

Thesis submitted in fulfilment of the requirements for the
award of the degree of Doctor in Sciences

INTEGRATING NETWORK DISTANCES INTO AN ACTIVITY BASED CELLULAR AUTOMATA LAND-USE MODEL

**Semi-automated calibration
and application to Flanders, Belgium**

TOMAS CROLS
Academic year 2016-2017

Promoters: Prof. Dr. Frank Canters
Prof. Dr. Roger White
Co-promoter: Dr. Lien Poelmans

Sciences & Bio-Engineering Sciences

This PhD project has resulted from a scientific collaboration between VUB and VITO, for which the grant was financed by VITO (the Flemish Institute for Technological Research, Mol, Belgium).

Members of the exam committee

Chair

Prof. Dr. Matthieu Kervyn

Vrije Universiteit Brussel

Secretary

Prof. Dr. Kobe Boussauw

Vrije Universiteit Brussel

Promoters

Prof. Dr. Frank Canters

Vrije Universiteit Brussel

Prof. Dr. Roger White

Memorial University of Newfoundland

Co-promoter

Dr. Lien Poelmans

VITO – Flemish Institute for Technological Research

Jury members

Prof. Dr. David Bassens

Vrije Universiteit Brussel

Prof. Dr. Cathy Macharis

Vrije Universiteit Brussel

Prof. Dr. Anton Van Rompaey

Katholieke Universiteit Leuven

Prof. Dr. Nuno Pinto

University of Manchester

Summary

A strong population growth, extended transportation networks and a lack of structured spatial planning have all caused urban sprawl in a large number of regions worldwide. The region of Flanders, Belgium, experienced strong urban sprawl, with until recently 6 ha of open space per day being transformed into non-natural land uses. Spatial modelling of these land-use changes can support the government in tackling the problems related to decreasing open space and in controlling the urban sprawl phenomenon.

Land-use changes in urban areas are often simulated with cellular automata (CA) models. CA models allow to explicitly handle spatial interactions between different land-use categories. The core component of the CA transition rules is the set of neighbourhood rules that capture the influence of nearby land uses on the potential of a cell changing from one state (land use) into another. Due to the complex spatial structure and the strong mixing of functions in a region like Flanders though, the different types of human activity and their intensity cannot be fully represented by models relying on change in dominant land-use categories (residential, industrial,...) only.

Recently, White, Uljee and Engelen (2012) proposed a more straightforward, activity-based model of urban dynamics, which directly models spatial changes in the density of different activities (population and employment in several economic sectors). The activity-based CA model (or ACA model) is an interesting alternative to handle mixed and multifunctional land use. The model uses a variable grid approach to deal

with long-distance spatial interactions without the need to couple the model to a regional component.

The main goal of this research was to further enhance the ACA model proposed by White et al. (2012) and to apply it to simulate the impact of alternative policy scenarios on the spatial development of a region, with Flanders as the key application. Initial research revealed that the usage of Euclidean distances in the CA rules was not well representing long-distance effects. Therefore, long Euclidean distances were replaced by travel times computed along a transportation network. As a consequence, the model can now also be used to test the impact of different transport scenarios on future land uses and activities. Next, changes were proposed to the model equations that determine the activity density of the cells, in order to better deal with the densification of Belgian cities in the model.

A historical calibration is necessary to thoroughly evaluate the model and its modifications. A second objective of the project was therefore to develop a method that allows reconstructing land-use and population maps of the past from available data. Input maps (land use and population) for the past, required for model calibration, were obtained by combining detailed land-use and population maps of the present with sealed surface cover data for different moments in the past, available from remote sensing. The resulting time series of population maps (1986, 2001, 2013) provides a good insight into how urban sprawl in Flanders evolved over the last decades. While in general sprawl seems to be decreasing and densification of urban centres starts to set in, built-up area growth in suburban areas and in some rural regions is still strong.

A third objective of the project was to develop a semi-automated framework for model calibration. CA models are often calibrated manually

because incorporating expert-based knowledge leads to good results, but this is a slow process. The new approach proposed in this research iteratively applies an automated optimisation routine (a genetic algorithm), in which the modeller can use his/her knowledge to determine the range of possible parameter values and to give more or less weight to specific land-use and activity categories in the calibration.

The final objective of the research was to apply the enhanced ACA model for Flanders to simulate the impact of a land-take neutral scenario, reducing pressure on open space of the future development of built-up area in the region. The scenario was defined in line with the recent white paper of the Spatial Policy Plan for Flanders, in which the built-up area only increases in locations that have a good public transport accessibility and a sufficiently wide variety of retail and public services in their neighbourhood. Results of the application show that the land-take neutral scenario preserves much better the remaining clusters of open space in Flanders than a business-as-usual scenario. Yet, it turns out that unprotected urban green space might be threatened by a densification of urban cores. Hence, policy makers should reflect about alternative densification strategies that would not or less affect urban green areas.

Samenvatting

Een combinatie van een sterke bevolkingsgroei, toegenomen transportmogelijkheden en een te weinig gestructureerde ruimtelijke planning heeft in heel wat regio's wereldwijd voor een ongeziene groei van de bebouwde oppervlakte gezorgd. Dit proces, beter bekend als “urban sprawl”, speelt ook sterk in Vlaanderen waar tot voor kort elke dag 6 ha open ruimte getransformeerd werd in niet-natuurlijke landgebruiken. Ruimtelijk modelleren van deze landgebruikveranderingen kan het beleid helpen om problemen gekoppeld aan de verdere afname van de open ruimte aan te pakken en naar oplossingen te zoeken om het probleem van “urban sprawl” in te dijken.

Voor de modellering van landgebruikveranderingen in verstedelijkte gebieden wordt vaak gebruikt gemaakt van Cellular Automata (CA) modellen. Deze laten toe om de ruimtelijke interactie tussen verschillende types van landgebruik expliciet te beschrijven. Ze doen dat ondermeer aan de hand van *distance decay* functies als onderdeel van de transitieregels waarop de ruimtelijke allocatie zich baseert. In gebieden met een complexe ruimtelijke structuur en een sterke vermenging van functies, zoals Vlaanderen, wordt de toepassing van CA-modellen op basis van landgebruikclassen echter bemoeilijkt door verschillen in de aard en de intensiteit van diverse vormen van menselijke activiteit die binnen eenzelfde type van dominant landgebruik (residentieel, industrieel,...) voorkomen.

Recent stelden White, Uljee en Engelen (2012) een meer directe, op activiteitsgraad gesteunde modellering van stedelijke dynamiek voor, waarbij veranderingen in de intensiteit van diverse activiteiten

(bevolkingsdichtheid, tewerkstelling in belangrijke economische sectoren) ruimtelijk gemodelleerd worden. Het concept van activiteitsgebaseerde CA-modellering (ACA) biedt een interessant alternatief voor het omgaan met gemengd en multifunctioneel landgebruik. Het nieuwe ACA-model maakt ook gebruik van een variabele gridaanpak, zodat ruimtelijke interacties op grotere afstand in rekening gebracht kunnen worden en het niet meer nodig is te linken met een regionaal model.

De belangrijkste doelstelling van dit onderzoek bestond erin dit ACA-landgebruikmodel verder te ontwikkelen, zodat het toegepast kan worden om de impact van alternatieve beleidsscenario's op de ruimtelijke ontwikkeling van regio's te voorspellen, met Vlaanderen als belangrijkste toepassing. Uit initieel onderzoek bleek dat het gebruik van Euclidische afstanden in de CA-afstandsregels niet volstaat om lange-afstandsrelaties rechtstreeks te modelleren. Daarom werden de langere afstanden in de regels vervangen door reistijden berekend langs een transportnetwerk. Hierdoor kan het model ook ingezet worden om de impact van verschillende eenvoudige transportsenario's op de toekomstige landgebruiken en activiteiten uit te testen. Verder werden ook enkele aanpassingen voorgesteld aan de dichtheidsvergelijkingen die de activiteiten over de cellen verdelen, zodat ze beter met de recente sterke verdichting van Belgische steden kunnen omgaan.

Om het model en de aanpassingen grondig te evalueren is het noodzakelijk om het model in een historische calibratie over een periode in het verleden te laten lopen. Een tweede doelstelling van het project was dan ook om een methode te ontwikkelen die toelaat om landgebruik- en bevolkingskaarten voor het verleden te reconstrueren uit beschikbare data. Hiervoor werd gedetailleerde informatie over landgebruik en

bevolkingsspreiding vandaag gecombineerd met een tijdreeks van fractiekaarten van ondoorlaatbare oppervlakte, afgeleid uit remote sensing data. De hieruit bekomen tijdsreeks van bevolkingskaarten (1986, 2001, 2013) geeft een goed inzicht in hoe een aanvankelijk zeer sterk effect van urban sprawl in Vlaanderen langzaam aan het veranderen is in een proces van stedelijke verdichting, alhoewel de bebouwing verlopig toch nog vrij sterk toeneemt in suburbane gebieden en in sommige rurale regio's.

In een derde luik van het onderzoek werd een semi-automatische methode ontwikkeld om het verbeterde ACA model te calibreren. CA-modellen worden doorgaans manueel gecalibreerd omdat expertkennis belangrijk is, maar dit is een traag proces. De nieuwe aanpak bestaat uit het iteratief toepassen van een automatische optimalisatieroutine (een genetisch algoritme), waarbij de modelleerder zijn of haar kennis kan gebruiken om grenzen van parameterwaarden in te stellen en bepaalde categorieën van landgebruik of activiteit meer of minder gewicht te geven in de calibratie.

Tenslotte werd in een vierde deel van het onderzoek het verbeterde ACA-model ingezet om in Vlaanderen de impact van een ruimteneutraal scenario te simuleren, en zo de toekomstige inname van ruimte zo optimaal mogelijk te laten verlopen zodat de druk op de omgeving tot een minimum beperkt wordt. Het scenario werd gedefinieerd in lijn met het Witboek van het Beleidsplan Ruimte Vlaanderen, waarin de bebouwde oppervlakte enkel nog mag toenemen op locaties die goed ontsloten zijn door het openbaar vervoer en die in de nabijheid liggen van een voldoende groot aanbod van diensten. Resultaten van de toepassing tonen aan dat in het ruimteneutrale scenario de overblijvende clusters van de open ruimte veel beter bewaard blijven dan in een *business-as-usual* scenario. Toch blijkt dat de onbeschermdede stedelijke groene ruimte bedreigd kan worden door een

densificatie van de stedelijke kernen. Bijgevolg zullen beleidsmakers moeten nadenken over alternatieve strategieën zodat deze groene ruimtes niet of minder aangetast worden.

Acknowledgements

First of all, I want to express my deepest thanks to Frank. Frank is a great teacher and has already inspired me as a student to focus my studies on cartography, geographical information science and remote sensing. My master thesis was on the optimisation of interrupted map projections. As I have always loved maps from my early childhood, this lovely research field was just perfect for me. But after my studies, Frank found another research field for me thanks to cooperating in different projects with Guy Engelen of VITO, actually the fourth supervisor of this PhD project. From then on, I would study land-use change modelling with cellular automata models and I would live part-time in the small town of Mol for about five years. After some initial doubts (“Can you really model land-use changes? Can you really live in Mol?”), I have started to love my new subject, and a few years later my little room in Mol as well. In land-use change modelling, you can actually apply every possible geographical knowledge: GIS, spatial analysis, cartography, some remote sensing to make input maps, and knowledge about geographical phenomena in general. And policy makers are really interested in model applications, which gives you the feeling that you are performing useful research. Anyway, Frank and Guy, you have always been there to give (a lot of) useful comments and to organise progress meetings despite your busy agendas, so thank you for your big support and for sharing your experience with me!

The activity-based CA (ACA) model, as well as its predecessors without activities, wouldn't exist without the innovative ideas of Roger. Thanks to the long-during friendship and cooperation between Guy and

Roger, the ACA model was further developed in Mol, and I could enjoy Roger's annual visits to Belgium in which he enjoyed in turn bicycle rides, Belgian beers, and meanwhile greatly improved my PhD research and other projects of VITO. Roger, I hope that we can still exchange research ideas in the future or just enjoy a nice beer!

