\ 0.70 & 1.00 & 0.70 & 0.40 \\ 0.40 & 0.70 & 1.00 & 0.70 \\ 0.10 & 0.40 & 0.70 & 1.00 \end{pmatrix}$$

Since the choice of values in the matrix is subjective, it has to be selected on the basis of a priori insights. We advise experimenting with different settings and evaluating the sensitivity of map similarity (Kappa's) to these settings. If more than one pair of maps is involved, it could be advantageous to determine one matrix setting for all pairs of maps.

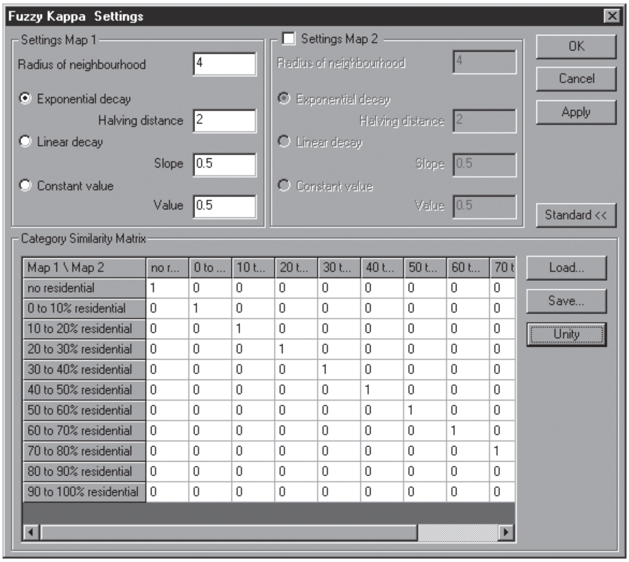


Figure 6.4 Advanced settings window for the Fuzzy Set algorithm. In the upper left part we choose between three distance functions ('Exponential decay', 'Linear decay', 'Constant value'). The 'radius of neighbourhood' is set to 4, meaning that we compare the value of one specific grid cell in map 1 with grid cells in map 2 that lie within a circle with a radius of four grid cells. The Category Similarity Matrix is used to define the similarity between categories (1= total similarity, 0= total dissimilarity). Here, we can introduce fuzziness in category definitions. The category names in the matrix refer to an ordinal map with a percentage residential for each grid cell. There are 11 categories: from 'no residential' up to '90 – 100% residential'. Setting off-diagonal elements to values between 0 and 1, we may employ the ordinal character of the maps.

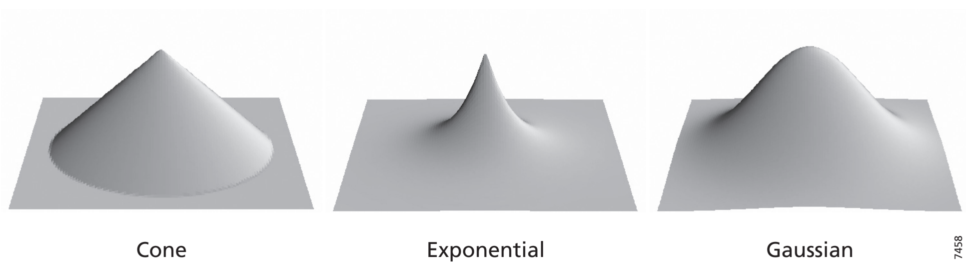


Figure 6.5 Three distance-decay functions. The cone shape ('linear decay') and the exponential shape are part of the advanced fuzzy screen, shown in Figure 6.4.

When we are dealing with maps on interval or ratio scales, we may choose a method from methods 5 to 10 (Figure 6.3) and exploit the character of specific continuous variables: subtracting or dividing values for identical grid cells. We then obtain result maps showing map differences on a continuous scale. If the values in both maps are denoted by $y_{i,j}$ and $x_{i,j}$, with i,j being the coordinates of a specific grid, the differences $d_{i,j}$ for methods 5 through 10 are defined as:

- Method 5: $d_{i,j} = y_{i,j} - x_{i,j}$
- Method 6: $d_{i,j} = |y_{i,j} - x_{i,j}|$
- Method 7: $d_{i,j} = (y_{i,j} - x_{i,j})/\max \forall_{i,j} (|y_{i,j} - x_{i,j}|)$
- Method 8: $d_{i,j} = |y_{i,j} - x_{i,j}|/\max \forall_{i,j} (|y_{i,j} - x_{i,j}|)$
- Method 9: $d_{i,j} = y_{i,j}/x_{i,j}$
- Method 10: $d_{i,j} = |y_{i,j}/x_{i,j}|$

We note that these six operations are not exclusive to the MCK software and can be performed by most GIS packages. The methods have been added to the MCK software package to make it more complete.

As an alternative approach, we may transform the continuous variable to disjunct categories and, subsequently, explore map similarity with the Fuzzy Set algorithm (by choosing a proper Category Similarity Matrix). Of course, some information will be lost in the transformation.

6.3 Map comparison methods

Here we describe the map comparison ideas underlying the methods contained in the MCK software and described in Hagen (2002b). For more detail on the theoretical aspects of the MCK software, we refer the reader to Pontius (2000), Power et al.(2001) and Hagen (2002a, 2003).

6.3.1 Visual versus automated map comparison

For most purposes, a visual, human comparison of maps outperforms automated procedures. When comparing maps, the human observer automatically takes many aspects into consideration, recognising local, but also global similarities, logical coherence and patterns. Map-comparison software methods usually capture single aspects, while overlooking others, and they generally lack the flexibility to switch from one aspect to the other when required by the data. An example of this rigidity is the simple cell-by-cell comparison of the two checkerboards described in the Introduction. However, despite the clear disadvantages of automated map comparison, automated procedures are preferred to visual map comparison in the following situation:

- Where time and human effort can be saved.
- Where objectivity and repeatability are desired, since automated procedures are explicitly defined. The method can therefore be analysed and evaluated and the results verified. A visual map comparison will always be subjective and often intuitive, so that the outcome of a visual map comparison may depend on the person performing the comparison.
- Where most comparison techniques fall short, recent developments in fuzzy-based map comparison have shown that these methods highly reflect the way humans look at maps.

The general philosophy behind MCK software is that instead of visual and automated map comparisons excluding each other, they actually yield complementary insights. The same holds for map comparison based on cell-by-cell similarity (section 6.3.2) and map comparison based on fuzzy-set calculation rules (section 6.3.3).

6.3.2 Cell-by-cell map comparison

The Kappa statistic for nominal maps was introduced by Cohen (1960). It is often used to assess the similarities between observed and predicted results and is not only applied to geographical problems (Monserud and Leemans, 1992, Pontius, 2000) but to many other disciplines, such as the medical and social sciences. This has led to a great deal of information being available on the Kappa statistic, including an extensive discussion on its functionality (Maxwell, 1977, Carletta, 1996, Foody, 2002, Foody, 2004).

Kappa is a measure of the similarity between two maps based on a contingency table (shown in the lower part of the right panel of Figure 6.10). In essence, Kappa is based on the percentage of agreement between two maps, corrected for the fraction of agreement that can be expected by pure chance. Recent extensions of Kappa are the 'Kappa location' (Pontius, 2000), which measures respective differences in location and in the histogram shape of all the categories. In the software, 'Kappa location' is abbreviated as KLoc and 'Kappa histo' as KHisto. Kappa, KLoc and KHisto are connected through the multiplicative relation

$$Kappa = KLoc \cdot KHisto \qquad \text{(eq. 6.1)}$$

Thus, given a fixed value for Kappa, KLoc will increase if KHisto decreases, and vice versa. Furthermore, if categories in both maps lie at identical locations, KLoc = 1.0 and Kappa = KHisto. If the frequency of categories in both maps is equal, KHisto = 1.0 and Kappa = KLoc. Examples from practice have been given in Visser (2004).

After the calculation of Kappa's, one may want to determine whether the observed value was greater than the value which would be expected by chance. Or one may want to compare Kappa's from different map comparisons. To perform either of these procedures, we need to know the significance of Kappa. Formulae for obtaining the significance of Kappa are given in Siegel and Castellan (1988). Confidence limits for Kappa (and KHisto/KLoc) are not incorporated in the present version of the MCK software, but we will be adding confidence limits to Kappa, KHisto and KLoc in an update of the software, based on the Monte Carlo simulation of randomised maps which are derived from the original maps. We note that Kappa confidence limits should be handled with care, as discussed by Foody (2004).

6.3.3 Fuzzy-based map comparison

The fuzzy-based map-comparison method was primarily developed for use in the calibration and validation process of cellular models for land use dynamics. The method is based on fuzzy-set calculation rules (Zadeh, 1965), and several authors have addressed the potential of fuzzy-set theory for geographical applications (Fisher, 2000, Cheng et al., 2001). In the past, fuzzy-set theory was used to assess the accuracy of map representations and for map comparisons (Metternicht, 1999, Power et al., 2001).

The objective of fuzzy-based map comparison is to find a method that to some extent mimics human comparison and provides a detailed assessment of similarity. The method is

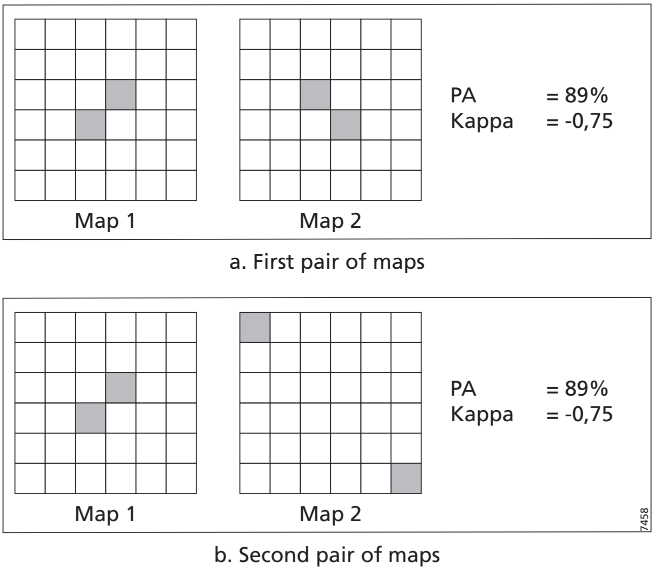


Figure 6.6 Two pairs of maps: There is a perceivable difference in similarity, but identical results are obtained for Percentage-of-Agreement and the Kappa statistic.

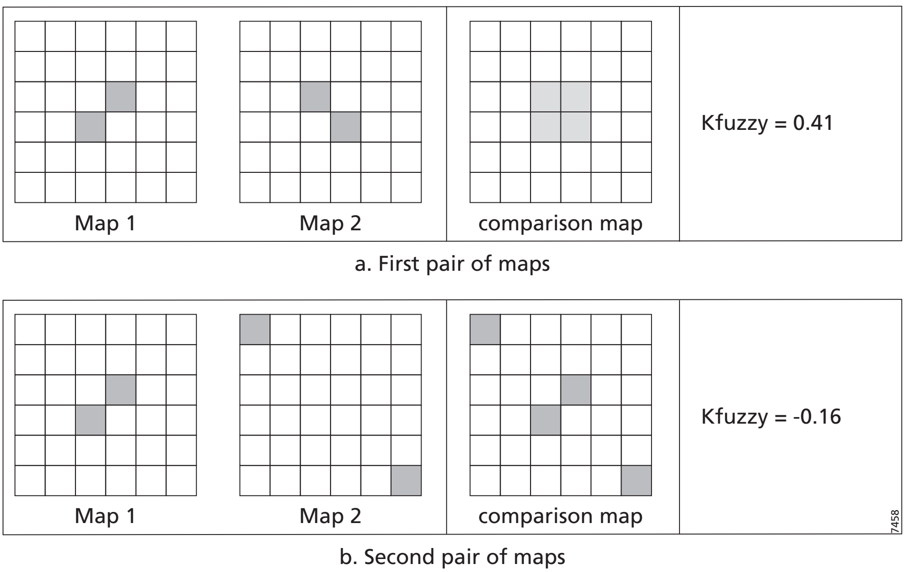


Figure 6.7 The Fuzzy Set Map Comparison of the two pairs of maps shown in Figure 6.6. Greyscales in the comparison map indicate the level of similarity. The value of Kfuzzy is based upon a distance decay function, with a constant value of 0.5 for directly adjacent cells and 0 for all other cells in the neighbourhood.

primarily directed at comparing nominal and ordinal raster maps, as illustrated in section 6.4. The assessment results are spatial, with (dis)similarities presented on a gradual scale.

Additionally, an overall figure for similarity, the Fuzzy Kappa statistic, is aggregated from the detailed spatial results (fourth method in Figure 6.3 and right panel of Figure 6.11). The Fuzzy Kappa (or Kfuzzy for short) is similar to the traditional Kappa statistic: the expected percentage of agreement between two maps is corrected for the fraction of agreement statistically expected from randomly relocating all cells in both maps.

A simple illustration of the difference between the traditional Kappa and Kfuzzy is given in Figures 6.6 and 6.7. Figure 6.6 demonstrates this with two pairs of maps. The Kappa statistic and the Percentage-of-Agreement are identical for both pairs, even though a human observer would consider the first pair (Figure 6.6a) to be more similar than the second pair (Figure 6.6.b). The effect of applying Kfuzzy is illustrated in Figure 6.7, which shows that the tolerance for small spatial differences indeed leads to a higher similarity of the first pair of maps. For details, the reader is referred Hagen and Uljee (2003) and Hagen (2003).

Another promising approach is that of Hierarchical fuzzy pattern matching (third method in Figure 6.3). This approach is based on the comparison of overlapping polygons (i.e. groups of concatenated cells belonging to the same category). The polygons are then compared on the basis of a number of properties, such as size and overlap. These comparisons are based on a so-called Fuzzy Inference System (FIS). A global agreement value can be derived by the fuzzy summation of the local matchings. This global agreement, comparable to the Kfuzzy statistic, is denoted as Global matching index. The theory behind Hierarchical fuzzy pattern matching is described in detail by Power et al. (2001). The fuzzy inference approach currently misses the functionality shown in Figure 6.4 for the fourth method. Therefore, we prefer to use Kfuzzy statistics (fourth method in Figure 6.3) rather than the FIS approach (third method in Figure 6.3).

6.3.4 Map similarity measures other than Kappa

The concept of the Kappa statistic (and variants thereon) plays a central role in the MCK software. The reasons for choosing this statistic are threefold:

- It has been extensively applied in research fields such as land use modelling, remote Sensing and ecological modelling.
- The chance correction in Kappa's is an attractive option: because of this correction, we are able to compare Kappa's (and its variants as well) across pairs of maps. Suppose, for example, that we have two binary map comparisons yielding an identical percentage of agreement (PA) of, say, 70%. Are both pairs of maps equally similar? Yes, they are equally similar if the frequencies of categories are equal across all four maps. If category 1 and 2 in the first pair both have frequencies of 50% and if, in the second pair, category 1 has a frequency of 10% and category 2 a frequency of 90%, the chance correction is 50% for the first pair and 82% for the second pair. Thus, a PA of 70% in the first pair is better than what could be expected by chance (50%). For the second pair, a PA of 70% is worse (82% by chance).
- Using a heuristic map comparison exercise from the Internet, Hagen (2002a) has shown that Kfuzzy estimates compare very well to human judgement (127 people rated the similarity of ten sets of maps with a number between 0.0 an 1.0). Thus, applying Kfuzzy, we have a measure which is close to the way humans compare maps.

However, there are other measures that can be used for map similarity comparison, and some authors doubt the practical value of Kappa. Indeed, a measure with a chance correction may not be advantageous under all circumstances. Other measures are discussed by Cohen (1968), Turk (2002) and Foody (2004). Pontius (2002) suggests choosing 'disagreement due to quantity' and 'disagreement due to location' as better alternatives for KHisto and KLoc.

However, as explained above, the rationale for choosing Kappa is not that it is the 'best' measure for all applications. 'Best' can not be seen as an absolute criterion; rather, it depends on the specific goals of the user. We have found the Kappa statistic to have attractive properties that are useful in our applications. Moreover, the Kfuzzy statistic scores very well if we define 'best' in terms of 'similar to human judgement'. Finally, the fuzzy approach is unique in providing the possibility of defining both vagueness in location and vagueness in category definitions.

6.4 Comparing land use maps

This section provides an application example of the use of the MCK software. We use two land use maps for the Netherlands, both on a 500×500-m grid. Both maps are on a nominal scale and include seven categories. One map is for a past situation (1995), and the other is a model prediction for the year 2020. More examples of map comparisons on other measurement scales are given in Visser (2004).

6.4.1 Simulating land use

To support planning systems in the Netherlands and to evaluate the consequences of spatial development, such as urbanisation, a decision was made in 1996 to jointly develop an instrument for the prediction of future land use patterns for valuable landscapes, nature areas, the environment and water systems. This led to the development of the Land Use Scanner. The aim of the Land Use Scanner is to elaborate different scenarios by integrating and allocating exogenous land use claims originating in sectoral models, such as those for housing and employment. The resulting future land use configurations may serve in turn as input for other models, such as hydrological and ecological models. The Land Use Scanner incorporates the traditional aspects of equilibrium models that use micro-economic-theory-based assumptions on supply, demand and price setting (Schotten et al., 2001, Borsboom – van Beurden et al., 2002).

Visser (2004) analyses two scenario simulations for 2020: the 'Country of Cities' scenario and the 'Country of Flows' scenario. According to the 'Country of Flows' policy concept, future land use will be directed towards the optimal functioning of international streams, both in economic and ecological contexts. This goal determines the choice of future locations for residential, work and nature areas. The economic 'streams' are directed to the road network and the ecological 'streams' to the water network. The main challenge here is combining these two streams. Urbanisation will be concentrated along a few main transport axes in the form of a 'string of beads'. In contrast, the 'Country of Cities' scenario foresees new residential areas with high intensities arising in and around existing cities. The actual land use map in 1995 and the Land Use Scanner simulation for 2020 according to the 'Country of Flows' scenario are given in Figure 6.8.

6.4.2 Map comparison

The 'Per category' comparison method is illustrated in Figure 6.9 for the categories 'Residential' and 'Agricultural'. The blue colour in the left panel shows the predicted 'string of beads'

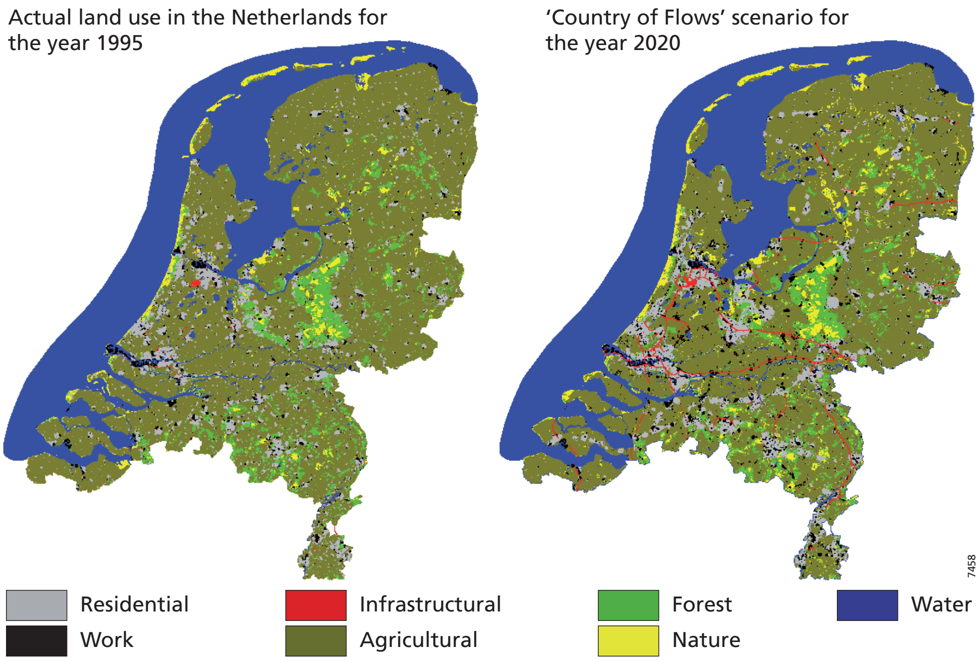


Figure 6.8 Actual land use for 1995 (left panel) and simulated land use for 2020 (right panel).

for 'Residential' in 2020 (= map 2), while the red colour in right panel shows (= map 1) that 'Agricultural' has disappeared at the expense of 'Residential'.

The equal–unequal map in Figure 6.10 is obtained as result of applying the 'Cell by cell' method. The colour green points to cells in both maps with identical categories, and the red colour to cells with unequal categories. The right panel shows the Kappa statistics. The overall statistics are shown first (Kappa, Fraction correct, KLoc and KHisto, as described in section 6.3); then Kappa statistics are shown per category. This means that we simplify both maps in Figure 6.8 to maps with only two categories: 'Residential' versus 'Other', 'Industrial/Work' versus 'Other', etc. The next step is to calculate Kappa's for these 'binary' maps. The lower matrix shows the contingency table for the seven categories, 'Residential', 'Industrial/Work', 'Infrastructure', 'Agricultural', 'Forest', 'Nature' and 'Water'.

From the overall Kappa's, we see that the KLoc value (0.87) is less than the KHisto value (0.92), signifying that both maps differ more in the over-all frequency of the various categories than in the location of categories.

The binary comparison of categories reveals minimal changes for 'Water', with all Kappa's being found near the maximum value 1.0 (Kappa = 0.97, KLoc = 0.99 and KHisto = 0.98). Clearly, no hydrological changes are foreseen. Kappa for 'Residential' is found to be relatively low (0.48), a result due entirely to the different locations of the categories (KLoc = 0.49 and KHisto = 0.99) and easily explained by the location of the scenario 'Country of Flows'. Finally, changes for 'Infrastructure' are maximal; all Kappa's are small (Kappa = 0.20, KLoc = 0.66 and KHisto =

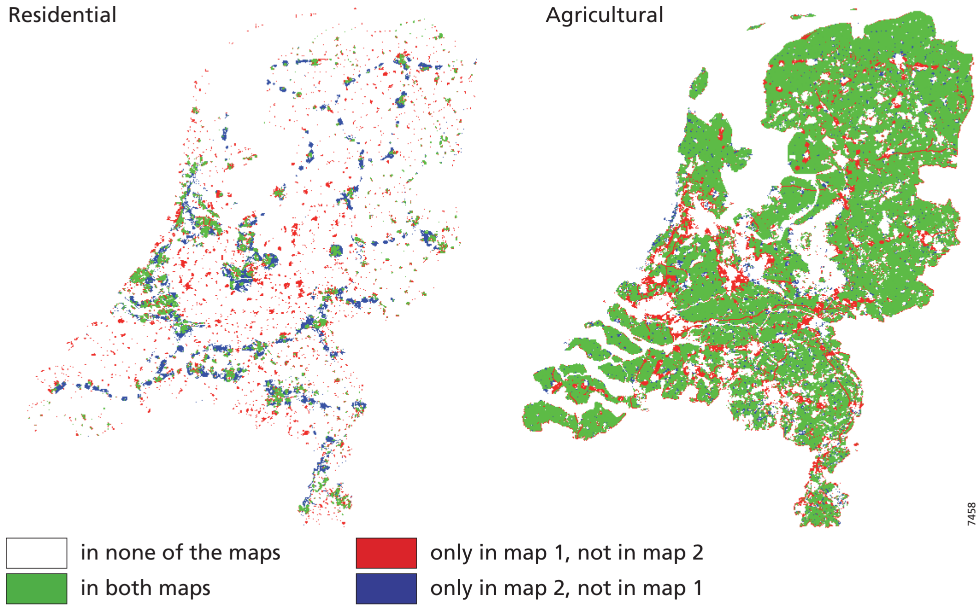
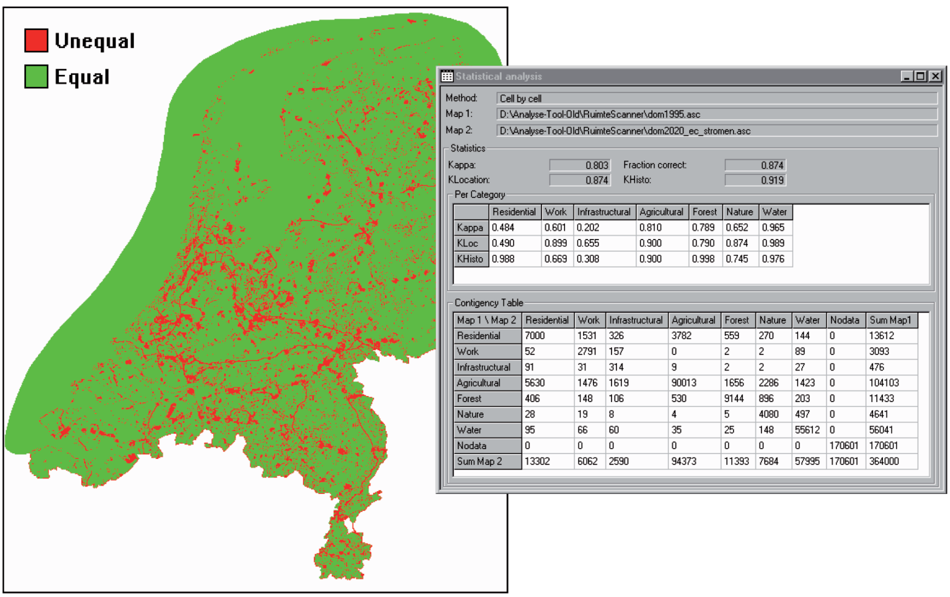


Figure 6.9 Map differences for two single land use categories using the first option ‘Per Category’, shown in Figure 6.3. The left panel is for ‘Residential’ and the right panel for ‘Agricultural’. Map 1 = situation 1995; map 2 = simulation 2020.



	Residential	Work	Infrastructural	Agricultural	Forest	Nature	Water
Kappa	0.484	0.601	0.202	0.810	0.789	0.652	0.965
KLoc	0.490	0.899	0.655	0.900	0.790	0.874	0.989
KHisto	0.988	0.669	0.308	0.900	0.998	0.745	0.976

Map 1 \ Map 2	Residential	Work	Infrastructural	Agricultural	Forest	Nature	Water	Nodata	Sum Map1
Residential	7000	1531	326	3782	559	270	144	0	13612
Work	52	2791	157	0	2	2	89	0	3093
Infrastructural	91	31	314	9	2	2	27	0	476
Agricultural	5630	1476	1619	90013	1656	2286	1423	0	104103
Forest	406	148	106	530	9144	896	203	0	11433
Nature	28	19	8	4	5	4080	497	0	4641
Water	95	66	60	35	25	148	55612	0	56041
Nodata	0	0	0	0	0	0	0	170601	170601
Sum Map 2	13302	6062	2590	94373	11393	7684	57995	170601	364000

Figure 6.10 Equal-unequal map and corresponding Kappa statistics, using the second option ‘Cell by cell’, shown in Figure 6.3.

0.31) due to such projected infrastructural changes in the Netherlands as the result of denser rail freight transport.

The left panel of Figure 6.11 shows the result map of the Fuzzy Set algorithm, using default settings for the MCK software (exponential decay with a neighbourhood radius of 4 cells). It is clear that the fuzzy approach reveals many more nuances between the two maps than the simple equal–unequal map. The right panel shows the Kfuzzy statistics corresponding to the maps being compared.

6.5 Final remarks

In this chapter we have presented and described map-comparison software in which both traditional and advanced techniques have been integrated. To the best of our knowledge, no other comparative software package is available at the present time. Statistical packages such as SAS and SPSS and a number of digital image processing packages have only incorporated the traditional Kappa statistic.

We have paid special attention to the handling of nominal and ordinal maps in view of the scarcity of existing methods in this area. In fact, the software implementation for ordinal map comparison techniques is unique. The handling of maps on interval/ratio scales is also part of the software and has been added for completeness (many GIS packages can perform the operations given in section 6.2.2 as well).

The application in Chapter 4 has illustrated that both a cell-by-cell comparison and fuzzy-set map comparison are able to grasp the characteristics of map differences, although in different

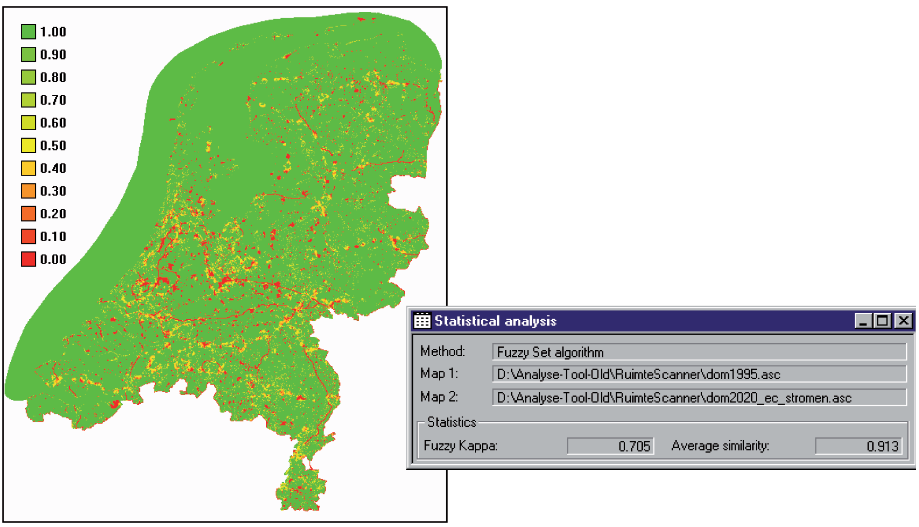


Figure 6.11 The fuzzy similarity map and corresponding Fuzzy Kappa statistic, using the fourth option 'Fuzzy Set algorithm', shown in Figure 6.3. A value of 1.0 (dark green) means total similarity and a value of 0.0, total dissimilarity (dark red). Values in the yellow shades (around 0.5) point to some degree of similarity in both maps.

ways. Based on our experience, the combination of these approaches yields an ideal manner for analysing maps and map differences (cf. examples in Visser, 2004).

Our experiences with the Map Comparison Kit software have been very positive. The software is easy to use, and limits are only set by the memory conditions of the PC. However, the software is not static, and improvements are planned. The software will be extended in the following ways:

- The creation of a batch environment around the MCK software for handling large amounts of maps would be helpful. In some environmental applications over 150 pairs of maps have to be compared, a process which is too time consuming to do 'by hand'.
- For some analyses, the ability to aggregate both maps to a coarser grid would prove interesting (Pontius, 2002, Pontius et al., 2004a). Following that step, we want to analyse differences through the use of map statistics. A different problem will also be treated – that of two maps with unequal grid sizes.
- In practice, we have to compare maps having unequal legends. In some cases, a re-classification of categories would be sufficient to make the legends the same (by use of the Category Similarity Matrix, shown in Figure 6.4), but this will not be sufficient in all cases.
- Kappa, KLoc, KHisto and Kfuzzy are not calculated with accompanying uncertainty limits within the MCK software. The advantage of having such limits is that two Kappa's' could be tested for significance [see Foody (2004) for examples and warnings]. Therefore, Kappa, KLoc, KHisto and Kfuzzy will be extended, together with their respective uncertainty estimates. To do this, we will choose a Monte Carlo approach, since this approach will apply to all four Kappa measures.
- It will be possible to apply the fuzzy approach directly to maps on an interval or ratio scale. The 'transform step' shown in Figure 6.3 will then become redundant, and information will no longer be lost when these maps are transformed to ordinal maps.

7 Measuring the performance of land use models. An evaluation framework for the calibration and validation of integrated land use models featuring cellular automata

Reprinted from *Proceedings of the 14th European Colloquium on Theoretical and Quantitative Geography.* Tomar, Portugal. Hagen-Zanker, A., van Loon, J., Maas, A., Straatman, B., de Nijs, A. C. M. & Engelen, G. (2005) Measuring Performance of Land Use Models. An Evaluation Framework for the Calibration and Validation of Integrated Land Use Models Featuring Cellular Automata.

Abstract

The automated calibration and validation of integrated land use models requires the objective quantification of model performance at different levels of abstraction. This paper presents a calibration and validation routine and specifically focuses on the procedures for evaluating the model output. We were able to determine that although a fully objective procedure is not yet available, a major portion of the analytical tasks can be automated.

7.1 Introduction

We consider a land use model proposed by Engelen et al. (2003) which combines a spatial interaction (gravity) model at the regional level with constrained cellular automata (CA) to disaggregate regional land use claims to cells. A feedback of cellular aggregates, such as the mean suitability of available land, is input to the spatial interaction model. Thus, a tightly integrated and dynamic system is established. This model has been applied for ex post analysis (Geurs et al., 2003) and ex ante analysis (de Nijs et al., 2005) of Dutch spatial policy, among others. A major challenge in the application of the model is its calibration and validation (White and Engelen, 2003, Straatman et al., 2004).

The reality of using the model learns that as time progresses, it is necessary to recalibrate the model, either because new data become available or new expectations of the model require adjustments. As always in such cases, limited time and resources are available, calling for a manageable procedure for setting parameter values. This paper describes a number of steps that have been taken towards establishing a fully automated calibration, which were part of a calibration exercise using the Dutch Environment Explorer model. The results of the exercise were not only a calibrated model but also guidelines for future calibrations. These guidelines were put to test when calibrated models had to be produced for two other regions – Estonia and northern Italy. This paper will focus on the evaluation framework applied in the calibration and validation procedures. Nevertheless, in accordance with the handbook Good Modelling Practice (van Waveren et al., 1999), other analytical tasks, such as verification, global behaviour analysis, sensitivity analysis and robustness tests, have been performed as well. The reader is referred to the calibration report (RIKS, 2004) for details on precise procedures for adjusting parameter values.

7.2 Methodology

7.2.1 Calibration procedure

The primary intended use of the model is to explore possible future land use for approximately 30 years in the future. The historical land use data that are available, however, do not stretch sufficiently back in time. Therefore, the calibration is not only aimed at a best fit with historical data but also at the best fit with general landscape structure or morphology that unrolls from the model dynamics when it is applied for periods that long surpass the available data.

It is important to realise that the model under consideration is an integrated model consisting of several model components that are dynamically linked. The difficulty in attributing discrepancies in the integrated model results to one or the other model component is an attractive quality, possibly an essential one, for calibrating the model components individually. Thus, the dynamic link between the model components is temporarily cut, and the models are applied in a chain. After the models have been calibrated individually, they are reconnected. It is then necessary to assess how the different goodness-of-fit measures are affected by the restored dynamics.

The model components are the national model, the regional model and the local model. The regional model can further be split into the gravity model and the density model. At the national level, the model comprises simply of timelines that are the driving force of the regional model. For the national model, historical data can be used, and calibration is not required. The regional model poses constraints on the local model in the form of land use demands per region. Therefore, in the decoupled state, these land use demands need to be supplied. A feedback to the regional model is provided by the regional aggregates of cellular characteristics as they transpire from the local model. These are logged once by the local model, after which they are used as input for the regional model. The local model is the constrained CA model, which is the most calculation intensive model and requires the most parameters.

The flowchart of Figure 7.1 illustrates the general calibration procedure. Six iteration loops are recognised (labelled in Roman numbers I -VI). Each iteration loop is started when the previous one is satisfied. The model starts by tuning the parameters of the CA model to best match the historical data (iteration I). The optimisation criterion is the agreement between the actual and simulated land use at the end of the calibration period, as measured by the Fuzzy Kappa metric (Hagen, 2003). The premise of this comparison method is that it not only credits exact cell-to-cell agreement but also near cell-to-cell agreement. Additional analysis of the maps helps to determine which land use categories require the most adjustment (Fuzzy Kappa per category, following Hagen-Zanker et al. (2005a), and also which interactions between land

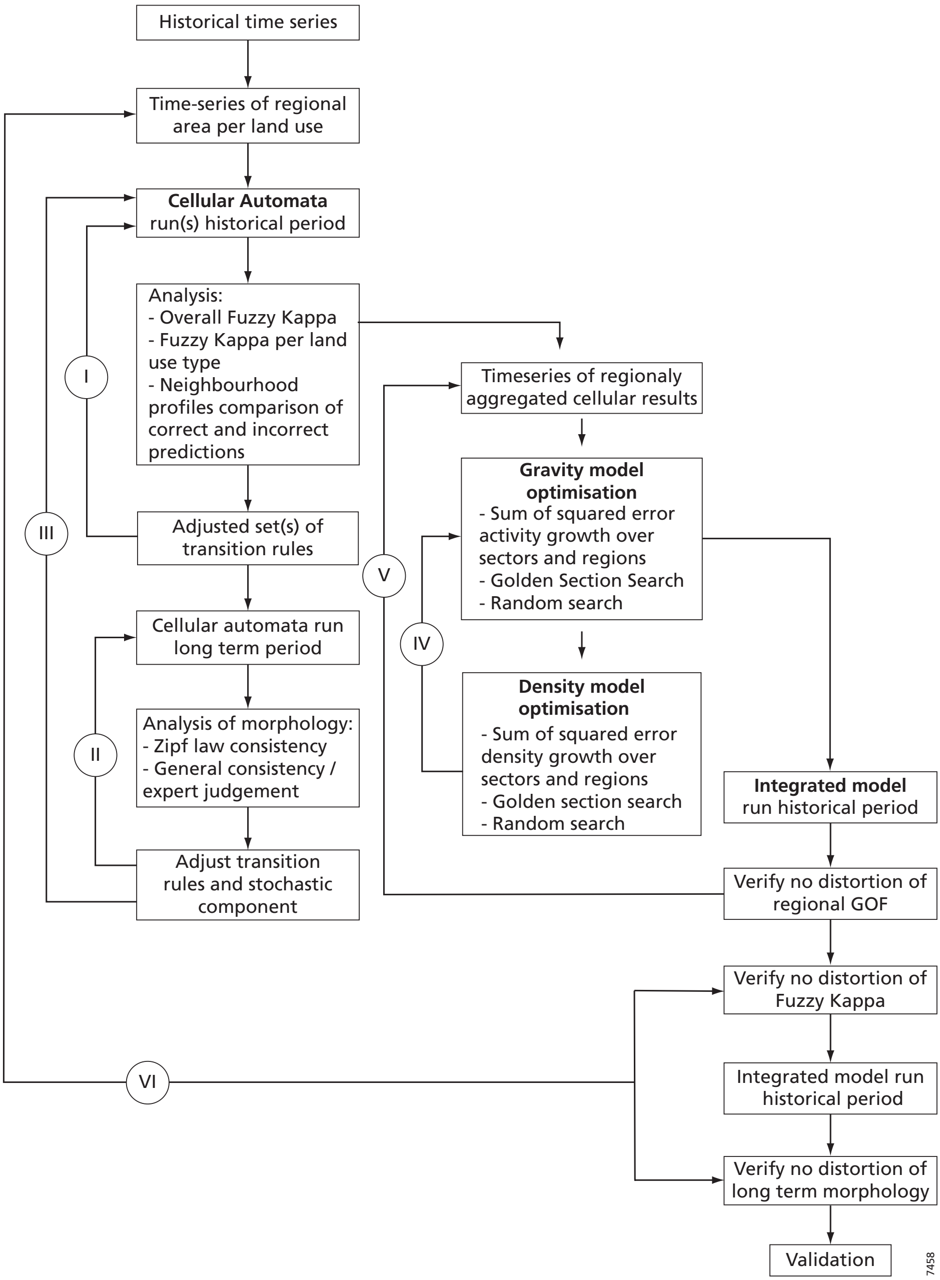


Figure 7.1 Flowchart overview of the full calibration procedure.

uses are likely to fix the discrepancies. This procedure is a variation on the one described in Straatman et al. (2004). This part of the calibration (iteration I) can be performed automatically, although we have found that the transition rules identified by the automatic procedure need to be scrutinised to exclude aberrations.

Even if the model performs well in reproducing the historical data, it is possible that unrealistic landscape morphology is obtained when these trends are followed into the future. The evaluation of morphology and the transition rule adjustment follows the expert judgment-based procedure of White and Engelen (2003). The experienced user of the model will have a good understanding of the relation between transition rules and resulting patterns. A regularity in the distribution of cluster sizes, known as Zipf's law, may be used to automatize this iteration. To date, Zipf's law has been used only for a single parameter (alpha), which has a strong influence on the morphology of the resulting landscape. The alpha parameter sets the stochastic volatility of the model and is thus responsible for the degree to which new clusters are formed.

Iteration III alternately invokes Iteration I and iteration II and thus provides an equilibrium between the morphological coherence and historical accuracy. The general philosophy is that iteration II (morphology coherence) is meant for a number of coarse corrective parameter adjustments, whereas iteration I is meant for more refined calibrations. Once both coherence and accuracy are satisfactory, the regional aggregates to be used inputs for the calibration of the regional model are calculated (such as the mean suitability of as yet undeveloped land). 'Satisfactory' is still a non-quantified property and is mainly based on the modeller's perception of whether a prolonged calibration effort of the CA model will yield substantial improvements.

Finally, one parameter that is essential to the morphology of the model is calibrated by an analysis of the cluster size distribution. It was found that the maps – over time – obeyed Zipf's law which assumes a linear relation between the log of the cluster size and the proportion of clusters exceeding that size. It was then found that this linear relation changes over time, and the parameter stimulating stochastic perturbation was adjusted to best fit a projection of the cluster size distribution

Iteration IV is the calibration of the regional model. The gravity model and the density model can be distinguished within the regional model. These have a reciprocal dynamical relation, as the density of activities in a region are a contributing factor to the attractivity of that region and, vice versa, the activity level of a region determines to a large extent the density of that activity (jobs or population per ha). The parameters of the two model components are calibrated separately, although the links between the components remain intact. The same search algorithm is used for the two components, Golden Section Search, which searches for local optimums, and Random Search, for escaping the local optimum in the search for the global optimum. The procedure has been described in depth by Van Loon (2004), and it is fully automated. The goodness-of-fit measure is different for both components; the gravity model is optimised with respect to levels of activity and the density model is automated with respect to density. These criteria are chosen (instead of, for example, the land use area demands that are the product of activity and density) because they do not allow inappropriate parameter values in both modules to cancel each other out.

Once both the local and the regional model are calibrated, the dynamic link between the models is re-established and the integrated model is run once more. An evaluation now takes place to ensure that the recoupling does not negatively distort the goodness of fit obtained in the decoupled modules. This evaluation is performed for both the regional model and the cellular model, but first for the regional model (iteration V) because recalibration of the regional model is a minor task compared to recalibration of the local model. If the goodness of fit has been found to have deteriorated, the run with the integrated model is used to generate a new time series of regional aggregates, and the regional model is recalibrated.

Once the regional model is stable, the same procedure is applied as that for the CA model. If, during the calibration period, the Fuzzy Kappa is found to have diminished or the landscape morphology in the long run to be inconsistent, then the whole procedure is started again (Iteration VI) with the difference being that the time series for regional land use demands are not based on historical data, but on the run with the integrated model.

7.2.2 Validation procedure

The model is applied over a validation period in order to assess its predictive value. The goodness-of-fit measures for the validation are essentially the same as those for the calibration, except that measures are chosen for the regional model that are more intuitive to interpret. This means that instead of the sum of squared errors of activity and density growth, the mean absolute error of activity growth and land use demands are calculated. The Fuzzy Kappa statistic is applied for the cellular model, but the spatial distribution of similarity is also considered.

In order to obtain a feeling for the meaning of the quantitative results, these are compared against the results obtained by naive predictors. Naive predictors are alternate models that satisfy the constraints put upon the actual model by minimally changing the initial situation. The general rationale behind these models is that 'the best prediction for the weather of tomorrow is the weather of today'. The naive predictor for the regional model is the constant share model. This model distributes the national growth of an activity by keeping the relative distribution over the different regions constant. This means that the same growth factor is applied for all regions. A second aspect of the constant share model is that the density of the different activities remains constant. In effect, this means that the growth factor for a given activity is also applied for the land use claim associated to that activity. For example, the area of a region taken in by the land use "Industrial" follows the national trend of "Employment in the industrial sector."

The naive predictor for cellular land use change is that of the minimum change needed to satisfy the constraints and random selection of the location for those changes. Thus, the naive predictor is used to satisfy the same regional constraints as the Constrained CA. The model starts with the initial map and randomly selects cells of land uses for a region that are overrepresented (compared to the regional constraints); it subsequently randomly assigns those cells to land uses that are underrepresented, repeating the process until all constraints are satisfied. An alternative would be to apply a 'no change at all' naive predictor (as in Hagen 2003), but the current approach has the advantage that the overall composition of the maps is identical, thereby increasing the possibility of being able to focus on the quality of the configuration. It would otherwise be difficult to separate composition (quantity) and configuration (location). In addition, the naive predictor and the CA model would not be subject to the same constraints, leaving them less comparable.

Table 7.1 Goodness of fit measures for the calibration and validation period for the model and the naïve predictor: Fuzzy Kappa displaying map similarity, Relative error in growth of activities and the Absolute error in regional land use claims defined in 25ha cells per land use type.

Measure		1989-1996	1996-2000
Fuzzy Kappa	Model	0.936	0.913
Fuzzy Kappa	Naive	0.926	0.922
Relative error	Model	3.90%	5.20%
Relative error	Naive	5.20%	3.90%
Absolute error	Model	3.3	7.7
Absolute error	Naive	5.7	6.4

Table 7.2 Left: contingency table summarising changes from 1989(rows) to 1996 (columns) and right: the overall changes

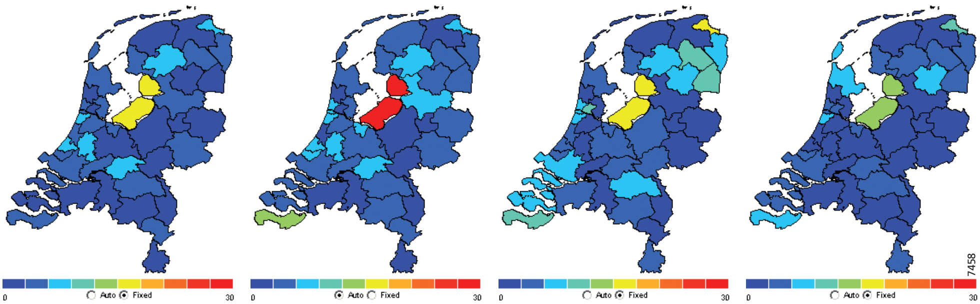
	Nature	Agriculture	Urban	Work	Features	Overall Changes	Cells
Nature	11458	333	202	196	64	Total cells	139681
Agriculture	1371	97097	674	1127	245	Unchanged	133963
Urban	231	252	3524	113	104	Changed	5718
Work	178	145	81	17282	44	Suspect	1621
Features	91	125	78	64	4602	Perc Susp Changes	28.3%

7.3 Results

7.3.1 The pilot case: Environment Explorer for the Netherlands.

Four land use maps, all raster maps with a cell size of 500 m, were available for the years 1989, 1993, 1996 and 2000. The map for 2000 has consistency issues because some of the definitions of a number of land use categories were changed. Therefore, we chose the calibration period to be 1989-1996, which allows the model to be calibrated over the longest available period of time that is not affected by the inconsistencies of 2000. The validation period is 1996-2000. Despite the problems associated with the 2000 map, it is used in the evaluation, under the rationale that the naive predictor is affected by this inconsistency as much as the CA model.

The results (Table 7.1) indicate that the calibration of the cellular model is successful in the sense that the model outperforms the naive predictors over the calibration period (1989-1996). The model does not outperform the naive predictors during the validation period of 1996-2000. Different reasons can be put forward to explain this, with the main one likely being the fact that the calibration period was too short to pick up large-scale spatial processes. Relatively little change in land use occurs in such a short period; in addition, the relative proportion of mapping errors over true land use change is also large. An analysis on the basis of the (thematically aggregated) contingency Table 7.2 shows that 25% of the changes are unlikely in the sense that they are transitions from urban area to natural or agricultural area or because they are changes in land use types that are not expected to change (mainly fresh water, but also salt water, airports and foreign country). This is a large percentage, but given the situation that only 3% of all cells



a. Model 1989-1996 b. Naive 1989-1996 c. Model 1996-2000 d. Naive 1996-2000

Figure 7.2. Distribution of activity growth error over the regions

Figure 7.3. Spatial distribution of disagreement in 2000

change, a mapping accuracy of just less than 99.6% in both maps can be sufficient to cause such a disturbance. Note that this does not imply that the accuracy is larger than 99.6 %, since an unknown number of cells may harbour identical inadequacies for both years.

A spatial distribution of the error of the regional model (Figure 7.2) demonstrates that much of the error can be attributed to a single region, the 'Flevopolder'. As this is a very young region – the land was only claimed from the sea between 1939 and 1968 – it is not surprising that this region is still developing at pace different from that of the rest of the Netherlands. This outlier may have contributed to an over-calibration of the regional model, indicating that the parameters have been adjusted too much in order to reproduce the behaviour of the 'Flevopolder' and that this has been at the cost of the other regions.

The spatial distribution of errors at the cellular level (Figure 7.3) indicates that errors are distributed more or less equally over the map, although they are, in general, found in connection with urban areas. This is not surprising considering that urban areas are the more dynamic areas, and it is more difficult to correctly predict change than to predict non-change.

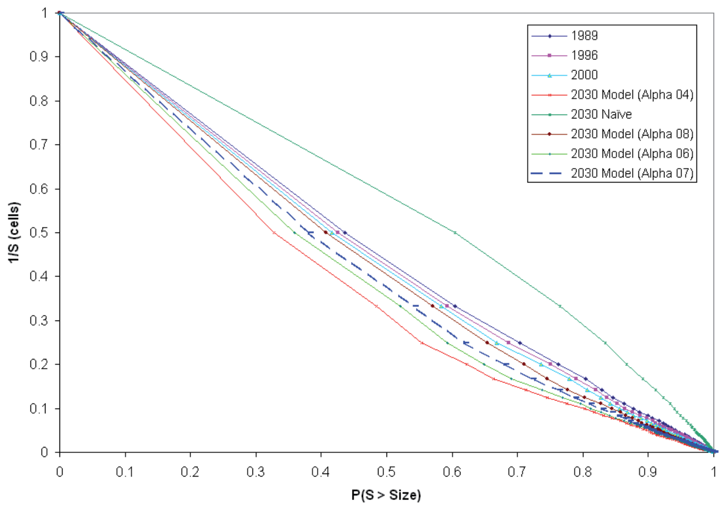


Figure 7.4. Cluster size distribution, for clusters of urban cells: the y-axis displays the inverse of the cluster size, and the x-axis shows the fraction of clusters equal or larger in size.

The cluster size distribution displayed in Figure 7.4 illustrates how the cluster size changes over time. It also shows how cluster size distributions are impacted by the value of the parameter alpha. On the basis of these results, alpha has been set at the value 0.7. Figure 7.5 shows the input and output land use maps where categories have been collapsed into main land use types. The outcome of the cluster analysis is that the clusters in the model map of 2030 fit the trend in cluster size distribution better than the naive predictor, a trend that can be recognised in the land use maps of 1989 and 2000.

7.3.2 Further experience with Estonia and northern Italy

The pilot case for the Netherlands is followed by two more cases for Estonia and northern Italy. The data on which the model is based is CORINE90 and CORINE2000 and regional economic indicators of the New Cronos REGIO database. The calibration of the regional spatial interaction model proceeded according to plan and is therefore not considered here. The true challenge has been the calibration of the CA model. Although maps of 2 years were available, the difference between the 1990 and 2000 map was small and, to a certain extent, erratic. For example, it is striking that between 1990 and 2000 the Milan area in northern Italy does not display any growth in terms of urban area whereas the Turin area does. The only major change that occurred in the maps of Estonia is a large-scale transition from pastures to intensive agriculture. If the land use categories are collapsed and only Agriculture, Nature, Industry, Urban and Features (water and roads) remain (Table 7.3), then less than 0.5% of the cells are subject to change, of which 9% are 'suspect' changes. In the case of Italy, fewer than 2 % of the cells changed, of which 15 % are suspect changes (Table 7.4)

An additional challenge of the Estonia and Italy cases is that the model is intended to provide spatially explicit explorations of scenarios that include land uses which in the past

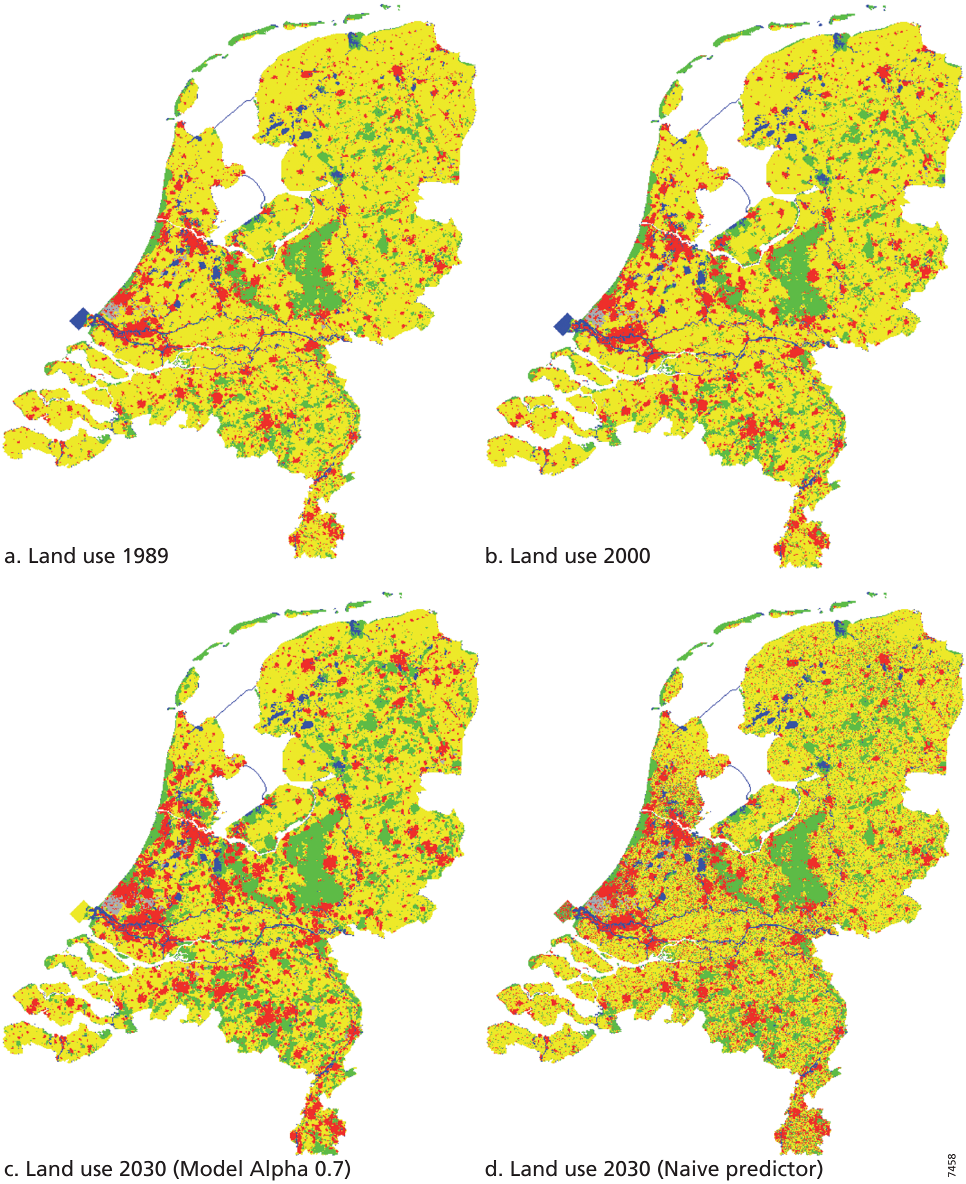


Figure 7.5 Land use maps used as input to the cluster analysis

were not even present, including bio-fuels, urban areas for specific population groups, gated communities for the affluent and thematic cities for alternative lifestyles. The purpose of the calibration here is not to ensure that the CA transition rules optimally mimic historical time-series but, rather, that the rules exhibit the behaviour in accordance with the scenarios as they were delivered. We interpreted a description of the daily activities of inhabitants of the new land

Table 7.3 Left: contingency table summarising changes in Estonia from CORINE90 (rows) to CORINE2000 (columns) and right: overall changes

	Nature	Agriculture	Urban	Industrial	Features	Overall Changes	Cells
Nature	111460	254	10	41	8	Total cells	184379
Agriculture	334	59254	65	5	2	Unchanged	183601
Urban	2	9	2161	2	0	Changed	778
Industrial	24	5	3	1288	0	Suspect	64
Features	11	1	0	2	9438	Perc Susp Changes	8.2%

Table 7.4 Left: contingency table summarising changes in Northern Italy from CORINE90 (rows) to CORINE2000 (columns) and right: overall changes.

	Nature	Agriculture	Urban	Industrial	Features	Overall Changes	Cells
Nature	274364	290	78	6	59	Total cells	573938
Agriculture	1086	257760	1786	93	142	Unchanged	563489
Urban	16	71	23023	703	3	Changed	10449
Industrial	28	49	4888	615	751	Suspect	1519
Features	31	24	14	331	7727	Perc Susp Changes	14.5%

use types as well as of their modes of transport into rules of spatial configuration. These rules were then expressed in terms of cluster sizes and their dependence on other land use types.

In the end, for lack of data, the transition rules as they were followed in the Netherlands case were used. To test the robustness of the rules for both Italy and Estonia, we developed four test-scenarios that were not based on story lines or trend extrapolations; instead, the aim was merely to force the model to cater to considerable changes in land use, while preserving land use structure. On the basis of these test scenarios, we tuned the rules to display realistic behaviour.

Finally, the models were applied on the actual story-line-based scenarios. The changes over time that these runs produced were evaluated on the basis of comparison methods that simultaneously account for structure and overlap (Hagen-Zanker et al., 2005a). These methods apply a distance-weighted moving window to obtain a spatial account of changes in structure (e.g. mean patch size, Shannon diversity and prevalence). The outcome of this validation exercise was that structural land use changes that appear from the change analysis are in line with expectations. Figure 7.6 displays several of the maps that were used as quantifications of structural change.

7.4 Conclusions

A typical approach to calibrating simulation models is to run them for a period in the past and then to adjust parameters in order to obtain the best historical fit. Although intuitively sound, there are some drawbacks to this approach, with the main one being the dependency on data availability and data quality. If a historical calibration is not possible, then a calibration aimed at the structuring quality of the model can be advised. A calibration based on historical data is a major challenge because of the non-linear relations between input and output and the difficulty

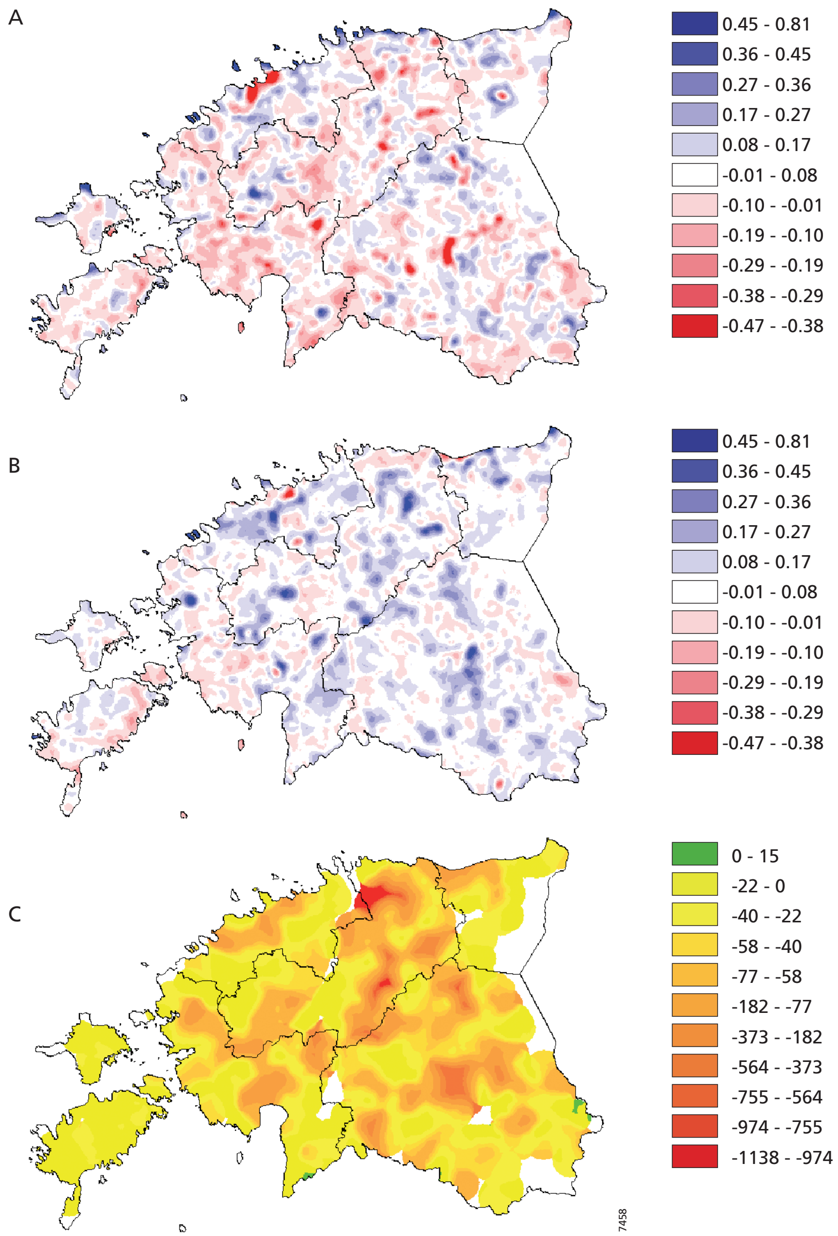


Figure 7.6 Moving window based-structure comparisons to validate the global behaviour of the Estonia model. A: Change in landscape diversity in Estonia over the period 2000-2030 for scenario 1; B. Change in landscape diversity in Estonia over the period 2000-2030 for scenario 2. C. Difference in patch size of agriculture, generally smaller clusters are found in Scenario 2 compared to Scenario 1. The results confirm the expectations that in scenario 1 agricultural specialisation will be stronger and the landscape will be less fragmented.

in quantifying the agreement between model output and actual data. When the evasive concept of structuring quality needs to be analysed, it is even more tempting to consider automatic calibration to be impossible and fully rely on expert judgement.

By making our calibration procedure explicit and by recognising the different iterations in the procedure, we were able to split the adopted procedure into a number of sub-tasks. The methodologies put forward in this paper are a mixture of automatic (objective) and human judgement (subjective) procedures that deal with the sub-tasks of the calibration. We found that although it is not yet possible to automatically calibrate the models, we can step-by-step seek further quantitative approaches to replace human judgement. This approach will not only relieve the researcher of a number of labour-intensive and often boring tasks, but it also results in the models being better transferable and transparent.

8 Spatial uncertainty in land use models. An alternative method to estimate uncertainty in logistic regression methods

Reprinted from *Proceedings of the 15th European Colloquium on Theoretical and Quantitative Geography*. Montreux, Switzerland. De Nijs, A. C. M. & Pebesma, E. J. (2007) Spatial Uncertainty in Land Use Models. An Alternative Method to Estimate Uncertainty in Logistic Regression Methods.

Abstract

We propose a novel method to estimate uncertainty in logistic regression models. Uncertainty in regression predictions is normally derived from the standard error of the model and the residual variance. However, using this approach, the uncertainty in the model of residential developments in the Netherlands is very large, mainly due to the large variance in the working residuals. An alternative method is based on the relationship between the logit and residuals. The residuals vary strongly only over a small range of logit values, and above and below certain threshold values, residual variability is low. We used the 2.5 and 97.5 percentiles of residuals, binned over logit values, as the lower and upper prediction limits. The results show that model uncertainty varies with location and land use category. Uncertainty is low if the initial land use is airports, water or building sites; it is high where agricultural or urban land is in the neighbourhood of residential area.

8.1 Introduction

During the last decade, many land use and land cover change models have been developed to assess future land use developments (Verburg et al., 2004c). We believe that decision-makers, who increasingly rely on the results of these models, should have some concept of the certainty associated with these models. Moreover, given the uncertainty in model predictions, research can be directed towards improving the model on these points. Methods to address uncertainty in land use change models are beginning to appear in the literature (Pontius and Spencer, 2005). It has been found that misclassification in land use maps may have a large effect on predicted land use (Fang et al., 2005, Fang et al., 2006). Model validation techniques have also been proposed and applied to several models (Pontius Jr et al., 2008), partially reflecting the uncertainty in model predictions. As many land use models use logistic regression, the standard error and residuals can be used to address uncertainty in the predicted probabilities. However, in our case, the large variance in the residuals poses a problem.

The objective of this study was to define the uncertainty in the predicted probabilities of a model of residential developments in the Netherlands. To this end, a novel method has been developed.

This article begins with a brief introduction of the model. We then introduce our proposed method for estimating uncertainty in the logistic regression model and provide some results. Finally, a number of conclusions are drawn.

8.2 Model

A binary logistic regression model was fit to predict the conversion to residential land use – or not – in the so-called Randstad region in the Netherlands over the period 1993-1996. The Randstad region forms the socio-economic heart of the Netherlands and includes the four major cities Amsterdam, Rotterdam, The Hague and Utrecht.

We used high-resolution land use data, rasterised to 25×25 m, that distinguished 33 different land use categories. The dataset included initial land use, land use in the neighbourhood at various distances and a range of policy- and suitability-related variables. It was divided in two equal parts, one part to fit and one part to validate the resulting regression models. The dataset to fit the model consisted of a random selection of 2580 cases (5%) that converted to residential area and a smaller random selection of 22597 cases that did not convert to residential area. Forward conditional stepwise weighted binary logistic regression, maximising the likelihood, was used. The final model included 87 explanatory variables. The ROC, based on the validation dataset, was 0.959.

8.3 Uncertainty

Homoscedastic uncertainty in a regression prediction for a single new observation can usually be described by an upper and lower prediction limit, defined by

$$[\hat{y} - a, \hat{y} + a] \qquad \text{(eq. 8.1)}$$

with a calculated by:

$$a \approx 2\sqrt{\left(SE^2 + Var\right)} \qquad \text{(eq. 8.2)}$$

where *SE* is the standard error of the regression model and *Var* is the variance of the residuals, both on a logit scale (Hastie et al., 2001). The approximation stems from the fact that the value 2 should be replaced by an appropriate t value. The variance of the residuals is included in these prediction limits to take account of the uncertainty by the regression residuals.

In our case, the uncertainty in the regression predictions according to this method is very large, with the upper limit for probability of land use change for the whole region being practically 1, and the lower limit being practically 0, thus rendering the prediction intervals useless. Land use changes may take place anywhere in the region. An underlying assumption is, however, that residual variance is constant under all circumstances.

Figure 8.1 suggests otherwise: the large uncertainty is mainly due to the large variance of the working residuals at a specific range of logit values. An alternative method to estimate model uncertainty has been developed based on the relationship between the logit and working

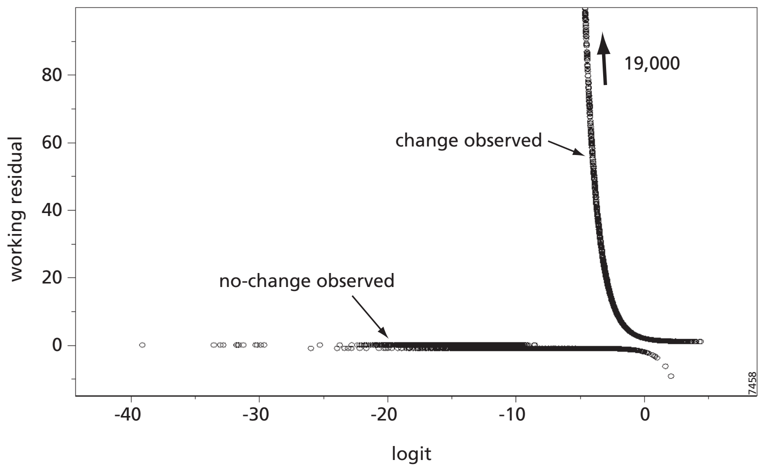


Figure 8.1 Relationship between the logit value and working residuals.

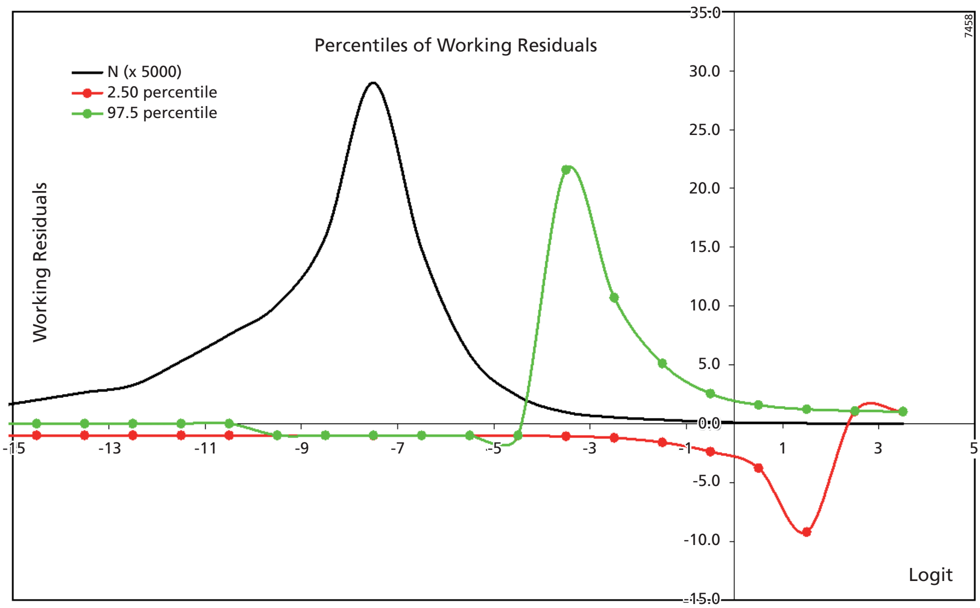


Figure 8.2 Percentiles of the working residuals and number of cases per binned logit value. For cases with a logit less than -5, the uncertainty is low.

residuals. These working residuals, which are response residuals that have been transformed from the response scale to the logit scale, vary from -10 up to 190,00. For logit values ranging from -40 to -10, the working residuals and model uncertainty are very small.

Locations where land use has changed occur at logit values starting from -10. At this point, the working residuals (and model uncertainty) are very high, but decrease rapidly with increasing logit value. Once the logit value exceeds zero, these working residuals are smaller, but those of locations with no-change start to increase. As the upper line in Figure 8.1 shows the residuals of observed land use changes, the upper prediction limit can be defined on the upper line. The lower prediction limit, in turn, can be defined on the lower line, which shows residuals for which no land use change has been observed.

Instead of using the variance to reflect the uncertainty in the residuals, the 2.5 and 97.5 percentile of residuals, binned over logit values, have been used to define the lower and upper prediction limits, respectively. Figure 8.2 shows the 2.5 and 97.5 percentile for binned logit values, with the bin size equal to one and N, the number of cases per bin. The 97.5 percentile drops from the maximum of 22.8 to -1 once the logit value becomes less than -4, while the 2.5 percentile decreases down to -9 and then jumps back to 1 at logit values above 2. Both these discontinuities are explained by the decreasing number of cases with observed land use change/no-change. Given the large number of cases, the uncertainty at the discontinuity in the 97.5 percentile is relative small.

The uncertainty in the 2.5 percentile is relatively higher and increases towards the discontinuity as the total number of cases diminishes. Below the discontinuity in the 97.5 percentile, uncertainty in working residuals is small. Therefore, the uncertainty in a major portion of the working residuals is small, as most cases lie below this discontinuity.

Based on the 2.5 and 97.5 percentile of the working residuals a lower (*LPL*) and upper prediction limit (*UPL*) can be defined, in conformity with:

$$LPL = -2 * SE + P_{wr}^{2..5}$$

$$UPL = 2 * SE + P_{wr}^{97.5} \qquad \text{(eq. 8.3)}$$

where SE is the standard error of the regression model and P_{wr} is the 2.5 or 97.5 percentile of the binned working residuals. For a logistic regression model, the upper and lower probability of land use change can be calculated according to:

$$P_{low} = \frac{1}{1+e^{-(Logit+LPL)}}$$

$$P_{up} = \frac{1}{1+e^{-(Logit+UPL)}} \qquad \text{(eq. 8.4)}$$

with *logit* being the value of the logit model. In order to obtain a continuous description, we fit a (fourth order) polynomial through the log-transformed 2.5 and 97.5 percentiles of the binned working residuals. This polynomial was used only over the relevant range of logit values – for example, above the discontinuity at logit -4 of the 97.5 percentile. At lower logit values, we used a constant of either -1 or 0, according to the value of the binned working residuals. Due to the uncertainty in the exact location of the discontinuity (breaking point) in the 2.5 percentile, we used the polynomial up to higher logit values of 4.

8.4 Results

In order to compare different sources of uncertainty, the *Logit*, 2 standard error, and 2.5 and 97.5 percentiles of the working residuals were computed.