Two other researchers at VITO have been particularly important in helping me to carry out the research. Lien, thanks to your experience in building policy-based spatial applications, I have been able to include an interesting final chapter in the thesis about scenario modelling, directly based on actual policies. You have well paid attention that the manuscript would not just be a technical description of the model, but would keep in touch with the applications of the real world. And then there was also Inge at VITO. What would I've been without Inge? You were always there to eat cookies or cakes during coffee breaks, to help me out with the model code and other applications, to listen to problems I encountered while performing the research, and again to share cookies. Maybe we should once make a model of the perfect "stroopwafel", I'm quite sure that Roger will also have some good ideas to make it!

I also want to thank all other colleagues of the Environmental Modelling Unit (RMA) of VITO. There are too many people to mention since sharing an open office with the whole department means that you have a lot of interesting or just funny conversations with everybody. But particularly I liked working in the Spatially Dynamic Modelling team (RDM) with next to Guy, Lien and Inge also Maarten (thanks for helping me with the GA or other coding problems!), Els (thanks for your help with the last chapter!), Karolien, Leen, Jean-Luc and Liliane. To the other former PhD students of our department (Oliver, Alistair, Jeremy, Dries, ...): it was

a big pleasure to work with you, and have breaks or lunches at the VITO restaurant!

Next to spending on average four days a week in Mol, I also had a weekly trip to the Vrije Universiteit Brussel, where I worked in the Cartography and GIS team of the Department of Geography. Next to Frank, I want to thank the fantastic people with whom I shared an office during the years. Kasper, Sophie, Sven (thanks for helping me with the sealed surface fractions!), Frederik, Koen and Kris: thanks to your enthusiasm, it was always a pleasure to come to VUB. I also want to thank our secretary Inge who is always there to help anyone who needs support!

Since a month, I have moved to the other side of the North Sea to continue my research career in the field of agent-based modelling at the University of Leeds. Nick Malleson, thank you very much for your trust in my capacities as a researcher! I hope that we will have a fantastic collaboration and after only one month, I already know that I ended up in a dynamic team (actually, two teams again) of nice and motivated researchers.

Finally, another very big thank you goes to my family. My grandparents, who all worked in the education sector, have been very important in who I have become. Unfortunately, none of them have reached the age to be able to still read this manuscript. My aunt Ingrid is one of my biggest fans and is always interested in everything I'm doing. And of course, my parents have been the best support team that you can imagine to complete a PhD. I will now have to prove that I can be a good researcher in another country with a little bit less support nearby...

Table of contents

Members of the exam committee	iii
Summary	v
Samenvatting	ix
Acknowledgements	xiii
List of figures	xxi
List of tables	xxvii
List of abbreviations	xxxii
1. Introduction	1
1.1 Context	2
1.2 The region of Flanders within the Belgian urban system	4
1.3 Goals of the PhD project	9
1.3.1 The activity-based CA model: limitations and improvements	9
1.3.2 Historical data	10
1.3.3 Calibration and validation	12
1.3.4 Simulation of future land use and activities in Flanders and Brussels	13
1.3.5 Summary of research objectives	14
1.4 Structure of the thesis	15
2. Spatially dynamic modelling of cities and regions	17
2.1 Theories of urban form	18
2.2 Two schools that influenced spatially dynamic modelling	23
2.3 The macro scale: urban systems models and LUTI models	26
2.4 Cellular automata models	28

2.4.1	Characteristics of CA models	29
2.4.2	Influential urban CA models and calibration techniques	31
2.5	Agent-based models	35
2.6	Discussion and conclusions	38
3.	The activity-based cellular automata model	41
3.1	Introduction	41
3.2	Key elements of the MOLAND model	42
3.3	Extending the neighbourhood: computations in a variable grid structure	47
3.4	The multiple activity cellular automata model	49
3.5	Robustness analysis: Application to the Antwerp area	57
3.6	Discussion and conclusions	63
4.	Integrating travel time into variable grid neighbourhood rules	67
4.1	Introduction	67
4.2	Travel time computations within the variable grid CA	69
4.3	Study area and implementation	77
4.4	Results	83
4.5	Discussion	90
4.6	Conclusions	95
4.A1	Appendix. Activity-based centres of variable grid supercells	95
5.	Downdating high-resolution population density maps using sealed surface cover time series	99
5.1	Introduction	99
5.2	Methods	101

5.2.1	Study area and land use	102
5.2.2	Sealed surface fraction mapping	105
5.2.3	High-resolution mapping of population density for 2013	107
5.2.4	Downdating of the population density map	109
5.2.5	Uncertainty propagation analysis	111
5.3	Results	112
5.3.1	Population trends between 1986 and 2013	112
5.3.2	Population change versus changes in built-up area	114
5.3.3	Uncertainty propagation analysis	117
5.4	Discussion	118
5.5	Conclusions	123
5.A1	Appendix. Downdating of land-use maps	124
5.A2	Appendix. Downdating of employment maps	126
6.	Semi-automated calibration	129
6.1	Introduction	129
6.2	Methods	132
6.2.1	The activity-based CA model	132
6.2.2	Semi-automated calibration with a genetic algorithm	133
6.2.3	Multiple-resolution error functions	136
6.2.4	Study area and data	139
6.2.4.1	Greater Dublin Region	139
6.2.4.2	Flanders and the Brussels Capital Region	142
6.2.5	Experimental design of the calibration	144
6.2.6	Validation of the global pattern of the residential land-use class	146
6.3	Results	147
6.3.1	Greater Dublin Region	147
6.3.1.1	Semi-automated calibration	147

6.3.1.2	Validation of the residential land-use pattern	154
6.3.2	Flanders and the Brussels Capital Region	159
6.3.2.1	Semi-automated calibration	159
6.3.2.2	Validation of the residential land-use pattern	174
6.4	Discussion	178
6.5	Conclusions	183
6.A1	Appendix. Additional tables of the results for Flanders and the BCR	184
7.	Simulation of future land use in Flanders and Brussels	193
7.1	Introduction	193
7.2	Methods	196
7.2.1	The activity-based CA model and input data	196
7.2.2	The business-as-usual (BAU) scenario	197
7.2.3	The land-take neutral (LTN) scenario	199
7.3	Results	205
7.4	Discussion	217
7.5	Conclusions	219
7.A1	Appendix. Protected natural and agricultural areas	220
8.	Conclusions and future perspectives	223
8.1	Network-based influence rules	223
8.2	Assumptions and equations of the model	226
8.3	Historical data	228
8.4	Calibration and validation	229
8.5	Future urban development in Flanders	231
8.6	Where do we go from here?	233
	References	237
	Publications	275

List of figures

Figure 1.1 The three regions of Belgium and the main cities.	5
Figure 1.2 Research objectives of this study with research output in green and input data in blue.	14
Figure 2.1 Central place system of Christaller based on a) markets, b) the transportation system, c) administrative regions (copied from White et al., 2015).	19
Figure 2.2 Cluster size-frequency diagrams for a) the Netherlands and b) Belgium (after White et al., 2015).	23
Figure 3.1 Structure of the variable grid.	48
Figure 3.2 Example of a neighbourhood influence function for the variable grid activity based CA model.	50
Figure 3.3 Actual land use around Antwerp in 2010.	58
Figure 3.4 Simulated land use around Antwerp in 2025 for (a) an extreme decrease of urban areas, followed by (b) an extreme increase in 2040.	60
Figure 3.5 Simulated land use around Antwerp in 2025 if residential activity repels itself at medium distances.	61
Figure 3.6 Simulated residential land-use change around Antwerp in 2040 with a high railway accessibility parameter, a zero value for other accessibility parameters, low inertia and diseconomies of scale, and no zoning status.	61
Figure 3.7 Simulated land use north of Antwerp in 2040 with a high random perturbation.	62
Figure 4.1 Detail of a fixed ‘network grid’ (NG) with resolution 3^{2R} . The model cells with resolution R are shown within the central grey NG cell only. When the distance is needed from a black unit cell to the centroid of its black variable grid neighbour, the distance from the red centroid of the fixed NG cell containing the black unit cell towards the centroid of the red NG cell containing the centroid of the black variable grid neighbour is used.	72
Figure 4.2 The road component and the public transport component can be combined or used separately to generate time distances.	73

Figure 4.3 Example of a neighbourhood influence function with a Euclidean part for local interactions and a network time-distance part for long-distance interactions. The cell distance axis is made continuous with equation (4.4) and then made logarithmic with equation (3.7).	77
Figure 4.4 CORINE land-use map of Belgium in 2000, with an indication of the regions (Flanders, Wallonia and the Brussels Capital Region (B)) and the largest cities (Brussels (B), Antwerp (A), Ghent (G), Leuven (Le), Liège (Li), and Charleroi (C)).	78
Figure 4.5 Accessibility index for different transport scenarios in Belgium: (a) the standard ‘road-based’ scenario, (b) the ‘congestion’ scenario, (c) the ‘train only’ scenario, and (d) the ‘choice’ scenario.	84
Figure 4.6 New discontinuous residential development in Belgium in the road-based scenario between 2000 and 2060. The rectangle shows the extent of the zooms in the next figures.	86
Figure 4.7 Differences in residential development between the Euclidean and the ‘road-based’ scenarios in 2060.	87
Figure 4.8 Differences in population values (people per cell) between the Euclidean and the ‘road-based’ scenarios in 2060.	87
Figure 4.9 Differences in residential development between the ‘road-based’ and the ‘congestion’ scenarios in 2060.	88
Figure 4.10 Differences in population values (people per cell) between the ‘road-based’ and the ‘congestion’ scenarios in 2060.	88
Figure 4.11 Differences in residential development between the ‘road-based’ and the ‘train only’ scenarios in 2060.	89
Figure 4.12 Differences in population values (people per cell) between the ‘road-based’ and the ‘train only’ scenarios in 2060.	89
Figure 4.A1 Location of the cell centre of a supercell: (a) geometric centre, (b) centre of gravity, and (c) hierarchic centre.	96
Figure 5.1 Land-use map of Flanders and Brussels in 2013 (10 m resolution map developed by VITO aggregated to a 100 m resolution), with an indication of both regions (thick black lines; Brussels Capital Region (B)), the Flemish provinces (thin black lines) and their capitals (Bruges (Bg), Ghent (G), Antwerp (A), Leuven (L) and Hasselt (H)).	104
Figure 5.2 Population density map for 2013 (inhabitants per 30 m pixel).	108
Figure 5.3 Population difference maps (aggregated to a 300 m resolution) for the periods (a) 1986 – 2001, and (b) 2001 – 2013.	114

Figure 5.4 Densification index maps (300 m resolution) for cells with a growing population and a minimum of 10 inhabitants for the periods (a) 1986 – 2001, and (b) 2001 – 2013.	116
Figure 5.5 Average densification index values per municipality for the period 2001 – 2013. The Flemish provincial capitals are indicated: Bruges (Bg), Ghent (G), Antwerp (A), Leuven (L) and Hasselt (H).	116
Figure 5.6 Standard deviation of the estimated number of inhabitants in a Monte Carlo analysis at 300 m resolution for 1986.	118
Figure 6.1 An influence function of activity J on activity K , defined by its inertia $p_{JK,1}$ (not in the figure) and 6 other parameters ($p_{JK,2} - p_{JK,7}$).	135
Figure 6.2 CORINE land-use map of the Greater Dublin region in 1990, aggregated to a 200 m resolution. The rectangle shows the extent of the zooms in Figures 6.4 and 6.5.	141
Figure 6.3 VITO land-use map of Flanders and the Brussels Capital Region in 2013, aggregated to a 100 m resolution, with an indication of the three largest cities (Brussels (B), Antwerp (A) and Ghent (G)). Regional borders in thick black lines. The rectangle shows the extent of the zooms in Figures 6.12 and 6.14.	143
Figure 6.4 Evolution of population (per 200 m cell) between 1990 and 2000 around the city of Dublin in (a) the actual maps, (b) simulation D3 (no network distances, error weight type 2), and (c) simulation D2 (network distances, error weight type 1).	152
Figure 6.5 Difference in residential development around the city of Dublin in the year 2000 between the actual CORINE land-use map and three simulations with network distances: (a) simulation D2 (error weight type 1), (b) simulation D4 (error weight type 2), and (c) simulation D6 (error weight type 2 with alternative density equation (3.21)).	153
Figure 6.6 Difference in residential development in the Greater Dublin Region between 1990 and 2050 in three different simulations with error weight type 2: (a) run D3 (no network distances), (b) run D4 (network distances), and (c) run D6 (network distances with alternative density equation (3.21)).	157
Figure 6.7 Cluster size-frequency graphs of residential land use of the GDR for the actual map of 1990 and three simulations (runs D3, D4 and D6). The open triangles are cluster sizes with a low frequency not used to compute trend lines.	158

Figure 6.8 Actual urban development in Flanders and the Brussels Capital Region between 2001 and 2013: (a) evolution of the population per ha, and (b) development of new residential cells.	159
Figure 6.9 Actual urban development in the Antwerp-Brussels area between 2001 and 2013: (a) evolution of the population per ha, and (b) development of new residential cells. Legend as in Figure 6.8.	160
Figure 6.10 Difference in (a) population and (b) residential development in central Flanders and the Brussels Capital Region in the year 2013 between the actual maps and simulation F2.	165
Figure 6.11 The Flemish districts and the Brussels Capital Region (“Brussel”).	168
Figure 6.12 Evolution of population per ha between 2001 and 2013 in western and central Flanders and the Brussels Capital Region in (a) simulation F4 (with density equation (3.20)), and (b) simulation F5 (with density equation (3.21)).	170
Figure 6.13 Difference in residential development in the Antwerp-Brussels area in the year 2013 between the actual land-use map and (a) simulation F7 (multiple resolution optimisation, density equation (3.21)), and b) simulation F10 (regional optimisation, density equation (3.20)).	172
Figure 6.14 Employment in the services sector in western and central Flanders and the Brussels Capital Region: (a) actual map in 2001; (b) actual change in employment between 2001 and 2013; (c) simulated change in run F3.	173
Figure 6.15 Difference in residential development in western and central Flanders and the Brussels Capital Region between 2001 and 2060 in (a) simulation F6 and (b) simulation F7.	176
Figure 6.16 Cluster size-frequency graphs of residential land use of Flanders and the BCR for the actual map (2001) and three simulations (runs F6, F7 and F10). The open triangles are cluster sizes with a low frequency not used to compute trend lines.	177
Figure 7.1 Daily urban land take in Flanders and the BCR in both scenarios.	199
Figure 7.2 Categorical scores per ha in Flanders and the BCR for public transport accessibility and proximity to services (based on data of Verachtert et al., 2016).	200
Figure 7.3 Suitability values for the land-use categories residential, retail and services in the LTN scenario.	203

Figure 7.4 Relative difference in population (%) per district (a) between 2013 and the projection for 2050 of the Federal Planning Bureau; and in 2050 between (b) the projection of the Federal Planning Bureau and the BAU scenario, and (c) the BAU and the LTN scenario.	206
Figure 7.5 Current and new residential development in Flanders and the BCR between 2013 and 2050 in both scenarios in (a) the western and (b) the eastern half of the study area.	208
Figure 7.6 Current and new residential development in central Flanders and the BCR between 2013 and 2050 in (a) the BAU scenario and (b) the LTN scenario. With an indication if the new cells are located inside the LTN zoning and if so, in which of the main categories of the accessibility-proximity matrix.	210
Figure 7.7 Decline of open space in Flanders and the BCR between 2013 and 2050 in the two scenarios.	211
Figure 7.8 Change in population per 100 m cell in Flanders and the BCR between 2013 and 2050 in (a) the BAU scenario and (b) the LTN scenario. Increasing population in blue, decreasing population in orange.	211
Figure 7.9 Change in natural land use in Flanders and the BCR between 2013 and 2050 in (a) the BAU scenario and (b) the LTN scenario.	216

List of tables

Table 4.1 Average modelling speeds for road segments in km/h.	80
Table 4.2 Land-use and activity growth scenarios, and parameter values of the application to Belgium that were copied from White et al. (2012).	82
Table 4.3 Comparison in 2060, in terms of discontinuous residential land use and population, between the ‘road-based’ scenario (‘R’) and the ‘congestion’ scenario (‘C’). The classes are defined by relative accessibility index values (division of ‘congestion’ and ‘road-based’ index values).	91
Table 4.4 Comparison in 2060, in terms of discontinuous residential land use and population, between the ‘road-based’ scenario (‘R’) and the ‘train only’ scenario (‘T’). The classes are defined by relative accessibility index values (division of ‘train only’ and ‘road-based’ index values).	92
Table 4.5 Comparison in 2060, in terms of discontinuous residential land use and population, between the ‘road-based’ scenario (‘R’) and the ‘choice’ scenario (‘Ch’). The classes are defined by relative accessibility index values (division of ‘choice’ and ‘road-based’ index values).	93
Table 5.1 Proportions of land uses (%) in Flanders and the Brussels Capital Region according to the 2013 land-use map of the Flemish Institute for Technological Research (VITO). Residential, industrial and commercial, and infrastructure are built-up land uses. Natural land covers in military terrains were moved from infrastructure to nature to produce this table.	105
Table 5.2 Proportion of built-up land use / land cover (%) in Flanders and the Brussels Capital Region for 1986, 2001, and 2013 according to the 2013 land-use map of VITO, the Belgian Cadastre, and the sealed surface time series ($SSF \times \text{pixel area}$) used in this study.	105
Table 5.3 Population of Flanders and the Brussels Capital Region for 1986, 2001, and 2013.	113
Table 5.4 Average (Avg) densification index values and standard deviations (SD) for different categories of population per 300 m cell in the population density map for 2013. Cells with a population below 10 were excluded as outliers.	115

Table 5.5 Standard deviation of sealed surface fraction (30 m cells), population estimate (inhabitants per 300 m cell) and resulting densification index (300 m cells), averaged over all cells, in a Monte Carlo uncertainty propagation analysis.	118
Table 6.1 Definition of calibration error function weight types for the Greater Dublin Region. The regional error weight h and the multiple resolution error weights g_L are given per activity or land-use category.	148
Table 6.2 Errors in population in 2000 of different calibration runs for the Greater Dublin Region with indication of the model version ('Netw': network-based influence rules or not), the number of the density equation ('Dens'), the error function weight type ('Type') according to Table 6.1, the regional error R , and the errors E_L at different resolutions. Underlined values are optimised together with the underlined values in Table 6.3 for residential land use.	149
Table 6.3 Errors in the number of residential cells in 2000 of different calibration runs for the Greater Dublin Region with indication of the model version ('Netw': network-based influence rules or not), the number of the density equation ('Dens'), the error function weight type ('Type') according to Table 6.1, the regional error R , and the errors E_L at different resolutions. Underlined values are optimised together with the underlined values in Table 6.2 for population.	150
Table 6.4 Spatial metrics of the residential land-use pattern of the actual land-use maps ('Act') and different calibration runs for the Greater Dublin Region.	155
Table 6.5 Overview of calibration runs of the Belgian application with indication of the model version ('Netw': network-based influence rules or not), the number of the density equation ('Dens'), the error function weight type ('Type') according to Tables 6.A1 and 6.A2, and the influence weights ($W_{i,j}$ is the influence of category i on category j) and parameters (p_K is a parameter for activity K) which were optimised by the calibration algorithm. Abbreviations are used for population (pop) and employment (ret = retail, ser = services).	162

Table 6.6 Errors in population in 2013 of different calibration runs for Flanders and the BCR with indication of the model version (‘Netw’: network-based influence rules or not), the number of the density equation (‘Dens’), the regional error R and the errors E_L at different resolutions. Underlined values are optimised together with the underlined values in Table 6.7 for residential land use, and in some cases with an optimisation of employment and economic land uses (see Tables 6.A3 and 6A4).	163
Table 6.7 Errors in the number of residential cells in 2013 of different calibration runs for Flanders and the BCR with indication of the model version (‘Netw’: network-based influence rules or not), the number of the density equation (‘Dens’), the regional error R and the errors E_L at different resolutions. Underlined values are optimised together with the underlined values in Table 6.6 for population, and in some cases with an optimisation of employment and economic land uses (see Tables 6.A3 and 6A4).	164
Table 6.8 Relative error (%) in activity K per district in 2013 of calibration run F2 for Flanders and the BCR. Abbreviations are used for population (pop) and economic sectors (ind = industry, whl = wholesale and logistics, ret = retail, ser = services).	166
Table 6.9 Relative error (%) in the number of cells of active land-use category i per district in 2013 of calibration run F2 for Flanders and the BCR. Abbreviations are used for residential land use (res) and economic sectors (ind = industry, whl = wholesale and logistics, ret = retail, ser = services).	167
Table 6.10 Spatial metrics of the residential land-use pattern of the actual land-use maps (‘Act01’ for 2001 and ‘Act13’ for 2013) and different calibration runs for the Belgian application.	175
Table 6.A1 Definition of calibration error function weight types for Flanders and the BCR without regional errors. The multiple resolution error weights g_L are given per activity K or land-use category i . Abbreviations are used for population (pop), residential land use (res) and economic sectors (ind = industry, whl = wholesale and logistics, ret = retail, ser = services).	186
Table 6.A2 Definition of calibration error function weight types for Flanders and the BCR with regional errors. The regional error weight h and the multiple resolution error weights g_L are given per activity K or land-use category i . Abbreviations are used for population (pop), residential land use (res) and economic sectors (ind = industry, whl = wholesale and logistics, ret = retail, ser = services).	187

Table 6.A3 Errors in activity K in 2013 of different calibration runs for Flanders and the BCR: the regional error R and the errors E_L at different resolutions. Underlined values are optimised together with the underlined values in Table 6.A4 for the active land uses. Abbreviations are used for population (pop) and economic sectors (ind = industry, whl = wholesale and logistics, ret = retail, ser = services).	188
Table 6.A4 Errors in the number of cells of active land-use category i in 2013 of different calibration runs for Flanders and the BCR: the regional error R and the errors E_L at different resolutions. Underlined values are optimised together with the underlined values in Table 6.A3 for the activities. Abbreviations are used for residential land use (res) and economic sectors (ind = industry, whl = wholesale and logistics, ret = retail, ser = services).	190
Table 7.1 Suitability values for each possible combination of accessibility and service proximity scores of Verachtert et al. (2016).	202
Table 7.2 Suitability correction factors for six different categorical potential scores to install district heating in Flanders, according to a study of Renders et al. (2015).	202
Table 7.3 Parameters of population in the two scenarios.	204
Table 7.4 Comparison of population and population density between actual values in 2013 and simulated values in both scenarios in 2050 in non-residential cells, residential cells already existing in 2013, and new residential cells in 2050.	212
Table 7.5 Total population and proportion of population (%) in the BAU scenario in the four generalised categories of the matrix of Verachtert et al. (2016).	213
Table 7.6 Total population and proportion of population (%) in the LTN scenario in the four generalised categories of the matrix of Verachtert et al. (2016).	213
Table 7.7 Definition of different functional types of green space that can be reached by the population, based on a study of Van Herzele and Wiedemann (2003).	215
Table 7.8 Accessibility of the population to different types of green space in the BAU scenario.	215
Table 7.9 Accessibility of the population to different types of green space in the LTN scenario.	215

List of abbreviations

ABM	Agent-based modelling
ACA	Activity-based cellular automata
AGIV	Agentschap voor Geografische Informatie Vlaanderen (Flanders Geographical Information Agency)
ANN	Artificial neural networks
BAU	Business-as-usual
BCR	Brussels Capital Region
BRV	Beleidsplan Ruimte Vlaanderen (Spatial Policy Plan for Flanders)
CA	Cellular automata
CPB	Centraal Planbureau (Dutch Bureau for Economic Policy Analysis)
CV RMSE	Coefficient of variation of the root mean square error
ED	Edge density
EDs	Electoral divisions
ETM+	Enhanced Thematic Mapper Plus
FPB	Federaal Planbureau (Federal Planning Bureau)
GA	Genetic algorithm
GAUL	Genetic Algorithm Utility Library
GDR	Greater Dublin Region
GIS	Geographical information system
GRB	Grootschalig Referentiebestand (Large-scale reference data set)
GroWaDRISK	Drought-related vulnerability and risk assessment of groundwater resources in Belgium
INBO	Instituut voor Natuur- en Bosonderzoek (Research Institute for Nature and Forest)
LOV	Leefomgevingsverkenner (Environment Explorer)
LPI	Largest patch index
LTN	Land-take neutral
LUTI	Land use transportation interaction
MAE	Mean absolute error
MESH	Effective mesh size
MIRA	Milieurapport (Flemish Environment Outlook)
MOLAND	Monitoring Land use / cover Dynamics