The logit in Figure 8.3 shows the largest variation in low and high values, ranging from +25 to -25. The standard error is positive by definition and has a maximum of 10.0. In Figure

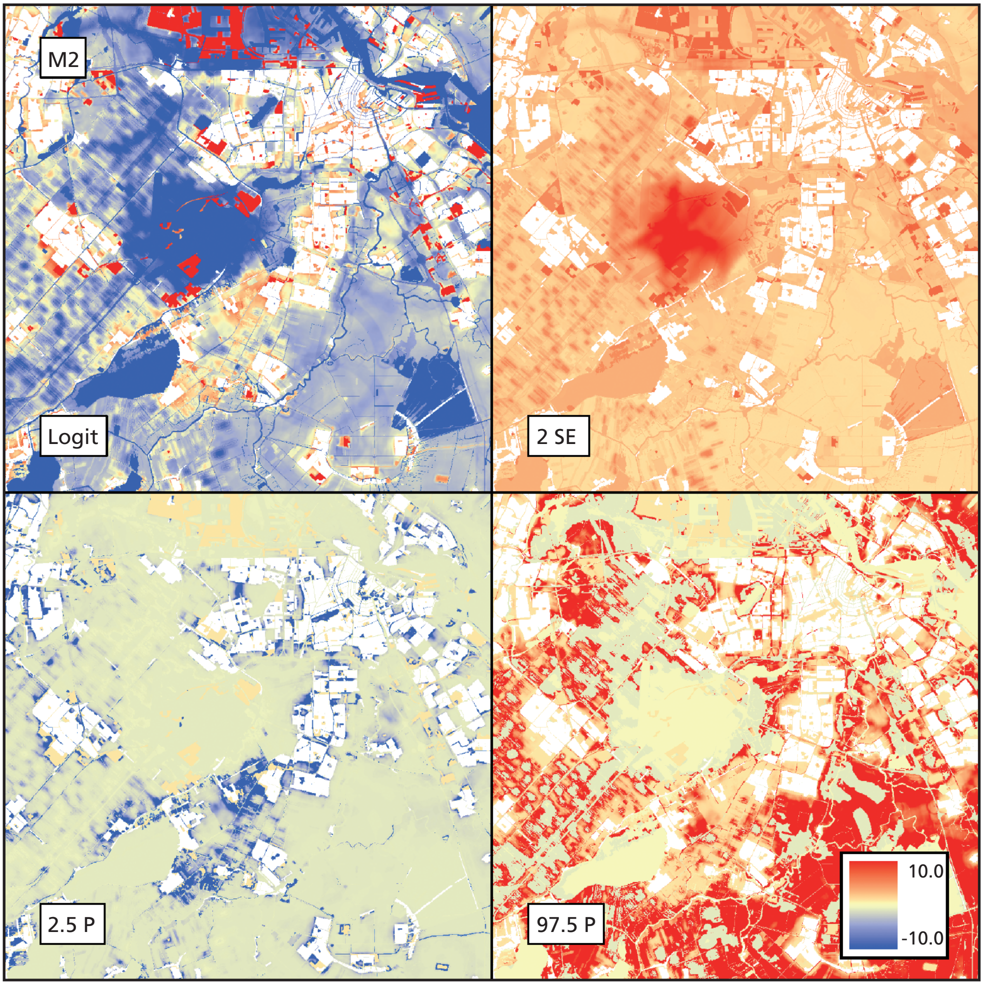


Figure 8.3 Logit, 2 ×standard error (2 SE), 2.5 and 97.5 percentile (P) of the working residuals. Current residential areas are white.

8.4 the 2.5 percentile is negative for most cases, with only a small fraction having values of less than -5. The 97.5 percentile is positive in most cases, ranging from -1.0 to 25.0. The lowest logit values and the highest standard error are found at Amsterdam Airport Schiphol. Open water surfaces, lakes, rivers and canals have a low logit value combined with a low standard error. The highest logit values are correlated with building sites, but these have an increased standard error. The 2.5 percentile has the lowest values in the neighbourhood of existing urban areas, but such areas have a relatively high logit value. The 97.5 percentile shows the opposite, with high values in agricultural areas which have low logit values. The 2.5 percentile shows the locations likely to transform, but this may not occur; in contrast, the 97.5 percentile shows locations unlikely to

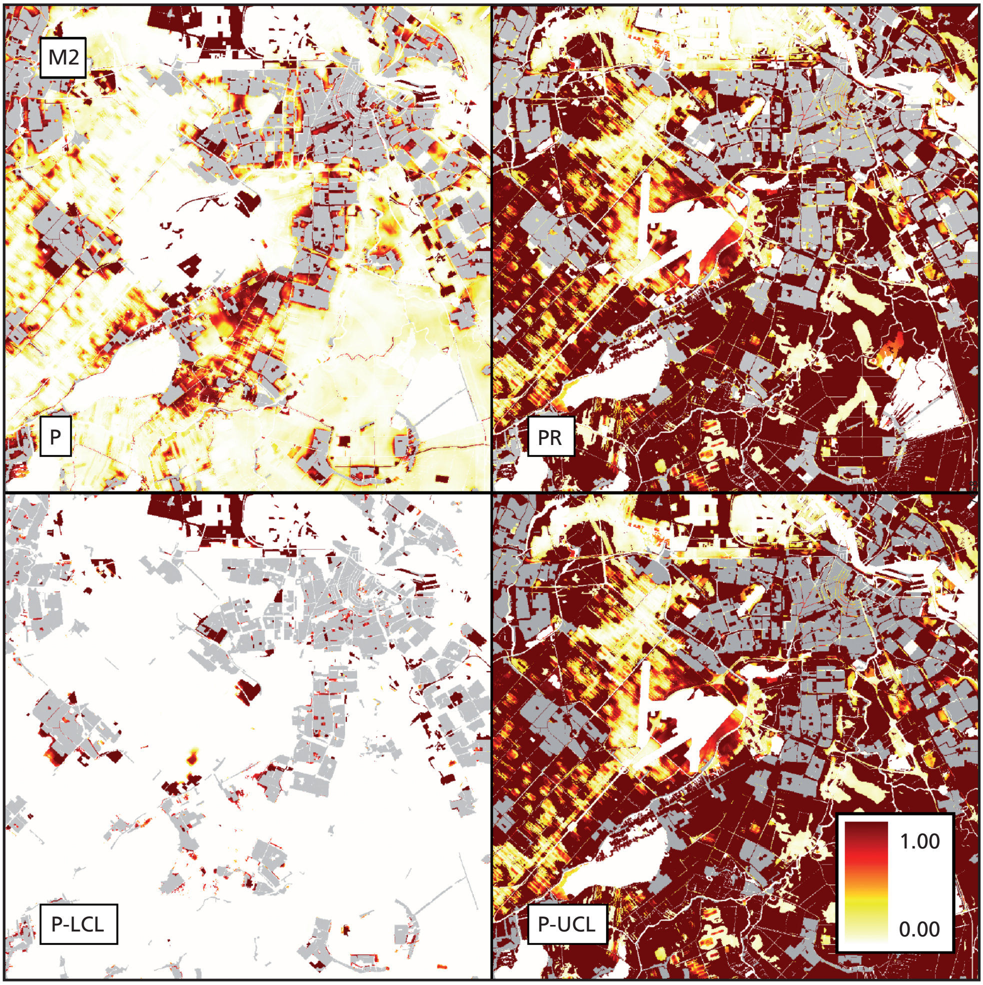


Figure 8.4 Probability (P), probability at lower and upper prediction limit (P-LPL, P-UPL) and the prediction range (PR) for the residential land use model. Current residential areas are in grey.

transform, but it may occur. Amsterdam Airport Schiphol and open water both have 2.5 and 97.5 percentile values close to zero, with a low uncertainty in the working residuals.

Figure 8.4 shows the probability of new residential area together with the probabilities at the lower and upper prediction limit and the prediction range. The prediction range is defined as the difference between the upper and lower probability and shows the uncertainty in the estimated probability of the model.

The estimated probability of new residential areas has relatively high values at the building sites and in the neighbourhood of existing urban areas. The probability at the lower prediction limit only shows building sites and locations within or very close to existing urban areas, while that at the upper prediction limit has high values for nearly all locations, with the exception of Amsterdam Airport Schiphol and surface waters. The prediction range reflects this large uncertainty in the probability of new residential area in most locations, with the uncertainty in the estimated probability being small only for building sites, Amsterdam Airport Schiphol and open water.

8.5 Discussion and Conclusions

An alternative method to estimate uncertainty in logistic regression was developed using the relationship between the residuals and logit value. The methodology is applicable when there are sufficient cases to estimate the upper and lower percentiles. The regression residuals are small below and above certain threshold logit values, and they vary widely only over a small range of logit values. Prediction limits have also been estimated.

The example shows that the uncertainty, as defined by the prediction range, is high for all locations, with the exception of Amsterdam Airport Schiphol, surface waters and building sites. While it is very unlikely that airports and surface waters will change to residential area, building sites are, by definition, likely to change in terms of land use. Logistic regression models minimise the residual variance. In general, larger models give rise to smaller residual variance, but they do not necessarily perform better in terms of prediction errors. Future research will focus on searching for an optimal model given the uncertainty in both the explanatory variables and residuals.

9 Future land use in the Netherlands: evaluation of the National Spatial Strategy

Reprinted from Chapter 3 of *Planning Support Systems: Best Practice and New Methods*. Geertman, S. and Stillwell, J. (Eds.) (2009) De Nijs, A. C. M. Future land use in the Netherlands: evaluation of the National Spatial Strategy. With kind permission of Springer Science and Business Media.

Abstract

This chapter presents the results of a study carried out by the National Environmental Assessment Agency in the Netherlands aimed at evaluating the National Spatial Strategy. It also reports the impact of this study on various stakeholders and the level of acceptance of the results among these stakeholders. A spatially detailed probability map of future urban developments in 2030 has been simulated by Environment Explorer, a land use simulation model that uses cellular automata. This map has been used to evaluate the potential effects of the new policy measures based on two scenarios. The highest urbanisation probabilities were found near existing urban areas, such as the 'Randstad' – in particular, around the major cities of Amsterdam, The Hague and Utrecht, where development pressures are high. The objectives of the new policy plan have been evaluated and potential problems areas identified. The simulation revealed large uncertainties surrounding future land use developments. Since its publication, the report has initiated discussions among scientists, planners and policymakers on spatial developments in the Netherlands.

9.1 Introduction

The Netherlands is a small but densely populated country with about 16 million people living in an area of 40,000 square kilometres. The 'Randstad' is a highly urbanised area of the country that comprises the major cities of Amsterdam, Rotterdam, Utrecht and The Hague, with a combined population of more than 5 million people. This conurbation is the economic heart of the Netherlands. Urban development threatens the natural and historical qualities of the landscape of the 'Green Heart', the central open area of this region. The first spatial plans for this area date from 1958 (WWdL, 1958, RPD, 1960) and these have been updated regularly since then (RPD, 1966, RPD, 1977, RPD, 1988, RPD, 1994). New plans have been drawn up for the creation of about 2,500 km^2 of forests and nature reserves by 2018 (LNV, 2002, MNP, 2002c). In addition, roughly 1,500 km^2 of land will be needed for new residential and industrial areas by 2030 (ABF, 2002). According to current policy, plans and trends, 10% of the land use will change over the next 30 years. The National Spatial Strategy (VROM et al., 2005), drawn up by the present Dutch Government and prepared to provide a spatial policy framework to guide all of these developments, was adopted by Parliament at the end of 2005.

Roughly speaking, the National Spatial Strategy comprises three major categories of spatial planning policies. These concerns:

- restrictive areas, where further urbanisation will be curbed in the interests of a range of objectives; these areas include the nature reserves and protected areas and the risk zones around Amsterdam Airport Schiphol and other centres of hazardous activities;
- National Landscapes, where development is possible within these areas as long as the core qualities of the landscape are conserved or enhanced; within these landscapes, local authorities may only build new housing to meet the demand resulting from local population growth;
- concentration areas, where new urban development will be clustered.

By order of the Dutch Ministry of Housing, Spatial Planning and Environment, the National Environmental Assessment Agency has evaluated the effects of this new National Spatial

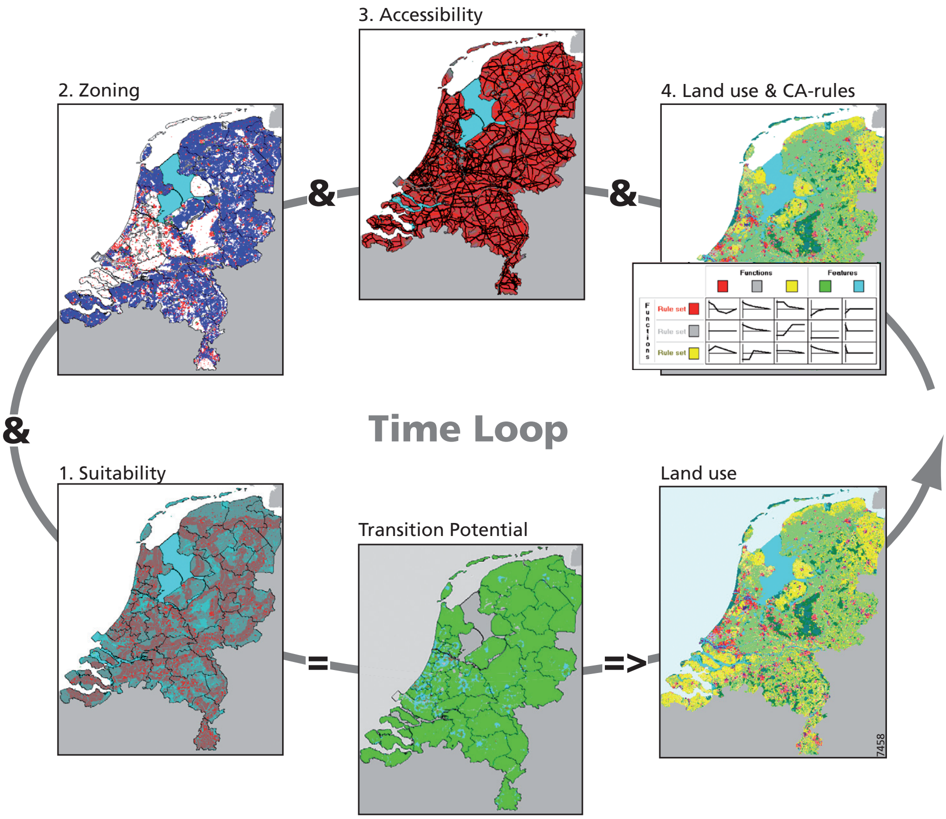


Figure 9.1 The Environment Explorer simulates future land use as a function of the suitability of each area, the relevant spatial planning policies, accessibility, current land use and the influence of the neighbourhood

Strategy, giving particular attention to a number of questions. How will the National Spatial Strategy affect future land use patterns in the Netherlands? Where, despite restrictions, is the pressure so high that new developments will ultimately be allowed to occur? What effect will the Strategy have on national landscapes? Will there be sufficient room in the concentration areas to accommodate all new developments? To answer these questions, future land use developments have been simulated by the Environment Explorer in accordance with a 'Trend' scenario and a 'Sprawl' scenario (de Nijs et al., 2004, de Nijs et al., 2005). Limiting urban spatial sprawl is one of the main political topics in the Netherlands.

This chapter takes a closer look at the nature of the land use model, at the Trend and Sprawl scenarios and at future land use developments in general and restrictive areas, national landscapes and concentration areas in particular. The results of this study are discussed and conclusions are drawn that relate to the impacts of the policies, potential problems and the application of the land use model itself. In closing, the effect of the study itself on the discussion of spatial developments in the Netherlands is discussed.

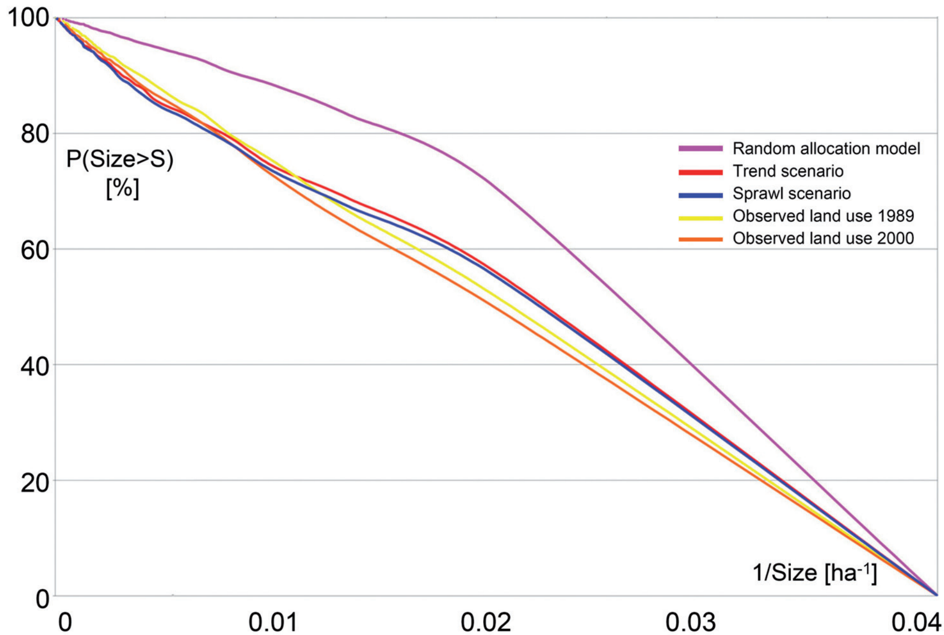


Figure 9.2 Size distributions of urban clusters according to Zipf's Law for the observed land use in 1989 and 2000, the simulated land use according to the two scenarios and the random allocation model in 2030

9.2 Environment Explorer

The Environment Explorer (White and Engelen, 1997a, de Nijs et al., 2001b, Engelen et al., 2003, de Nijs et al., 2004, MNP, 2005), a land use simulation model, was used to simulate future land use in the Netherlands. This model attempts to plot regional spatial developments as accurately as possible on a land use map with a grid size of 500 m. The model recognises 16 different land use classes, including residential use, industrial use, offices, facilities, recreation, greenhouse horticulture, forest and nature conservation. The location of land use change is determined by current land use, spatial policies, suitability, accessibility and the effect of current land uses in the neighbourhood (Figure 9.1). The amount of land use development is controlled in the scenarios. Zoning, suitability and accessibility maps have been compiled for each land use category. The zoning map shows the areas where development is permitted and where it is prohibited. The suitability map shows how suitable each cell is for that particular land use. Accessibility by car and public transport is calculated by a dynamic traffic model in the Environment Explorer. This model computes changes in accessibility given the changes in land use, population, employment and infrastructure growth (RIKS, 2002).

The neighbourhood potential is the influence exerted by the surrounding cells on the allocation of the various land uses. This neighbourhood potential is calculated by a cellular automata (CA) model in which a set of transition rules describes how the various land uses interact – by either attracting or repelling each other. For example, over a short distance, airfields and industrial sites can have a negative impact on residential development, but the latter in turn is positively influenced by the proximity of green space. A set of transition rules determines the spatial interactions between each combination of land use over a number of distances. Where the various types of development will take place is relatively uncertain. This uncertainty is built into the allocation model in the form of a stochastic parameter and it can be visualised in a probability map of future land use by running the model not once but 1000 times in a Monte-Carlo simulation.

For the evaluation of the National Spatial Strategy, the spatial interactions between the various land uses in the Environment Explorer were calibrated according to the observed spatial developments over the period 1989-1996 (MNP, 2005). Therefore, a semi-automatic calibration routine was added to the model with an 'error selection and correction' routine (Hagen-Zanker et al., 2005b). The criterion to select the combination of land uses that gives the largest error is based on the Fuzzy Kappa statistic (Hagen, 2003). The selected error is reduced by adjusting the transition rules and errors that cannot be resolved are placed on a taboo list. This procedure iterates until the model error is minimised.

Analyses of the observed land use changes in the period 1989-1996 revealed that 25% of the changes can be considered to have been improbable. These unlikely land use changes consist of changes from urban to agricultural uses or forest and nature conservation and they can be

Table 9.1 Results of the calibration and validation of the Environment Explorer based on the Fuzzy Kappa

	Environment Explorer	Random
Calibration 1989-1996	0.936	0.926
Validation 1996-2000	0.913	0.922

attributed to changes in the definitions of the land use classes over the years. For example, the grass bordering the roads is included in category roads in the map of 1989, while it is excluded from the map of 1996. In general, this is a major problem in the development of land use models (Fang et al., 2006). The model was subsequently validated for the period 1996-2000, based on the Fuzzy Kappa. The results indicate that the Environment Explorer performed better during the calibration period than a random allocation model (Visser and De Nijs, 2006), whereas during the validation period it did not (Table 9.1). Two factors can mostly explain this performance: (1) the validation period was rather short; (2) further changes in the definitions of the land use classes in the map of 2000. These latter changes have a number of effects, among which is a net decrease in residential area in the map of 2000 compared to the map of 1996, which can not be reproduced by the model.

The validation results over the short term are not so important; in contrast, the simulation results for the long term are very important. Nevertheless, it is impossible to verify these long-term simulations on observed land use changes. These simulated land use maps should be realistic, transparent and easy to explain, but it is not that easy to develop objective criteria to evaluate these characteristics. The first question that arises is just what are realistic land use patterns. At the very least they should all have the same morphological characteristics. One very tight morphological constraint on models of urban growth is Zipf's law (Zipf, 1949, Gabaix, 1999, Gabaix and Ioannides, 2003). Zipf's law for cities is one of the most striking empirical facts in geography and economics. For most countries, the size distribution of cities strikingly fits a power law: the number of cities with populations greater than S is proportional to 1/S.

Therefore, we verified the long-term spatial developments of the model against the size distribution of urban clusters. The stochastic parameter in the Environment Explorer, which determines the number of new clusters over the long term, was used to accurately fine-tune this relation (Hagen-Zanker et al., 2005a). The resulting size distributions for both the Trend and Sprawl scenarios in 2030 obey Zipf's Law (Figure 9.2) and both approximately overlap with the city size distribution of observed land use in 1989 and 2000. The size distribution of the random allocation model does not obey Zipf's Law as a large number of small urban clusters have been allocated randomly on the map. The relative percentage of 50-ha clusters has increased from 60% to nearly 80%. In the long term, the Environment Explorer outperforms the random allocation model.

9.3 The Trend and Sprawl scenario

Although the names suggest otherwise, the Trend and Sprawl scenarios are almost identical. There is only one essential difference between both scenarios: smaller settlements will grow relatively faster in the Sprawl scenario than in the Trend scenario. Why and how this has been implemented will be described at the end of this section. Here, the common aspect will be discussed. Implementation of a scenario in the Environment Explorer includes the definition of:

- initial land use;
- amount of land use developments;
- zoning, suitability and accessibility maps.

The spatial planning policies contained in the National Spatial Strategy have to be translated into either the input of these scenarios or the calibrated set of transition rules. In both scenarios,

Table 9.2 National targets per land use category for the period 2000-2030

Land use	Target (ha)	Land use	Target (ha)
Residential	87,675		
Industrial	35,975	Recreation	17,625
Services	6,225	Forest	51,525
Social/Cultural	13,450	Natural grasslands	173,025
Greenhouse horticulture	5,425	Nature conservation	11,300

initial land use is based on the land use/cover map from the Bodemstatistiek 2000 (CBS, 2003). Agricultural land use on this map is further broken down into 'grassland', 'arable' and 'other agricultural land' using a second land-cover map of the Netherlands, Landgebruikskaart Nederland, 2000 (Alterra, 2003). Residential land use is split into two categories based on the number of inhabitants: 'low-density population' and 'high-density population'.

The amount or quantity of land use developments can be defined regionally in the scenarios of the Environment Explorer. For both scenarios, the growth of residential and employment uses is based on the so-called High Land use Pressure Trend (HLPT) scenario (ABF, 2002). The growth of recreational uses and greenhouse horticulture was derived by extrapolating the regional developments from 1989 to 2000. Regional developments in forest, nature and natural grassland areas are based on the 'Nature Reference Map 2020' (Referentiebeeld Natuur 2020) (Goetgeluk et al., 2000). Table 9.2 summarises the national targets for each land use for the period 2000-2030.

The same suitability maps have been used in both scenarios. These are based largely on a study (Verburg et al., 2004a, Verburg et al., 2004d) that shows that soil conditions are not important in determining residential, employment and recreational land uses. The suitability maps for forest and nature conservation were also taken from this study. The suitability for natural grassland is based equally on the suitability for nature conservation and the presence of grassland in the 'Land use Map' of the Netherlands 2000. We assumed that suitability for agricultural uses is the highest in those areas where these uses are currently found. Areas under other agricultural uses are slightly less suitable and areas under all other urban uses are not suitable at all for agricultural use. The suitability map for greenhouse horticulture was derived from a specific study 'Potential options for greenhouse development' (LEI-DLO, 1997), which is based on the amount of light received per year and the proximity of distribution centres.

For each land use, the expansion locations and restrictions are defined in the zoning map. Expansion locations are locations where a specific land use change has been planned in accordance with to local, regional or national policy plans. Restrictions are locations where a specific land use is not allowed to develop. The expansion locations for the various land uses are based on the plans contained in 'Plans in the Netherlands' (Nederland in Plannen) and the 'New Map of the Netherlands' (NIROV, 2004). The expansion locations for forest, natural grassland and nature conservation marked on the zoning map were derived from the 'net boundaries' of the National Ecological Network 2003 (NEN). All of the restrictive policies in the National Spatial Strategy have been reproduced on the zoning map. The restrictions for each land use are listed in Table 9.3. The policies for concentration areas and National Landscapes policies in the National Spatial Strategy, as described in the Introduction, are expected to have less effect on spatial development than the expansion locations and restricted areas. In both scenarios, it is

Table 9.3 Definition of restrictive areas per land use

Restrictive areas	Source	Residential, Services, Facilities	Industry, Greenhouse horticulture	Forest, Nature conservation
20-ke contour Schiphol 2004, noise disturbance contour	NR[1]: PKB 3	X		
Coastal Foundation Zone & Weak Spots	NR: PKB 4	X	X	
Net boundaries NEN 2003	NR: PKB 5	X	X	
Protected areas under the Birds and Habitats Directives and the Nature Conservancy Act	NR: PKB 6	X	X	
Existing and New Nature 2003	(MNP, 2002b)	X	X	
Space for the rivers	(VenW, 2005)	X	X	
National buffer zones/Regional parks	(LNV, 2002)	X	X	
Risk contours for companies requiring an external safety report	MNP (in preparation)	X		
Schiphol bird protection zone	(VenW, 2004)			X

1 NR: PKB, National Spatial Strategy: Spatial Planning Key Decision (VROM et al., 2005)

assumed that the impact of expansion locations and restrictive areas is 20-fold greater than that of the concentration areas and the National Landscapes policies. This figure is based on expert judgement. The road network, motorway entrances and exits and the locations of the stations are specified in the Environment Explorer to allow the Traffic module to calculate accessibility (RIKS, 2002). Accessibility was calculated for the period 2000-2010 using the standard Dutch National Model System network of 1995 (HCG, 1997). The assumption was made that after 2010 all extensions to the infrastructure will be completed, as planned in the Multi-annual Programme for Infrastructure and Transport (VenW, 2003) and that no further extensions will be made to the road network before 2030.

Up to this point, the Trend and Sprawl scenarios do not diverge. However, the policies in the National Spatial Strategy give the municipal and provincial councils more freedom to approve new developments in the National Landscapes than they had previously. Whereas the municipal councils in the Green Heart formerly had virtually no opportunities for expansion, they are now permitted to build new homes to meet demands resulting from local population growth. It is most likely that the smaller settlements in the Green Heart will grow relatively faster under this new policy than larger towns and cities outside this National Landscape. In time, this change in policy will affect the transition rule for housing development in the model. These rules have been calibrated on historical land use changes. Therefore, to estimate the impact of this change in policy, we constructed a Sprawl scenario in which the transition rules for housing development were adjusted to allow the smaller settlements to grow slightly faster than they do in the Trend scenario. The effect of this change in policy and the size of the adjustment to the neighbourhood rules are difficult to determine. However, the actual difference between the rules should be minor. Therefore, the effect of this adjustment on urban clusters sizes in 2030 has been verified against Zipf's Law.

9.4 Results

9.4.1 Urbanisation probabilities in 2030

Figure 9.3 depicts the urbanisation probabilities in 2030 according the Trend scenario. The urban area comprises the following land use categories in the Environment Explorer: housing, employment, recreation, sports fields and greenhouse horticulture complexes. The map shows that a number of urban areas will probably expand towards each other and eventually merge. In the Randstad conurbation, the open spaces between The Hague and Rotterdam will close. The city of Utrecht will expand to the south, merging with several small towns. New development around Amsterdam will be more dispersed.

In this scenario, the expansion of residential and employment land uses is heavily determined by the restrictive areas, including, among others, nature reserves and protected areas, regional parks and land in the floodplains reserved for flood control, water retention and habitat development. Developments in the concentration areas are given a preference over other areas. The National Landscapes are kept free of new housing and employment land as much as possible.

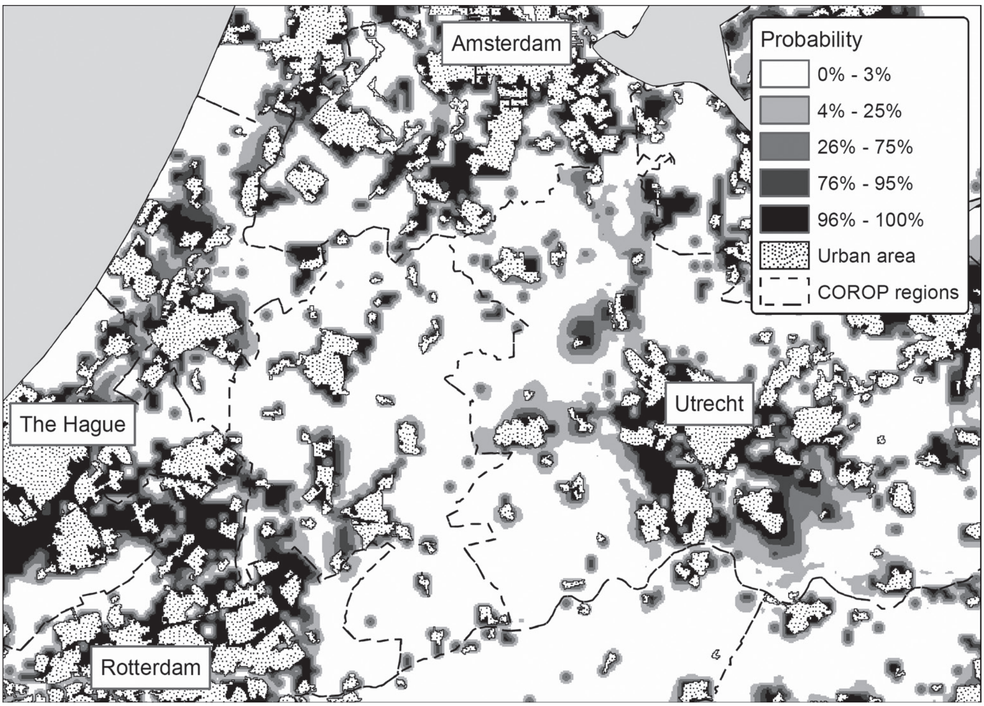


Figure 9.3 Urbanisation probabilities in the Randstad conurbation in 2030 under the Trend scenario

9.4.2 Pressures on land in the restrictive areas

The Trend scenario was used to determine the relative pressure on land in the restrictive areas. These restrictive areas are only built on if all other available land within the COROP region has already been developed. COROP regions are the same as the NUTS Level 3 regions. The 40 regions are sub-divisions of the provinces in the Netherlands, each consisting of a central town or city and catchments area. Only if growth exceeds the amount of available land will these last open areas be developed. The probability of urban development in the restrictive areas was mapped for all urban restrictions, with the exception of the 20-ke contour around Amsterdam Airport Schiphol. This safety zone around Amsterdam Airport Schiphol does not exclude industrial development and greenhouse horticulture.

There appears to be little pressure for urban development in the restrictive areas (Figure 9.4) and only in the Randstad is the pressure on land so high that there is a chance of urbanisation in restrictive areas. New urban development may occur in the restrictive areas north of The Hague, around Amsterdam and in the vicinity of small towns in the nature conservation area to the north-east of Utrecht.

9.4.3 Urban development in the National Landscapes

Figure 9.5 shows the differences in urbanisation probabilities between the Trend and Sprawl scenarios in the Green Heart, one of the National Landscapes mostly threatened by

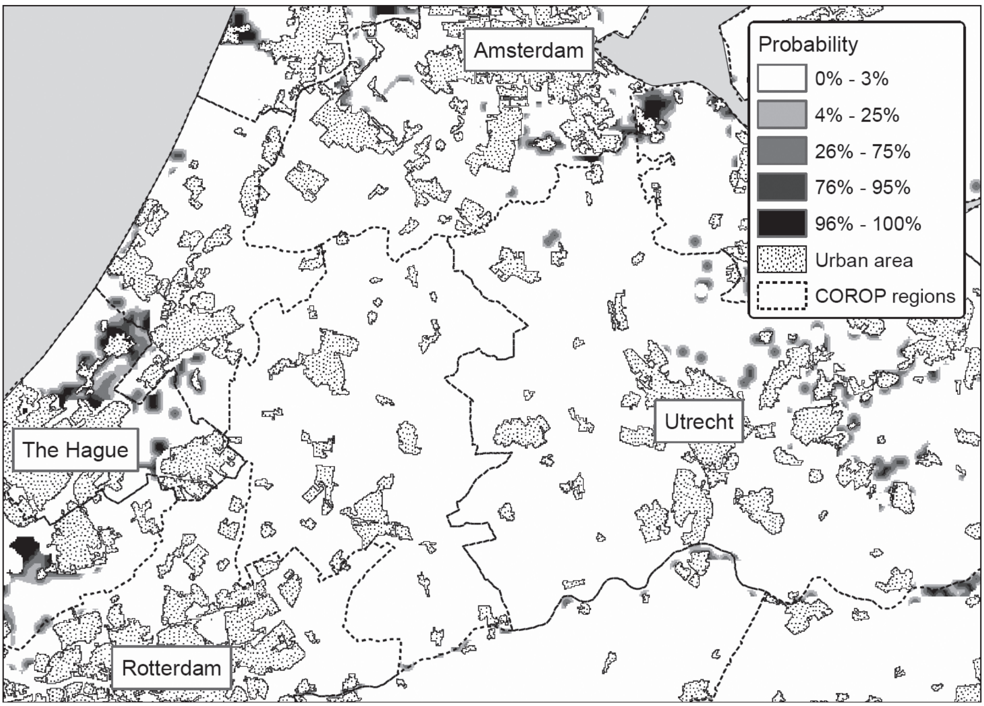


Figure 9.4 Probability of housing development in the restrictive areas

urbanisation. Here, the probability of urbanisation is slightly higher under the Sprawl scenario, particularly in the eastern part of the Green Heart. The probability of urbanisation decreases in the area south-east of Utrecht and increases at the western side of the city. The areas of different grey shading indicate sites where the probability of urbanisation under the Trend scenario is higher; the black areas indicate where the urbanisation probability under the Sprawl scenario is higher. In the Sprawl scenario new residential development is less concentrated around the main cities, but it will be located more often near small villages in the Green Heart.

9.4.4 Space in the concentration areas

One of the main principles of the National Spatial Strategy is to locate new urban development in the concentration areas. The National Spatial Strategy states that the 'concentration percentage' in these areas should – at the very least – remain the same, but at the same time it does not clearly define concentration percentage. In addition, land also has to be reserved for water, nature and landscape conservation, recreation areas, sports fields and agriculture. In this analysis the concentration percentage (*P*) is taken to be the percentage of the urban area that is located in the concentration area ($A_{Urban,\ Concentration\ Area}$) relative to the total urban area in the province ($A_{Urban,\ Province}$):

$$P = \frac{A_{Urban, ConcentrationArea}}{A_{Urban, Province}} * 100 \qquad \text{(eq. 9.1)}$$

Urban areas in this definition include residential and employment land, facilities, parks and green space, sports fields, recreation areas, roads, railways and greenhouse horticulture. It is

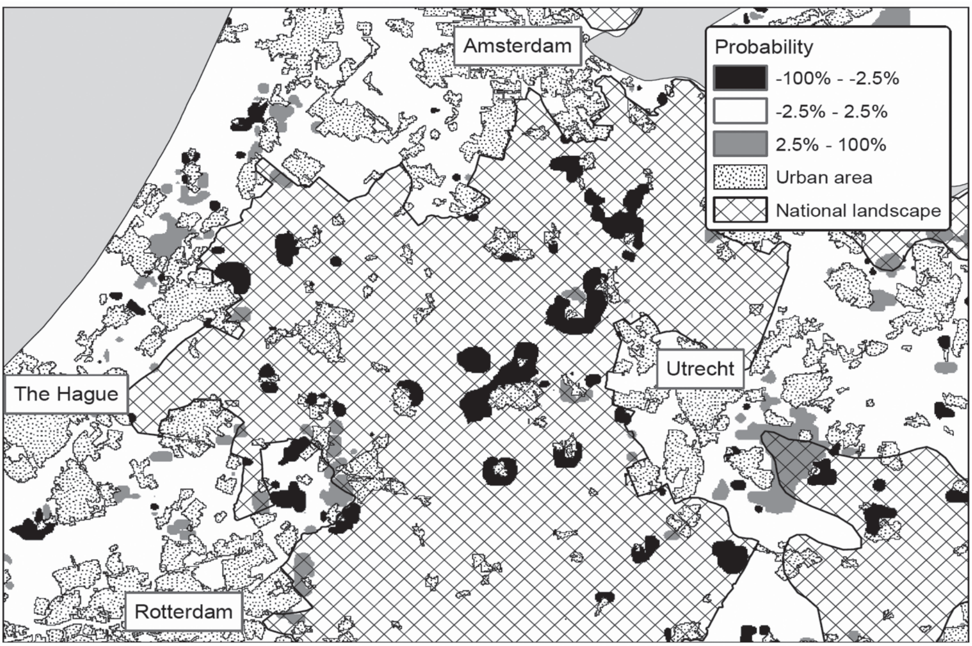


Figure 9.5 Difference in urbanisation probabilities between the Trend and Sprawl scenarios

Table 9.4 Available, required and developed land (ha) in the concentration areas and the concentration percentages in 2000 and 2030 under the Trend scenario

	Available land	Required land	Developed land	Concentration % 2000	Concentration % 2030
Groningen	8,400	989	942	21.0	20.0
Drenthe	5,319	758	682	10.3	9.2
Overijssel	15,543	2,665	2,596	26.9	26.2
Gelderland	15,028	4,782	5,456	19.1	21.8
Utrecht	10,734	8,838	8,045	57.9	52.7
Noord-Holland	9,113	10,357	7,966	60.4	46.5
Zuid-Holland	24,931	17,195	16,202	73.0	68.8
Noord-Brabant	33,561	11,646	11,174	43.0	41.3
Limburg	5,634	6,491	4,444	36.2	24.8
Flevoland	6,012	2,255	2,991	28.8	38.2

There are no concentration areas in the provinces of Friesland and Zeeland

questionable whether all spatial developments can be accommodated within these concentration areas. Therefore, for each province, the available, required and developed land has been determined for 2030. The available land is the land which will still be available for new urban developments; it is defined as the agricultural land and building sites in the land use map of 2000, barring all physical and spatial restrictions. The required land in 2030 will be the land needed to maintain – at the very least – the same concentration percentage. It is estimated by multiplying the growth of urban land per province, in accordance with the scenario, by the concentration percentage in 2000; as such, it is the amount of land that will be needed in the concentration areas for the development of urban land uses. The developed land is the land which will most probably be developed in 2030. The developed land in 2030 was calculated from the urbanisation probability in 2030 according to the Trend scenario.

On the basis of this analysis, there is not enough land available in the concentration areas in the Randstad and in Limburg to accommodate all new developments (Table 9.4). Accordingly, the concentration percentage in these provinces will decline sharply in the future. The differences between the available, required and (probably) developed land provides some indication of the policy inputs needed to cluster new urban development in the concentration areas. There are no concentration areas in the provinces of Friesland and Zeeland

9.5 Conclusions and future directions

For the ex ante evaluation of the National Spatial Strategy, future land use was simulated by the Environment Explorer. Observed development trends were used to comprehensively calibrate and validate the Environment Explorer for short-term projections. The Environment Explorer outperformed a random allocation model in the calibration period, but not over the validation period. Land use models generally perform poorly in validation studies when used for large-scale applications (Pontius Jr et al., 2008), while their performance in local-scale applications, such as at the city level, appears to be better (Hagoort, 2006). The calibration and validation of land use models are impeded by classification and aggregation errors in land use maps (Pontius et al., 2004b, Pontius and Spencer, 2005, Fang et al., 2006). A novel methodology has been applied to verify long-term projections. The distribution of urban cluster sizes in future land use maps was verified using Zipf's Law. The calibrated model was used to estimate the urbanisation probabilities in 2030 for a Trend and a Sprawl scenario using Monte Carlo simulations. The effects of policy on the restrictive areas, the National Landscapes and the concentration areas were examined.

The probability of urbanisation in 2030 is highest near existing urban areas and is influenced by spatial planning policy. Urban development in restrictive areas is avoided as much as possible. Only in the Randstad is the pressure on land so high that it is likely that urban development will take place in the restrictive areas. Urbanisation in the National Landscapes will probably increase as a result, particularly in the Randstad and Limburg where development pressures are the highest. There is sufficient space in most of the concentration areas to accommodate all new developments. Given the scenario assumptions, there will be insufficient land in the provinces of Noord-Holland and Limburg to accommodate all new urban developments in the future while retaining enough land for water, nature and landscape conservation. This analysis identifies those provinces for which problems may be expected over the long term with implementing national spatial planning policies.

This evaluation of the National Spatial Strategy using the Environment Explorer provides detailed insights into the feasibility of the policy goals and potential problems. The study shows how land use models can be used to determine the effects of spatial planning policies on future spatial development. Even so, the results of such modelling studies are inherently uncertain. The Monte-Carlo simulation enables the large uncertainties surrounding the location of developments to be visualised through the use of urbanisation probabilities. However, the uncertainty in the growth of land use and the effectiveness of current policies has not yet been incorporated in the simulated scenarios. These results provide policymakers with the opportunity to take potential problems into account beforehand. Monitoring programmes have been developed, especially those focussing on the simulated hot spots of urban developments in restricted areas. Additional measures can be defined to circumvent potential problems in the national landscapes and concentration areas. Moreover, the detailed spatial results have made it possible to determine various – potentially adverse – effects of these new policy plans; for example, to assess future noise levels and risk contours near airports in the Netherlands (Dassen, 2005).

For this study, the Environment Explorer was calibrated for a relatively short period, given the slow pace of spatial development. Calibration on a consistent dataset over a longer time-span could improve simulation results. Moreover, a major part of the land use changes in the calibration period was dubious: urban areas changing to forest, nature conservation or agricultural uses. Therefore, future research should be directed towards reducing the uncertainties in the results of the land use model by taking advantage of better monitoring data. For example, to circumvent the classification problems in land use maps, land use models should utilise the primary remote sensing data instead of the interpreted results in the land use maps. Moreover, all information available in subsequent remote sensing images (time-series) should be used to identify current land use developments and develop algorithms to simulate future land use developments. In terms of the allocation algorithm, the optimal spatial and temporal resolution needs to be determined. How well does the same model perform at different spatial or temporal scales? To what extent is it possible to aggregate land use categories? How well does the same model distinguish two or twenty land use categories? Finally, to measure these differences in model performance, it is of utmost importance to develop a set of sensitive and appropriate indicators.

As this study was first published in 2005 it is possible to look back on its impact in the Netherlands. Shortly after publication, the results of this study were presented to the relevant policymakers at the Ministry of Housing, Spatial Planning and Environment. The conclusions were quoted by various stakeholder groups in the Netherlands. One of these, the Netherlands Society for Nature and Environment (SNM, 2005), published an article in its newsletter calling on the Dutch population to stop further urbanisation. In general, the results of this study have been widely accepted. In Revolutionary Future for Housing Construction, Hugo Priemus (2005) explicitly supports the results of the study and states his opposition to the popular idea of a 'transformation of the Randstad', as suggested by planners and administrators. The results of this study have been used in several regional scenario studies in the Netherlands (SafeCoast, 2005, Witmond et al., 2006, Dijk, 2007). It is difficult to determine whether this study influenced policy, but the results have been discussed widely among scientists, stakeholders and policymakers.

10 Estimating the influence of the neighbourhood in the development of residential areas in the Netherlands

Reprinted from *Environment and Planning B Planning and Design*, forthcoming. De Nijs, A. C. M. & Pebesma, E. J. Estimating the influence of the neighbourhood in the development of residential areas in theNetherlands. Pages No. 125-150. Copyright (2009), with permission from Pion Limited, London

Abstract

Land use models have been developed for predicting future land use patterns. We describe a logistic regression model for residential developments in the Netherlands that takes the neighbourhood effects of land use and infrastructure into account. Hereto, a new methodology was developed, Single Parameter Impact Estimation (SPIE) for estimating the influence of land use and infrastructure in the neighbourhood on residential developments. The model also incorporates initial land use and a set of policy- and suitability-related variables. In essence, the neighbourhood effects typically simulated with cellular automata-based land-use models have been integrated into a logistic regression model. The model was calibrated and validated on historical land use developments in the Netherlands.

Based on the model, residential areas develop preferably at the fringe of the city, on agricultural land in the neighbourhood of residential areas. The probability of residential developments depends on the cluster size of the land use categories. Large clusters of allotments, cemeteries or sporting fields are relatively attractive for residential developments, while forest, nature and industrial areas become more attractive when present in small clusters.

10.1 Introduction

Land use change is directly related to the socio-economic development of a society, which in turn is affected by factors such as changing agricultural practices, industrialisation and urbanisation. It is therefore important to determine where these changes will take place in order to evaluate the potential adverse effects they may have on nature and the environment. To simulate the location of these socio-economic developments, Verburg et al. (2004c) distinguished two relevant approaches to quantify the association between the locations of these developments and their driving factors. These approaches are based on:

- statistical relations, often using logistic regression models to relate historical land use changes to changes in driving forces and location characteristics (Pijanowski et al., 2000, Aspinall, 2004, Verburg et al., 2004d, Zang and Huang, 2006), or
- a combination of calibration and expert knowledge, as is the case with many cellular automata (CA) models (White et al., 1997, Clarke and Gaydos, 1998, de Nijs et al., 2004, Torrens, 2006).

The CA-based land use models focus on spatial interactions with neighbouring land use. They predict future land use developments given the initial land use of a cell and the effect of land use in the neighbourhood as described by transition or neighbourhood rules. While these simple rules, which describe the relation between two land use types, can generate very realistic land use developments, their quantification is a fundamental yet challenging step in the creation of CA-based land use models (Kocabas and Dragicevic, 2006) Diverse methods have been employed to acquire the relevant transition rules (see Wu and Webster, 1998, Li and Yeh, 2000, Yeh and Li, 2002, de Almeida et al., 2003, Yeh and Li, 2003, Hagen-Zanker et al., 2005b). The neighbourhood interactions, which form the essence of CA-based models, are increasingly being used in statistical and spatial econometric models. Dendoncker et al. (2007) compared the variance explained by the non-neighbourhood-based variables with that explained by the neighbourhood-based variables in a logistic regression model, concluding that "accounting for neighbourhood interactions makes an essential contribution to analysing spatial patterns of land use"

Given this tendency in regression-based land use models towards CA modelling, the question arises whether it is possible to build a hybrid model that incorporates the neighbourhood interactions from a CA-based model into a logistic regression model.

The main objective of this study was to build a logistic regression model for residential developments in the Netherlands that includes a wide range of neighbourhood interactions accounting for effects from policy- and suitability-related variables. As will be shown, the major challenge in achieving this objective was coping with spatial autocorrelation and multicollinearity in the data.

This paper describes the selection of explanatory variables, the new methodology developed to deal with multicollinearity in the neighbourhood variables and to estimate the neighbourhood rules and the calibration and validation of the model. The impact of initial land use, policy and suitability variables and the neighbourhood rules are discussed. Finally, a number of conclusions are drawn on the methodology and determinants of land use change in the Netherlands.

10.2 Data

A wide range of potential variables has been selected from different sources to account for the development of residential areas in the Netherlands between 1993 and 1996. The data set can be classified in four broad categories: land use data, neighbourhood composition, spatial policy measures and suitability.

10.2.1 Land use

Land use maps of the Netherlands are available for 1989, 1993, 1996 and 2000 (CBS, 1997, CBS, 2000). Those on land use developments between 1993 and 1996 proved to be the most

Table 10.1 Description of land use categories.

Land use	Description
Agriculture	Grasslands, arable land, horticulture and orchards
Greenhouse farming	Greenhouses
Forest	Forests
Nature	Natural areas, heather and swamps
Residential	Residential areas including local roads
Industrial areas	Industrial areas including harbours
Offices & services	Offices, schools, hospitals, cultural facilities, churches, shopping mall, prisons, etc.
Building sites	All building sites, residential, industrial, harbours, offices, etc.
Parks	Parks, public gardens
Cemeteries	Cemeteries
Sporting fields	Sporting fields
Allotments	Allotments
Recreation	Camping sites, amusement parks, zoos
Railroads	Rail, tram and subway lines
Roads	Paved roads
Airports	Passenger centers, run ways including neighbouring grasslands
Other land uses	Mining, unpaved and semi-paved roads, dump and wreck sites
Fresh water	Rivers, lakes and streams
Sea water	Sea water
Foreign country	Belgium and Germany

Table 10.2 Radius and lower and upper limit of the ring-shaped neighbourhoods.

Radius [km]	0.1	0.15	0.2	0.3	0.4	0.6	0.8	1.2	1.6	2.4	3.2	4.8	6.4
Lower [km]	0.09	0.135	0.18	0.27	0.36	0.54	0.72	1.08	1.44	2.16	2.88	4.32	5.76
Upper [km]	0.11	0.165	0.22	0.33	0.44	0.66	0.88	1.32	1.76	2.64	3.52	5.28	7.04

consistent because there were no changes in the definition of land use categories and in data collection procedures during this period. The model was therefore calibrated for this time interval and validated on land use developments between 1996 and 2000. We reclassified land use maps of these years to comparable sets of twenty land use categories (Table 10.1). The category "Building site" encompasses a wide range of land uses, including residential and industrial developments, office buildings and harbour areas. As only about 20% of building sites are developed as residential areas, this variable must be included in the logistic regression model in order to predict the conversion of building sites into residential areas.

10.2.2 Composition of the neighbourhood

The composition of the neighbourhood is described by the enrichment factor according to Verburg et al. (2004a) or by the location quotient (Moineddin et al., 2003). The enrichment factor is defined as the occurrence of a land use type in the neighbourhood of a location relative to the occurrence of this land use type in the study area as a whole:

$$F_{i,l,d} = \frac{n_{i,l,d} / n_{i,d}}{N_l / N} \qquad \text{(eq. 10.1)}$$

where
- $F_{i,l,d}$ is the enrichment factor of neighbourhood *d* on location *i* with land use *l*,
- $n_{i,l,d}$ is the number of cells in neighbourhood *d* at location *i* with land use *l*,
- $n_{i,d}$ is the total number of cells in neighbourhood *d* at location *i*,
- N_l is the total number of cells in the study area with land use *l* and
- *N* is the total number of cells in the study area

The relative enrichment of each land use in the neighbourhood was calculated for ring-shaped neighbourhoods in terms of thirteen distance classes, starting at a distance of 100 m, i.e. the immediately neighbouring cells, up to 6400 m (Table 10.2).

The relative accessibility of the neighbourhood by three main transport nodes – intercity stations, railway stations and slip roads – was expressed as enrichment factors. Consequently, the composition of the neighbourhood was defined at thirteen distances for all twenty land uses and three transport nodes, resulting in 299 neighbourhood variables

10.2.3 Policy and suitability maps

Government policy on spatial planning is a strong determinant of residential developments in the Netherlands. The very first spatial policy was developed in 1958 for the so-called "Randstad" region, a highly urbanised area that forms the socio-economic heart of the Netherlands. The four major cities in the Netherlands, Amsterdam, Utrecht, Rotterdam and The Hague, define this area; consequently, the natural and historically valuable inner space of this region, the so-called Green Heart, is highly threatened by urbanisation. Government policy on spatial planning is updated regularly (RPD, 1988, VROM et al., 2004).

Various national policy measures and a number of EC directives affecting spatial developments were implemented between 1993 and 1996 (Table 10.3), among which were national restrictions to urban developments in the buffer zones between the major cities, the so-called 20-Ke zone around Amsterdam Airport Schiphol and the Green Heart (RPD, 1988). We also included the townships designated as growth centres during this period in our analysis. While only 1% of these townships have been developed as residential areas, this variable must be included in the model to determine the effect of such policies on residential developments. The EC-Bird (EU, 1979) and EC-Habitat directives (EU, 1992) as well as the Dutch Ecological Main Structure and the Nature Protection Areas (VROM et al., 2004) are designated nature protection areas. Also, the direct proximity of power lines and national and regional gas pipes impose restrictions on urban developments. All of these factors have been included in the data set. All policy maps are binary maps.

Many suitability-related variables, such as altitude, soil composition and ground water level, have been shown to be relatively insignificant in terms of affecting residential development in the Netherlands (Verburg et al., 2004d). Consequently, only "subsidence in peat areas" and the "urban area in 1993" have been included in the model, as both factors have been found to affect development costs in the Netherlands. In most peat areas, piles under the houses provide adequate structural support, but in heavily subsiding areas extra costs are involved because of subsidence of the sewer system. Peat subsidence ranges from 0.0 to -13.81mm/year in the Netherlands.

Table 10.3 Description of policy and suitability maps.

Policy Map	Description
Designated growth centers	Twenty-two municipalities designated to accommodate new residential developments
Green Heart region	Policy to avoid urbanisation in the Green Heart region
Buffer zones	Policy to keep the major cities in the Randstad separated
20-Ke zone 1993	Noise disturbance and safety zone around Amsterdam airport
Bird directive areas	EU protection areas for the conservation of wild birds
Habitat directive areas	EU protection areas on the conservation of natural habitats
Gross ecological structure	National policy plan to conserve and develop natural habitats
Nature protection areas	National Act protecting specified nature conservation areas
Groundwater protection areas	National act for groundwater areas
Drinking water areas	National act for drinking water areas
National gas pipes	Direct neighbourhood (25 m) of national gas lines is restricted for urban developments
Regional gas pipes	Depending on the diameter of the gas pipes, the direct neighbourhood of regional gas lines is restricted for urban developments.
Power lines > 111 kV	Direct neighbourhood of power lines greater than 111 kV is partly restricted for residential developments
Suitability Map	
Subsidence in peat areas	Map with the subsidence in mm/year in peat areas
Urban areas 1993	Map with the delineation of urban areas in 1993 based on the land use map of 1993

Table 10.4 Number of cells and weights between brackets in the logistic regression for cells changing to residential area and cells not changing to residential area.

	Total	To residential	Not (To residential)
Cases (Weight)	100,003	16,938 (1)	83,065 (116)

10.2.4 Study area and data sampling

The study area is the Netherlands in its entirety, including a 2-km-wide zone along the national borders and coastline. All data were collected at the 100 × 100-m grid level. All grid cells with residential area development were included in the data set. To limit the size of the data set, we included only a limited selection of cells not changing to residential area (cells not changing or changing to one of the other land uses); these cells were weighted in the logistic regression to keep the ratio of changes constant. Table 10.4 shows the number of cells and the weights used for the binary logistic regression.

10.3 Methods

All neighbourhood variables and the peat subsidence maps had to be standardised to prevent a slow fit or the non-convergence of the regression algorithm.

The impact of one land use category over distance on residential developments can be described by a neighbourhood rule, which is approximated by estimating the coefficients of the thirteen neighbourhood variables for each land use category or transport node relative to those of the other explanatory variables – initial land use, policy and suitability.

As many of these neighbourhood variables are highly correlated, the regression coefficients are highly variable and poorly determined (Hastie et al., 2001). Multicollinearity inflates the variances of the parameter estimates, resulting in estimates of the regression coefficients with the wrong signs and magnitudes. The model then becomes unstable and, consequently, incorrect conclusions are drawn on the relationships between independent and dependent variables.

To reduce these model instabilities, we used a logistic ridge regression (Venables and Ripley, 2002) to explain the development of residential areas using all policy, suitability, initial land use and neighbourhood variables as explanatory variables. Ridge regression shrinks the regression coefficients by imposing a penalty on their size. To this end, the logit model,

$$Logit(P_i) = \log[p_i/(1-p_i)] = \beta_0 + \sum_{j=1}^{J} \beta_j x_{i,j} \qquad \text{(eq. 10.2)}$$

where

- P_i is the probability that the land use at location i changes to residential land use,
- β_0 is the intercept,
- $x_{i,j}$ is the j-th explanatory variable at location i,
- β_j is the j-th coefficient to be estimated,
- J is the number of explanatory variables

is estimated. The weighted sum of the residual sum of squares and the sum of the squared regression coefficients are minimised as

$$\sum_{i=1}^{N} (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^{J} \beta_j^2 \qquad \text{(eq. 10.3)}$$

where

- y_i is the i-th observation,
- $\hat{y}_i$ is the i-th prediction value and
- λ is the decay factor.

The decay factor, λ, controls the amount of shrinkage: the larger the value of λ, the greater the amount of shrinkage, thereby reducing the sum of the squared regression coefficients. The regression coefficients have been estimated in R language for statistical computing (http://www.r-project.org/) using the function "multinom" in package "nnet".

The optimal decay factor, λ, was determined based on the agreement between the reference map and the simulated map, as defined by Pontius et al. (2004a). In this method, correspondence between the observed and simulated map is compared at increasing spatial resolutions. The reference map was a binary map that included all observed residential developments between 1993 and 1996. The simulated map was obtained by allocating the total observed developments in one step to the locations with the highest probability, given the land use in 1993. As this agreement starts to decrease over all spatial resolutions for λ larger than one, a decay factor of one was chosen.

A model that included all variables in the data set was fit using this logistic ridge regression, but the coefficients still remained highly variable, making it impossible to perform consistent estimations of the influence of the neighbourhood variables. Multicollinearity therefore remained a problem to be solved.

Variables causing multicollinearity occasionally can be combined into a single variable. For example, the variables bodyweight and height may be combined into a body mass index. For our model, the pertinent question was how to combine these neighbourhood variables while preserving their impact at each distance?

10.3.1 Single-parameter impact estimation.

Although the variables cannot be estimated in one regression, there is a possibility of estimating them in separate regressions. The assumption is that the impact of each neighbourhood variable can be estimated in a separate logistic regression given a basic model (BM) that, to reduce model instabilities, includes all variables except those with correlation coefficients larger than 0.5. Therefore, correlation coefficients were calculated between all explanatory variables. If the correlation coefficient between two explanatory variables was larger than 0.5, one of these variables was arbitrarily dropped. Once the influence of the neighbourhood variables at the various distances has been estimated, the variables can be combined into a single neighbourhood variable for each land use. In the basic model the development of residential area is explained by all policy, suitability and initial land use variables and a subset of the neighbourhood variables (Table 10.5).

Table 10.5 Neighbourhood variables included in the basic model.

Distance [km] **Land use**	**0.1**	**0.15**	**0.2**	**0.3**	**0.4**	**0.6**	**0.8**	**1.2**	**1.6**	**2.4**	**3.2**	**4.8**	**6.4**
Agriculture					x								
Greenhouse farming					x								
Freshwater							x						
Seawater							x						
Foreign country							x						
Residential areas	x								x				
Forest	x								x				x
Nature	x										x		x
Industrial areas	x						x				x		
Commerce & offices	x				x				x				
Airports	x								x				x
Building sites	x						x				x		x
Parks	x				x		x		x				
Sporting fields	x				x		x		x				
Allotments	x				x		x		x		x		
Railroads	x				x		x		x		x		
Cemeteries	x				x		x		x		x		x
Recreation	x				x		x		x		x		x
Roads	x				x		x		x		x		x
Other land uses	x				x		x		x		x		x
Intercity stations	x	x	x	x	x	x	x	x	x	x	x	x	x
Railway stations	x	x	x	x	x	x	x	x	x	x	x	x	x
Slip roads	x	x	x	x	x	x	x	x	x	x	x	x	x

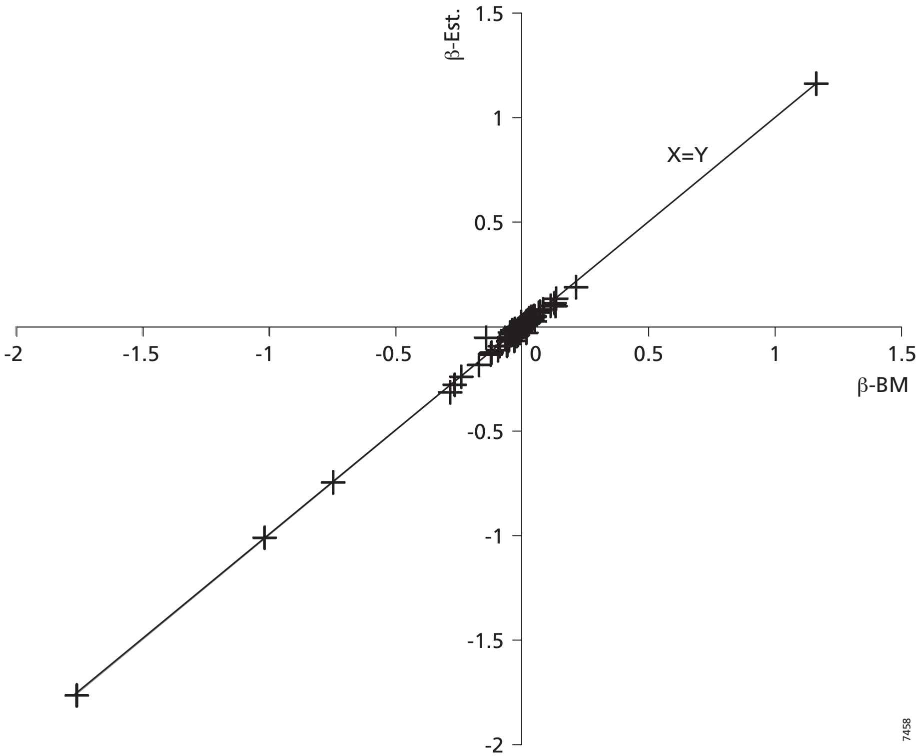


Figure 10.1 Comparison of the estimated regression coefficients for the neighbourhood variables with regression coefficients of the basic model.

Land use categories with a large cluster size are represented by at least one neighbourhood variable, such as agriculture, while land use categories with a relatively small cluster size are represented by more neighbourhood variables.

The neighbourhood rules of a certain land use category were approximated by estimating the regression coefficients of all thirteen neighbourhood variables in thirteen separate consecutive logistic regressions. To estimate regression coefficients of one neighbourhood variable (industry at 100 m), we fit a logistic regression model that incorporated all explanatory variables of the basic set but only one neighbourhood variable from the neighbourhood rule of that land use category (industry). In total, 299 regressions were performed to estimate all thirteen neighbourhood variables for all 23 neighbourhood categories.

Validation of this methodology consists of comparing the estimated regression coefficients with the values fitted for the basic model. Figure 10.1 shows that the estimated values are practically the same as those fitted for the neighbourhood variables included in the basic model.

These estimated regression coefficients were used to combine the influence of the neighbourhood at all thirteen distances per land use category into one variable. Therefore, aggregated neighbourhood variables were calculated according to

$$x_l^{Ag} = \sum_{d=1}^{D} \left(\beta_{l,d} x_{l,d} \right) \qquad \text{(eq. 10.4)}$$

where

- $x^{Ag}{}_l$ is the aggregated neighbourhood effect of land use l,
- $x_{l,d}$ is the neighbourhood variable of land use l at distance d,
- $\beta_{l,d}$ is the estimated regression coefficient of the neighbourhood variable of land use l at distance d.

Finally, the neighbourhood model was estimated using a ridge regression with a decay factor of 1, as derived for the basic model. The explanatory variables of this model include the aggregated neighbourhood variables as well as initial land use, policy and suitability.

10.4 Model validation

The model results were compared by determining the Receiver Operator Characteristic (ROC) on an independent data set obtained from the calibration period (Pontius and Schneider, 2001). The ROC for both the basic and neighbourhood model was 0.96.

For all models, developments of residential land use in the Netherlands were simulated over the calibration and validation period by allocating the total observed developments to the locations with the highest probability on the land use maps of 1993 and 1996.

The agreement between the reference map and the simulated map for all these simulations was determined according to Pontius et al. (2004a), in which the percentage correct is estimated between the reference map and simulated map at multiple resolutions. This resolution doubles in each comparison: 1 × 1, 2 × 2, 4 × 4 up to 1024 × 1024 pixels. An estimation of model performance is obtained by comparing this percentage correct with the initial land use map and the results of a random land use allocation model (Hagen-Zanker and Lajoie, 2008).

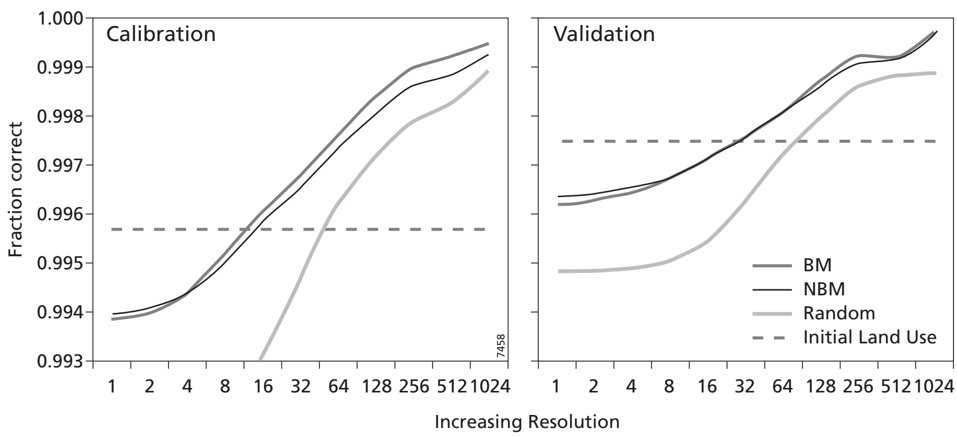


Figure 10.2 Agreement between the simulated land use map and the reference map for the basic model (BM), neighbourhood model (NBM), a random land use allocation model (Random) and the maximum fraction correct for the initial land use map for the calibration and validation period with increasing resolution.

The differences between the basic model and the neighbourhood model are small compared to those with the random land use allocation model (Figure 10.2). Over the calibration period, the agreement for the basic model is slightly better than that for the neighbourhood model, while over the validation period the agreement for the neighbourhood model is slightly better. In the validation period, the estimated models outperform the initial land use map at a resolution of 32 cells while the random allocation model becomes better at a resolution of 128 cells.

It can be argued that the model performs well because many of the changes that occur are "announced" by the inclusion of building sites and designated growth centres in the set of explanatory variables. However, the majority of residential developments do not occur on building sites (73%) or in designated growth centres (90%). Moreover, according to the observed developments only 20% of buildings sites and 1% of the designated growth centres should develop into residential area while in the simulation 42% of buildings sites and 2% of designated growth centres develop into residential areas, irrespective of whether or not these locations overlap with the observed locations. The influence of the assumed "announced" urbanisation is therefore relatively unimportant.

10.5 Model results

10.5.1 Policy and suitability

The effect of the various policy and suitability maps is given in Table 10.6. Most of these maps define restrictions on the development of residential areas and have negative coefficients, with only designated growth centres, regional gas pipes and the ground water protection areas showing a positive effect. However, the significance of the maps with regional gas pipes and ground water protection areas is relatively low because the restricted zone beside the gas pipes is – relatively – too small to influence the developments at a resolution of 100 m, while the ground water protection areas were officially implemented too late (1995) to have affected new residential developments. The designated growth centres have a positive effect on residential developments. The influence of the Green Heart region on residential developments is not significant, most likely because the delineation of this region has been adjusted over the years to accommodate the urban growth of the Randstad conurbation. The gross ecological main structure, nature protection areas, habitat and bird directive areas also partly overlap. Of these policy maps, the Bird directive areas seem to have the highest (negative) impact on residential developments.

Power lines and national gas pipes have a negative effect in their neighbourhood because of associated safety regulations. The influence of the buffer zones is highly significant, as the land in these zones has been acquired by the national government to maintain an open "green" space between the major cities in the Randstad conurbation. The 20-Ke zone around Amsterdam Airport Schiphol has a relatively strong effect due to noise disturbance and safety regulations.

The suitability map with urban areas in 1993 has been added because the development of sites in urban areas is, in general, more expensive than that at locations outside, along the fringe, of urban areas. The subsidence in peat areas has a significant negative influence on residential developments.

Table 10.6 Estimated regression coefficients (β) and significance (Pr(>|z|)) of the policy and suitability maps in the neighbourhood model.

Policy map	**β**	**Pr(>|z|)**	**Policy map**	**β**	**Pr(>|z|)**
Designated growth centres	0.439	< 0.000001	20- Ke zone 1993	-0.490	< 0.000001
Regional gas pipes	0.067	0.363620	Nature protection areas	-0.493	0.012897
Groundwater protection areas	0.018	0.721028	National gas pipes	-0.788	0.000027
Green Heart region	-0.009	0.835099	Drinking water areas	-0.793	0.000070
Gross ecological structure	-0.120	0.000060	Bird directive areas	-1.386	< 0.000001
Power lines > 111 kV	-0.150	0.064184	**Suitability Map**	**β**	**Pr(>|z|)**
Habitat directive areas	-0.370	0.001882	Urban areas 1993	-0.849	< 0.000001
Buffer zones	-0.405	< 0.000001	Subsidence in peat areas	0.043	< 0.000001

Table 10.7 Estimated regression coefficients (β) and significance (Pr(> |z|)) of initial land uses in the neighbourhood model.

Initial land use	**β**	**Pr(>|z|)**	**Initial land use**	**β**	**Pr(>|z|)**
Building sites	4.893	< 0.000001	Recreation	0.718	0.026940
Roads	1.733	< 0.000001	Forest	0.482	0.120153
Other land use	1.335	0.000045	Greenhouse farming	0.191	0.566318
Agriculture	1.327	0.000013	Airports	-0.005	< 0.000001
Freshwater	1.213	0.000083	Sporting fields	-0.321	0.302134
Nature	1.138	0.000773	Commercial & offices	-0.404	0.192974
Industrial	1.116	0.000368	Parks	-0.516	0.097960
Allotments	0.990	0.001989	Cemeteries	-1.348	0.000050
Railroads	0.745	0.017869	Residential area	-8.182	< 0.000001

10.5.2 Initial land use

The development of residential areas is strongly linked to the initial land use of the location (Table 10.7). Building sites are clearly the preferred locations, while residential areas are practically excluded from new residential developments. Agricultural areas are also more likely to be developed than recreational areas, while developed locations, such as commerce and offices and/or parks, are less likely to change into residential areas. Redevelopment of cemeteries is highly unlikely.

The coefficient of airports is lower than might be expected because the neighbourhood effects are very strong (see following section). The significance of land uses with a relative small regression coefficient – such as greenhouse farming, sporting fields, parks, forest and commerce and offices – is low.

10.5.3 Influence of the composition of the neighbourhood

Table 10.8 shows the results of the aggregated neighbourhood variables in the model. The value of the regression coefficient is difficult to interpret as it reflects the influence of the weighted sum of the thirteen neighbourhood variables. More important is the high significance of most of these aggregated neighbourhood variables.

Table 10.8 Estimated coefficients (β^{Ag}) and significance (Pr(>|z|)) of the aggregated neighbourhood variables.

Aggregated neighbourhood	β^{Ag}	**Pr(>|z|)**	**Aggregated neighbourhood**	β^{Ag}	**Pr(>|z|)**
Intercity station	0.767	< 0.000001	Nature	0.069	0.002488
Railway stations	0.722	< 0.000001	Other land uses	0.256	< 0.000001
Slip roads	0.894	< 0.000001	Parks	0.205	< 0.000001
Roads	0.280	< 0.000001	Recreation	0.419	< 0.000001
Agriculture	0.032	0.000190	Railroads	0.197	< 0.000001
Industrial areas	0.141	< 0.000001	Sporting fields	0.277	< 0.000001
Cemeteries	0.391	< 0.000001	Airports	0.356	< 0.000001
Forest	0.066	< 0.000001	Allotments	0.297	< 0.000001
Building sites	0.470	< 0.000001	Residential area	0.245	< 0.000001
Foreign country	0.072	< 0.000001	Fresh water	0.039	0.000104
Greenhouse farming	1.877	< 0.000001	Sea water	0.037	0.004759
Commerce & Offices	0.343	< 0.000001			

The aggregated neighbourhood variables have been "disaggregated" to facilitate our understanding of the influence of the neighbourhood on the development of residential area. Therefore, the thirteen separately estimated regression coefficients of the neighbourhood variables for each land use were multiplied with the regression coefficient of the aggregated neighbourhood variable of that land use. However, it remains relatively more difficult to interpret the influence of the neighbourhood than policy or initial land use because the former is dependent on the value of the neighbourhood variable. It is therefore more straightforward to look at the combined effect of the regression coefficient and the neighbourhood variable in the regression model. This contribution to the logit model or partial logit is defined as

$$pLogit_{l,d} = \beta_{l,d}\beta_l^{Ag}x_{l,d} \qquad \text{(eq. 10.5)}$$

where

- $pLogit_{l,d}$ is the partial logit of land use l at distance d
- $\beta_{l,d}$ is the estimated regression coefficient of land use l at distance d
- β^{Ag}_l is the estimated regression coefficient of the aggregated neighbourhood variable of land use l and
- $x_{l,d}$ is the explanatory variable of land use l at distance d.