NACE	Nomenclature Statistique des activités économiques dans la Communauté européenne (Statistical Classification of Economic Activities in the European Community)
NARA	Natuurrapport (Flemish Nature Outlook)
NDVI	Normalised Difference Vegetation Index
NG	Network grid
NMBS/SNCB	Belgian National Railway Company
NP	Number of patches
NSGA-II	Non-dominated Sorting Genetic Algorithm II
PD	Patch density
PROX_AM	Proximity index distribution with area-weighted mean
RMSE	Root mean square error
RSV	Ruimtelijk Structuurplan Vlaanderen (Spatial Structure Plan for Flanders)
SHAPE_AM	Shape index distribution with area-weighted mean
SSF	Sealed surface fraction
TM	Thematic Mapper
VITO	Vlaamse Instelling voor Technologisch Onderzoek (Flemish Institute for Technological Research)

1. Introduction¹

*The last defender of the sprawl
Said "Well, where do you kids live?"
Well, sir, if you only knew what the answer's worth
Been searching every corner of the earth*

(from the song "Sprawl I (Flatland)" by Arcade Fire, on the album "The Suburbs")

The population of countries, regions and cities grows and declines as time proceeds. The last two centuries global industrialisation has caused an exponential population growth (Cleland, 2013), that is associated with spatial expansion of urban areas. Suburbanisation is a global process that drives today's urban governance, land markets, retail system and way of life (Keil, 2013). Since the end of the Second World War, mass production, greater wealth of the middle class and increased access to cheap transportation systems have caused a massive growth of urban sprawl (Newman & Kenworthy, 1989; Litman, 2015). Cities and their surrounding rural areas have always influenced each other. However, in recent times the border between urban and rural land uses is gradually disappearing in many regions, turning urban growth into an increasingly complex phenomenon (White et al., 2015). To acquire more insight into processes of urban development, increasing use is made of urban growth models. In this thesis, urban growth and its consequences are modelled with a new, multifunctional approach, an activity-based cellular automata model,

¹ Sections 1.1 and 1.2 of this chapter are partially based on the introduction and study area sections of: Crols, T., Vanderhaegen, S., Canters, F., Engelen, G., Poelmans, L., Uljee, I., & White, R. (2017). Datedating high-resolution population density maps using sealed surface cover time series. *Landscape and Urban Planning*, 160, 96–106.

initially developed by White et al. (2012), and further enhanced during this PhD project.

1.1 Context

Many environmental problems worldwide arise from the ever growing area on Earth occupied by human activities. This growth often manifests itself in the form of urban sprawl, especially in large parts of North America and Europe (Kasanko et al., 2006; Ewing, 2008; Ravetz et al., 2013). The main causes of urban sprawl are a growing population, increased income and new fast transportation networks with decreased travel costs (Anas et al., 1998; Brueckner, 2000; European Environment Agency, 2016). Other important forces leading to the diffuse growth of urban areas are economies and diseconomies of agglomeration (Richardsson, 1995): while a concentration of population and companies produces advantages of proximity, a too high concentration of activities in a single location can also lead to high land prices, reduced accessibility and environmental pollution. Extended transportation networks and diseconomies of agglomeration can both lead to a growth of suburban housing and employment. This, in turn, causes new health concerns (Jackson, 2003) and environmental problems due to high commuting distances and related congestion (Camagni et al., 2002), increased energy consumption (Newman & Kenworthy, 1989), altered natural landscapes and loss of valuable agricultural land (Antrop, 2000). In recent times the average home-to-work travel distance has grown significantly (Boussauw et al., 2011). For many reasons residential location choices are often determined by more than the current job location: two-worker households need housing in between both jobs, job uncertainty plays a major role, and moving costs can be substantial (Crane, 1996; Anas et al.,

1998). Job accessibility has become less important for land prices since non-work travel for shopping, social interactions and other purposes has increased and is in many regions more important than commuting (Balmer, 2007; Giuliano et al., 2010). Owning a suburban house is therefore often perceived as an 'ideal' living location (Batty et al., 2003). In the new millennium, however, a number of cities and regions around the world have started to recognise the need for more compact urban forms (Dieleman & Wegener, 2004; Hassan & Lee, 2015).

To gain a spatially differentiated insight in the future evolution of urban land use and the associated population growth or decline, land-use change modelling can be beneficial (Batty, 2005; Volkery et al., 2008; Haase & Schwarz, 2009). Cellular automata (CA) models explicitly deal with the dynamic spatial interactions between different land uses (Poelmans & Van Rompaey, 2010; Santé et al., 2010; White et al., 2015). The core of a CA land-use model is the set of neighbourhood rules. For each cell, they weigh the influences exerted by the land uses present in a specified neighbourhood of the cell to compute the potential for its future land-use state. The constrained cellular automata model of White, Engelen, and Ulfje (1997), also referred to as the MOLAND model, is one of the best known implementations of CA modelling with many worldwide applications. The model couples a local CA model to a spatial interaction based regional model that handles dynamics of population and economic activity. In Flanders, it is known as the RuimteModel Vlaanderen (Gobin et al., 2009), and was used to run different scenarios to 2050 (de Kok et al., 2012) with 12 urban land uses at a 100 m resolution. Recently, some publications have proposed an integration of population and economic data within the CA structure itself by using an activity-based approach (White, 2006a; van Vliet

et al., 2012; White et al., 2012). This approach has the advantage that no separate regional component has to be coupled to the micro-scale CA model. All spatial scales are represented by using a variable grid structure (see chapter 3). The model of White et al. (2012), which defined the framework for urban growth modelling in this thesis, is a multiple activity-based cellular automaton: the simulation allocates a land use and one or more ‘activities’ to each cell at each time step. Activities are defined in a quantitative way: they allow the model to express the density of different functions within each cell. Although each type of activity is associated with one specific active land use, that activity may also be present on cells with other land uses. In principle any ‘activity’ could be considered, but at present the activities are population and the employment in various economic sectors. This PhD research had as its main goal to further develop and improve the current version of the activity-based CA model (or ACA model), and then to apply it to the region of Flanders, Belgium.

1.2 The region of Flanders within the Belgian urban system

Belgium is divided into three regions (Figure 1.1) and three language communities. The Dutch-speaking Flemish region in the north of Belgium is the principal study area of this thesis, together with the French-Dutch bilingual Brussels Capital Region (BCR), which is an enclave in the south of Flanders. The southern half of the country consists of the region of Wallonia, which is mainly French-speaking but has a small German-speaking community in the east. The legislative power of the regions and language communities have consistently been enlarged since the 1980s (Saey et al., 1998). The regions fully determine spatial development and

housing policies, as well as many economic, environmental and transportation policies.

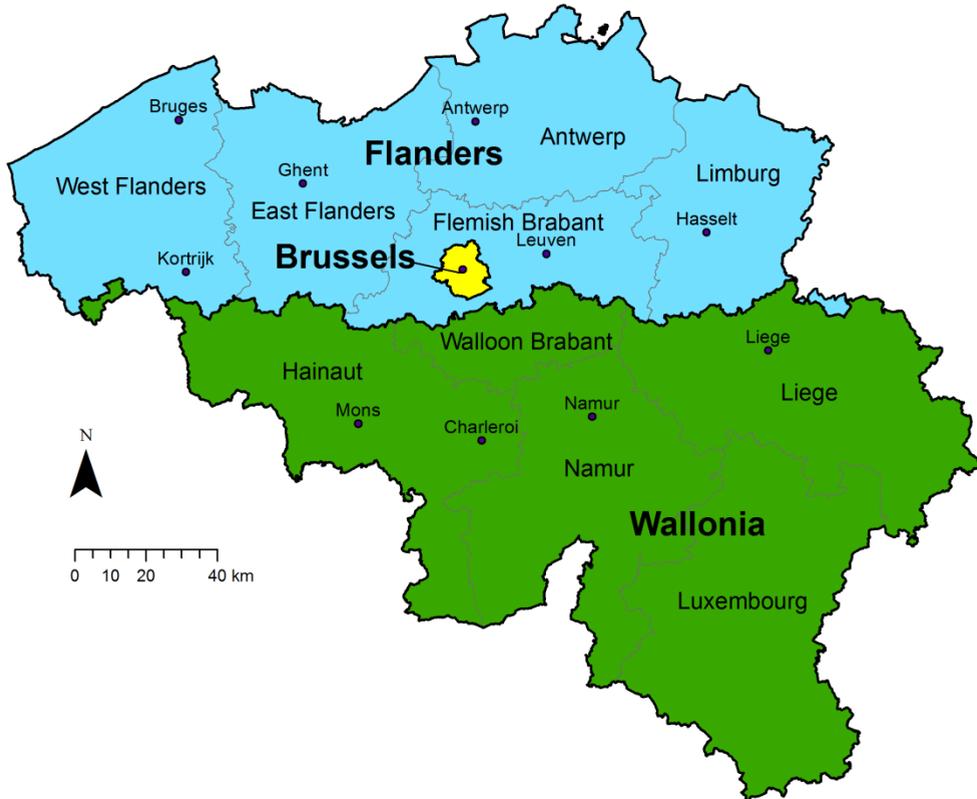


Figure 1.1 The three regions of Belgium and the main cities.

The location of the main cities and of most regional centres in Belgium was already laid out in the medieval period as most of them have a strategic position along the waterways (Vandermotten & Vandewattyne, 1985). Together they constitute a system that more or less can be described as a market city system with Christaller's central place theory (see chapter 2; Christaller, 1966) (Van Nuffel & Saey, 2005). Urban and peri-urban growth from 1840 until the present day have been heavily influenced by several economic transitions linked to Kondratieff cycles (Vandermotten,

1998; van Meeteren et al., 2016). A number of cities and smaller towns expanded greatly and transformed themselves into manufacturing centres during the Industrial Revolution, mainly along the Sambre and Meuse rivers in Wallonia where natural resources could be found (e.g. Liege and Charleroi), but also in Flanders (e.g. Ghent). The main development of the first half of the 20th century can be found along the ‘ABC’ axis, starting at the port of Antwerp, and going along Brussels (the political and financial capital) towards the industrial city of Charleroi. In the post-war period a second major transition occurred, when the older industries collapsed and a new industrial and service economy appeared in suburban zones, often close to motorways and canals. During this period, Flanders had a clearly stronger economic growth than Wallonia. Today the highest concentration of population and economic activity can be found in the so called ‘Flemish Diamond’ between Brussels, Ghent, Antwerp and Leuven (see Van Meeteren et al., 2016 for a recent overview).

Today, northern and central Belgium is a region with one of the highest proportions of built-up surfaces and fragmented landscapes in Europe, with extensive urban sprawl and ribbon development with mixed urban and rural functions (Antrop, 2004; Poelmans & Van Rompaey, 2009; Prokop et al., 2011; Verbeek et al., 2014; European Environment Agency, 2016; Vanderhaegen & Canters, 2016). The spatial characteristics of this peri-urban development pattern actually find their origin in two 19th century government policies that were decided when the Belgian cities were rapidly industrialising but the political elite wanted to inhibit massive rural-urban migration towards the cities in order to avoid social unrest and health problems (De Decker, 2008). Firstly, a dense light railway network was built to enable fast and cheap travel from rural areas to the city (De Block &

Polasky, 2011). Secondly, home ownership was promoted with beneficial tax policies for mortgages. While the light railway network largely disappeared in the 1950s, suburban housing policies continued and resulted in the development of a car-based society (De Meulder et al., 1999). Frequent moving has always been discouraged by high transaction costs, and has gradually become rather exceptional in a society where the majority of the population has become used to owning a suburban house with a large garden, close to several cities and job possibilities (Meeus & De Decker, 2015). During recent years, the urban population is growing again, though mainly triggered by immigration, still, the rate of urban sprawl is only slightly decreasing (see chapter 5).