Figure 10.3 shows the partial logit of residential areas, industrial areas, building sites as well as the total contribution of the neighbourhood on the development of residential areas in the region of Amsterdam. For residential areas, locations with the highest partial logit coincide with current residential areas. Larger residential clusters have higher values than small clusters and values decrease at the edge of these residential clusters. Industrial areas "behave" in the opposite direction in that locations with the lowest partial logit coincide with current industrial areas. Building sites display a different behaviour: small clusters have a high partial logit, while large clusters have a low partial logit. In general, the small clusters are building sites for residential areas while the larger clusters are industrial building sites or new harbour areas. The lowest values for the total partial logit of all neighbourhood variables are in the neighbourhood of the

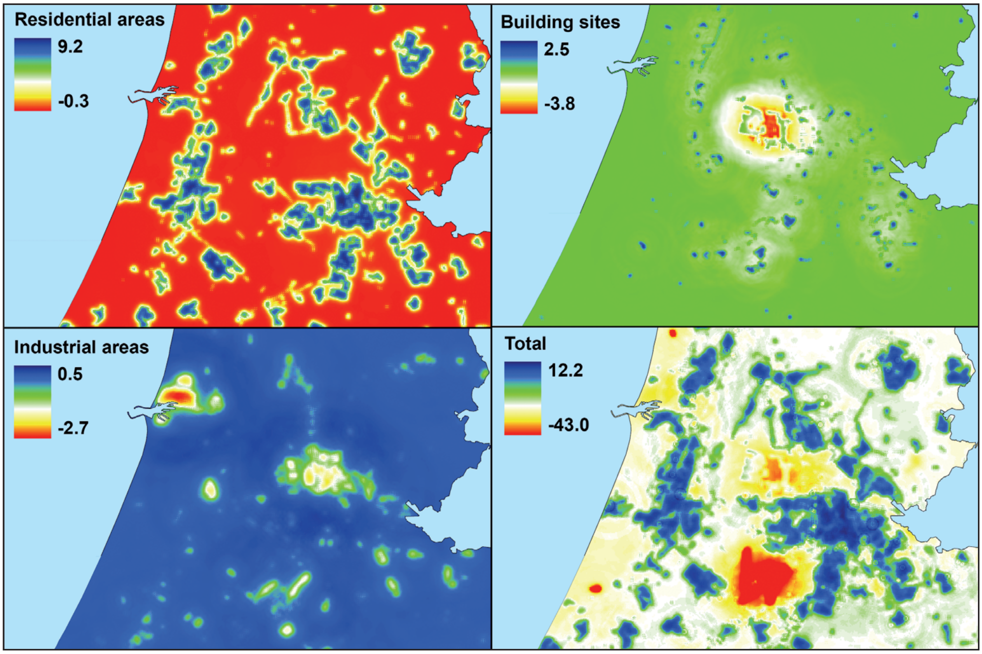


Figure 10.3 Contribution of the neighbourhood of residential areas, industrial areas, building sites and total of all neighbourhood categories in the logit model for the region of Amsterdam, the Netherlands.

airports. The negative influence of the industrial areas and large building sites is less than that of airports, but still noticeable in the total influence of the neighbourhood.

The influence of all neighbourhood variables, as defined by the partial logit, has been plotted over distance for all land use types and transport nodes in figures 10.4 and 10.5. Because locations near a large industrial area may have a different effect on residential developments than those near a small industrial area, each figure displays three lines, reflecting the effect of slightly, median and highly enriched neighbourhoods, respectively. This relative enrichment of the neighbourhood has been defined by estimating the 0.05, 0.5 and 0.95 percentile of the neighbourhood variable $x_{l,d}$ (eq. 5). Locations without land use l in the neighbourhood, for which $x_{l,d} = 0$, have no influence on the development of residential area and have been excluded from the data set prior to estimating the percentiles.

The uncertainty in the partial logit is shown by the error bars for the highly enriched neighbourhoods ($\alpha = 0.95$) in figures 10.4 and 10.5. This uncertainty is based on both the standard errors of both the regression coefficient of the aggregated neighbourhood variables and the regression coefficient of the neighbourhood variables at the specific radii. Assuming that correlations can be neglected, the overall uncertainty can be estimated according to

$$\varepsilon_{total} = \sqrt{\beta_1^2 \varepsilon_2^{\,2} + \beta_2^{\,2} \varepsilon_1^{\,2}} \qquad \text{(eq. 10.6)}$$

where:

- ε_{total} is the total standard error,
- β_1, β_2 are the estimated regression coefficient of the neighbourhood variable at a specific distance and the aggregated neighbourhood variable, respectively,
- ε_1, ε_2 are the standard errors in the estimated regression coefficients.

Figures 10.4 and 10.5 show the influence of urban land use categories on residential developments and rural land uses, the transport nodes and links, respectively. Rules displaying similar trends have been grouped together. Only those neighbourhood rules displaying specific

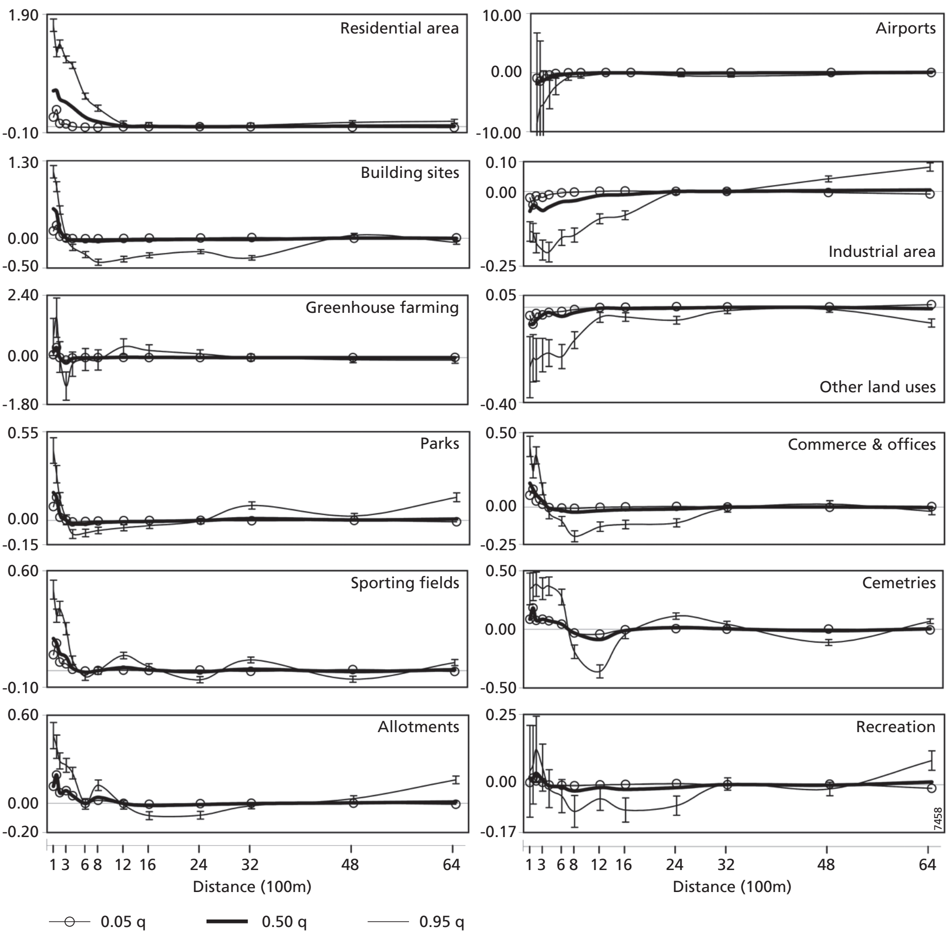


Figure 10.4 Neighbourhood rules of the neighbourhood model, expressed as the partial logit (eq. 5), showing the influence of various land uses on residential developments with increasing distance. Each graph shows the partial logit for three neighbourhoods with a slight, median and high enrichment of the specific land use defined by 0.05, 0.50 and 0.95 percentile of the neighbourhood variable, respectively. Error bars show ± 2 standard errors of the regression coefficients.

and clear effects are discussed in detail here. In most cases, the logic of the rules can be related to basic geographical knowledge: residential areas are attracted by residential areas and repelled by most other land uses, with the effects becoming weaker with distance. The effects of most rules increase with the relative presence of the land use in the neighbourhood. The significance of the neighbourhood variables may be low at specific distances, but all aggregated neighbourhood variables are significant at $p < 0.005$.

The strong positive influence of nearby residential area and building sites in the neighbourhood on new residential developments is shown in Figure 10.4. In terms of building sites, however, the influence of highly enriched neighbourhoods reverses and becomes negative

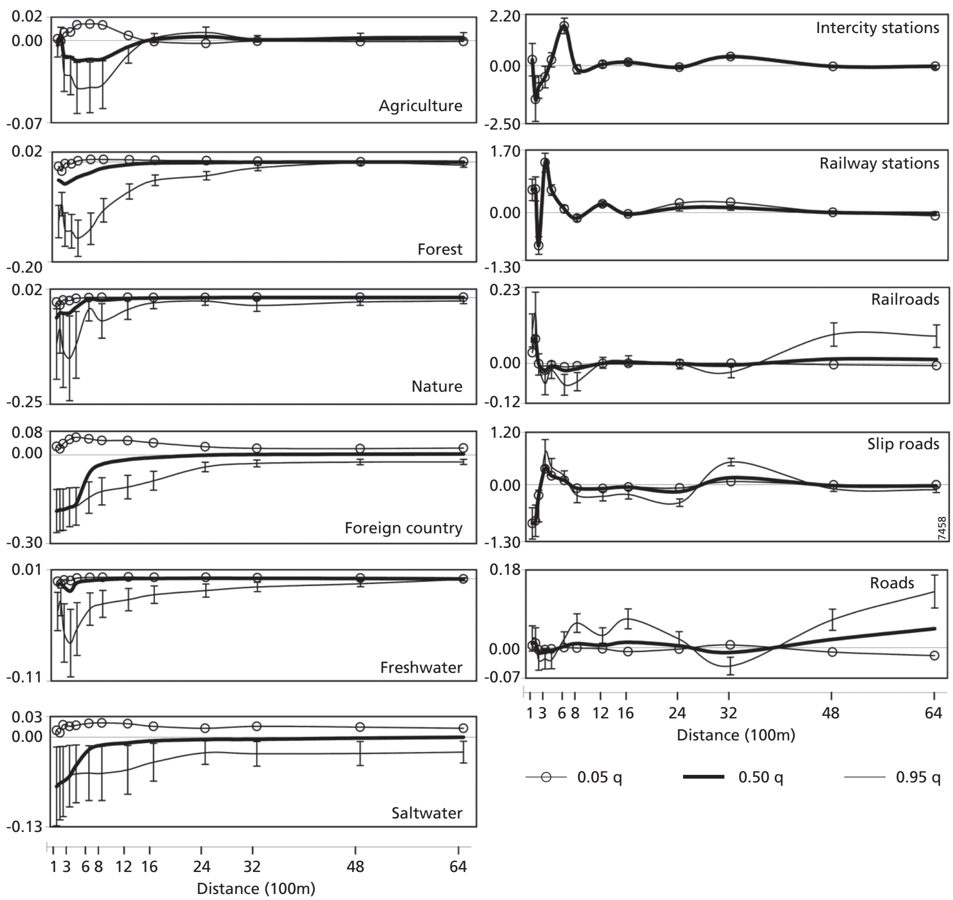


Figure 10.5 Neighbourhood rules of the neighbourhood model, expressed as the partial primary influence of most of these urban land use categories extends for logit (eq.5), showing the influence of various land uses on residential developments with increasing distance. Each graph shows the partial logit for three neighbourhoods with a low, median and high enrichment of the specific land use defined by 0.05, 0.50 and 0.95 percentile of the neighbourhood variable, respectively. Error bars show ± 2 standard errors of the regression coefficients.

at about 400 m. Most of these sites in the land use map of 1993 are designated new harbour or industrial building sites. The differences between these sites are reflected in the neighbourhood rules. Airports, industrial areas and "other land uses" have a negative influence on residential developments. The class "other land use" includes dumpsites, automotive wrecking/salvage sites and mining sites, the presence of which is not likely to attract the development of a residential area. It is notable that the impact of large industrial areas increases at large distances; for example, the large industrial areas near the Rotterdam and Amsterdam harbours have a positive influence at large distances on the development of residential areas. These locations are more attractive because of the employment they offer at acceptable commuting times.

In general, the rules of land uses for parks, sporting fields and cemeteries tend to fluctuate more over distance than those for the major land use categories agriculture, forest and water. The uncertainty in the estimated regression coefficients is relatively higher for the former, especially at larger distances. The primary influence of most of these urban land use categories extends for a distance of 400-800 m.

Figure 10.5 shows that the influence of the rural land uses on residential developments is relatively small. However, on average, the range of influence of the rural land uses, including water and foreign countries, is larger than that for the urban and transport-related land use categories. The influence of agriculture depends strongly on the relative enrichment of the neighbourhood. At small distances, all partial logits are practically zero; by 200 m, however

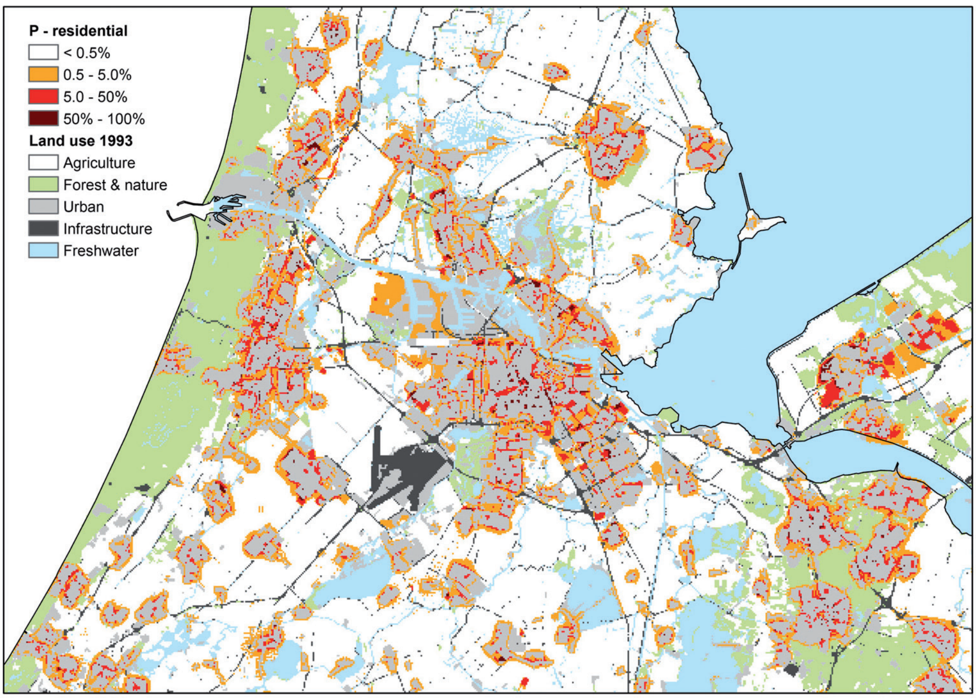


Figure 10.6 Probability of residential developments according the neighbourhood model near Amsterdam, the Netherlands.

they start to diverge. Slightly enriched neighbourhoods have a positive influence, while highly enriched neighbourhoods have a negative influence. The rules diverge up to about 600 m, but the trends then converge with increasing distances. Small parcels of forest and nature are comparable to agriculture in being relatively more attractive for residential developments. The land use category foreign countries shows a negative influence on residential developments, and, in general, developments within 500 m of the national border are not allowed (VNG, 1992). For seawater, the development of the median neighbourhood rule is comparable to that for foreign countries: in most places along the shore, the land is protected from the sea by dunes or dikes and new developments are subject to stringent restrictions. In terms of freshwater, only large water bodies have a negative influence on new residential developments. The primary influence of these rural land use categories extends a distance of approximately 800 m up to 1600 m. For highly enriched neighbourhoods, however, the effect may extend yet further.

Railway stations and slip roads have a strong impact on residential development in their neighbourhood. These are represented by a single cell on the map and there is therefore no difference in the relative enrichment of the neighbourhood at small distances; at larger distances, however, more locations may be included in the same neighbourhood. The influence of both intercity and ordinary railway stations fluctuates strongly at close distance. Slip roads have a relatively greater negative influence than railway stations on residential developments at close distance, with this trend reversing with increasing distance, showing a slight maximum at

Table 10.9 Size of small ($A_{0.05}$), median ($A_{0.5}$) and large ($A_{0.95}$) clusters and median partial logit ($pLogit_{0.5}$) of these, defined as 0.05, 0.5 and 0.95 percentile of cluster size distribution respectively.

Land use	$A_{0.05}$	$A_{0.5}$	$A_{0.95}$	$pLogit_{0.5}$ $A_{0.05}$	$pLogit_{0.5}$ $A_{0.5}$	$pLogit_{0.5}$ $A_{0.95}$
Residential area	12	216	1700	2.21	5.89	7.52
Cemeteries	1	5	38	0.09	0.74	3.77
Allotments	1	4	37	0.13	0.76	1.86
Sporting fields	2	12	70	0.22	1.15	1.89
Greenhouse farming	1	135	5300	0.11	0.85	1.64
Parks	2	13.5	86	0.22	0.95	1.21
Commerce & offices	2	17	153	0.11	0.63	1.03
Railroads	1	4	57	0.03	0.12	0.42
Recreation	2	22	103	0.01	0.19	0.29
Roads	1	8	448	0.09	0.09	0.18
Seawater	2500000	2500000	2500000	-0.53	-0.53	-0.53
Foreign countries	2500000	2500000	2500000	-1.40	-1.40	-1.40
Agriculture	4100	240000	1100000	-0.15	-0.14	-0.16
Freshwater	3	62000	210000	-0.01	-0.33	-0.51
Nature	6	997	6200	-0.11	-0.37	-0.74
Building sites	2	31	863	0.10	1.17	-0.65
Forest	5	667	65000	-0.13	-0.83	-1.09
Industrial	4	58	464	-0.18	-0.88	-1.78
Other land use	1	11	242	-0.10	-0.66	-3.36
Airports	9.4	133	1100	-6.10	-13.58	-34.66

about 300 m. The effect of these transport nodes and land use categories extends mostly up to a distance of about 800 m.

10.5.4 Influence of the cluster size

The median partial logit of each land use for locations in small, median and large clusters has been determined to display the effect of cluster size (Table 10.9). Clusters have been defined as cells with the same land use that have been grouped together if they are within the immediate eight-cell neighbourhood of each other. Small, median and large clusters of each land use have been determined as the 0.05, 0.5 and 0.95 percentile, respectively, of the cluster size distribution. Most clusters consist of a large number of 100 × 100-m grid cells and, depending on the land use, a large number of these clusters may exist on the map. Therefore, the median partial logit has been determined over all locations falling in the categories of small, median or large clusters of each land use.

Most urban land uses become more attractive for residential developments as cluster size increases, while the rural land uses become less attractive with increasing cluster size. Building sites display an optimal cluster size, with median-scale sites being the most attractive because the

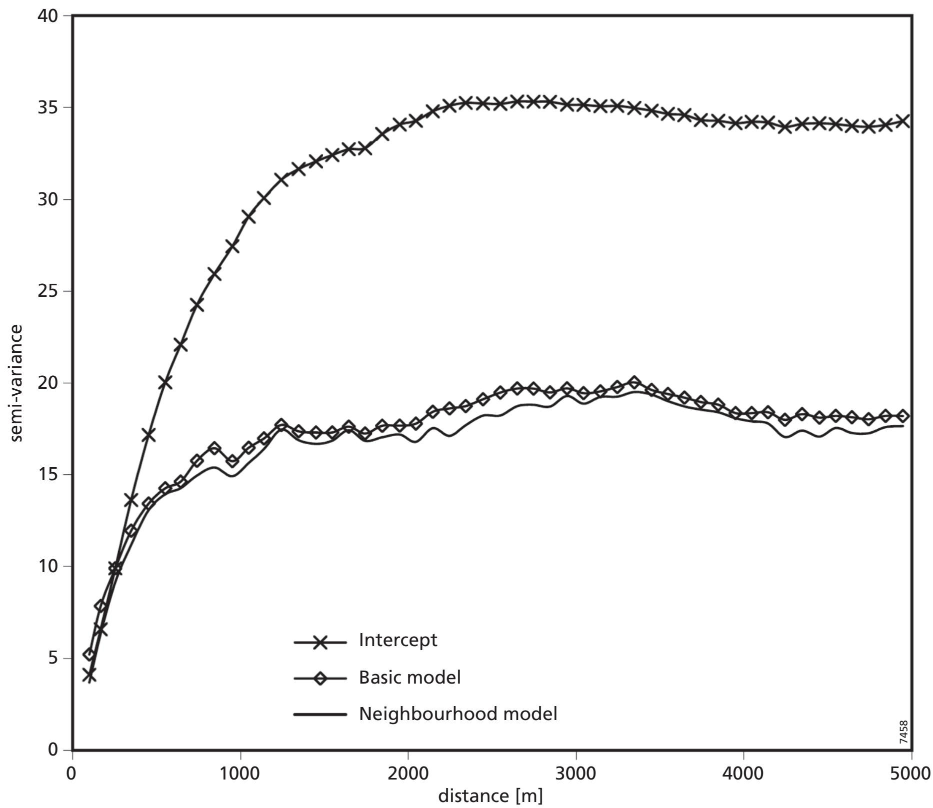


Figure 10.7 Variogram of Pearson residuals of a null model with only an intercept, the basic model and the neighbourhood model.

larger clusters are designated industrial or harbour building sites. Airports and "other land uses" become highly unattractive for residential developments as cluster size increases. Small airports may be replaced, but it is highly unlikely that Amsterdam Airport Schiphol will be relocated. Seawater and foreign countries comprise only one cluster in the map and, therefore, display the same value for all cluster sizes.

10.5.5 Probability map

Figure 10.6 shows the probability of residential developments near Amsterdam according to the neighbourhood model. Building sites and undeveloped locations at the urban fringe have the highest probabilities due to the positive influence of residential area and other urban land uses.

10.5.6 Spatial autocorrelation

Logistic regression assumes that the values of observations in each sample are independent of one another; however, due to spatial autocorrelation, the error terms in logistic regression models are not fully independent (Anselin, 2002). The inclusion of neighbourhood variables decreases the spatial autocorrelation in the residuals. To compare the spatial autocorrelation in the residuals, variograms of the Pearson residuals (Bio et al., 2002) have been made of both regression models and a null-model with only an intercept (Figure 10.7). The null model shows spatial autocorrelation in the residuals up to a distance of 2500 m, while the spatial autocorrelation in the residuals in the neighbourhood model starts to decrease at a distance of about 250 m, more or less disappearing as the curve flattens at a distance of about 750 m.

As apparent from Figure 10.7, residential developments are, according to the neighbourhood model, most likely to be located near existing areas. At large distances – beyond 500 m – the model correctly predicts that residential developments are highly improbable, all residuals will be zero and a large part of the spatial autocorrelation is included in the model; in contrast, within a distance of 300 m, both the probability of residential developments and the number of incorrect predictions are high and the spatial autocorrelation increases. It is difficult for the model to pinpoint the exact location, A or B (Figure 10.8).

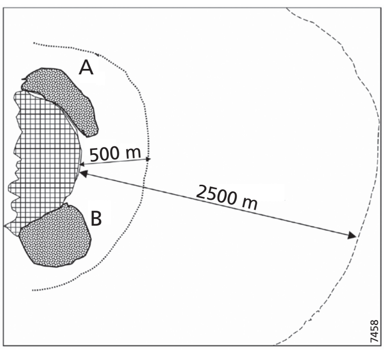


Figure 10.8 Spatial autocorrelation decreases with increasing distance from existing urban locations. The model correctly predicts the absence of residential developments at larger distances. At closer distances, the model predicts mostly new residential developments but often at the wrong location, i.e. A instead of B.

10.6 Discussion and conclusions

We have developed a logistic regression model that incorporates both initial land use, land use in the neighbourhood at various distances and policy- and suitability-related variables to explain residential developments in the Netherlands.

10.6.1 Methodological aspects

We present a novel methodology for estimating the influence of land use in the neighbourhood in which the influence of each land use category in the neighbourhood is described by a series of neighbourhood variables at a specified range of distances. Most of these neighbourhood variables are highly correlated, resulting in highly variable regression coefficients. Therefore, a basic model was defined that included only those neighbourhood variables with correlation coefficients smaller than 0.5. A "Single Parameter Impact Estimation" (SPIE) was then used to estimate the regression coefficient of each neighbourhood variable, $x_{l,d}$. Using this approach, all 299 neighbourhood variables were estimated in 299 logistic regressions. The influence of all thirteen neighbourhood variables in each rule was then combined in one aggregated neighbourhood variable, x_l^{Ag}. In the final neighbourhood model, these aggregated neighbourhood variables, together with the initial land use, policy and suitability, were used to explain residential developments.

The coefficients estimated with SPIE compare well with the regression coefficients from the basic model. However, as suggested by the reviewers, it would be very useful to see how well the method manages to recover the parameter values of a known model given the output of one model simulation.

At first glance, other models, such as autoregressive models, may appear to be more appropriate (Dendoncker et al., 2007); however, the inclusion of these large neighbourhoods in autoregressive models is prohibited due to limitations in computer memory.

Ultimately, it is the performance of the model in the validation that justifies the application of logistic regression in this study. The ROC is 0.96 for both the basic and the neighbourhood model. Although these results are difficult to compare with those from other studies, our values are relatively high. Verburg et al. (2004c) report a ROC of 0.90 for a binomial logistic regression model of residential developments in the Netherlands with neighbourhood interaction, while Dendoncker et al. (2007) reports a value of 0.88 for a pure autoregressive model for residential areas in Belgium.

Our neighbourhood model outperforms a random land use allocation model at all resolutions and has a null model resolution of 32 cell sizes, or 3.2 km – a high resolution compared to the null model resolution of the models reported by Pontius et al. (2008).

10.6.2 Determinants of residential growth

The resulting neighbourhood model shows that various associations are important to the development of residential areas in the Netherlands: initial land use, composition of the neighbourhood, policy and suitability. In general, the results correspond with our general understanding of residential developments in the Netherlands.

Most policy variables in the model have a negative influence on residential development, with the exception of designated growth centres, which are highly significant. The suitability of the location is of minor importance in the Netherlands as only locations affected by subsidence in peat areas are less prone to residential development. These results are in accordance with those of previous studies by Verburg et al. (2004c). The inclusion of initial land use in the model is of major importance, as there are great differences between developing residential areas on building sites, allotments, nature areas or land occupied by airports or forest. The influence of land use in the neighbourhood changes with cluster size, with small- to median-sized building sites being highly attractive for residential development and large building sites less attractive. This effect is incorporated into the neighbourhood variable

Most of the neighbourhood rules, expressed as partial logits, are quite logical. The effect of most land uses decays with distance. In general, the effect of rural land uses, forest, nature and water has a wider range than that of the urban land use categories. The main attractors of residential developments are residential areas and building sites, while new developments are mostly discouraged by the presence of airports and industrial areas. Railway stations and slip roads have a negative influence at short distances, but this trend reverses and becomes positive at larger distances. The aggregated neighbourhood variables are highly significant for most land use categories

It is difficult to compare the results of this study with data reported elsewhere. Many real estate studies that have used hedonic pricing report neighbourhood effects for green areas, water, roads or dump sites on residential values. However, the sites had already been developed in these cases, while our question was whether it would be developed at all. The relevant explanatory variables differ among all of these studies, especially in terms of policy-related variables. Most studies report comparable neighbourhood effects on residential developments (Barredo et al., 2003, Verburg et al., 2004d, Hagoort et al., 2008), although a comparison of the exact figures is difficult given differences in land use classifications, cell size and quantification of the rules..

The results of the regression model are in line with the development costs of residential areas. Locations in urban areas are generally more expensive than elsewhere; demolition costs should be included. Undeveloped agricultural lands near existing urban areas are favoured because the development of infrastructure, roads, gas, water, electricity, sewage systems and internet connections becomes cheaper.

The methodology described here, including the single-parameter impact estimation, can be applied to calibrate the rules of CA-based land use models. The procedure is straightforward and relatively fast compared to other calibration techniques. The estimated coefficients and standard errors facilitate interpretation of the results.

Calibration through logistic regression does not affect the emergent properties of the model. Emergence develops because the locations of the developments are highly uncertain (de Nijs and Pebesma, 2007). As the calibration explains only a part of the developments, the residual error should be used to define the randomness in the dynamic allocation algorithm of the simulation model. Once residential land use has developed, the probability of new residential developments in the neighbourhood of that location increases.

More knowledge is required if we are to gain an understanding of the spatial processes that occur and eventually lead up to land use change. This will necessitate studies that report on the influence of neighbourhood, suitability, policy and accessibility on the development of residential or industrial areas in other locations and for other periods in time. Future research should test the sensitivity of the model and neighbourhood rules in terms of scale, number of land use categories and temporal resolution.

11 Synthesis

11.1 Introduction

In the General Introduction to this thesis, I have described the societal and scientific relevance of studying land use change and have defined research priorities according to the four steps involved in the development of a simulation model according to Refsgaard and Henriksen (2004). I have also defined the overall objective of the thesis within the context of these research priorities. As such, this thesis combines a selection of papers that all have one general theme to improve the quality of land use simulation models and the prediction of future land use patterns. The papers address a wide range of research questions related to the improvement of land use models. Each particular question has resulted in a peer reviewed paper as presented in the previous chapters of this thesis.

Here, I will integrate and discuss the achievements according the consecutive stages in model development: Conceptual model, Model and Application (Figure 11.1). The focus of this thesis was not on programming and code verification, but on the improvement of the land use models and the predictions of future developments of land use. The framework does not define

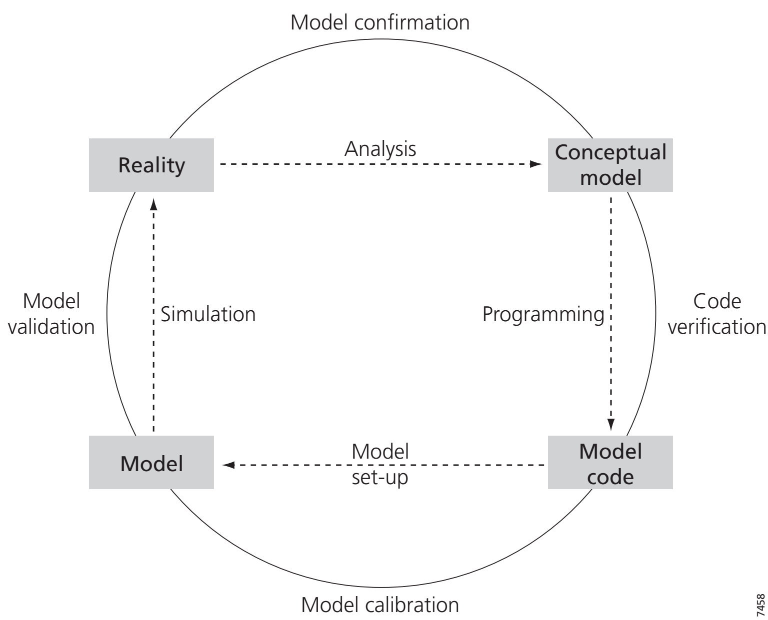


Figure 11.1 Definition and terminology of the four steps in the development of a simulation model and the necessary quality assurance as proposed by Refsgaard and Henriksen (2004).

the final product of the simulation nor the object of the validation. As it certainly is not Reality it is suggested here that the term Application be used

In the Introduction, I have classified the research questions according to these aforementioned stages in model development. However, the resulting papers are not restricted to only one step in the development of the model. For example, in the analysis of the neighbourhood interactions, a statistical model is fitted to historical land use developments in order to determine the influence of the various neighbourhood interactions. Fitting statistical models is a form of calibration.

Moreover, the papers on the application of the Environment Explorer specifically address all three steps in the model development. Therefore, in this section, I will describe the results obtained in the consecutive stages of model development. As such, I will not limit myself to answering the research questions specified in the Introduction but will integrate all of the relevant elements from all of the papers in this thesis. Where possible, I will also suggest recommendations for further research.

11.2 Improving the Conceptual Model

This stage focuses on improving the definition of the conceptual model through analysis of changes in land use maps.

In chapter 4, 'A method to analyse neighbourhood characteristics of land use change patterns', the 'enrichment factor' is introduced to analyse the neighbourhood interactions from observed changes in land use maps. This factor reveals those locations more prone to change in land use. The method can be used to analyse (1) the regional and temporal variation in neighbourhood interactions; (2) the influence of spatial and temporal resolution in land use maps; (3) the effects of aggregating land use categories; (4) the shape, extent and resolution of the neighbourhood.

Most important, however, the 'enrichment factor' can be used to select the relevant neighbourhood characteristics in models of land use change. Models addressing neighbourhood interactions by cellular automata can use the enrichment factors that are identified from analysis of historical land use changes as a first estimation of the transition rules. The influence of the transitions, however, still needs to be calibrated. Statistical models of land use change can include the enrichment factor as an explanatory variable. The paper comprising chapter 4 shows that the predictability of locations of land use change based on the neighbourhood characteristics alone is relatively high. On the other hand, neighbourhood characteristics are not the only factor defining the locations of land use change; other factors, such as accessibility, suitability and policy, are also important.

Chapter 5, 'Determinants of land use change patterns in the Netherlands' presents an empirical study, the results of which reveal the main determinants of land use change for a wide range of land use types. The main objective of this study is to increase our understanding of the land use change processes and to present a basis for the specification of a land use change model. The paper describes the impact of a wide range of determinants of land use change, all of which are derived from various land use allocation theories.

Spatial analysis of the land use map of 1989 reveal that the location of agricultural and natural areas can, to a large degree, be explained by the suitability of the soil for agricultural use. Residential land use patterns in 1989 are not well explained by the location characteristics. However, residential as well as industrial land use developments over the period 1989 to 1996 can be explained by neighbourhood interactions, accessibility and policy. The results of these empirical spatial analyses should be enhanced with in-depth studies to reveal causalities and define the process leading to land use change. The methodology can be used to improve the (conceptual) models of land use change. However, because of collinearity, the number of neighbourhood interactions in these analyses has been limited to eight.

Chapter 10, 'Estimating the influence of the neighbourhood in the development of residential areas in the Netherlands', focuses on the neighbourhood interactions involved in the development of residential areas. In this study, the composition of the neighbourhood is defined at 13 distances for all 20 land uses and three transport nodes, resulting in 299 neighbourhood variables. Also included in the logit model were policy and suitability variables as well as initial land use.

The results show that the development of residential areas in the Netherlands is primarily determined by initial land use, neighbourhood interactions and policy. Initial land use and the neighbourhood are highly significant factors in any explanation of residential developments in the Netherlands. The influence of urban land uses and transport nodes extends for a distance of 400 up to 800 m, while that of rural land uses extends further, up to 1600 m. The influence and significance of spatial policies varies. In terms of suitability, only locations impacted by strong subsidence due to the oxidation of peat are less prone for residential development due to the location-related construction problems. These results are in line with the development costs of the residential areas.

The estimated regression coefficients for initial land use and aggregated neighbourhood variables are strongly related. The influence of the neighbourhood depends, for various land use categories, on the cluster size of the land use, with the result that locations near larger residential areas have a higher probability to change to residential area than locations near small residential areas.

Chapter 7, 'Measuring performance of land use models. An evaluation framework for the calibration and validation of integrated land use models featuring cellular automata', describes a semi-automated procedure to calibrate the Environment Explorer. In an attempt to shorten the calibration time, the 30 parameters defining the neighbourhood rules in the previous version of the model were reduced to three, thereby changing the conceptual model. A linear relation is assumed between these three points, thereby highly simplifying the shape of the neighbourhood rule; this is definitely not in accordance with the shapes found in chapter 10 for the neighbourhood of residential developments.

Future research should focus on empirical spatial analysis of land use developments in order to gain a better understanding of the spatial processes which, ultimately, lead up to land use change. To improve the conceptual model comparable studies are necessary for other land uses and from other geographic regions, also influenced by different spatial processes. The sensitivity of the conceptual model and its neighbourhood rules should be tested for spatial and temporal resolution and extent, aggregation of land use categories and the shape and size of the neighbourhood (Dietzel and Clarke, 2004, Verburg and Veldkamp, 2004, Beneson, 2007). These results indicate that spatial resolution is especially important as most neighbourhood rules display the largest effect at close distances, as shown in figures 10.4 and 10.5 and by various authors (Tobler, 1979, Menard and Marceau, 2005, Hagoort et al., 2008). The temporal resolution is important in order to reveal the spatial behaviour at a certain 'moment' in time. On the one hand, this time span should not be too large as the rules will reveal a mixture of spatial behaviour during the period included; on the other hand, the time span should not be too short in order to allow the effect of the spatial behaviour of the land use categories to be revealed as land use changes in the map (Fang et al., 2006).

11.3 Improving the Model

New calibration tools and performance measures have been developed with the aim of improving the model.

Chapter 6, 'The Map Comparison Kit', gives an overview of the map comparison techniques that have been combined in one software package. These map comparison techniques can be applied as objective performance measures to calibrate and validate land use models. Simulated and observed land use maps can be compared by category, cell-by-cell and according to two techniques based on Fuzzy Sets: the Fuzzy Inference System and Fuzzy Set algorithms. A tool is now available in the free domain to quantitatively compare nominal and ordinal maps.

Chapter 7, 'Measuring performance of land use models. An Evaluation framework for the calibration and validation of integrated land use models featuring cellular automata', describes a semi-automated procedure to calibrate the Environment Explorer. The Environment Explorer consists of a spatial interaction (gravity) model at the regional level and cellular automata based land use model at the local level. The regional model defines the amount of land use change in the land use model; in turn, the land use model feeds back regional aggregates of the cellular model. During the calibration procedure, these submodels are separated and calibrated according to their own goodness-of-fit measures. After being individually calibrated, the models are once again put back together to restore the dynamics and the overall goodness-of-fit measures are assessed. This process is repeated until the overall goodness-of-fit measures are optimal. The procedure is complex and time-consuming as the modeller has to check the neighbourhood rules resulting from the calibration routine.

The calibration and validation of land use models are impeded by existing classification and aggregation errors in land use maps. An analysis of the land use maps used in this study showed that 25% of the changes are unlikely, such as changes from urban area to natural or agricultural area.

In chapter 10, 'Estimating the influence of the neighbourhood in the development of residential areas in the Netherlands', a logistic regression model is fitted for residential developments comparable to that used in the cellular automata model in the Environment Explorer, including the influence of the initial land use and the neighbourhood. A new methodology, Single Parameter Impact Estimation (SPIE), has been applied to circumvent the problems associated with the collinearity of variables. The model is estimated in three steps. First, the influence of each neighbourhood variable is estimated using SPIE. Next, the influence of each land use is aggregated over the various distances in the neighbourhood into one neighbourhood variable. Finally, a logistic regression model is fitted, including all relevant policy and suitability variables, initial land use and the aggregated neighbourhood variables.

Fitting a logistic regression model is, in principle, not different from calibrating a dynamic land use model on observed land use developments. If the influence of the new land use developments in the allocation mechanism is negligible or small, the estimated coefficients of the logistic regression can be used to parameterise a comparable dynamic land use model. In this study, only 0.15% of the observed land use changes to residential area, which is small enough to be able to neglect the influence of these changes in the logistic regression model. The allocation mechanism of the Environment Explorer has been adapted to include the results of a logistic regression. In this way a 'dynamic logistic cellular automata model' has been developed. The procedure used to estimate the parameters is also complex and time-consuming, but it has the potential to be fully automated. Moreover, in contrast to the exponential increase in computation time in the semi-automated calibration procedure, the computation time in this procedure increases linearly with the number of neighbourhood variables. Therefore, compared to the calibration of dynamic cellular automata-based land use models, the need to limit the number of neighbourhood variables in order to restrict computation time is less stringent.

Future studies should test the methodology in a complete land use model, including all relevant land use categories. For calibration of the model, 'consistent' land use maps over a short time span are necessary as well as the development of good performance measures (Hagen-Zanker, 2008, Pontius Jr et al., 2008). The errors in the land use map should be minor compared to the actual changes. Once calibrated, the model should be validated on land use maps over a large time span (>25 years).

Future research should also be directed towards reducing the uncertainties in the results of the land use model by taking advantage of improved monitoring data (Pontius and Spencer, 2005, Fang et al., 2006). For example, as a means to circumvent the classification problems in land use maps, land use models could use the primary remote sensing data and derived spatial patterns instead of the interpreted results in the land use maps in which errors and uncertainty have already accumulated (de Almeida et al., 2005, Chowdhury, 2006, Jat et al., 2008, Yuan, 2008). Moreover, all information available in subsequent remote sensing images (time-series) should be used to identify current land use developments and develop algorithms to simulate future land use developments.

11.4 Improving the Application

The application of the model has been improved by applying the model in policy applications and by validating the model on historical land use developments and according to Zipf's Law. Finally, the uncertainties have been characterised for logistic regression model.

Chapter 3, 'Constructing land use maps of the Netherlands in 2030', describes the application of the Environment Explorer model in the National Nature Outlook 2 (RIVM, 2002). This model was used to create spatially detailed maps of the Netherlands in 2030 with the aim of evaluating the effects of alternative socio-economic developments on nature and the landscape. The application of a land use model ensures that all relevant aspects of a scenario, such as economic and demographic development, zoning policies and urban growth, will be integrated in one consistent framework. Each of these elements can be defined on the basis of historical trends, scenario assumptions or urban designs, thereby linking prescriptive design methods with descriptive simulation tools. In this application, the model was still 'hand calibrated' using the Fuzzy Kappa metric and the model was not yet validated.

In chapter 9, 'Future land use in the Netherlands: evaluation of the National Spatial Strategy', the Environment Explorer was applied to evaluate the effect of the spatial strategy on future land use patterns in the Netherlands. In this case a 'Trend and a Sprawl' scenario were developed to assess the effect of the national strategy on the urban concentration areas, the national landscapes and the restricted areas. In this application the Environment Explorer was calibrated and validated using the semi-automatic calibration procedure and the performance measures discussed in chapter 7. In this case, the distribution of urban cluster sizes in future land use maps was verified using Zipf's Law. Next, a series of Monte Carlo simulations was applied to define the probability of urbanisation in the future land use maps. Comparable probability maps are presented in chapters 5 and 10. Figure 5.3 in chapter 5, 'Determinants of land use change patterns in the Netherlands', shows the importance of including the neighbourhood for residential and industrial/commercial developments. In both cases the variability explained by the model increases considerably with the inclusion of neighbourhood characteristics, resulting in more reliable probability maps. For new residential areas, the Receiver Operator Characteristic (ROC) value increases from 0.81 to 0.90. Figure 10.6 in chapter 10 shows the probability map of residential developments according to the neighbourhood model near Amsterdam. The ROC value of this model is 0.96. All of these maps look similar, with the highest probabilities for residential or urban developments being found near existing urban areas. The results of the Monte Carlo simulations could have been used to quantify the accuracy and the upper and lower confidence levels of the urbanisation probability.

In chapter 8, 'Spatial uncertainty in land use models. An alternative method to estimate uncertainty in logistic regression methods', a novel method to estimate the uncertainty in logistic regression models is proposed. The uncertainty is normally estimated from the standard error and the residual variance. However, in our case, this method results in an upper limit of 1 and a lower limit of 0 for probability of land use change. Therefore, the 2.5 and 97.5 percentiles of residuals, binned over the logit values, were used as the lower and upper prediction limits, respectively. The results show that the model uncertainty varies with location and land use category. The uncertainty is high for all locations, with the exception of Amsterdam Airport Schiphol, surface waters and building sites. While it is very unlikely that airports and surface waters will change to residential areas, building sites are, by definition, likely to change.

Logistic regression models minimise the residual variance. In general, larger models lead to a smaller residual variance, but they do not necessarily perform better in terms of prediction errors (Zhou and Kockelman, 2008). Future research should focus on searching for an optimal model given the uncertainty in both the explanatory variables and residuals.

As suggested in chapter 10, the stochastic perturbation in the Environment Explorer could be defined given the residual error in the logistic regression. Moreover, chapter 8 shows that model uncertainty and, consequently, the residual error depend on the initial land use considered. As such, the stochastic perturbation should depend on initial land use.

11.5 Main findings and developments

The research presented in this thesis has and will improve the quality of land use simulation models and the prediction of future land use patterns by:

- the development of a method to analyse neighbourhood characteristics;
- the development of the Map Comparison Kit, a software toolkit, to compare maps with a main focus on nominal, land use maps;
- the development of a semi-automated calibration tool for dynamic cellular automata-based land use models, such as the Environment Explorer;
- the development of a methodology to define the uncertainty in logistic regression models of land use change;
- the development of a methodology to cope with spatial autocorrelation in logistic regression models, such as Single Parameter Impact Estimation (SPIE) and to analyse the influence of the neighbourhood on land use developments;
- the development of a hybrid land use model integrating the influence of the neighbourhood in a logistic regression model or, posed differently, the calibration of a dynamic cellular automata-based land use model using logistic regression;
- the determination of the main explanatory variables of land use change – initial land use, neighbourhood, policy and suitability;
- the application of Zipf's Law to validate simulated future land use maps;
- and the application of the Environment Explorer in two scenario studies – the Nature Outlook 2 and the evaluation of the National Spatial Strategy in the Netherlands.

Finally, in my opinion, the development of land use models and their application would significantly benefit from model comparison (Pontius Jr et al., 2008), especially if these applications were to be based on the same data. Standard comprehensive datasets should be made freely available to encourage researchers to develop, calibrate and validate land use models for various regions, together with a set of performance measures. Future research should focus on empirical research, the analysis of land use data on various spatial scales and the development of relative simple statistical models as opposed to complex multi-agent-based land use models which are difficult to validate.

Summary

Modelling land use change
Improving the prediction of future land use patterns

Introduction

Man has been altering his living environment since prehistoric times. With the advent of the agricultural revolution, forests were cleared to make space for crops and pastures. The industrial and resulting demographic revolution, in turn, marked the beginning of urban expansion. The world's population has grown to 6.7 billion people and this growth has impacted on our environment in many ways. Agricultural and urban land use expands globally at the cost of forest and nature. Economic and demographic growth will give way to future land use changes and these will be very evident in the Netherlands where each square meter of land is exploited for agriculture, housing, industry, office, commerce, recreation, transport, forest, nature.

It is predicted that by 2030 about 90,000 ha will be needed for residential developments in the Netherlands and 55,000 ha for industry, offices and commerce. Moreover, nearly 250,000 ha of agricultural land will be required for the development of the Ecological Main Structure. Where will these developments take place? What effect will these developments have on our environment, natural landscapes, ecosystems and human health?

Land use simulation models are used to depict future spatial developments, ideally revealing conflicting interests between urbanisation, nature conservation, industrialisation, water management and agriculture. However, decision makers relying on the results of these models should have some concept of the validity and accuracy of these models.

In this thesis, I have compiled a selection of papers *to improve the quality of land use simulation models and the prediction of future land use patterns*. The following sections summarise the main findings according to the main stages in model development, progressing from a conceptual model to a site-specific model and the application of that model. In order to improve the definition of the conceptual model one trys to discover the main determinants and processes describing land use change. To improve the site-specific model the research foused on the calibration of land use models and selection of the performance criteria. In order to apply land use models, to set out spatial policy, the results, the simulated land use in 2030, should be reliable. This stage in model development concentrated on validation of the land use model and assessment of the uncertainty in the predictions.

The Environment Explorer was used in this thesis as the main land use model to simulate future land use developments. The development of the Environment Explorer started at the National Institute of Public Health and the Environment (RIVM) in the Netherlands in collaboration with the Research Institute of Knowledge Systems (RIKS) to evaluate the impacts of future spatial developments. The Environment Explorer consists of a national and regional module, to account for developments on both the national and regional scale and a geographically detailed, dynamic land use model. Land use developments are allocated to those locations on the land use map with the highest probability. Optionally, the potential impacts of the future land use developments on nature and landscape can be estimated with Environment Explorer for a wide range of spatially detailed indicators.

The dynamic land use model in the Environment Explorer applies cellular automata to describe the influence of the neighbourhood in the land use developments. The calibration of the parameters in this cellular automata model is one of the aspects addressed in this thesis. Other land use models use logistic regression to estimate the probability of land use change. These models are calibrated on the observed land use changes. Both models have been applied in this thesis to improve the simulation of future land use developments. Finally both methods have combined into one land use model using cellular automata to describe the influence of the neighbourhood that is calibrated using logistic regression.

Improving the Conceptual Model

The definition of the conceptual model has been improved through analysis of historical land use developments. An empirical study to reveal the main determinants of land use change in the Netherlands, as derived from various land use allocation theories, was carried out with the aim of increasing our understanding of land use change processes. Therefore, a new method for analysing neighbourhood characteristics has been applied with the aim of examining the composition of the neighbourhood of various land uses and land use developments. This method can be used to explore the relevant neighbourhood characteristics in models of land use change. The empirical study shows that both residential and industrial developments can be explained by neighbourhood interactions, accessibility and policy. Therefore a second more detailed study focused on the influence of the neighbourhood including initial land use in the development of residential areas in the Netherlands. The results of this study show that the development of residential areas is primarily determined by initial land use and the land use in its neighbourhood. Spatial policies affect these primary drivers of land use change by shifting the probability of residential developments according to its nature – either restrictive or supportive. In the Netherlands, locations strongly subsiding due to the oxidation of peat are less suitable for residential developments.

This study also shows that the influence of the neighbourhood on the probability of new residential developments depends on cluster-size. Locations near large residential areas are more likely to change than locations near small residential areas, while locations near large airports are less likely to change to residential area than those near small airports.

To improve our understanding of the spatial processes leading to land use change, future research should focus on empirical spatial analysis of land use developments for various land use categories, various regions in the world and at different spatial and temporal resolutions.

Improving the Model

The knowledge developed during the definition of the conceptual model has been applied in the development of the Environment Explorer. After the implementation of the model for the Netherlands the Environment Explorer has been improved by calibration of the model on the historical land use developments. During calibration the parameters of model are adjusted to reproduce the historical land use developments as best one can.

The research in this thesis was initially focused on the development of the correct measures to calibrate the land use model. The main research question was: How to compare maps? 'Cel by cel' comparison of the simulated and observed land use developments does not fully reflect the performance the land use model. If the model does not predict the exact location of land use change but in a directly neighbouring position it's performance is quite well. New techniques have been developed comparing simulated and observed land use within a larger region. These new techniques have been implemented in the 'Map Comparison Kit' together with a large number of existing map comparison algorithms.

One of these new map comparison techniques, the Fuzzy Kappa, has been used to develop a semi-automated procedure to calibrate the Environment Explorer. To decrease the computation time for the calibration, the number of parameters describing the neighbourhood rule was limited to a minimum of four per land use, thereby highly simplifying the shape of the neighbourhood rule. This semi-automated procedure is complex, time-consuming and strongly hindered by classification and aggregation errors in existing land use maps.

This thesis also shows how logistic regression can be used to calibrate a cellular automata-based land use model for residential developments in the Netherlands. To cope with the multicollinearity, a new methodology has been developed: Single Parameter Impact Estimation (SPIE). This procedure for estimating the parameters is also complex and time-consuming, but it has the potential to be fully automated. The computation time in this procedure increases linearly with the number of neighbourhood variables and, therefore, the need to limit the number of neighbourhood variables in order to restrict computation time is less stringent. The methodology should be tested in the future for a complete land use model, including all relevant land use categories.

In the future, it should be possible to improve the calibration of land use models using successive remote sensing images. In terms of calibrating the model, these spatial time series provide more information on the land use developments over time and they also avoid the effect of the classification errors made in the translation of the remote sensing images into land use maps.

Improving the Application

Once the model has been set up for a certain location and calibrated it can be applied to gain insight into reality and obtain predictions that can be used by managers. Therefore, the resulting maps of the future land use should be reliable. The reliability of the land use model can be improved by testing the validity of that model under various conditions and by definition of the uncertainties in the results.

For the evaluation of the National Spatial Strategy, the Environment Explorer has been calibrated and validated on historical land use developments. To validate the long-term developments, the distribution of urban cluster sizes in future land use maps was verified using Zipf's Law, a power law describing the relation between the occurrence and size of urban clusters.

To depict the uncertainties in the future spatial developments, a Monte Carlo simulation was applied to define the probability of urbanisation in the future land use map of 2030. This probability map (Figure 9.3) is, to large extent, comparable with the probability maps presented in this thesis based on the logistic regression models that include the neighbourhood composition (Figure 5.3, 8.4, 10.6). The validity of these logistic regression models has been tested with the Receiver Operator Characteristic (ROC). The neighbourhood-based logistic regression model for residential developments in the Netherlands showed the highest ROC value of 0.96.

The uncertainty in the land use models is so large that the customary methods to define the uncertainty in the prediction were not applicable. Therefore, a novel methodology has been developed to estimate the model uncertainty in the prediction of logistic regression models. Application of this methodology to the region of Amsterdam showed that the model uncertainty varies spatially and depends for the most part on the initial land use category. The uncertainty is low for all land uses that are unlikely to change, such as airports and surface waters, or those which are very likely to change, such as building sites. The uncertainty is high for allotments, sporting fields and agricultural land on the outskirts of the towns.

More insight into the uncertainties associated with the results of land use change simulations is needed to improve the application of land use models. Better insight is also needed to improve the most uncertain processes and parameters in the models themselves. To improve the models it is best to start with the most uncertain processes and parameters.

The residual error of the logistic regression forms an essential part of logistic regression models. It should be used to define the stochastic component of the model that, if applied in a dynamic simulation, will endow the model with its emergent properties.

Conclusions

The research described in this thesis shows that the development of land use is determined by the initial land use, the land use in the neighbourhood and spatial policy. The suitability of the soil seems to be less important for urban developments in the Netherlands except for the peat areas where the soil subsides due the oxidation of the peat. Except for building sites the probability of urban developments is highest on the undeveloped sites in towns and on allotments, sporting fields and agricultural land on the outskirts of the towns.

Calibration of land use models is a necessary, complex and time-consuming business in which Single Parameter Impact Estimation (SPIE) should be used to estimate the parameters of the land use model. The simulated urban land use developments can and should be validated using Zipf's Law describing the relation between the occurrence and the size of the urban land use clusters.

The uncertainty in the simulated urban land use developments is large. The uncertainty is low for airports and water as these land uses are unlikely to change as well as for building sites because these sites will change most likely. The uncertainty is highest for the undeveloped sites at the outskirts of the towns and cities.

Samenvatting

Modellering van landgebruiksveranderingen
Verbetering van de voorspelling van toekomstige landgebruikspatronen

Inleiding

Sinds de prehistorie verandert de mens zijn omgeving. Met de ontwikkeling van de landbouw werden bossen gekapt om plaats te maken voor akkers en weilanden. Vanaf de tweede helft van de 18^{e} eeuw begon de bevolking te groeien door verbeterde agrarische methoden, nieuwe gewassen zoals maïs en aardappel, door verbetering van de hygiëne, de aanleg van drinkwater en rioleringsstelsels en de industriële revolutie. Met de demografische en industriële revolutie begonnen de steden te groeien.

Nu telt de wereld bijna 6.7 miljard mensen, meer dan de helft daarvan woont de laatste jaren in steden. De ontwikkeling van de wereldbevolking en de daaraan gerelateerde economische groei heeft zijn weerslag op het gebruik van onze leefomgeving. Globaal gezien neemt het agrarische en stedelijke landgebruik toe ten koste van bos en natuur.

Economische en demografische ontwikkelingen zullen ook in de toekomst aanleiding zijn om het landgebruik te veranderen. Dat is zeker waar voor Nederland waar iedere vierkante meter al wordt gebruikt voor landbouw, industrie, woonwijken, kantoren, winkels, recreatie, wegen, bos en natuur. Zo verwacht men tot 2030 in Nederland 90.000 ha land nodig te hebben voor nieuwe woonwijken, bijna 55.000 ha voor industrie, kantoren en winkels. Volgens de Nota Mobiliteit wordt het wegennet uitgebreid met 3.700 km strooklengte. Daarnaast is er een kleine 250.000 ha grond nodig voor de ontwikkeling van bos en natuur in de Ecologische Hoofdstructuur. Waar zullen deze ontwikkelingen plaatsvinden? Wat betekent dit voor onze leefomgeving, ruimte, milieu, natuur en gezondheid?

Landgebruiksmodellen worden gebruikt om inzicht te krijgen in de toekomstige ruimtelijke ontwikkelingen en de conflicterende belangen tussen verstedelijking, industrialisatie, landbouw, natuur en waterbeheer. Omdat de resultaten van deze modellen worden gebruikt om beleidsmatige keuzes te maken over de toekomstige ruimtelijke ontwikkelingen is het van belang dat deze modellen ook betrouwbaar zijn.

In dit proefschrift heb ik mijn onderzoek om de voorspelling van het toekomstige landgebruik te verbeteren, gebundeld. De resultaten van dit onderzoek worden in onderstaande paragrafen samengevat op basis van de drie relevante stappen in de ontwikkeling van een model: de definitie van het conceptueel model, de uitwerking van het landgebruiksmodel voor een bepaald gebied en de toepassing van het landgebruiksmodel. Bij de definitie van het conceptueel model tracht men te achterhalen welke processen en parameters de ontwikkeling van het landgebruik het beste beschrijven. Bij de uitwerking van het model heeft het onderzoek zich met name gericht op de kalibratie van landgebruiksmodellen en de selectie van de juiste kalibratieparameters. In de derde stap, bij de toepassing van het landgebruiksmodel, is het van belang om te weten of men de resultaten van het model, het voorspelde landgebruik in 2030, ook kan vertrouwen. Deze stap richt zich op de validatie van het model en bepaling van de onzekerheid in de voorspellingen.

In dit proefschrift is met name de LeefOmgevingsVerkenner gebruikt om de landgebruiksveranderingen te modelleren. De LeefOmgevingsVerkenner is een ruimtelijk dynamisch model dat eerst bij het Rijksinstituut voor Volksgezondheid en Milieu (RIVM) en later bij het Milieu en Natuurplanbureau (MNP) is ontwikkeld in samenwerking met Research Institute for Knowledge Systems (RIKS) uit Maastricht. Het model verdeelt de nationale demografische en economische ontwikkelingen naar het regionale, COROP, niveau waarna de regionale vraag per type landgebruik wordt berekend. Deze regionale landgebruiksontwikkelingen worden vervolgens vertaald naar de verandering van het landgebruik op locale schaal. De ontwikkelingen worden toebedeeld aan die locaties op de landgebruikskaart die de meeste kans maken om te veranderen. Welke locaties dat zijn hangt af van het type landgebruik, wonen, industrieterrein, recreatie, glastuinbouw, etc. Op basis van het landgebruik kan de LeefOmgevingsVerkenner de potentiële effecten op een groot aantal milieu- en natuurindicatoren berekenen zoals de belasting met geluid, de versnippering van de natuur of de veranderingen in de piekafvoer van de rivieren.

De LeefOmgevingsVerkenner maakt gebruikt van cellulaire automaten om de invloed van de omgeving mee te nemen in de kans dat een locatie van landgebruik zal veranderen. De kalibratie van de parameters van deze cellulaire automaten is een van de aspecten die in dit proefschrift aan de orde komt. Daarnaast bestaan er landgebruiksmodellen die logistische regressie gebruiken om de kans dat een locatie van landgebruik zal veranderen te bepalen. Deze modellen worden gekalibreerd op de veranderingen in het waargenomen landgebruik. Beide typen modellen zijn in dit proefschrift gebruikt om de voorspelling van het toekomstige landgebruik te verbeteren en uiteindelijk te integreren tot een landgebruiksmodel op basis van cellulaire automaten dat gekalibreerd kan worden met logistische regressies.

Verbetering van het Conceptueel Model

De definitie van het landgebruiksmodel is verbeterd door de historische landgebruiks-veranderingen te analyseren. Zo zijn op basis van de verschillende theoretische concepten met logistische regressie de belangrijkste factoren bepaald die de ontwikkeling van het landgebruik in Nederland bepalen. Voor deze analyse is een nieuwe indicator ontwikkeld om de samenstelling van het landgebruik in de omgeving van een locatie te beschrijven. Deze analyse liet zien dat de locatie waar het landgebruik veranderd voor een belangrijk gedeelte verklaard wordt door de samenstelling van het landgebruik in de omgeving. Een tweede, uitgebreidere analyse liet zien dat daarnaast ook het huidige landgebruik van belang is. Nieuwe woonwijken ontwikkelen zich meestal op locaties die nog niet bebouwd zijn zoals volkstuinen, sportvelden en landbouwgrond. Zoals te verwachten bleek uit deze analyse dat locaties die door het beleid zijn aangewezen als groeikern of uitleggebied relatief meer kans maken om in een woonwijk te veranderen en dat locaties die in de Ecologische Hoofdstructuur of de Natura-2000 gebieden liggen relatief minder kans maken om in een woonwijk te veranderen.

Ook liet deze studie zien dat een locatie nabij een grote stad relatief meer kans maakt in een woonwijk te veranderen dan een locatie nabij een klein dorpje terwijl een locatie nabij een groot vliegveld zoals Schiphol relatief minder kans maakt om in wonen te veranderen dan een locatie nabij een klein vliegveld.

Om het inzicht in de processen en parameters die leiden tot verandering van het landgebruik verder te ontwikkelen wordt het aanbevolen om vergelijkbare analyses van de historische veranderingen in het landgebruik te maken voor andere gebieden, andere landgebruiksklassen en op andere ruimtelijke en temporele schaalniveaus

Verbetering van het Model

De kennis die in de vorige stap is ontwikkeld, is vervolgens gebruikt bij de uitwerking van de LeefOmgevingsVerkenner voor Nederland. Nadat het model was opgezet en alle gegevens ingevoerd is het model verbeterd door het te kalibreren op de historische landgebruiksontwikkelingen in Nederland. Bij de kalibratie worden de parameters van het model zo ingesteld dat het model de historische ontwikkeling van het landgebruik zo goed mogelijk imiteert.

Het onderzoek in deze stap van de modelontwikkeling heeft zich in eerste instantie gericht op de ontwikkeling van de juiste maten om het model te kalibreren. De centrale vraag daarbij was: hoe vergelijk je twee kaarten? Het 'cel per cel' vergelijken van gesimuleerde en waargenomen landgebruiksveranderingen geeft namelijk niet juist weer hoe goed het model is. Als het model de verandering van het landgebruik niet precies op de juiste locatie maar in naastgelegen locatie voorspelt dan is dat zo slecht nog niet. Om het gesimuleerde en waargenomen landgebruik met elkaar te vergelijken zijn nieuwe methoden ontwikkeld die de landgebruiksveranderingen binnen een groter gebied vergelijken. Deze nieuwe methoden om zijn samen met een groot aantal bestaande methoden geïmplementeerd in de Map Comparison Kit.

Op basis van de 'Fuzzy Kappa', een van de nieuwe methoden om kaarten te vergelijken, is vervolgens een semi-geautomatiseerde kalibratie procedure ontwikkeld om de LeefOmgevingsVerkenner te kalibreren. Om de rekentijd van de kalibratie te beperken is het aantal parameters in de LeefOmgevingsVerkenner sterk verminderd ten koste van de manier waarop de invloed van de omgeving in het model wordt beschreven. De semi-geautomatiseerde kalibratie procedure blijft complex, kost veel tijd en wordt sterk gehinderd door fouten en wijzigingen in de classificatie van de landgebruikskaarten.

Dit proefschrift laat tevens zien hoe een vergelijkbaar landgebruiksmodel, gebaseerd op cellulaire automaten, gekalibreerd kan worden door middel van logistische regressies. Om de problemen met multi-collineariteit in de logistische regressie te vermijden is een nieuwe methode ontwikkeld om de parameter waarden te bepalen: 'Single Parameter Impact Estimation' (SPIE). Deze methode om het model te kalibreren is ook complex en tijdsintensief maar kan in principe volledig geautomatiseerd worden. Een ander voordeel is dat de rekentijd van deze methode lineair toeneemt met het aantal variabelen om de omgeving te beschrijven. De noodzaak om het aantal omgevingsvariabelen te beperken vanwege de lange rekentijden is daarom kleiner. De methode is in dit proefschrift uitgewerkt voor het landgebruikstype wonen en dient in de toekomst nog getest te worden voor een volledig landgebruiksmodel met alle relevante landgebruiksklassen.

In de toekomst zou de kalibratie van landgebruiksmodellen verbeterd kunnen worden door opeenvolgende remote-sensing beelden te gebruiken. Deze beelden leveren enerzijds meer informatie om een landgebruiksmodel te kalibreren anderzijds worden de problemen door classificatiefouten, die bij de interpretatie van deze beelden tot landgebruikskaarten worden gemaakt, vermeden.

Verbetering van de Toepassing

Als het landgebruiksmodel gekalibreerd is, kan het worden toegepast om meer inzicht te krijgen in de toekomstige ruimtelijke ontwikkelingen of om beleidsmakers te ondersteunen bij de uitwerking van het toekomstige ruimtelijke beleid door verschillende scenario's door te rekenen. Het is daarbij belangrijk dat de resultaten van het landgebruiksmodel betrouwbaar zijn. De betrouwbaarheid van het landgebruiksmodel kan worden verbeterd door de validiteit van het model in verschillende omstandigheden te testen en door de onzekerheid in de uitkomsten te bepalen.

Voor de evaluatie van de Nota Ruimte is de LeefOmgevingsVerkenner gekalibreerd en gevalideerd op basis van de historische landgebruiksveranderingen. Om de modelresultaten op de lange termijn te valideren is de gesimuleerde landgebruikskaart van 2030 getoetst op basis van de Wet van Zipf, een zogenaamde 'power law' die de relatie tussen het voorkomen en de grootte van de stedelijke gebieden beschrijft. Naarmate het oppervlak van de stedelijke gebieden groter is, komen ze relatief minder vaak voor.

Om inzicht in de onzekerheid van de gesimuleerde landgebruiksontwikkelingen met de LeefOmgevingsVerkenner te krijgen is een Monte Carlo simulatie uitgevoerd om de kans op verstedelijking in 2030 te bepalen. Deze kansenkaart (Figuur 9.3) lijkt sterk op de kansenkaarten die uit de logistische regressie modellen komen in dit proefschrift waarbij de invloed van de omgeving wordt meegenomen (Figuren 5.3, 8.4, 10.6). De geldigheid van deze logistische regressie modellen kan getoetst worden met de Receiver Operator Characteristic (ROC). Het logistische regressie model voor de ontwikkeling van woonwijken waarbij het huidige landgebruik en het landgebruik in de omgeving wordt meegenomen toonde de hoogste ROC waarde van 0.96.

De onzekerheid in de voorspelling van de logistische landgebruiksmodellen is zo groot dat de gebruikelijke methoden om de onzekerheid te bepalen niet toegepast konden worden. Daarom is een nieuwe methode ontwikkeld om de onzekerheid in de voorspelling van deze logistische regressie modellen te bepalen. Toepassing van deze nieuwe methode voor de regio Amsterdam laat zien dat de onzekerheden in de modelvoorspellingen ruimtelijk variëren en sterk afhankelijk zijn van het huidige landgebruik. De onzekerheid is laag voor alle landgebruiken die hoogstwaarschijnlijk wel (bouwterreinen) en hoogstwaarschijnlijk niet in wonen zullen veranderen (vliegvelden en water). De onzekerheid is het grootst voor volkstuinen, sportterreinen en landbouwgrond aan de rand van de stad.

Er is meer inzicht nodig in de onzekerheid in de gesimuleerde landgebruikskaarten. Dit is nodig om de toepassing van deze modellen te verbeteren maar ook om de modellen zelf te verbeteren. Om de onzekerheid in de voorspellingen van het model te verminderen wordt het aanbevolen om eerst de meest onzekere processen en parameters in het model te verbeteren.

De residuele fout van de logistische regressie vormt een essentieel onderdeel van logistische landgebruiksmodellen. Deze fout moet gebruikt worden om de stochastische component van het model te definiëren dat, indien toegepast in een dynamische simulatie, het model zijn emergente eigenschappen verleent.

Conclusies

Uit het onderzoek dat in dit proefschrift wordt beschreven blijkt dat de ontwikkeling van het landgebruik wordt bepaald door het huidge landgebruik, het landgebruik in de omgeving en door het ruimtelijk beleid. De geschiktheid van de bodem blijkt in Nederland nauwelijks een rol te spelen behalve in veengebieden waar de bodem snel daalt door oxidatie van het veen. De kans op stedelijke uitbreidingen is naast locaties op bouwterreinen het grootst op onbebouwde locaties in de stad en op volkstuintjes, sportterreinen en landbouwgrond aan de rand van de stad.

Kalibratie van landgebruiksmodellen is een noodzakelijke, complexe en tijdrovende aangelegenheid waarbij de invloed van de verschillende parameters het beste met "Single Parameter Impact Estimation" (SPIE), zoals beschreven in dit proefschrift bepaald kan worden. Validatie van het gesimuleerde stedelijke landgebruik in de toekomst is mogelijk op basis van de Wet van Zipf die de relatie beschrijft tussen het voorkomen en de grootte van de stedelijke gebieden.

De onzekerheid in de gesimuleerde stedelijke landgebruiksveranderingen is groot. Alleen als het huidige landgebruik vliegvelden, water of bouwterrein is dan is de onzekerheid in de modelresultaten relatief klein. De onzekerheid is het grootst voor de onbebouwde locaties aan de rand van de stad.

Dankwoord

Het is raar maar het is klaar. Met dit proefschrift komt er een einde aan een lang verhaal. Een verhaal met zijn 'ups and downs' maar, en dat is wel het belangrijkste, met een goed einde.
Het verhaal begint in 1997 toen ik bij Wilbert Slooff en Raymond de Niet terecht kwam. Wilbert moest een Leefomgevingsbalans ontwikkelen, een beleidsinstrument om de integrale gevolgen van ingrepen in het fysieke milieu op de leefomgeving te beoordelen. In deze balans werd naast de economische waarde ook de ecologische en sociaalpsychologische waarden van de omgeving onderkend: people, planet & profit. In aanvulling op die balans mocht ik een LeefOmgevingsVerkenner gaan ontwikkelen die een beeld moest kunnen schetsen van de ontwikkeling van deze waarden in de toekomst. Als eerste wil ik dan ook Wilbert en Raymond bedanken voor de ontdekking van het "Bruto Nationaal Geluk" en de jaren die we samen aan dat thema hebben gewerkt.

Met de ontwikkeling van deze verkenner begon de samenwerking met Guy Engelen, Roger White en alle andere mensen van RIKS. Met veel plezier denk ik terug aan de discussies in Maastricht die altijd uitliepen, waardoor ik er als een haas vandoor moest om alsnog de trein te missen. Guy, Roger, Alex, Hedwig, Inge, Maarten, Bernhard, Patrick, Bas en alle andere mensen van RIKS bedankt voor alle mooie momenten.

Toen het onderzoeksprogramma Networks in the Delta startte aan de Universiteit Utrecht kon ik 2 dagen per week bij Geowetenschappen werken. Daar maakte ik kennis met Martin Dijst, Paul Schot, Stan Geertman, Peter Verburg, Kor de Jong, Michel Hagoort, Ivo Thonon, Anette Weterings, Saim Mohammed, Jan Ritsema van Eck en vele anderen. Binnen dit universitaire gezelschap groeide mijn kennis van de ruimtelijke modellering gestaag. Met de een had ik meer contact dan met de ander. Michel kwam altijd langs met vragen over de LeefOmgevingsVerkenner, Kor ontwikkelde software waar ik nooit wat van begreep, Stan is uiteindelijk een van mijn co-promotoren geworden. Peter Verburg heeft mij veel geleerd onder andere het gebruik van logistische regressies en, niet te vergeten, we hebben samen een aantal artikelen gepubliceerd die ook in dit proefschrift zijn opgenomen en het is leuk dat hij nu in de commissie zit.

Kor is ook degene geweest die me naar Edzer Pebesma verwees, bij fysische geografie waar ik hem nog steeds dankbaar voor ben. Edzer was en is een ruimtelijk statisticus 'pur sang', maar heeft me ook gesteund in de discussie over de kwaliteit van de ruimtelijke modellen. Het is toch wonderlijk dat onze wegen elkaar precies op het juiste moment kruisten. Het is jammer dat hij naar Duitsland is vertrokken maar ik kan hem geen ongelijk geven.

In die tijd werkte ik binnen het Milieu- en Natuurplanbureau binnen het team Ruimte. Ik dank Karst Geurts, Leon Crommentuin, Rienk Kuiper, Arno Bouwman, Hans Visser en alle anderen waar ik toen mee samenwerkte.