Flanders had until recently a daily loss of open space of close to 6 ha on a total area of 13,522 km² (Ruimte Vlaanderen, 2016). Based on land-use simulations with the RuimteModel Vlaanderen assuming no changes in spatial policy, the Flemish Environment Outlook 2009 (MIRA 2009) already indicated that valuable natural and agricultural land would be lost at the same if not an increasing rate until 2030 (Van Steertegem et al., 2009). Furthermore, the growth of urban sprawl triggers more traffic, worse air quality, inefficient energy consumption and more risk of floods. More recently, De Kok et al. (2012) analysed a set of five different “world views” with the RuimteModel, based on global and local trends, and how future policy could deal with these trends. Such model-based outlooks can help policy makers to adapt regional spatial planning in a sustainable way. However, Flanders has in many locations mixed and multifunctional land uses containing different urban activities. These are not compatible with the discrete land-use categories of a simple land-use change model. Therefore, the activity-based CA model of White et al. (2012) seemed to be a good

choice as a tool to analyse the future development of the Flemish urban structure at a high resolution (100 m). The simulated values of population and employment can also be used as an input to different spatial indicators (e.g. air quality, energy consumption, waste production, labour force, etc.). Initial simulations with the model applied to Flanders proved this potential, but also indicated that the model had to be further refined.

Although Flanders is the principal study area, some of the applications in this thesis have focused on other regions too. First of all, the Brussels Capital Region was always included in the applications to Flanders, since its suburban periphery lies largely within the Flemish Region. It would have been even better to also include generalised information of bordering regions or cities (e.g. Wallonia, Maastricht, Lille) in the modelling. Due to the extensive preprocessing that would be needed to get comparable data, this was considered to be beyond the scope of the project. However, the transportation scenarios defined in chapter 4 were applied to the entire country of Belgium since this initial part of the research was building further on already available data from the study of White et al. (2012) for Belgium as a whole. Finally, a calibration of the model in chapter 6 was evaluated for both Flanders and the Greater Dublin Region in Ireland. Since the activity-based model (White et al., 2012) had already been applied to Dublin too, this permitted a critical assessment of the relative performance of the different versions of the ACA model in both a monocentric (Dublin) and a polycentric setting (Flanders).

1.3 Goals of the PhD research

As indicated before, the main goal of this PhD research was to further enhance the activity-based CA model of White et al. (2012). This involved the development of a method for introducing network based distance measures in the context of the variable grid defining the model. A secondary goal in terms of model enhancement was to validate and improve the activity density equations used in the model. In addition to model improvements, this PhD research had three other specific goals: (1) to develop a remote sensing based approach for producing historical activity and land use maps of Flanders, as these are required for model calibration, (2) to define suitable methods for calibration and validation, including automated methods, and (3) to simulate different scenarios of future land use and activity distributions in Flanders.

1.3.1 The activity-based CA model: limitations and model improvements

White et al. (2012) concluded that their ACA model results for population estimation outperformed previous results of the MOLAND model. However, some model assumptions about activity density growth and long-distance activity interactions could still be improved. Moreover, the model was not yet tested with spatially detailed employment data and the initial application to Belgium had a rather coarse resolution (300 m). One of the goals of this thesis therefore was to develop an enhanced version of the model that would be able to generate simulation results for Flanders at a 100 m resolution for different urban land uses and activities, just as the RuimteModel Vlaanderen (Gobin et al., 2009) does. The model should thus be able to realistically simulate the functional and morphological

polycentric structure of the Flemish urban landscape, the widespread urban sprawl and the typical ribbon development. The activity density equations proposed by White et al. (2012) therefore needed to be evaluated to check if the Flemish growth pattern could be successfully reproduced, and, where necessary, how model specifications could be improved. Also, while the model directly handles regional economic and demographic relations in its CA rules, the distances between locations are Euclidean in the model of White et al. (2012); hence, after an initial robustness analysis, substituting travel times via a transportation network was seen as one of the most promising ways to improve the performance of the model.

1.3.2 Historical data

A historical calibration strategy is necessary to improve the model equations and obtain reliable parameter estimates for business-as-usual runs of the activity-based CA model. In many applications, however, an important problem is the lack of consistent historical land-use data to serve as input for the calibration (Engelen & White, 2007). In Flanders, several sets of land-use and activity data are available from GIS databases. The Flemish Institute for Technological Research (VITO) combined a large number of GIS datasets to develop a land-use map with many urban categories at a resolution of 10 m for the reference year 2013 (Poelmans et al., 2016). Yet high-resolution data on land use for earlier years were not available.

The activity-based model also strongly depends on high-resolution spatial data of population and employment. Belgian population data of 2013 were provided at a resolution of 100 m by the Federal Department of the Interior. Employment data of individual establishments of large companies were available from a database on home-to-work travel of the Federal

Department of Mobility. General employment data per economic sector per municipality can be obtained from the Federal Department of Economics. Both datasets were combined to create a high-resolution employment map for the same reference year 2013 that has enough detail to be used as input in the model. An important drawback, however, is that for previous years data on population and employment are only available at the level of statistical units (municipalities or at best statistical sectors).

Therefore, one of the goals of the PhD research was to use remote sensing based products to derive land-use and activity maps for the past. The remote sensing data were not used to directly determine different urban land uses (or population), given that earlier studies had proven that it is difficult to identify different urban land-use classes from medium-resolution remote sensing data (Van de Voorde et al., 2011; Grinblat et al., 2016). Instead, the idea was to start from the most recent high-resolution maps of land use and population and then define an approach to identify changes in land use and population from a time series of sealed surface maps obtained through remote sensing. In this ‘downdating’ methodology, observed differences in sealed surface cover were associated with differences in land use and population. Employment maps of the past cannot be directly inferred from changes in sealed surface cover. Therefore, rescaling the maps with the available statistics and the downdated land-use maps was the best possible strategy to follow.

With the downdating methodology land-use and activity maps of the years 1986 and 2001 were computed, but due to concerns about the quality of the 1986 land-use and employment maps, only the 2001 maps were finally used as an input for the historical calibration.

1.3.3 Calibration and validation

White et al. (2012) manually calibrated their version of the ACA model with expert-based knowledge gained in previous modelling work, yet manual calibration has the disadvantage of being extremely time consuming. In the past, several researchers (Straatman et al., 2004; van der Kwast et al., 2012; van Vliet et al., 2013b) have developed automated calibration algorithms for the neighbourhood rules of the CA model of White et al. (1997). The disadvantage of these algorithms though is that in these approaches the influence rules are highly simplified, and unrealistic sets of influence weights may be obtained. One of the goals of this PhD research therefore was to develop a semi-automated calibration methodology that works for the different parameters and rules needed in the activity-based model. ‘Semi-automated’ means that the algorithm needs to be able to incorporate the modeller’s knowledge concerning the likely range of parameter values or the combinations of parameters that can iteratively be changed.

The model produces both land-use maps (categorical data) and activity value maps (continuous data). The errors in both outputs for the reference year 2013 were thus evaluated with one consistent validation algorithm. This algorithm had to compute errors at different spatial scales (cellular errors, focal errors, regional errors...) or even better for a continuous range of scales, such as in the multi-scale error measure defined by Costanza (1989). Next to using historical data for model calibration and validation, models should also be tested by evaluating how realistic scenario results of future land use are. This can be done by checking if their fractal properties are constant in time. Future regional activity totals in Belgium of business-as-usual simulations were validated by comparing them to the prognoses of the Federal Planning Bureau.

1.3.4 Simulation of future land use and activities in Flanders and Brussels

As a final stage of the research, it was a specific goal of the PhD project to run policy scenarios with the model, adapted to the context of an ongoing project of the Flemish government. In the Spatial Policy Plan for Flanders (Beleidsplan Ruimte Vlaanderen, BRV), under development since 2010, the Flemish Department of Spatial Planning (“Ruimte Vlaanderen”) aims at reducing land take. In its BRV white paper, approved by the Flemish Government on 30th November 2016, the plan is proposed to rapidly reduce the expansion of the settlement area on agricultural and natural land, and halt it completely by the year 2040 (Ruimte Vlaanderen, 2016). Moreover, expansion of the settlement area until 2040 and the densification of the existing settlement area should be in those locations with an optimal accessibility to the public transport system and in the proximity of services of various kinds. Verachttert et al. (2016) came up with a methodology to calculate an index that gives a score to both criteria for each location in Flanders. The results of the study have gained media attention, and investigating the possible impact on future spatial policies of a further densification of areas with good access to transport and services, was therefore thought to be highly relevant. Hence, the specific objective of the simulation work of the PhD research was defined as running a what-if scenario that uses the scores obtained by Verachttert et al. (2016) as an input in the zoning and suitability components of the ACA model. As such, the model actually becomes a tool in which the most likely locations (of a limited set of predetermined suitable locations) of future urban growth are chosen, while the largest part of population increase actually will take place in existing residential areas, depending on the potential of these areas to host

more people. This land-take neutral (LTN) scenario had then to be compared to a business-as-usual (BAU) scenario, where no changes in policy are assumed to take place.

1.3.5 Summary of research objectives

A graphical overview of the research objectives can be found in Figure 1.2.

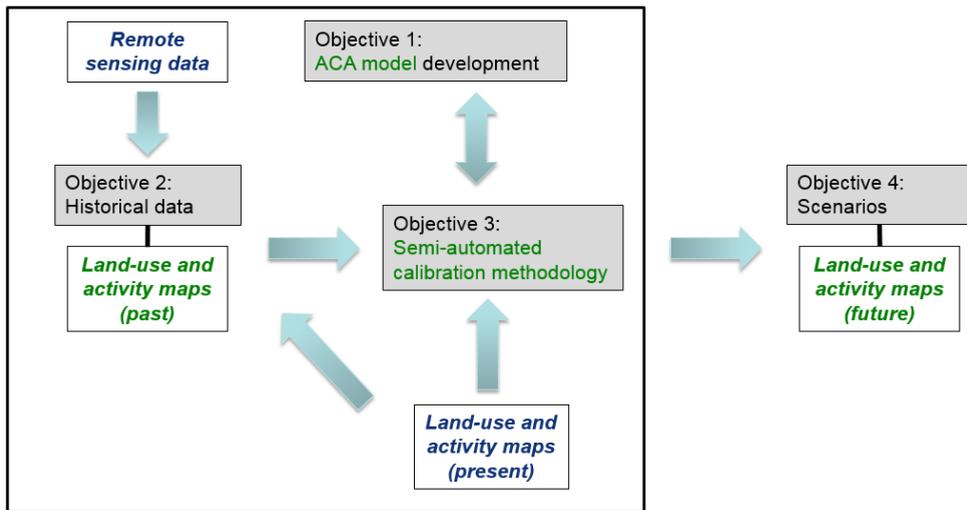


Figure 1.2 Research objectives of this study with research output in green and input data in blue.

- **Objective 1.** Enhancing the activity-based CA model by replacing Euclidian distances with time distances over a network and modifying the model’s activity density equations, so that it can better deal with land-use and activity transitions in Flanders (chapters 3 and 4).
- **Objective 2.** Developing a methodology to derive land-use and activity maps of the past, using both the available high-resolution maps of the present and remote sensing data of the past (chapter 5).

- **Objective 3.** Developing a semi-automated calibration methodology that combines an optimisation routine with expert-based knowledge (chapter 6).
- **Objective 4.** Running different possible scenarios of future land-use and activity development in Flanders, based on recent Flemish spatial policies (chapter 7).