Inmiddels is het 2007 en ik ga terug naar mijn roots, het Laboratorium voor Effecten en Risico's binnen het RIVM waar ik na mijn studie ben begonnen. Ton Breure wilde mij graag hebben, ik ga aan de slag voor Wilko Verweij en Leo Posthuma en tussen de bedrijven door mag ik mijn promotie afmaken. Ton, ik dank jou, Aart en Dick voor het vertrouwen en de ruimte die jullie me hebben gegeven.

Het heeft even geduurd maar uiteindelijk is het me dan toch gelukt. Ik dank mijn promotor, Steven de Jong die me in de laatste fase heeft geholpen en mijn begeleiders, Edzer en Stan, voor hun geduld, advies en inzicht. Ik dank ook alle andere mensen binnen Geowetenschappen die me direct danwel indirect hebben geholpen om deze mijlpaal te bereiken zoals Paul Hiemstra, Elisabeth Addink, Derek Karssenberg, Oliver Schmitz, Harm Deters, Gerlach van Beinum, Margot Stoete en niet te vergeten Juul Beltman.

Een groot aantal van bovengenoemde mensen hebben als auteur of co-auteur direct bijgedragen aan de totstandkoming van dit proefschrift waarvoor mijn dank.

Ik dank mijn huidge collega's bij het Laboratorium voor Ecologische Risicobeoordeling en de afdeling Water. Het zijn er teveel om allemaal te vermelden maar een paar wil ik toch noemen omdat ze me op de een of andere manier geholpen hebben bij mijn promotie Esther, Dik, Wilko, Dick, Job, Marga en Sally.

I also would like to extend a special word of thanks to my copy-editor, Ann Jenks, who has re-written my English and made it easier to read. Ann, I sincerely thank you for playing an essential role in the production of this thesis.

Ik dank mijn vrienden voor de gezelligheid, ontspanning en relativerende momenten. Hier wil ik vooral Rene maar ook Roland bedanken voor de geweldige discussies in Binnen Best.

Maar de meeste uren heb ik toch thuis zitten ploeteren, lezen, rekenen, schrijven, herschrijven, meestal tot in de vroege uurtjes als het rustig was en de kinderen in bed lagen. Het gevolg was wel dat ik 's morgens niet altijd even fris en vrolijk uit mijn bed stapte. Ik dank Monique voor alle compassie in het ochtendgloren, de steun en het vertrouwen. Ik dank de kinderen dat ze er waren, voor de liefde, de genegenheid en de warmte die ze me hebben gegeven.

Ton

De Bilt, februari 2009

Curriculum Vitae

Ton de Nijs, born 1960, graduated in Geochemistry (cum laude) in 1986 at Utrecht University, the Netherlands. After his studies he started working on the risk assessment of chemical substances at the National Institute of Public Health and Environment (RIVM) and was closely involved with the development of the European Uniform System for Evaluation of Substances. In 1992 he became involved in water quality modelling at the local and national scale. In 1997 he started working on the Environment Explorer, thereby becoming involved in spatial analysis and research aimed at improving land use models in general and the Environment Explorer in particular. He has applied this model in various projects, including in the evaluation of the National Spatial Strategy of the Ministry of Housing, Spatial Planning and Environment. He participated in the organisation of ERSA 2005 at the Free University of Amsterdam and Framing Land Use Dynamics I and II at Utrecht University. He has edited a special issue on land use simulation and analysis of the Annals of Regional Science. He is currently working at the Laboratory for Ecological Risk Assessment (RIVM).

Publications

de Nijs, A. C. M. & Pebesma, E. J. (accepted) Estimating the influence of the neighbourhood in the development of residential areas in the Netherlands. Environment and Planning B Planning and Design.

Zijp, M., de Nijs, A. C. M. & Verweij, W. (2009) Using exemptions of the WFD. Two cases in the Netherlands. HydroEco 2009: 2nd International Multidisciplinary Conference on Hydrology and Ecology: Ecosystems Interfacing with Groundwater and Surface Water. Vienna.

de Nijs, A. C. M., de Zwart, D. & Peijnenburg, W. (2009) Risk of mixture toxicity in Dutch surface waters. SETAC Europe. Göteborg.

de Nijs, A. C. M. (2009) Future land use in the Netherlands: evaluation of the National Spatial Strategy. IN GEERTMAN, S. & STILLWELL, J. (Eds.) Planning Support Systems: Best Practice and New Methods. Springer Publishers.

de Nijs, A. C. M., Zijp, M. & Verweij, W. (2008) Experiences in the Assessment of Groundwater Chemical Status in the Netherlands. IN GRUNDVAND, A. J. O. (Ed.) GroPro – Groundwater Protection Plans and Implementation in a North European Perspective. Comwell Klarskovgaard, Korsør.

de Nijs, A. C. M., De Zwart, D. & Peijnenburg, W. (2008) Toxiciteit van oppervlaktewater blijft een probleem. H2O, 21, 29-31.

Koomen, E., Rietveld, P. & de Nijs, T. (2008) Modelling land use change for spatial planning support. Annals of Regional Science, 42, 1-10.

Pontius Jr, R. G., Boersma, W., Castella, J. C., Clarke, K., de Nijs, T., Dietzel, C., Duan, Z., Fotsing, E., Goldstein, N., Kok, K., Koomen, E., Lippitt, C. D., McConnell, W., Mohd Sood, A., Pijanowski, B., Pithadia, S., Sweeney, S., Trung, T. N., Veldkamp, A. T. & Verburg, P. H. (2008) Comparing the input, output and validation maps for several models of land change. Annals of Regional Science, 42, 11-37.

Addink, E., Barendregt, A., Ettema, D., Karssenberg, D. & de Nijs, A. C. M. (2007) Abstracts of keynote lectures, oral presentations and posters. Abstract book, Utrecht University, The Netherlands.

de Nijs, A. C. M. & Pebesma, E. J. (2007) Spatial Uncertainty in land use models. An alternative method to estimate uncertainty in logistic regression methods. 15th European Colloqium on Theretical and quantitave Geography. Montreux, Switserland, Institute of Geography of the University of Lausanne, Switzerland.

Pontius, G. R., Castella, J., de Nijs, T. C. M., Duan, Z., Fotsing, E., Goldstein, N., Kok, K., Koomen, E., Lippitt, C. D., McConnell, W., Mohd Sood, A., Pijanowski, B., Veldkamp, A. T. & Verburg, P. H. (2007) Lessons and challenges in land change modeling as revealed by map comparisons. Proceedings for conference on the science and education of land use: a transatlantic, multidisciplinary and comparative approach. Washington DC USA.

Visser, H. & de Nijs, A. C. M. (2006) The Map Comparison Kit. Environmental Modelling and Software, 21, 346-358.

Hagen-Zanker, A., van Loon, J., Maas, A., Straatman, B., de Nijs, T. C. M. & Engelen, G. (2005) Measuring Performance of Land Use Models. An Evaluation Framework for the Calibration and Validation of Integrated Land Use Models Featuring Cellular Automata. European Colloquium on Theoretical and Quantitative Geography. Tomar, Portugal.

Crommentuijn, L. E. M., de Nijs, A. C. M. & Kuiper, R. (2004) Bouwen eigen groei platteland uiteenlopend in te vullen – (RIM); 11)17-9 ROM, 22, 17-9.

de Nijs, A. C. M., De Niet, R. & Crommentuijn, L. (2004) Constructing land use maps of the Netherlands in 2030. Journal of Environmental Management, 72, 35-42.

Groen, J., Koomen, E., Piek, M., Eck, J. v., Tisma, A., Bouwman, A., Bakema, A., de Nijs, T. C. M., Kuhlman, T., Blijie, B. & Bok, M. d. (2004) Scenario's in kaart. Model- en ontwerpbenaderingen voor toekomstig ruimtegebruik. Ruimtelijk Planbureau. Den Haag

Kuiper, R. & de Nijs, A. C. M. (2004) Nota Ruimte leidt tot verdichting stedelijke netwerken. ROM, 6.

Kuiper, R., Niet, R., de & de Nijs, T. C. M. (2004) Deltametropool langs buitenrand verstedelijken. Quick scan groen/blauwe effecten woningbouwlocaties deltametropool 2004;. ROM 1, 8-11.

Verburg, P. H., de Nijs, T. C. M., Ritsema van Eck, J., Visser, H. & Jong, K. d. (2004a) A method to analyse neighbourhood characteristics of land use patterns. Computers, Environment and Urban Systems, 28, 667-690.

Verburg, P. H., Van Eck, J. R. R., de Nijs, T. C. M., Dijst, M. J. & Schot, P. (2004b) Determinants of land use change patterns in the netherlands. Environment and Planning B Planning and Design, 31, 125-150.

Engelen, G., White, R. & de Nijs, T. C. M. (2003) The Environment Explorer: Spatial Support system for Integrated Assessment of Socio-economic and Environmental Policies in the Netherlands. Integrated Assessment, 4, 97-105.

Beurskens, J. E. M., de Nijs, A. C. M. & Teodorascu, D. M. (1996a) Predicting microbial PCB dehalogenation pathways. In: Abstracts of the IBC's international symposium Biological dehalogenation. Annapolis, USA, Southborough: International Business Communications.

Beurskens, K., de Nijs, A. C. M. & Traas, T. (1996b) Coupling long term in situ microbial degradation of pollutants to thresholds for wildlife, a modeling approach. Sixth SETAC-Europe annual meeting. Sicily, Italy, Society of Environmental Toxicology and Chemistry.

Kramer, P. R. G., de Nijs, A. C. M. & Aldenberg, T. (1995) Modelling combined effects of nutrients and toxicants in a branch of the Rhine delta.In: Global environmental protection: science, politics and common sense. 2nd SETAC world congress Vancouver, Canada, Society of Environmental Toxicology and Chemistry,.

Kramer, P. R. G., de Nijs, A. C. M. & Aldenberg, T. (1994) Bayesian uncertainty analysis in water quality modelling. IN GRASMAN, J. & VAN STRATEN, G. (Eds.) Predictability and nonlinear modelling in natural sciences and economics. Kluwer, Dordrecht.

de Nijs, A. C. M., Toet, C., Vermeire, T., Poel, P. v. d. & Tuinstra, J. (1993) Dutch risk assessment system for new chemicals: DRANC. Sci Total Environ Suppl, 1729-48

de Nijs, A. C. M. & de Greef, J. (1992) Ecotoxicological risk evaluation of the cationic fabric softener DTDMAC. II. Exposure modelling. Chemosphere, 24, 611-27

Leeuwen, K. v., Roghair, C., de Nijs, T. C. M. & de Greef, J. (1992) Ecotoxicological risk evaluation of the cationic fabric softener DTDMAC. III. Risk assessment. Chemosphere 24, 629-39.

de Nijs, A. C. M. & de Greef, J. (1991) Verdunning van effluent van rioolwaterzuiveringsinstallaties in het Nederlandse oppervlaktewater. H2O 24, 725-30

Leeuwen, K. v., Roghair, C., de Greef, J. & de Nijs, A. C. M. (1990) Wasverzachters II. Resultaten van aanvullend onderzoek. H2O 23, 295-9.

De Kruijf, H. A. M., Gaag, M. A. v. d., de Nijs, A. C. M. & Knoop, J. M. (1988) The Rhine as a european sewage system and as a source of drinking water: coping with emergencies. Hum Toxicol, 7, 58-60.

References

ABF (1998) Woonmilieudatabase 98. Delft, The Netherlands

ABF (2002) Ruimtevraag wonen, werken en voorzieningen herberekend. Verkenning 2000-2030 voor deel 3 van de Vijfde Nota Ruimtelijke Ordening. R 2001-0097LM. ABF Onderzoek en Informatie. Delft, The Netherlands

Al-Kheder, S., Wang, J. & Shan, J. (2007) Cellular automata urban growth model calibration with genetic algorithms. 2007 Urban Remote Sensing Joint Event, URS.

Allen, P. M. & Sanglier, M. (1979a) A Dynamic Model of a central Place System-II. Geographical Analysis, 13, 149-164.

Allen, P. M. & Sanglier, M. (1979b) A Dynamical Model of Growth in a Central Place System. Geographical Analysis, 11, 256-272.

Alonso, W. (1964) Location and land use, Cambridge, Mass., Harvard University Press.

Alterra (2003) Landelijk grondgebruiksbestand LGN4 2000. Alterra Wageningen, The Netherlands

Angelsen, A. & Kaimowitz, D. (1999) Rethinking the causes of deforestation: Lessons from Economic Models. The World Bank Observer, 14, 73-98.

Anselin, L. (1988) Spatial Econometrics: Methods and Models, Boston, Kluwer.

Anselin, L. (2002) Under the hood – Issues in the specification and interpretation of spatial regression models. Agricultural Economics, 27, 247-267.

Arthur, W. B. (1994) Urban systems and historical path dependence. IN ARTHUR, W. B. (Ed.) Increasing returns and path dependence in the economy. Ann Arbor The University of Michican Press.

Aspinall, R. (2004) Modelling land use change with generalized linear models – a multi-model analysis of change between 1860 and 2000 in gallatin valley, montana. Journal of Environmental Management, 72, 91-103.

Bandemer, H. & Gottwald, S. (1995) Fuzzy sets, fuzzy logic, fuzzy methods with applications, New York, Wiley & Sons.

Barredo, J. I., Kasanko, M., Mccormick, N. & Lavalle, C. (2003) Modelling dynamic spatial processes: simulation of urban future scenarios through cellular automata. Landscape and Urban Planning, 64, 145-160.

Batty, M. (1994) A chronicle of scientific planning. The Anglo-American modeling experience. Journal of the American Planning Association, 60, p7, 10p.

Batty, M. & Longley, P. (1994) Fractal cities: a geometry of form and function, San Diego, CA, USA, Academic Press Professional, Inc.

Batty, M. & Xie, Y. (1994) From Cells to Cities. Environment and Planning B-Planning & Design, 21, S31-S38.

Bell, K. P. & Bockstael, N. E. (2000) Applying the generalized-moments estimation approach to spatial problems involving microlevel data. The Review of Economics and Statistics, 82, 72-82.

Beneson, I. (2007) Warning! The scale of land-use CA is changing! Computers, Environment and Urban Systems, 31, 107-113.

Bio, A. M. F., De Becker, P., De Bie, E., Huybrechts, W. & Wassen, M. (2002) Prediction of plant species distribution in lowland river valleys in Belgium: modelling species response to site conditions. Biodiversity and Conservation, 11, 2189-2216.

Black, A. E., Strand, E., Wright, R. G., Scott, J. M., Morgan, P. & Watson, C. (1998) Land-use history at multiple scales: implications for conservation planning. Landscape and Urban Planning, 43, 49-63.

Bockstael, N. E. (1996) Modeling economics and ecology: the importance of a spatial perspective. American Journal of Agricultural Economics, 78, 1168-1180.
Bonham-Carter, G. F. (1994) Geographic Information Systems for Geoscientiests: Modelling With GIS, Oxford, Pergamon.
Borsboom – van Beurden, J. A. M., de Regt, W. & Schotten, K. (2002) Land Use Scanner: the continuous cycle of application, evaluation and improvement in land use modeling. 42th congress of the European Regional Science Association. Dortmund, Germany.
Bouwman, A., Kram, T. & Klein Goldewijk, K. (2006) Integrated modelling of global environmental change. An overview of IMAGE 2.4 Report Number 500110002. Netherlands Environmental Assessment Agency (MNP). Bilthoven
Brown, D. G., Page, S., Riolo, R., Zellner, M. & Rand, W. (2005) Path dependence and the validation of agent-based spatial models of land use. International Journal of Geographical Information Science, 19, 153-174.
Bürgi, M. & Turner, M. G. (2002) Factors and Processes Shaping Land Cover and Land Cover Changes Along the Wisconsin River. Ecosystems 5.
Burrough, P. A. & McDonell, R. A. (1998) Principles of Geographical Information Systems, New York, Oxford University Press.
Candau, J. (2000) Calibrating a cellular automaton model of urban growth in a timely manner. IN PARKS, B. O., CLARKE, K. M. & CRANE, M. P. (Eds.) 4th International Conference on Integrating Geographic Information Systems and Environmental Modeling: Problems, Prospects, and Needs for Research. Boulder, Sep 2-8, University of Colorado.
Carletta, J. (1996) Assessing agreement on classification tasks: the kappa statistic. Computational linguistics, 22.
Castells, M. (2000) The rise of the network society, Oxford, Blackwell Publishers.
CBS (1993) Bodemstatistiek 1989. Central Bureau of Statistics. Voorburg, Netherlands
CBS (1997) Bodemstatistiek 1993 Central Bureau of Statistics. Voorburg, Netherlands
CBS (2000) Bodemstatistiek 1996 Central Bureau of Statistics. Voorburg, Netherlands
CBS (2003) Bodemstatistiek 2000. Central Bureau of Statistics. The Hague, The Netherlands
CBS (2008) StatLine. All figures within reach. website: www.statline.nl Laatst geopend op august 26, 2008
Cheng, T., Molenaar, M. & Lin, H. (2001) Formalizing fuzzy objects from uncertain classification results. International Journal of Geographical Information Science, 15, 27-42.
Chomitz, K. M. & Gray, D. A. (1996) Roads, land use, and deforestation: a spatial model applied to belize. The World Bank Economic Review, 10, 487-512.
Chowdhury, R. R. (2006) Driving forces of tropical deforestation: the role of remote sensing and spatial models. Singapore Journal of Tropical Geography, 27, 82-101.
Christaller, W. (1933) Central Places of Southern Germany. Edition 1966, London, Prentice Hall.
Clark, W. A. V. (2000) Monocentric to polycentric: new urban forms and old paradigms. IN BRIDGE, G. & WATSON, S. (Eds.) A companion to the City. Oxford, Blackwell Publishers.
Clarke, K. C. & Gaydos, L. J. (1998) Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. International Journal of Geographical Information Science, 12, 699-714.
Clarke, K. C., Hoppen, S. & Gaydos, L. (1997) A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. Environment and Planning B: Planning and Design, 24, 247-261.
Clarke, K. C., Hoppen, S. & Gaydos, L. J. (1996) Methods and Techniques for Rigorous Calibration of a Cellular Automaton Model of Urban Growth. Third International Conference/Workshop Integrating GIS and Environmental Modeling. Santa Fe, NM, January 21-25, 1996.

Cohen, J. (1960) A coefficient of agreement for nominal scales. Educational and Psychological Measurement, XX, 37-46.

Cohen, J. (1968) Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychological Bulletin, 70, 213-220.

Couclelis, H. (1985) A Framework for Modelling Micro-Macro Dynamics. Environment and Planning A, 17, 585-596.

CPB (1996) Omgevingsscenario's lange termijn verkenning 1995-2020. The Hague, The Netherlands, CPB.

Dai, E., Wu, S. H., Shi, W. Z., Cheung, C. K. & Shaker, A. (2005) Modeling change-pattern-value dynamics on land use: an integrated gis and artificial neural networks approach. Environmental Management, 36, 576-591.

Dassen, A. G. (2005) Evaluatie Regelgeving Burgerluchthavens en Militaire Luchthavens. 500047002. Milieu- en Natuurplanbureau. Bilthoven

de Almeida, C. M., Monteiro, A. M. V., Soares, G., Cerqueira, G. C., Pennachin, C. L. & Batty, M. (2005) Gis and remote sensing as tools for the simulation of urban land-use change. International Journal of Remote Sensing, 26, 759-774.

de Almeida, M. C., Batty, M., Vieira Monteiro, A. M., Camara, G., Soares-Filho, B. S., Cerqueira, G. C. & Pennachin, C. L. (2003) Stochastic cellular automata modeling of urban land use dynamics: empirical development and estimation. Computers, Environment and Urban Systems, 27, 481-509.

de Jong, T. & Ritsema van Eck, J. R. (1996) Location profile based measures as an improvement on accessibility modelling in GIS. Computers, Environment and Urban Systems, 20, 181-190.

de Nijs, A. C. M. (2009) Future land use in the Netherlands: evaluation of the National Spatial Strategy. IN GEERTMAN, S. & STILLWELL, J. (Eds.) Planning Support Systems: Best Practice and New Methods. Springer Publishers.

de Nijs, A. C. M., Crommentuijn, L., Farjon, J. M. J., Leneman, H., Ligtvoet, W., de Niet, R., Schotten, K. & Schilderman, C. (2002) Vier scenario's van het landgebruik in 2030. Achtergronddrapport bij de Nationale Natuurverkenning 2, 2000-2030. 408764003. RIVM. Bilthoven, The Netherlands

de Nijs, A. C. M., de Niet, R. & Crommentuijn, L. (2004) Constructing land-use maps of the Netherlands in 2030. Journal of Environmental Management, 72, 35-42.

de Nijs, A. C. M., de Niet, R., de Hollander, G., Filius, F. & Groen, J. (2001a) De LeefOmgevingsVerkenner: Kaartbeelden van 2030. Een verkenning van de inzet bij beleidsondersteuning (in dutch). RIVM rapport 408505 004. National Institute of Public Health and the Environment. Bilthoven

de Nijs, A. C. M., Engelen, G., White, R. & van Delden, H. (2001b) De LeefOmgevings Verkenner. Technische Documentatie. RIVM-rapport 408505007. Rijksinstituut voor Volksgezondheid en Milieu. Bilthoven, The Netherlands

de Nijs, A. C. M., Kuiper, R. & Crommentuin, L. E. M. (2005) Het landgebruik in 2030: Een projectie van de Nota Ruimte 711931010. Rijksinstituut voor Volksgezondheid en Milieu. Bilthoven, The Netherlands

de Nijs, A. C. M. & Pebesma, E. J. (2007) Spatial Uncertainty in land use models. An alternative method to estimate uncertainty in logistic regression methods. 15th European Colloqium on Theretical and quantitave Geography. Montreux, Switserland, Institute of Geography of the University of Lausanne, Switzerland.

de Nijs, A. C. M. & Pebesma, E. J. (accepted) Estimating the influence of the neighbourhood in the development of residential areas in the Netherlands. Environment and Planning B Planning and Design.

de Wit, A. J. W., van der Heijden, T. G. C. & Thunnissen, H. A. M. (1999) Vervaardiging en nauwkeurigheid van het LGN3-grondgebruiksbestand (in dutch). Report no. 663. DLO-Staring Centrum. Wageningen, The Netherlands

Deadman, P., Robinson, D., Moran, E. & Brondizio, E. (2004) Colonist household decisionmaking and land-use change in the amazon rainforest: an agent-based simulation. Environment and Planning B Planning and Design, 31, 693-709.
Dendoncker, N., Rounsevell, M. & Bogaert, P. (2007) Spatial analysis and modelling of land use distributions in Belgium. Computers Environment and Urban Systems, 31, 188-205.
Dieleman, F. M., Dijst, M. & Burghouwt, G. (2002) Urban form and travel behaviour: Micro-level household attributes and residential context. Urban Studies, 39, 507-527.
Dieleman, F. M., Dijst, M. J. & Spit, T. (1999) Planning the compact city: the Randstad Holland experience. European Planning Studies, 7, 605-620.
Dietzel, C. & Clarke, K. C. (2004) Replication of spatio-temporal land use patterns at three levels of aggregation by an urban cellular automata. Cellular Automata, Proceedings.
Dietzel, C. & Clarke, K. C. (2007) Toward optimal calibration of the SLEUTH land use change model. Transactions in GIS, 11, 29-45.
Dijk, v. T. (2007) Scale dependent synergy between risk management and open space preservation. AESOP 2007 International Conference 'Planning for the Risk Society'.
Dolmans, J. H. J., Hamminga, N. A. R., Jabben, J., Hoffmans, W. H. & Nijland, H. A. (2000) Background report to the fifth Environmental Outlook: Noise. 408129009. RIVM. Bilthoven, The Netherlands
Eckhardt, K., Breuer, L. & Frede, H. G. (2003) Parameter uncertainty and the significance of simulated land use change effects. Journal of Hydrology, 273, 164-176.
Engelen, G., White, R. & de Nijs, T. C. M. (2003) The Environment Explorer: Spatial Support system for Integrated Assessment of Socio-economic and Environmental Policies in the Netherlands. Integrated Assessment, 4, 97-105.
Engelen, G., White, R., Uljee, I. & Drazan, P. (1995) Using cellular automata for integrated modelling of socio-environmental systems. Environmental Monitoring and Assessment, 34, 203-214.
Engelen, G., White, R., Uljee, I. & Wargnies, S. (1996) Numerical Modeling of Small Island Socio-Economics to Achieve Sustainable Development. IN MAUL, G. A. (Ed.) Small Islands: Marine Science and Sustainable Development. Washington, American Geophysical Union.
EU (1979) Council Directive 79/409/EEC of 2 April 1979 on the conservation of wild birds
EU (1992) Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora.
Faludi, A. & Van der Valk, A. J. (1990) De groeikernen als hoekstenen van de Nederlandse ruimtelijke planningsdoctrine, Assen/Maastricht, Van Gorcum.
Faludi, A. & Van der Valk, A. J. (1994) Rule and Order: Dutch Planning Doctrine in the Twentieth Century, Dordrecht, Kluwer Academic Publishers.
Fang, S. F., Gertner, G., Wang, G. X. & Anderson, A. (2006) The impact of misclassification in land use maps in the prediction of landscape dynamics. Landscape Ecology, 21, 233-242.
Fang, S. F., Gertner, G. Z., Sun, Z. L. & Anderson, A. A. (2005) The impact of interactions in spatial simulation of the dynamics of urban sprawl. Landscape and Urban Planning, 73, 294-306.
Faust, K., Entwisle, B., Rindfuss, R. R., Walsh, S. J. & Sawangdee, Y. (1999) Spatial arrangements of social and economic networks among villages in Nang Rong District, Thailand. Social Networks, 21, 311-337.
Fischer, G., Velthuizen, H., Shah, M. & Nachtergaele, F. (2002) Global Agro-ecological Assessment for Agriculture in the 21st century: Methodology and Results, Laxenburg, Austria, IIASA and FAO.
Fisher, P. (2000) Sorites paradox and vague geographies. Fuzzy Sets and Systems, 113, 7-18.
Foody, G. M. (2002) Status of land cover classification accuracy assessment. Remote Sensing of Environment, 80, 185-201.

Foody, G. M. (2004) Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. Photogrammetric Engineering and Remote Sensing, 70, 627-633.

Fujita, M., Krugman, P. & Mori, T. (1999) On te evolution of hierarchical urban systems. European Economic Review, 43, 209-251.

Gabaix, X. (1999) Zipf's Law for Cities: An Explanation. Quarterly Journal of Economics, 114, 739-67.

Gabaix, X. & Ioannides, Y. M. (2003) The Evolution of City Size Distributions. IN THISSE, H. A. (Ed.) Handbook of Urban and Regional Economics. Volume IV: Cities and Geography. Amsterdam, North-Holland Publishing Company.

Geoghegan, J., Villar, S. C., Klepeis, P., Mendoza, P. M., Ogneva Himmelberger, Y., Chowdhury, R. R., Turner, B. L. & Vance, C. (2001) Modeling tropical deforestation in the southern yucatan peninsular region: comparing survey and satellite data. Agriculture Ecosystems and Environment, 85, 25-46.

Geoghegan, J., Wainger, L. A. & Bockstael, N. E. (1997) Spatial landscape indices in a hedonic framework: an ecological economics analysis using GIS. Ecological Economics 23, 251-264.

Geurs, K., Hagen, A., Hoen, A. & van Wee, B. (2003) Ex post evaluation of Dutch spatial planning and infrastructure policies. European Transport Conference. Strasbourg, France.

Goetgeluk, R. W., Louter, P. J., Beurden, J. A. M. B.-v., Waals, J. F. M. v. d. & Geurs, K. T. (2000) Wonen en werken ruimtelijk verkend. Waar wonen en werken we in 2020 volgens een compacte inrichtingsvariant voor de Vijfde Nota Ruimtelijke Ordening? RIVM-rapport 711931001. RIVM. Bilthoven.

Gottmann, J. & Harper, R. A. (1967) Metropolis on the Move, John Wiley & Sons Inc

Hagen-Zanker, A. (2008) Measuring the performance of geosimulation models by map comparison. Maastricht, Univ. of Maastricht.

Hagen-Zanker, A. & Lajoie, G. (2008) Neutral models of landscape change as benchmarks in the assessment of model performance. Landscape and Urban Planning, 86, 284-296.

Hagen-Zanker, A., Straatman, B. & Uljee, I. (2005a) Further developments of a fuzzy set map comparison approach. International Journal of Geographical Information Science 19, 769-785.

Hagen-Zanker, A., van Loon, J., Maas, A., Straatman, B., de Nijs, T. C. M. & Engelen, G. (2005b) Measuring Performance of Land Use Models. An Evaluation Framework for the Calibration and Validation of Integrated Land Use Models Featuring Cellular Automata. European Colloquium on Theoretical and Quantitative Geography. Tomar, Portugal.

Hagen, A. (2002a) Approaching human judgement in the automated comparison of categorical maps. 4th International Conference on Recent Advances in Soft Computing. Nottingham Trent Unversity, United Kingdom.

Hagen, A. (2002b) Multi-method assessment of map similarity. IN RUIZ, M., GOULD, M. & RAMON, J. (Eds.) 5th AGILE Conference on Geographic Information Science. Palma, Spain.

Hagen, A. (2003) Fuzzy set approach to assessing similarity of categorical maps. International Journal for Geographical Information Science, 17, 235-249.

Hagen, A. & Uljee, I. (2003) Map Comparison Kit. User manual. RIKS. Maastricht

Hagoort, M., Geertman, S. & Ottens, H. (2008) Spatial externalities, neighbourhood rules and CA land-use modelling. Annals of Regional Science, 42, 39-56.

Hagoort, M. J. (2006) The neighborhood rules: Land-use interactions, urban dynamics, and cellular automata modelling. Utrecht, Utrecht.

Hastie, T., Tibshirani, R. & Friedman, J. (2001) The Elements of Statistical Learning Data Mining, Inference, and Prediction.

HCG (1997) LMS 5.0 Deel 1 – 7. Hague Consulting Group. The Hague, Netherlands

He, H. S., Ventura, S. J. & Mladenoff, D. J. (2002) Effects of spatial aggregation approaches on classified satellite imagery. Int.J.Geographical Information Science, 16, 93-109.
Helling, A. (1998) Collaborative Visioning: Proceed with Caution! Results from Evaluating Atlanta's Vision 2020 Project. Journal of the American Planning Association, 64, 335-349.
Hohenberg, P. M. & Lees, L. H. (1985) The making of urban Europe 1000-1950, Cambridge (Mass.) and London, Harvard University Press.
Hoshino, S. (1996) Statistical analysis of land-use change and driving forces in the Kansai District, Japan. WP-96-120. IIASA. Laxenburg
Huigen, M. G. A. (2004) First principles of the mameluke multi-actor modelling framework for land use change, illustrated with a philippine case study. Journal of Environmental Management, 72, 5-21.
Huigen, M. G. A., Overmars, K. P. & de Groot, W. T. (2006) Multiactor modeling of settling decisions and behavior in the San Mariano watershed, the Philippines: a first application with the MameLuke framework – art. no. 33. Ecology and Society, 11, NIL_537-NIL_569.
IPCC (2000) Special Report on Emissions Scenarios, Cambridge, Cambridge University Press.
Irwin, E. G. & Geoghegan, J. (2001) Theory, data, methods: developing spatially explicit economic models of land use change. Agriculture Ecosystems and Environment, 85, 7-23.
Jat, M. K., Garg, P. K. & Khare, D. (2008) Monitoring and modelling of urban sprawl using remote sensing and GIS techniques. International Journal of Applied Earth Observation and Geoinformation, 10, 26-43.
Jenerette, G. D. & Wu, J. G. (2001) Analysis and simulation of land-use change in the central arizona-phoenix region, usa. Landscape Ecology, 16, 611-626.
Karssenberg, D., Burrough, P. A., Sluiter, R. & de Jong, K. (2001) The PCRaster Software and Course Materials for Teaching Numerical Modelling in the Environmental Sciences. Transactions in GIS, 5, 99-110.
Kasperson, J. X. & Kasperson, R. E. (2001) A Workshop Summary prepared on behalf of workshop participants.. International Workshop on Vulnerability and Global Environmental Change. 17-19 May 2001 Stockholm, Sweden, Stockholm Environment Institute.
Kasperson, J. X., Kasperson, R. E. & Turner II, B. L. (1995) Regions at Risk: Comparisons of Threatened Environments, Tokyo, United Nations University,.
Kempen, R. v., Hooimeijer, P., Atzema, O., Dijst, M. & Verhoeff, R. (2000) Sociale en culturele ontwikkelingen en ruimtelijke configuraties (in Dutch), Assen, Van Gorcum.
Kitamura, T., Kagatsume, M., Hoshino, S. & Morita, H. (1997) A Theoretical Consideration on the Land Use Change Model for the Japan Case Study Area. Interim Report IR-97-064. International Institute for Applied Systems Analysis. Laxenburg, Austria
Kocabas, V. & Dragicevic, S. (2006) Coupling Bayesian networks with GIS-based cellular automata for modeling land use change. Geographic Information Science, Proceedings.
Kokla, M. & Kavouras, M. (2001) Fusion of top-level and geographic domain ontologies based on context formation and complementarity. International Journal of Geographical Information Science, 15, 679-687.
Krugman, P. (1998) Space: The Final Frontier. The Journal of Economic Perspectives 12, 161-174.
Krugman, P. (1999) The Role of Geography in Development. International Regional Science Review, 22, 142-161.
Kuhnert, M. (2002) Map comparison survey. Retrieved form the Internet at http://www.uvm.edu/giee/AV/Maps/ (June 26, 2002).
Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., Coomes, O. T., Dirzo, R., Fischer, G., Folke, C., George, P. S., Homewood, K., Imbernon, J., Leemans, R., Li, X., Moran, E. F., Mortimore, M., Ramakrishnan, P. S., Richards, J. F., Skånes, H., Steffen, W., Stone, G. D., Svedin, U., Veldkamp, T. A., Vogel, C. & Xu, J. (2001a) The causes of land-use and land-cover change: moving beyond the myths. Global Environmental Change, 11, 261-269.

Lambin, E. F., Turner II, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., Coomes, O., Dirzo, R., Fischer, G., Folke, C., George, P. S., Homewood, K., Imbernon, J., Leemans, R., Li, X. B., Moran, E. F., Mortimore, M., Ramakrishnan, P. S., Richards, J. F., Skanes, H., Stone, G. D., Svedin, U., Veldkamp, A., Vogel, C. & Xu, J. C. (2001b) The Causes of Land-Use and Land-Cover Change: Moving Beyond the Myths. Global Environmental Change, 4.

Landis, J. & Zhang, M. (1998a) The second generation of the California urban futures model. Part 1: Model logic and theory. Environment and Planning B: Planning and Design, 25, 657-666.

Landis, J. & Zhang, M. (1998b) The second generation of the California urban futures model. Part 2: Specification and calibration results of the land-use change submodel. Environment and Planning B: Planning and Design, 25, 795-824.

Landis, J. D. (1994) The Californian Urban Futures Model: A New Generation of Metropolitan Simulation Models. Environment and Planning B: Planning and Design, 21, 399 – 420.

Lee, D. B. (1973) Requiem for Large-Scale Models. Journal of the American Institute of Planners, 39, 163-178.

LEI-DLO (1997) Kansen voor Kassen. LEI-DLO, The Hague.

Li, H. & Reynolds, J. F. (1997) Modeling effects of spatial pattern, drought, and grazing on rates of rangeland degradation: A combined Markov and cellular automaton approach. IN QUATTROCHI, D. A. & GOODCHILD, M. F. (Eds.) Scale in Remote Sensing and GIS. New York, Lewis Publishers.

Li, X. & Yeh, A. G.-O. (2001) Calibration of cellular automata by using neural networks for the simulation of complex urban systems. Environment & Planning A, 33, 1445.

Li, X. & Yeh, A. G. (2002) Neural-network-based cellular automata for simulating multiple land use changes using gis. International Journal of Geographical Information Science, 16, 323-343.

Li, X. & Yeh, A. G. O. (2000) Modelling sustainable urban development by the integration of constrained cellular automata and GIS. International Journal of Geographical Information Science, 14, 131-152.

Ligtenberg, A., Bregt, A. K. & Van Lammeren, R. (2001) Multi-actor-based land use modelling: spatial planning using agents. Landscape and Urban Planning, 56, 21-33.

Ligtenberg, A., Wachowicz, M., Bregt, A. K., Beulens, A. & Kettenis, D. L. (2004) A design and application of a multi-agent system for simulation of multi-actor spatial planning. Journal of Environmental Management, 72, 43-55.

Liu, X. P., Li, X. & Yeh, A. G. O. (2006) Multi-agent systems for simulating spatial decision behaviors and land-use dynamics. Science in China Series D Earth Sciences, 49, 1184-1194.

LNV (2002) Structuurschema Groene Ruimte 2. Samen werken aan groen Nederland. Ministerie van Landbouw, Natuurbeheer en Visserij. The Hague, The Netherlands

Luttik, J., Alkemade, J. R. M., van de Berg, L., Berkhout, P., van der Hamsvoort, C., Egmond, P. v., Kuindersma, W., Luijt, J. & Timmermans, W. (2002) Trends en scenario's voor de Nationale Natuurverkenning 2. Natuurplanbureau. Vestiging Wageningen. Wageningen, The Netherlands

Manson, S. M. (2005) Agent-based modeling and genetic programming for modeling land change in the southern yucatan peninsular region of mexico. Agriculture Ecosystems and Environment, 111, 47-62.

Maxwell, W. E. (1977) Coefficients of agreement between observers and their interpretations. British Journal of Psychiatry, 130, 79-83.

Menard, A. & Marceau, D. J. (2005) Exploration of spatial scale sensitivity in geographic cellular automata. Environment and Planning B Planning and Design, 32, 693-714.

Menard, S. (1995) Applied Logistic Regression Analysis, London, Sage.

Mertens, B., Poccard Chapuis, R., Piketty, M. G., Lacques, A. E. & Venturieri, A. (2002) Crossing spatial analyses and livestock economics to understand deforestation processes in the brazilian amazon: the case of sao felix do xingu in south para. Agricultural Economics, 27, 269-294.

Messina, J. P. & Walsh, S. J. (2001) 2.5d Morphogenesis: modeling landuse and landcover dynamics in the ecuadorian amazon. Plant Ecology, 156, 75-88.
Metternicht, G. (1999) Change detection assessment using fuzzy sets and remotely sensed data: an application of topographic map revision. ISPRS Journal of Photogrammetry and Remote Sensing, 54, 221-233.
Milne, B. T. & Johnson, A. R. (1993) Renormalization relations for scale transformation in ecology. Lectures on Mathematics in the Life Sciences, 23, 109-128.
MNP (2001) Who's afraid of red green and blue? Toets van de Vijfde Nota Ruimtelijke Ordening op ecologische effecten. 711931005. RIVM. Bilthoven, The Netherlands
MNP (2002a) Natuurbalans 2002, Alphen aan de Rijn, The Netherlands, Kluwer.
MNP (2002b) Natuurverkenning 2, Alphen aan den Rijn, The Netherlands, Kluwer.
MNP (2002c) Natuurverkenning 2. RIVM. Bilthoven, The Netherlands
MNP (2005) Kalibratie en validatie van de LeefOmgevingsVerkenner. Rapportnr. 550016006. MNP. Bilthoven, The Netherlands
Moineddin, R., Beyene, J. & Boyle, E. (2003) On the Location Quotient Confidence Interval. Geographical Analysis, 35, 249-256.
Monserud, R. A. & Leemans, R. (1992) The Comparison of Global Vegetation Maps. Ecol. Modell, 62, 275-293.
Moody, A. & Woodcock, C. E. (1994) Scale-dependent errors in the estimation of land-cover proportions: implications for global land-cover data sets. Photogrammetric Engineering & Remote Sensing, 60, 585-594.
Morita, H., Hoshino, S., Kagatsume, M. & Mizuno, K. (1997) An Application of the Land Use Change Model for the Japan Case Study Area. Interim Report IR-97-065. International Institute for Applied Systems Analysis. Laxenburg, Austria
Munroe, D. K., Southworth, J. & Tucker, C. M. (2002) The dynamics of land-cover change in western Honduras: exploring spatial and temporal complexity. Agricultural Economics, 27, 355-369.
Myers, D. & Kitsuse, A. (2000) Constructing the future in planning: a survey of theories and tools. Journal of Planning Education and Research 19, 221-231.
Nelson, G. C. & Geoghegan, J. (2002) Deforestation and land use change: sparse data environments. Agricultural Economics, 27, 201-216.
Nelson, G. C., Harris, V. & Stone, S. W. (2001) Deforestation, Land Use, and Property Rights: Empirical Evidence from Darien, Panama. Land Economics, 77, 187-205.
Nelson, G. C. & Hellerstein, D. (1997) Do roads cause deforestation? Using satellite images in econometric analysis of land use. 79(12): 80-88. American Journal of Agricultural Economics, 79, 80-88.
Nemmour, H. & Chibani, Y. (2006) Multiple support vector machines for land cover change detection: An application for mapping urban extensions. Isprs Journal of Photogrammetry and Remote Sensing, 61, 125-133.
NIROV (2004) De nieuwe Kaart van Nederland... NIROV. The Hague, Netherlands
NRC (1999) Dispuut over milieucijfers, RIVM laakt kritiek van medewerker' NRC Handelsblad.
O'Sullivan, D. (2001) Graph-cellular automata: a generalised discrete urban and regional model. Environment and Planning B, 28, 687-705.
Odum, H. T. (1983) Systems Ecology, New York, John Wiley.
Otter, H. S., Van Der Veen, A. & De Vriend, H. J. (2001) Abloom: location behaviour, spatial patterns, and agent-based modelling. Jasss the Journal of Artificial Societies and Social Simulation, 4, Nil_28-Nil_54.
Overmars, K. P., De Koning, G. H. J. & Veldkamp, A. (2003) Spatial autocorrelation in multi-scale land use models. Ecological Modelling, 164, 257-270.
Pahl, R. E. (1975) Whose city and further essays on urban society, Harmondsworth, Penguin Books.

Parker, D. C., Manson, S. M., Janssen, M. A., Hoffmann, M. J. & Deadman, P. (2003) 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.
Penning de Vries, F., Teng, P. & Metsalaar, K. (1992) Systems approaches for agricultural development, London, Kluwer Academic Publishers.
Pfaff, A. S. P. (1999) What Drives Deforestation in the Brazilian Amazon? Evidence from Satellite and Socioeconomic Data:. Journal of Environmental Economics and Management, 37, 25-43.
Pijanowski, B. C., Brown, D. G., Shellito, B. A. & Manik, G. A. (2002) Using neural networks and GIS to forecast land use changes: a Land Transformation Model. Computers, environment and urban systems, 26, 553-575.
Pijanowski, B. C., Gage, S. H., Long, D. T. & Cooper, W. C. (2000) A land transformation model for the Saginaw Bay Watershed., Lewis Publishing.
Pijanowski, B. C., Pithadia, S., Shellito, B. A. & Alexandridis, K. (2005) Calibrating a neural network-based urban change model for two metropolitan areas of the upper midwest of the united states. International Journal of Geographical Information Science, 19, 197-215.
Pontius, G. R., Malanson, J. & Wang, H. (2003) Methods to compare the accuracy of different types of land-use change models. Proceedings. Framing Land-use Dynamics. Utrecht University, The Netherlands.
Pontius Jr, R. G., Boersma, W., Castella, J. C., Clarke, K., Nijs, T., Dietzel, C., Duan, Z., Fotsing, E., Goldstein, N., Kok, K., Koomen, E., Lippitt, C. D., McConnell, W., Mohd Sood, A., Pijanowski, B., Pithadia, S., Sweeney, S., Trung, T. N., Veldkamp, A. T. & Verburg, P. H. (2008) Comparing the input, output, and validation maps for several models of land change. Annals of Regional Science, 42, 11-37.
Pontius, R. G. (2000) Quantification error versus location error in comparison of categorical maps. Photogrammetric Engineering and Remote Sensing, 66, 1011-1016.
Pontius, R. G. (2002) Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. Photogrammetric Engineering and Remote Sensing, 68, 1041-1049.
Pontius, R. G., Huffaker, D. & Denman, K. (2004a) Useful techniques of validation for spatially explicit land-change models. Ecological Modelling, 179, 445-461.
Pontius, R. G. & Schneider, L. C. (2001) Land-cover change model validation by an roc method for the ipswich watershed, massachusetts, usa. Agriculture Ecosystems and Environment, 85, 239-248.
Pontius, R. G., Shusas, E. & Mceachern, M. (2004b) Detecting important categorical land changes while accounting for persistence. Agriculture Ecosystems and Environment, 101, 251-268.
Pontius, R. G. & Spencer, J. (2005) Uncertainty in extrapolations of predictive land-change models. Environment and Planning B Planning and Design, 32, 211-230.
Power, C., Simms, A. & White, R. (2001) Hierarchical fuzzy pattern matching for the regional comparison of land use maps. International Journal for Geographical Information Science, 15, 77-100.
Priemus, H. (2005) Wat staat ons allemaal te wachten? Revolutionaire toekomst voor de woningbouw. Building Bussiness.
Rabin, M. (1998) Psychology and Economics. Journal of Economic Literature, 36, 11-46.
Raziei, A. & Evers, W. (2001) Afstemming ruimtelijke gegevens voor de L(eef)O(mgevings)V(erkenner) en R(uimte)S(canner) (in dutch). Intern CIM-report M009/01. National Institute of Public Health and the Environment. Bilthoven
Refsgaard, J. C. & Henriksen, H. J. (2004) Modelling guidelines – terminology and guiding principles. Advances in Water Resources, 27, 71-82.
Ricardo, D. (1817) Principles of Political Economy and Taxation.
RIKS (2002) Verkeer in de LeefOmgevingsVerkenner: Ontwikkeling en bouw van een verkeersmodule in de LOV.. RIKS, Maastricht, The Netherlands.

RIKS (2004) Kalibratie van de LeefOmgevingsVerkenner (Calibration of the " Environment Explorer", a dynamic land use model). RIKS rapport 179708. Research Institute for Knowledge Systems. Maastricht
RIVM (2002) Natuurverkenning 2. Rapport No. 9014093292; ISBN 9014093292. RIVM. Bilthoven, The Netherlands
RIVM (2004) The environmental consequences of the National Policy Document on Spatial Planning. RIVM Rapport 711931009 RIVM. Bilthoven
Robinson, W. (1950) Ecological correlations and the behaviour of individuals. American Sociological Review, 15, 351-357.
Rodriguez, M. A., Egenhofer, M. J. & Rugg, R. D. (1999) Assessing semantic similarities among geospatial feature class definitions. IN VCKOVSKI, A., BRASSEL, K. E. & SCHEK, H.-J. (Eds.) Interoperating Geographic Information Systems. Lecture Notes in Computer Science,1580. Berlin, Springer.
RPD (1960) Nota inzake de ruimtelijke ordening in Nederland (Memorandum on spatial planning in the Netherlands). The Hague, Netherlands
RPD (1966) Tweede nota over de ruimtelijke ordening in Nederland (Second memorandum on spatial planning in the Netherlands). The Hague, Netherlands
RPD (1977) Derde nota over de ruimtelijke ordening. Deel 2: Verstedelijkingsnota. (Third memorandium on spatial planning in the Netherlands. Part 2: Memorandum on urbanization). The Hague, Netherlands
RPD (1988) Vierde nota over de ruimtelijke ordening. Eenheid in verscheidenheid (Fourth memorandium on spatial planning. Uniformity in diversity). The Hague, Netherlands
RPD (1994) Vierde nota over de Ruimtelijke Ordening Extra (Fourth memorandum on spatial planning extra). The Hague, Netherlands
SafeCoast (2005) Inventory of spatial and economical scenarios available in the North Sea region. WL | Delft Hydraulics & Grontmij.
SC-DLO (1999) Vervaardiging en nauwkeurigheid van het LGN3-grondgebruiksbestand. Report no. 663. DLO-Staring Centrum. Wageningen, The Netherlands
Schneider, L. C. & Pontius, R. G. (2001) Modeling land-use change in the ipswich watershed, massachusetts, usa. Agriculture Ecosystems and Environment, 85, 83-94.
Schotten, K., Goetgeluk, R., Hilferink, M., Rietveld, P. & Scholten, H. (2001) Residential construction, land use and the environment. Simulations for the netherlands using a gis-based land use model. Environmental Modeling and Assessment, 6, 133-143.
Schröter, D., Cramer, W., Leemans, R., Prentice, I., Araújo, M., Arnell, N., Bondeau, A., Bugmann, H., Carter, T., Gracia, C., de la Vega-Leinert, A., Erhard, M., Ewert, F., Glendining, M., House, J., Kankaanpää, S., Klein, R., Lavorel, S., Lindner, M., Metzger, M., Meyer, J., Mitchell, T., Reginster, I., Rounsevell, M., Sabaté, S., Sitch, S., Smith, B., Smith, J., Smith, P., Sykes, M., Thonicke, K., Thuiller, W., Tuck, G., Zaehle, S. & Zierl, B. (2005) Ecosystem Service Supply and Vulnerability to Global Change in Europe. Science, 310, 1333-1337.
Serneels, S. & Lambin, E. F. (2001) Proximate causes of land-use change in narok district, kenya: a spatial statistical model. Agriculture Ecosystems and Environment, 85, 65-81.
Siegel, S. & Castellan, N. J. J. (1988) Nonparametric statistics for the behavioral sciences, New York, McGraw-Hill.
Silva, E. A. & Clarke, K. C. (2002) Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal. Computers, Environment and Urban Systems, 26, 525-552.
Smith, D. M. (1975) Patterns in Human Geography, Newton Abbot, David and Charles.
SNM (2005) RUIMTE – Verstedelijking gaat steeds meer ten koste van natuur, recreatie en water. website: Laatst geopend op

Soares-Filho, B. S., Cerqueira, G. C. & Pennachin, C. L. (2002) DINAMICA – a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. Ecological Modelling, 154, 217-235.
Sonneveld, M. P. W., Bouma, J. & Veldkamp, A. (2002) Refining soil survey information for a Dutch soil series using land use history. Soil Use and Management, 18, 157-163.
Soule, D. C. (2006) Urban Sprawl: A Comprehensive Reference Guide, Greenwood Press.
Steur, G. G. L. & Heijink, W. (1991) Bodemkaart van Nederland schaal 1: 50 000. Algemene begrippen en indelingen. DLO Staring Centrum. Wageningen
Straatman, B., Engelen, G. & White, R. (2004) Towards an Automatic Calibration Procedure for Constrained Cellular Automata. Computers Environment and Urban Systems, 28, 149-170.
Sui, D. Z. & Zeng, H. (2001) Modeling the dynamics of landscape structure in Asia's emerging desakota regions: a case study in Shenzhen. Landscape and Urban Planning, 53, 37-52.
Swets, J. A. (1986) Measuring the accuracy of diagnostic systems. Science, June, 1285-1293.
Thompson, D. M., Serneels, S. & Lambin, E. F. (2002) Land use strategies in the Mara ecosystem: a spatial analysis linking socio-economic data with landscape variables. IN WALSH, S. J. & CREWS-MEYER, K. A. (Eds.) Linking People, Place, and Policy. A GIScience Approach. Boston Kluwer Academic Publishers.
Thornton, P. K. & Jones, P. G. (1998) A Conceptual Approach to Dynamic Agricultural Land-Use Modelling. Agricultural Systems, 57, 505-521.
Tobler, W. R. (1979) Cellular geography.
Tomlin, C. D. (1990) Geographic Information Systems and Cartographic Modeling, Englewood Cliffs, N.J., Prentice Hall.
Torrens, P. M. (2006) Simulating sprawl. Annals of the Association of American Geographers, 96, 248-275.
Torrens, P. M. & O'Sullivan, D. (2001) Editorial. Cellular automata and urban simulation: where do we go from here? Environment and Planning B, 28, 163-168.
Tranmer, M. & Steel, D. G. (1998) Using census data to investigate the causes of the ecological fallacy. Environment and Planning A, 30, 817-831.
Turk, G. (2002) Map evaluation and "chance correction. Photogrammetric Engineering and Remote Sensing, 68, 123-125.
Turner, B. L., Lambin, E. F. & Reenberg, A. (2007) The emergence of land change science for global environmental change and sustainability. Proceedings of the National Academy of Sciences of the United States of America (PNAS), 104, 20666-20671.
Turner II, B. L., Kasperson, R. E., Meyer, W. B., Dow, K. M., Golding, D., Kasperson, J. X., Mitchel, R. C. & Ratick, S. J. (1990) Two types of global environmental change; definitional and spatial-scale issues in their human dimensions. Global Environmental Change, 1, 14-22.
Turner II, B. L., Ross,R.H. & Skole,D.L. (1993) Relating land use and global land cover change. HDP report no. 5. IGDP report no. 24.
Tyson, P., Steffen, W., Mitra, A. P., Fu, C. & Lebel, L. (2001) The earth system: regional-global linkages. Regional Environmental Change, 2, 128-140.
Uljee, I. (2003) Map Comparison Module. Reference manual. RIKS. Maastricht, the Netherlands
Vafeidis, A. T., Koukoulas, S., Gatsis, I. & Gkoltsiou, K. (2008) Forecasting land-use changes with the use of neural networks and GIS. International Geoscience and Remote Sensing Symposium (IGARSS). Barcelona.
van Delden, H., van Vught, J. & van Strien, C. (2003) De toekomst aangekaart. Een integrale, ambtelijke zoektocht met de LeefOmgevingsVerkenner naar woon- en werklocaties in het kader van het Interreg-IIc project REGIS, Case: streekplan Utrecht 2005-2015. Final Report. Province of Utrecht, the Netherlands.

van der Sluijs, J. P., Risbey, J. S., Kloprogge, P., Ravetz, J. R., Funtowicz, S. O., Quintana, S. C., Guimar˜aes Pereira, A., De Marchi, B., Petersen, A. C., Janssen, P. H. M., Hoppe, R. & Huijs, S. W. F. (2003) RIVM/MNP Guidance for Uncertainty Assessment and Communication. Detailed Guidance, Utrecht, The Netherlands, Copernicus Institute for Sustainable Development and Innovation. Utrecht University.

van der Vlist, A. J., Gorter, C., Nijkamp, P. & Rietveld, P. (2002) Residential mobility and local housing-market differences. Environment and Planning A, 34, 1147-1164.

van Loon, J. (2004) Towards model-specific automated calibration procedures for the LeefOmgevingsVerkenner model. Unpublished Master's. University Maastricht. Maastricht

van Waveren, R. H., Groot, S., Scholten, H., van Wösten, J. H. M., Koeze, R. D. & BNoort, J. J. (1999) Vloeiend modelleren in het waterbeheer. Handbook Good Modelling Practice. 99-05. RIZA/STOWA. Lelystad, The Netherlands

Veldkamp, A. & Fresco, L. O. (1996) CLUE-CR: an integrated multi-scale model to simulate land use change scenarios in Costa Rica. Ecological Modelling, 91.

Veldkamp, A. & Fresco, L. O. (1997) Reconstructing land use drivers and their spatial scale dependence for Costa Rica. Agricultural Systems, 55, 19-43.

Venables, A. J. (1999) But Why Does Geography Matter and Which Geography Matters? International Regional Science Review 22, 238-241.

Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S, Springer.

VenW (2003) Meerjarenprogramma Infrastruktuur en Transport. VenW. The Hague, The Netherlands

VenW (2004) Wijziging Luchthavenverkeerbesluit Schiphol en Wijziging Luchthavenindelingbesluit Schiphol. Ministerie van Verkeer en Waterstaat – Luchtvaart. The Hague, Netherlands

VenW (2005) Planologische Kernbeslissing Ruimte voor de Rivier Deel 3, Kabinetsstandpunt. Rijkswaterstaat. The Hague, Netherlands

Verburg, P. H., de Koning, G. H. J., Kok, K., Veldkamp, A. & Bouma, J. (1999) A spatial explicit allocation procedure for modelling the pattern of land use change based upon actual land use. Ecological Modelling, 116, 45-61.

Verburg, P. H., De Koning, G. H. J., Kok, K., Veldkamp, A. & Priess, J. (2001) The clue modelling framework: an integrated model for the analysis of land use change. Land-Use-and-Cover-Change. 2001: 11-25.

Verburg, P. H., de Nijs, T. C. M., Ritsema van Eck, J., Visser, H. & Jong, K. d. (2004a) A method to analyse neighbourhood characteristics of land use patterns. Computers, Environment and Urban Systems, 28, 667-690.

Verburg, P. H., Eickhout, B. & Meijl, H. (2008) A multi-scale, multi-model approach for analyzing the future dynamics of European land use. Annals of Regional Science, 42, 57-77.

Verburg, P. H., Overmars, K. P. & Witte, N. (2004b) Accessibility and land-use patterns at the forest fringe in the northeastern part of the philippines. Geographical Journal, 170, 238-255.

Verburg, P. H., Schot, P., Dijst, M. J. & Veldkamp, A. (2004c) Land use change modelling: current practice and research priorities. Geojournal, 61, 309-324.

Verburg, P. H., Van Eck, J. R. R., De Nijs, T. C. M., Dijst, M. J. & Schot, P. (2004d) Determinants of land-use change patterns in the netherlands. Environment and Planning B Planning and Design, 31, 125-150.

Verburg, P. H. & Veldkamp, A. (2004) Projecting land use transitions at forest fringes in the philippines at two spatial scales. Landscape Ecology, 19, 77-98.

Visser, H. (2004) The Map Comparison Kit: methods, software and applications. RIVM report 550002005. RIVM. Bilthoven, the Netherlands

Visser, H. & De Nijs, A. C. M. (2006) The Map Comparison Kit. Environmental Modelling and Software, 21, 346-358.

VNG (1992) Bedrijven en Milieuzonering (Businesses and environment zoning). Vereniging van Nederlandse Gemeenten. The Hague, Netherlands
Von Neumann, J. (1966) Theory of Self-Reproducing Automata, Illinois, University of Ilinois Press.
Von Thünen, J. (1966) Der Isolierte Staat in Beziehung der Landwirtschaft und Nationalökonomie. IN HALL, P. (Ed.) Von Thünen's Isolated State. Oxford, Pergamon.
VROM (1999) Nederland in Plannen 1999. Dossier p4713-01-00. Ministerie van Volkshuisvesting, Ruimtelijke Ordening en Milieubeheer. The Hague
VROM (2001) Ruimte maken, ruimte delen. Vijfde nota over de ruimtelijke ordening 2000/2030. Ministerie van Volkshuisvesting, Ruimtelijke Ordening en Milieubeheer. The Hague
VROM, LNV, VenW & EZ (2004) Nota Ruimte. Den Haag, Netherlands
VROM, LNV, VenW & EZ (2005) Nota Ruimte (Memorandum on space). The Hague, Netherlands
Waddell, P. (2000) A behavioral simulation model for metropolitan policy analysis and planning: residential location and housing market components of urbansim. Environment and Planning B Planning and Design, 27, 247-263.
Walsh, S. J., Crawford, T. W., Crews-Meyer, K. A. & Welsh, W. F. (2001) A Multi scale Analysis of Land use Land cover change and NDVI Variation in Nang Rong District, Northeast Thailand. Agriculture, Ecosystems and Environment, 85.
Ward, D. P., Murray, A. T. & Phinn, S. R. (2000) A stochastically constrained cellular model of urban growth. Computers, Environment and Urban Systems, 24, 539-558.
Ward, D. P., Murray, A. T. & Phinn, S. R. (2003) Integrating spatial optimization and cellular automata for evaluating urban change. Annals of Regional Science, 37, 131-148.
Webster, C. & Wu, F. L. (2001) Coase, spatial pricing and self-organising cities. Urban Studies, 38, 2037-2054.
White, R. (1977) Dynamic Central Place Theory: Results of a Simulation Approach. Geographical Analysis, 9, 227-243.
White, R. (1978) The simulation of Central Place Dynamics: Two Sector Systems and the Rank-Size Distribution. Geographical Analysis, 10, 201-208.
White, R. & Engelen, G. (1993) Cellular-Automata and Fractal Urban Form – a Cellular Modeling Approach to the Evolution of Urban Land-Use Patterns. Environment and Planning A, 25, 1175-1199.
White, R. & Engelen, G. (1994) Urban Systems Dynamics and Cellular-Automata – Fractal Structures between Order and Chaos. Chaos Solitons & Fractals, 4, 563-583.
White, R. & Engelen, G. (1997a) Cellular automata as the basis of integrated dynamic regional modelling. Environment and Planning B-Planning & Design, 24, 235-246.
White, R. & Engelen, G. (1997b) Cellular automata as the basis of integrated dynamic regional modelling. Environment and Planning B: Planning and Design, 24, 235-246.
White, R. & Engelen, G. (2000) High-resolution integrated modelling of the spatial dynamics of urban and regional systems. Computers, Environment and Urban Systems, 24, 383-400.
White, R. & Engelen, G. (2003) A Calibration Procedure for Constrained Large Neighbourhood Cellular Automata based Land Use Models. 13th European Colloquium on Theoretical and Quantitative Geography. Lucca, Italy.
White, R., Engelen, G. & Uljee, I. (1997) The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics. Environment and Planning B-Planning & Design, 24, 323-343.
Wilson, A. G. (1974) Urban and Regional Models in Geography and Planning, London, John Wiley and Sons.
Winter, S. (2000) Location similarity of regions. ISPRS Journal of Photogrammetry & Remote Sensing, 55, 189-200.
Witmond, B., Vos, W., Voogt, S. & Vervoort, K. (2006) Wetlands in het IJsselmeer. Ecorys BV. Rotterdam

Wright, J. R. (1990) ISIS: Toward an Integrated Spatial Information System. IN KIM, T. J., WIGGINGS, L. L. & WRIGHT, J. R. (Eds.) Expert Systems: Applications to Urban Modelling. New-York, Springer Verlag.

Wu, F. (1998) An experiment on the generic polycentricity of urban growth in a cellular automatic city. Environment and Planning B-Planning & Design, 25, 731-752.

Wu, F. (1999) GIS-based simulation as an exploratory analysis for space-time processes. Journal of Geographical Systems, 1, 199-218.

Wu, F. & Webster, C. J. (1998) Simulation of land development through the integration of cellular automata and multicriteria evaluation. Environment and Planning B-Planning & Design, 25, 103-126.

Wu, F. L. & Webster, C. J. (2000) Simulating artificial cities in a GIS environment: urban growth under alternative regulation regimes. International Journal of Geographical Information Science, 14, 625-648.

WWdL (1958) De ontwikkeling van het westen des lands. Werkcommissie Westen de Lands van de Rijksdienst voor het Nationale Plan. The Hague, Netherlands (The development of the Western Netherlands)

Yang, Y. & Billings, S. A. (2000a) Extracting Boolean Rules from CA Patterns. IEEE Transactions on systems, man, and cybernetics – Part A: systems and humans, 30, 573-581.

Yang, Y. & Billings, S. A. (2000b) Neighborhood Detection and Rule Selection from Cellular Automata Pattern. IEEE Transactions on systems, man, and cybernetics – Part A: systems and humans, 30, 840-847.

Yeh, A. G.-O. & Li, X. (2002) A cellular automata model to simulate development density for urban planning. Environment & Planning B: Planning & Design, 29, 431.

Yeh, A. G. O. & Li, X. (2003) Simulation of development alternatives using neural networks, cellular automata, and GIS for urban planning. Photogrammetric Engineering and Remote Sensing, 69, 1043-1052.

Yuan, F. (2008) Land-cover change and environmental impact analysis in the Greater Mankato area of Minnesota using remote sensing and GIS modelling. International Journal of Remote Sensing, 29, 1169-1184.

Zadeh, L. (1965) Fuzzy sets. Information and Control, 8, 338-353.

Zaehle, S., Bondeau, A., Carter, T. R., Cramer, W., Erhard, M., Prentice, I. C., Reginster, I., Rounsevell, M. D. A., Sitch, S., Smith, B., Smith, P. C. & Sykes, M. (2007) Projected Changes in Terrestrial Carbon Storage in Europe under Climate and Land-use Change, 1990–2100. Ecosystems, 10, 380-401.

Zang, S. Y. & Huang, X. (2006) An aggregated multivariate regression land-use model and its application to land-use change processes in the daqing region (northeast china). Ecological Modelling, 193, 503-516.

Zhou, B. & Kockelman, K. M. (2008) Neighborhood impacts on land use change: A multinomial logit model of spatial relationships. Annals of Regional Science, 42, 321-340.

Zipf, G. K. (1949) Human Behaviour and the Principle of Least Effort, Cambridge, MA, Addison-Wesley.

NEDERLANDSE GEOGRAFISCHE STUDIES / NETHERLANDS GEOGRAPHICAL STUDIES

355 W BORREN Carbon exchange in Western Siberian watershed mires and implication for the greenhouse effect; A spatial temporal modeling approach -- Utrecht 2007: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 125 pp, 36 figs, 17 tabs. ISBN: 978-90-6809-396-4, Euro 13,00

356 S O NEGRO Dynamics of technological innovation systems; The case of biomass energy -- Utrecht 2007: Knag/Copernicus Institute. 166 pp, 24 figs, 17 tabs. ISBN: 978-90-6809-397-1, Euro 18,00

357 R NAHUIS The politics of innovation in public transport; Issues, settings and displacements -- Utrecht 2007: Knag/Copernicus Institute. 184 pp, 9 figs, 40 tabs, 4 box. ISBN 978-90-6809-398-8, Euro 20,00

358 M STRAATSMA Hydrodynamic roughnesses of floodplain vegetation; Airborne parameterization and field validation -- Utrecht 2007: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 180 pp, 55 figs, 28 tabs. ISBN: 978-90-6809-399-5, Euro 23,00

359 H KRUIZE On environmental equity; Exploring the distribution of environment quality among socio-economic categories in the Netherlands -- Utrecht 2007: Knag/Copernicus Institute. 219 pp, 76 figs, 25 tabs. ISBN: 978-90-6809-401-5, Euro 25,00

360 T VAN DER VALK Technology dynamics, network dynamics and partnering; The case of Dutch dedicated life sciences firms -- Utrecht 2007: Knag/Copernicus Institute. 143 pp, 23 figs, 25 tabs. ISBN: 978-90-6809-402-2, Euro 15,00

361 M A SCHOUTEN Patterns in biodiversity; Spatial organisation of biodiversity in the Netherlands -- Utrecht 2007: Knag/Copernicus Institute. 152 pp, 16 figs, 20 tabs. ISBN: 978-90-6809-403-9, Euro 20,00

362 M H J W VAN AMSTEL – VAN SAANE Twilight on self-regulation; A socio-legal evaluation of conservation and sustainable use of agrobiodiversity by industry self-regulation -- Utrecht 2007: Knag/Copernicus Institute. 167 pp, 15 figs, 13 tabs, 5 box. ISBN: 978-90-6809-404-6, Euro 20,00

363 S MUHAMMAD Future urbanization patterns in the Netherlands, under the influence of information and communication technologies -- Utrecht 2007: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 187 pp, 82 figs, 20 tabs. ISBN: 978-90-6809-405-3, Euro 20,00

364 M GOUW Alluvial architecture of the Holocene Rhine-Meuse delta (The Netherlands) and the Lower Mississippi Valley (U.S.A.) -- Utrecht 2007: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 192 pp, 55 figs, 14 tabs. ISBN: 978-90-6809-406-0, Euro 22,00

365 E HEERE & M STORMS Ormeling's cartography; Presented to Ferjan Ormeling on the occasion of his 65th birthday and his retirement as Professor of Cartography -- Utrecht 2007: Knag/Faculteit Geowetenschappen Universiteit Utrecht. ISBN: 978-90-6809-407-7, Euro 20,00

366 S QUARTEL Beachwatch; The effect of daily morphodynamics on seasonal beach evolution -- Utrecht 2007: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 125 pp, 39 figs, 7 tabs. ISBN: 978-90-6809-408-4, Euro 12,50

367 R O VAN MERKERK Intervening in emerging nanotechnologies; A CTA of Lab-on-a-chip technology regulation -- Utrecht 2007: Knag/Copernicus Institute. 206 pp, 19 box, 35 figs, 12 tabs. ISBN: 978-90-6809-409-1, Euro 20,00

368 R M FRINGS From gravel to sand; Downstream fining of bed sediments in the lower river Rhine -- Utrecht 2007: Knag/Faculteit Geowetenschappen Universiteit Utrecht. ISBN: 978-90-6809-410-7, Euro 25,00

369 W IMMERZEEL Spatial modelling of the hydrological cycle, climate change and agriculture in mountainous basins -- Utrecht 2008: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 147 pp, 54 figs, 12 tabs. ISBN: 978-90-6809-411-4, Euro 25,00

370 D S J MOURAD Patterns of nutrient transfer in lowland catchments; A case study from northeastern Europe -- Utrecht 2008: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 176 pp, 44 figs, 19tabs. ISBN: 978-90-6809-412-1, Euro 20,00

371 M M H CHAPPIN Opening the black box of environmental innovation; Governmental policy and learning in the Dutch paper and board industry -- Utrecht 2008: Knag/Copernicus Institute. 202 pp, 41 figs, 30 tabs. ISBN: 978-90-6809-413-8, Euro 22,50

372 R P ODDENS & M VAN EGMOND Ormelings atlassen; Catalogus van atlassen geschonken aan de Universiteit Utrecht door de hoogleraren F.J. Ormeling sr. en jr. – Utrecht 2008: Knag/Faculteit Geowetenschappen Universiteit Utrechts. ISBN: 978-90-6809-415-2, Euro 15,00

373 R VAN MELIK Changing public space; The recent redevelopment of Dutch city squares – Utrecht 2008: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 323 pp, 47 figs, 32 tabs. ISBN: 978-90-6809-416-9, Euro 20,00

374 E ANDRIESSE Institutions and regional development in Southeast Asia; A comparative analysis of Satun (Thailand) and Perlis (Malaysia) -- Utrecht 2008: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 250 pp, 42 figs, 43 tabs, 18 box. ISBN: 978-90-6809-417-6, Euro 25,00

375 E HEERE GIS voor historisch landschapsonderzoek; Opzet en gebruik van een historisch GIS voor prekadastrale kaarten -- Utrecht 2008: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 231 pp, 73 figs, 13 tabs. ISBN: 978-90-6809-418-3, Euro 30,00
376 V D MAMADOUH, S M DE JONG, J F C M THISSEN & J A VAN DER SCHEE Dutch windows on the Mediterranean; Dutch Geography 2004-2008 -- Utrecht 2008: Knag/International Geographical Union Section The Netherlands. 104 pp + cd-rom, 38 figs, 9 tabs. ISBN: 978-90-6809-419-0, Euro 10,00
377 E VAN MARISSING Buurten bij beleidsmakers; Stedelijke beleidsprocessen, bewonersparticipatie en sociale cohesie in vroeg-naoorlogse stadswijken in Nederland -- Utrecht 2008: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 230 pp, 6 figs, 31 tabs. ISBN: 978-90-6809-420-6, Euro 25,00
378 M DE BEER, R C L BUITING, D J VAN DRUNEN & A J T GOORTS (Eds.) Water Wegen; Op zoek naar de balans in de ruimtelijke ordening -- Utrecht 2008: Knag/VUGS/Faculteit Geowetenschappen Universiteit Utrecht. 91 pp, 18 figs, 1 tab. ISBN: 978-90-6809-421-3, Euro 10,00
379 J M SCHUURMANS Hydrological now- and forecasting; Integration of operationally available remotely sensed and forecasted hydrometeorological variables into distributed hydrological models -- Utrecht 2008: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 154 pp, 65 figs, 12 tabs. ISBN 978-90-6809-422-0, Euro 15,00
380 M VAN DEN BROECKE Ortelius' *Theatrum Orbis Terrarum* (1570-1641); Characteristics and development of a sample of *on verso* map texts -- Utrecht 2009: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 304 pp + cd-rom, 9 figs, 65 tabs. ISBN 978-90-6809-423-7, Euro 30,00
381 J VAN DER KWAST Quantification of top soil moisture patterns; Evaluation of field methods, process-based modelling, remote sensing and an integrated approach -- Utrecht 2009: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 313 pp, 108 figs, 47 tabs. ISBN 978-90-6809-424-4, Euro 30,00
382 T J ZANEN 'Actie, actie, actie…'; De vakbeweging in regio Noord-Nederland, 1960-1992 -- Utrecht/ Groningen 2009: Knag/Faculteit Ruimtelijke Wetenschappen Rijksuniversiteit Groningen. ISBN 978-90-6809-425-1, Euro 30,00
383 M PERMENTIER Reputation, neighbourhoods and behaviour -- Utrecht 2009: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 146 pp, 10 figs, 19 tabs. ISBN 978-90-6809-426-8, Euro 15,00
384 A VISSER Trends in groundwater quality in relation to groundwater age -- Utrecht 2009: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 188 pp, 47 figs, 24 tabs. ISBN 978-90-6809-427-5, Euro, 20,00
385 B A FURTADO Modeling social heterogeneity, neighborhoods and local influences on urban real estate prices; Spatial dynamic analyses in the Belo Horizonte metropolitan area, Brazil -- Utrecht 2009: Knag/Faculteit Geowetenschappen Universiteit Utrecht. 236 pp, 50 figs, 48 tabs. ISBN 978-90-6809-428-2, Euro 25,00
386 T DE NIJS Modelling land use change; Improving the prediction of future land use patterns -- Utrecht 2009: Knag/Faculteit Geowetenschappen Universiteit Utrecht. ISBN 978-90-6809-429-9, Euro 25,00