1.4 Structure of the thesis

The next chapter gives an overview of theories of urban form and of spatial models that have been important in studying urban growth during the last half century. Chapter 3 introduces the components and equations of the MOLAND model and the activity-based CA model, and includes a first discussion of the limitations of the current ACA modelling approach as it has been applied to Flanders and Belgium, together with the modifications envisioned in this research. The most important change, making the influence rules partially dependent on travel time instead of Euclidean distance, is discussed in chapter 4. Chapter 5 describes a new technique to produce historical activity and land-use maps with the help of remote sensing data. These historical maps are then used in chapter 6 in a semi-automated calibration of the model. In chapter 7, the new version of the model is applied to simulate a business-as-usual scenario and a land-take neutral future for Flanders. Finally, concluding remarks and an outlook on possible future work are presented in chapter 8.

2. Spatially dynamic modelling of cities and regions

Although formal theories of urban structure were already developed by a number of geographers and economists in the 19th and 20th centuries, it was not until the last decades of the 20th century that computers and a progressive availability of urban spatial data led to the rise of many theories and applications of urban and regional modelling. At the same time urban sprawl became a worldwide environmental threat which can only be better understood with spatial models capable of capturing the underlying mechanisms of urban morphological development at high spatial resolutions. In post-industrial countries the peri-urban space is no longer simply the outer part of a city in development but rather a chaotic mix of urban functions that will never evolve to become completely urbanised (Ravetz et al., 2013). The search for solutions for better understanding and managing urban growth, including the development and enhancement of land-use models, is supported by many governments, e.g. by the European Union (Batista e Silva et al., 2013b). Modellers try to describe, map and simulate the state and evolution of urban form and urban activities, including the processes causing their evolution.

This chapter firstly gives a brief explanation of formal theories of urban form, and continues with the work of two schools (Brussels and Santa Fe) that have played a pioneering role in computer-based spatially dynamic modelling of cities and regions. Next, an overview will be given of different types of regional and/or spatially explicit models that emerged during the last decades and that have strongly influenced the research presented in this thesis. These are urban systems models and land use transportation

interaction (LUTI) models at the macro scale, and cellular automata (CA) and agent-based modelling (ABM) at the micro scale. Cellular automata will be dealt with in greater detail since the most important goal of this thesis is the development of an improved CA model. For the other types of models the focus will be on what CA modelling can learn from these various alternatives.

2.1 Theories of urban form

The first formal land-use theory in history was developed by Prussian social theorist and land owner Johann Heinrich von Thünen in 1826 in his book *The Isolated State*. Von Thünen (1966 [1826]) states that every land owner wants to maximise his profit which depends on the possible yield and the distance to a central city. Due to the distance effect, the von Thünen model consists of concentric zones of different types of agriculture around a city. It assumes a flat and uniform landscape and a market centered transportation system. In the same period, David Ricardo developed a similar theory of land rent, but did not address the spatial consequences for land use (Ricardo, 2001 [1821]).

The first theory of a multiple city system was central place theory, independently developed in Germany by geographer Walter Christaller and economist August Lösch in the 1930s. In these central place systems, cities emerge in regular hierarchic hexagonal patterns. In Christaller's theory the hierarchical principle can be based on the economy (markets), the transportation system, or politics (administrative regions) (Figure 2.1). The system contains a limited number of higher-order centres that offer higher-order goods and services, i.e. those that require a larger number of consumers in order to be profitably provided. Low-order goods and services

are offered by all centres, including higher-order centres. Assuming an isotropic space, i.e. a uniform landscape and distribution of customers, the

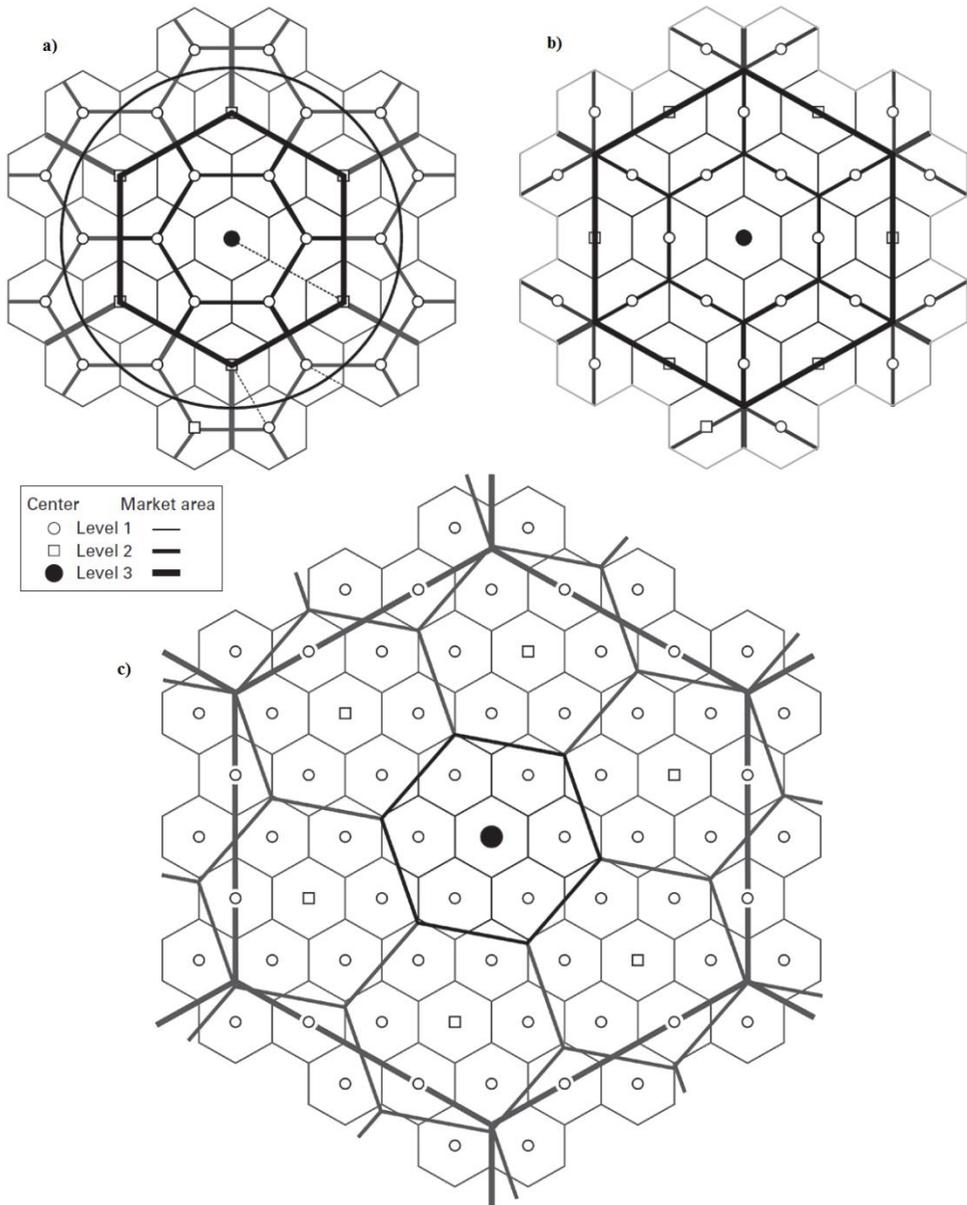


Figure 2.1 Central place system of Christaller based on a) markets, b) the transportation system, c) administrative regions (copied from White et al., 2015). © MIT, 2015.

resulting pattern will be a regular hexagonal geometry of market cities (Lösch, 1954 [1940]; Christaller, 1966 [1933]). While central place theory describes an optimal trade and service economy landscape, German economist Alfred Weber (1957 [1909]) developed in 1909 a mathematical theory for the ideal location of manufacturing establishments based on minimising shipping costs, again assuming an isotropic space and Euclidean distances.

The von Thünen model, the Ricardian land-rent theory, central place theory and the Weber model can be seen as the geometrical (and to an extent mathematical) foundations of spatial economics and regional science. They inspired several classic but simple models of land rent (most notably the Alonso-Muth model) and urban economics (e.g. Fujita, 1989). The Alonso-Muth model (Alonso, 1964; Muth, 1969) states that households want to maximise their utility and in order to do so, they can pay a maximum rent at a given distance of the centre of a city (the central business district). This maximum rent is given by the “bid-rent function”, which decreases with distance. The model therefore predicts increases in land consumption, and thus lower densities at higher distances from the centre (Briassoulis, 2000; Barthelemy, 2016). This model and the land-use models directly inspired by it, again use an isotropic plane and radial symmetry, and unfortunately fail to represent the behaviour of our anisotropic world in a realistic way. These models still produce a concentric zone pattern of monofunctional land uses and treat the urban system as a system in equilibrium. They can be useful as a component of a larger framework (e.g. a land rent model can be an input for a more general urban model), but are too simple to use to simulate spatial patterns themselves. In contrast, most urban modellers agree today that urban form is the expression of a system

that is far-from-equilibrium (White et al., 2015), described in the theories of complex self-organising systems (see next section).

Complex systems theories have led to a new mathematical theory of urban form (Batty & Longley, 1986, 1987; Frankhauser, 1988, 1994; White & Engelen, 1993) since all complex self-organising, far-from-equilibrium systems seem to have fractal structures (Bak, 1994). Fractals are “self-similar” and “scale-free structures” since they appear to have the same structure and properties regardless of the scale (Mandelbrot, 1982). A fractal can be generated through the recursive process of an algorithm, and often takes the form of some power law distribution. Many natural or human systems can be characterised by fractals: coastlines, river systems, trees, ant colonies, or innovation in cities are only a few examples. In a city, “the extremely convoluted edge of the built-up area, the spatial distribution of land uses, the size distribution of internal clusters, the transportation networks, and various local features are all fractal” (White et al., 2015). There are two types of fractal analysis that are often used to characterise urban systems: the radial dimension and the cluster size-frequency dimension.

The radial dimension D_R of a city, or more generally of a land-use cluster can be defined as:

$$D_R = \log A / \log \varepsilon \quad (2.1)$$

with A the area of the cluster and ε the radius of the comprising circle of the cluster. For cities, the centre of the circle equals the centre of the city. Frankhauser (1994) examined the radial dimension of many cities and came to the conclusion that they all have similar values between 1.9 and 2.0. Based on experiments on real cities, White and Engelen (1993) make a difference between the radial dimension of the inner city and an outer zone

in which urban growth is taking place, and where D_R is considerably lower than in the city core. As such, a city has a “bifractal structure”.

A second type of fractal analysis involves plotting a graph of the cluster size-frequency dimension (Figure 2.2). The logarithm of the number of patches of urban land (or more specifically residential land) of a specific size is a decreasing linear function of the logarithm of the size of the patches, or:

$$N = ks^{-D_C}$$
$$D_C = \frac{\log k - \log N}{\log s} \quad (2.2)$$

with N the number of clusters of size s , k a constant, and D_C the cluster size-frequency dimension. This rule, firstly observed by Zipf (1949), has often been used as a best possible validation of urban models (e.g. White & Engelen, 1993; White, 2006b; de Nijs, 2009; White et al., 2012), although it is – just as with radial analysis – only an indication that the model results probably make sense since the model produces cities with forms, cluster distributions, and fractal properties as observed in the cities of the study area. Moreover, these fractal properties cannot always be observed in an urban system: e.g. the cluster size-frequency graph in Belgium is less linear since there is a shortage of small clusters due to excessive linear development (Figure 2.2 b). But a well-built and well-calibrated urban model should be able to reproduce fractal properties if the urban system being studied has them too (White et al., 2015).

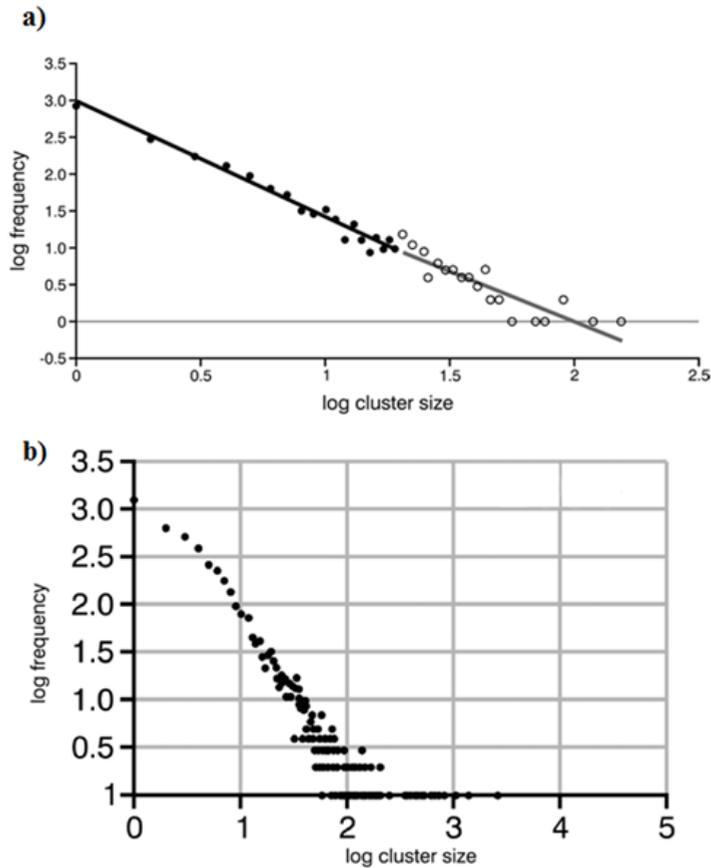


Figure 2.2 Cluster size-frequency diagrams for a) the Netherlands and b) Belgium (after White et al., 2015). © MIT, 2015.

2.2 Two schools that influenced spatially dynamic modelling

The research group of Ilya Prigogine, a Belgian chemist and Nobel prize winner at the Université Libre de Bruxelles, studied far-from-equilibrium systems, including social systems. Unlike classical science, the science of being, where time is reversible, Prigogine (1980) studied the emergence of new macro-scale structures: the science of becoming, where time is not reversible. To create new macroscopic structures, a high amount of energy is needed, and as a result entropy (or “pollution”) will increase elsewhere, as stated by the second law of thermodynamics (Prigogine & Stengers, 1984).

Self-organising systems are pushed away from thermodynamic equilibrium by inflows of energy, and they export entropy. An emerging regional system of cities can also be considered as such a self-organising system: the needed energy consists of food, electricity, construction materials, etc., while all types of waste are the entropy (White et al., 2015). The resulting cities have a specific macroscopic pattern of built-up land use. But the emergence as a result of energy input could lead to different patterns, depending on knowledge, physical and social constraints, and the city's history. And to an extent also by chance. The collection of all possible emergent urban patterns is generated by a series of bifurcation points. Each possible pattern would have its own history, of which only one has really happened, or will really happen in the future (White et al., 2015). Hence, urban models should ideally be able to reproduce this bifurcation structure by including stochasticity. Only in that way they are fully able to simulate different possible futures and to leave the path of deterministic classical science.

Alternatively, an algorithmic approach was developed at the Los Alamos National Laboratory. As White et al. (2015) state, "The key to modelling cities as self-organising systems is to treat them not as artifacts but as processes – which means embedding the model in time. The natural language of modelling is the algorithm ... which output can only be known by executing it, step by step, *in time*." Stanislaw Ulam and John von Neumann created in the late 1940s with the cellular automaton (CA) a simple but powerful algorithm that is able to generate self-reproducing structures with non-linear dynamics (von Neumann, 1966). The most famous CA is the "Game of Life", invented by British mathematician John Conway around 1970. It has only two states (dead or alive) and three simple transition rules (Gardner, 1970). After a short initialisation time, the cells in

the system quickly organise themselves as either stable or mobile structures which cause transformations when they collide. CA models gained popularity in urban modelling and other fields after the discovery that universal computation was possible with the “Game of Life” (Poundstone, 1985).

British physicist and computer scientist Stephen Wolfram, and American computer scientist Christopher Langton made major contributions to the field by defining types of CA rule sets that can generate ordered, chaotic, or highly structured evolving patterns. Wolfram (2002) defined four classes of cellular automata: stable (I), stable but oscillating (II), chaotic (III) and complex interactions (IV). At Santa Fe, Langton (1992) experimented with one-dimensional CA with many cell states and a large cell neighbourhood. He defined a parameter λ which corresponds to the fraction of rules in the rule set that generates a living cell in the next time step. Low values of λ lead to ordered patterns, high values to chaos. The most interesting behaviour appears when λ approaches a critical value: then the CA generates extremely long and highly structured transients with fractal properties; these transients ultimately collapse into either a simple stable structure or chaotic churn. The critical value of λ is not a constant, but depends on the system. As such, good rules for cellular automata create complex structures that are “at the edge of order and chaos” (Langton, 1992).

Stuart Kauffman (1993), a biologist, introduced an alternative for cellular automata at Santa Fe: boolean networks. The spatial system – rather an organism in Kauffman’s research to model life in general – is a collection of nodes that have explicit connections between each other. Inspired by genetic theories, the nodes (cfr. genes) have active or frozen states (cfr.

alleles) which can change to improve their fitness and to obtain an optimal coadaptation with other nodes. Coadaptation is not explicitly present in CA, but is implicitly forced when calibrating the rules (White et al., 2015). The disadvantage of Boolean networks is that the rules and calibration process are less straightforward and further away from the obvious real world logics of urban systems. This also applies to some computer science techniques such as artificial neural networks (see Section 2.4.2).

The Brussels school inspired the most dynamic urban systems models, while Santa Fe fostered the development of a variety of abstract simulation models (but not urban models) using cellular automata. A few modellers also tried to mix both approaches with remarkably interesting results, as in the work of White, Engelen, and Uljee (e.g. White & Engelen, 1997; Engelen et al., 2007) where a cellular automata model is coupled to a regional model inspired by the Brussels school. A third and today highly popular group of urban spatial models originated in computer science and starts from the interactions between individuals or agents to simulate spatial behaviour in multi-agent systems or agent-based models. The next sections will describe these three methodologies of urban modelling in detail.

2.3 The macro scale: urban systems models and LUTI models

System dynamics theory, initiated by a rather abstract model of Forrester (1969), has produced a number of possible models, mostly spatially non-explicit, to simulate the evolving attractiveness of economic centres or regions based on positive and negative feedbacks in the system (White et al., 2015). Alternative system models were created by Roger White (1977), by Alan Wilson and colleagues (e.g. Coelho & Wilson, 1976), by Paul Krugman (1998), and in the agent-based SIMPOP models (with cities as

agents) by Denise Pumain and colleagues (see Pumain (2012) for an overview). In general, growth or decline of a centre or region is a function of the relative attractiveness of a centre and of the distance between centres.

Roger White (1977) developed a regional model of urban trade centres, based on the assumption that their size depends on their profitability, which in turn depends on their actual size, the distance of population to the centre, and competition with other centres. Spatial competition for customers is represented by either a gravity equation or a negative exponential equation. In either case, the equation includes a spatial interaction parameter n_j specific to economic sector j , because spatial interaction is different for high-order and low-order goods and services. Customers are prepared to travel further to reach the first category, as such centrality is essential for high-order goods. For low-order goods proximity is more important, and the presence of many centres near each other will restrict the growth of all of them.

Land use and transportation interaction (LUTI) models examine feedbacks between the land-use system and the transportation system with urban or traffic zones as spatial units. Many early LUTI models have their origins in the *Model of Metropolis*, a double gravity model for urban zones of Lowry (1964), who coupled a population gravity model to a retail gravity model (Wegener, 2004). Other LUTI models have their roots in random utility theory and regional economics, and more recently they are turning often into agent-based approaches (Iacono et al., 2008). LUTI models became a class of frequently used urban spatial models (e.g. ITLUP, MEPLAN or UrbanSim). Although their complexity and level of detail have greatly increased, they unfortunately often still extend existing feedbacks towards the future. Some LUTI models are dynamic (e.g. IRPUD and

DELTA) but they only operate at the regional scale and some important dynamic processes are exogenous. A good recent overview of LUTI models and their equations is given by Vorel (2015).

A quite different approach was developed by Peter Allen and colleagues in Brussels (Allen & Sanglier, 1979; Allen et al., 1984), inspired by Prigogine's theories. In the models of Peter Allen, population grows with the introduction of new technologies and the growth of sectors that support the population until a certain carrying capacity of the region is reached (Allen et al., 1984; Engelen, 1988; Allen, 1997). Local changes might lead to different futures. As such, this transformation model is much more dynamic than system dynamics and LUTI models since its dynamics go beyond built-in feedbacks of the system.

Although all of these models capture the dynamics of regional competition, they lack the spatial detail that is expected by current potential users who are familiar with high resolution GIS data (White et al., 2015). However, when combined with spatially dynamic simulation models at the micro scale, a regional model of demographics and the economic system still has its benefits. Ideally, a simulation model at the micro scale should incorporate demographic and economic long-distance competition more directly, as is succeeded to a certain extent in this thesis.

2.4 Cellular automata models

Cellular automata (CA) models have interesting properties to be used as urban models. Waldo Tobler (1979) firstly suggested to use a cellular grid to simulate geographical dynamics. A general framework for urban CA models was formalised by Helen Couclelis (1985). Other early urban CA models

were developed by Roger White and Guy Engelen (1993), and Michael Batty and Yichun Xie (1994).

CA models account for spatial properties and proximity in a direct way, and they are perfectly compatible with today's high resolution raster data, available from GIS databases and remote sensing products (Poelmans & Van Rompaey, 2010; Santé et al., 2010; White et al., 2015). A CA model is a transformation model, a dynamic geographical information system that is able to track and simulate evolutions in time of raster-based spatial data of complex systems. As such, it is much more capable to deal with gradual urban growth than other raster-based land-change models which are calibrated using descriptive statistical analysis, such as the often applied (Dyna-)CLUE(-s) models of Verburg and colleagues (Verburg et al., 2002, 2008; Verburg & Overmars, 2009).

Computations in a CA model can be relatively fast due to the simple structure, although the speed highly depends on the extent of the study area and the spatial resolution. The modeler should reflect on the trade-off between realism and speed to determine the best suited resolution of the application, also knowing that generalisation of the data to a certain extent might have a positive effect if local errors are wiped out.

2.4.1 Characteristics of CA models

CA models are characterised by the cell space, the cell states, the cell neighbourhood, the time step, and the transition rules.

The *cell space* is the collection of units that contain and exchange information. This is typically a 2D raster of square cells, although early and theoretical models were rather one-dimensional (e.g. Wolfram, 1984). Additionally, the cell space can also be a vector of raster layers where every

cell is characterised by different physical and socio-economic properties (e.g. White et al., 1997, 2012): next to the land use these can be suitabilities for different land uses, zoning maps with planning constraints, accessibility maps and maps with socio-economic activities (population, employment per sector, etc.). A raster of square cells is not required (Couclelis, 1985). Some modellers experimented with polygons as units to enable the direct usage of cadastral information, polygon-based land-use and zoning maps, or buildings as units (Batty & Xie, 1994; O’Sullivan, 2001; Stevens et al., 2007; Moreno et al., 2008, 2009; Pinto & Antunes, 2010; Barreira-González et al., 2015), and with 3D modelling including the height of buildings (Sembolini, 2000).

The *cell states* are a set of two (alive or dead; built or unbuilt) or mostly more categorical values (all possible land uses in a land-use change model). Normally, each cell has exactly one specific state at each time step, except in a few examples of fuzzy CA in which cells have a fuzzy membership of land-use categories (Wu, 1998; Liu & Phinn, 2003; Al-Ahmadi et al., 2009). In more recent models, such as activity-based CA, there can be a vector of cell states and the values can be continuous instead of categorical, or both (White et al., 2011, 2012; van Vliet et al., 2012; Crols et al., 2015).

The *neighbourhood* is a collection of cells around a cell. Standard neighbourhoods in simple or theoretical CA models are often limited to the von Neumann neighbourhood (the 4 adjacent cells in the rook directions) or the Moore neighbourhood (the 8 adjacent cells in the rook and bishop directions). The models of White, Engelen, and Uljee use a large circular neighbourhood (Engelen et al., 1995), which has often a radius of 800 m (e.g. in the MOLAND model: Engelen et al., 2007). The most advanced

neighbourhoods have a calibrated size (Liao et al., 2016), have dynamic size properties (e.g. Moreno et al., 2009; Pinto & Antunes, 2010), are graph-based (O’Sullivan, 2001; Barreira-González et al., 2015) or are even as large as the whole study area (van Vliet et al., 2009; White et al., 2012; Crols et al., 2015).

Changes are evaluated after each discrete *time step*. Most CA models use 1 year as the time step. Variable time steps were suggested to treat specific events with distinct chronologies differently (Couclelis, 1997; Cecchini & Rizzi, 2001). Some CA models calibrated with neural networks immediately jump from the initial to the final state of a calibration period (e.g. Pijanowski et al., 2002; Basse et al., 2014). As such, these are essentially non-dynamic since they do not simulate the gradual evolution towards the final state.

Finally, the *transition rules* are the rule sets that are used to evaluate the impact of the neighbourhoods on the state of all cells in the cell space. In most models, all cell states are updated simultaneously at the end of each time step. While mostly uniform in space and time, some models have specific rules for different subregions (Li et al., 2008), or the rules evolve in time (Clarke et al., 1997), or both (Geertman et al., 2007). Fuzzy CA models use fuzzy transition rules (Wu, 1998; Liu & Phinn, 2003; Al-Ahmadi et al., 2009).

2.4.2 Influential urban CA models and calibration techniques

Several groups of modellers have developed advanced urban cellular automata models during the last decades. Many of them also combine CA with other urban modelling techniques, such as models of regional economics (e.g. White et al., 1997; Lauf et al., 2012) or agent-based

approaches (e.g. Portugali et al., 1997). The two most applied model families are the MOLAND model and its variations by Roger White, Guy Engelen, and Inge Uljee, and the SLEUTH model of Keith Clarke and colleagues. The groups of Xia Li, and of Juval Portugali and Itzhak Benenson, also performed extensive research on all possible aspects of urban CA variants, data sources, calibration and validation.

White, Engelen, and Uljee developed a family of models that has been widely applied. Important test regions included Cincinnati (White et al., 1997), the Caribbean island of Saint Lucia (White et al., 2000), and the Dublin region (e.g. Barredo et al., 2003; Engelen et al., 2007). Stanilov and Batty (2011) have used the model to simulate the development of West London during a period of 130 years (1875 – 2005). While the full model is internationally known as the MOLAND model, some of the more regularly used national applications have different names such as “RuimteModel” in Flanders (de Kok et al., 2012; White et al., 2015) or “Leefomgevingsverkenner” (“LOV” or “Environment Explorer”) in the Netherlands (White & Engelen, 2000; Engelen et al., 2003; Engelen & White, 2007; de Nijs, 2009). The modellers avoided building a purely deterministic model by including stochastic perturbation. As such, they created an approximation of Langton’s edge-of-chaos behaviour that can be observed in real urban systems. This random component introduces heterogenous behaviour: the transition rules can be slightly different in all cells. More precisely, transition potentials V are the core of the model to determine the likelihood for a cell i to turn into a land use j in the next time step. They are based on the neighbourhood effect N , physical suitability S , planning constraints Z , accessibility to the transport network X , and a random perturbation r which deals with uncertainty in actor behaviour:

$$V_{i,j} = r X_{i,j} S_{i,j} Z_{i,j} N_{i,j} \quad (2.3)$$

A macro-scale model coupled to the CA model handles regional demographics and economics, and also constrains regional land-use demands. Most applications of the MOLAND model were manually calibrated, although Straatman et al. (2004) investigated automated calibration techniques. The full model will be discussed in more detail in the next chapter since many of its components and equations are incorporated too in the activity-based CA model used in this thesis.

Another urban CA model with numerous applications is SLEUTH, which was first proposed by Clarke et al. (1997). Urban growth starts from randomly chosen cells, the growth rule depends on the suitability and neighbours of the chosen cell. There are four types of urban growth: spontaneous, diffusive, organic and road influenced. Cells have a suitability for growth depending on the topography (slopes) and the accessibility to highways. The model determines growth rates endogenously and transition rules are updated after each time step. SLEUTH was mostly calibrated with brute-force calibrations (Silva & Clarke, 2002; Dietzel & Clarke, 2007), but a few studies also experimented with genetic algorithms (Goldstein, 2004; Clarke-Lauer & Clarke, 2011). Clarke et al. (2007) compared the calibration results of different applications of SLEUTH, but still comparable cities do not seem to have comparable parameters. Chaudhuri and Clarke (2013) give an overview of all SLEUTH experiments and applications.

In China, a country in which rapid urban growth is omnipresent, Xia Li, Anthony Gar-On Yeh and colleagues improved input data, and calibration and validation techniques for cellular automata land-use models. They extracted past land-use maps using remote sensing and analysed past land-use changes with spatial metrics (Li & Yeh, 2004). They used artificial

neural networks (ANN) (Li & Yeh, 2002), support vector machines (Yang et al., 2008), and kernel-based learning techniques (Liu et al., 2008) to calibrate non-linear transition rules for CA models, concluding that all approaches perform well but that the last two are more transparent than ANN. Other researchers also adopted an ANN to calibrate CA rules (Guan et al., 2005; Almeida et al., 2008), while others directly simulated land-use change with an ANN model, sometimes with CA-like properties (White, 1989; Pijanowski et al., 2002; Basse et al., 2014), which has the disadvantage of creating black box transition rules. Still, neural networks lead to better calibration results than logistic regression (Lin et al., 2011). Li and colleagues also evaluated ant colony optimisation (Liu et al., 2007) and artificial immune systems (Liu et al., 2010b) as calibration techniques, with the latter technique being successful to perform external interventions when testing different policy scenarios. Moreover, they experimented with genetic algorithms: firstly to calibrate different rules for different subregions (Li et al., 2008), and secondly to do a pattern-based calibration with landscape metrics as error measure (Li et al., 2013). Calibrations with landscape metrics were also tested by Xin Yang and colleagues for a Markov chain-calibrated CA (Yang et al., 2014) and an ANN-CA model (Yang et al., 2016).

Most CA models focus on abstract land-use categories and less on how (groups of) individuals move in an urban context. Portugali and Benenson however, developed high-resolution CA models in which cells are inhabited by moving agents with well-defined economic, ethnic and socio-cultural properties (e.g. Portugali et al., 1997; Benenson, 1998; Benenson et al., 2002). During the years the models and the agents have gained complexity so that the models rather became agent-based models

although the cell remains the unit of computation. From the work of these authors, it is only a small step to put the actors themselves at the heart of the model in a proper agent-based model, which creates numerous opportunities but also some new difficulties.

2.5 Agent-based models

Agent-based modelling (ABM) originated in computer science and more specifically in object-oriented programming (An, 2012) which allows the agents to be defined as objects with specific properties and actions. The first urban ABM was a model of residential segregation by Schelling (1971), but ABMs only became really popular during the last two decades thanks to the strongly increased availability of computer power and individual-based data.

Crooks and Heppenstall (2012) define three essential characteristics of agents: (1) they are autonomous units that can store and process information; (2) (groups of) agents need to have heterogeneous properties; and (3) they actively influence a simulation: they can e.g. achieve goals, observe, communicate, move and/or learn. Especially agent heterogeneity is essential for the quality of an ABM (Matthews et al., 2007). Nevertheless, modellers should be cautious when introducing stochasticity in the behaviour of human agents since human decisions are never purely random (Kennedy, 2012). There is still no general guideline on how to incorporate this human agent heterogeneity (Huang et al., 2014). Agent rules are mostly conditional and can evolve as agents gain intelligence (Crooks & Heppenstall, 2012).

Main advantages of ABM are the high level of detail that naturally enables the simulation of complex systems, and the flexibility of the object-oriented structure. An ABM can be scaleless, or different types of

objects and behaviour can be represented at different scales (Crooks & Heppenstall, 2012). Although the agents generally do not have a fixed or regular geometry, the background landscape (the biophysical environment, the land use, the road network, etc.) can be fixed and can even be cellular as in a CA model (Parker et al., 2003; Stanilov, 2012). As such, the benefits of ABM and CA are combined. Raster structures are computationally more efficient and can directly use large-scale raster data such as remote sensing derivatives. However, the cell size of raster-based input data can influence the results, and full ABM flexibility can only be obtained with a completely object-based framework (Benenson et al., 2005; Stanilov, 2012). Major challenges are still how to handle efficiently neighbourhood effects, how to update (the shape of) objects in time, and the introduction of the third dimension as an extra source of information (Stanilov, 2012). Benenson et al. (2005) proposed the idea of a general framework called “Object-Based Environment for Urban Simulation” (OBEUS), that could both host ABMs, CA models and combinations of both.

An important disadvantage of ABM is the lack of general applicability of specific ABMs that need specific data (Couclelis, 2002). Heavy computations are also needed to run an ABM for a larger study area, unless agents are grouped into super-individuals which reduces then the power of the above-mentioned advantages (Parry & Bithell, 2012). Next, calibration of ABMs is not straightforward since, just as a CA model, an ABM needs to be able to predict many possible futures. Path dependence is as such an essential element of realism in urban simulation models (Brown et al., 2005).

It is not a goal of this chapter to give an overview of existing urban ABMs, but ABM studies can clearly influence CA theory and applications

in several ways. Most optimisation algorithms for ABM calibration are useful for CA models too, e.g. many modellers adapted genetic algorithms to find suitable ABM parameters (Ngo & See, 2012). Applications with a large agent heterogeneity are particularly useful to examine. In the SOME model of the SLUCE project the residential agents do not only want to be near market centres and jobs, they also have different aesthetic and socio-cultural neighbourhood preferences (Rand et al., 2002; Brown & Robinson, 2006). The authors conclude that urban sprawl is a direct consequence of the heterogeneous preferences of the residents. While implicitly present in the form of stochasticity, most CA models do not account for heterogeneity in residential preferences. According to Parker et al. (2012), CA and ABM models should also directly incorporate the economic importance of land market factors. These factors are “credit availability, interest rates, the strength of demand relative to supply, and institutional details of the land market function”. Most models include no, or only simplified or indirect land market effects. On the other hand, some land market factors can obviously also be an effect instead of a cause of other land-use demand factors such as neighbourhood effects. There still remains a gap to be filled between new modelling methodologies including CA and agent based models and the traditional models based on economic and/or physical principles.

In summary, agent-based models have a structure that is well adapted to simulate all details of urban complexity, but still they are often complex to use themselves. They need heavy computations and are difficult to calibrate when applied to real-world large study areas. Using them in an operational sense therefore still requires more research work.

2.6 Discussion and conclusions

Through the years, urban spatial models have evolved from purely theoretical assumptions of urban form into highly-detailed data-driven simulation models. Increased computer power, an enhanced availability of spatial databases and technological progress in remote sensing were the main driving forces in this evolution. Several types of models emerged out of the pioneering work of early urban modellers. Cellular automata and agent-based models account for most applications of the last decades.

Cellular automata models have a simple but efficient grid structure to simulate spatial phenomena such as the evolution of land use in urban areas. The neighbourhoods of many CA models are rather limited in area. Therefore it is not straightforward to develop applications to larger regions with these models. Furthermore, not all spatial data are raster-based. As Torrens and O'Sullivan (2001) state, the [structural] simplicity of CA models is both their largest quality and their most important constraint. Therefore, CA models can become much more powerful if (1) they are coupled to other models, or (2) if strengths of other model types are used to improve the CA structure and rules.

According to Verburg (2006), (urban) land-use models can and should be linked to related socio-economic and natural system models, model components or input data. When fully linked, the models can together constitute a comprehensive systems model in which bidirectional feedbacks can be simulated. Such comprehensive models are difficult to build, calibrate and run if they need to preserve all the details of their components. Depending on the desired application, urban modellers often choose to simplify or rule out several modules, or to use the output of other specific models (e.g. demographic models) as input in a land-use model without

direct interactions (White et al., 2015). Or vice versa, you do not need a complex intra-urban land-use model with many built-up classes if you only want to study the impact of urban growth on agriculture. Nevertheless, there are several successful examples of land-use models that were linked to natural system models (e.g. Engelen, 2003; Oxley et al., 2004; Hansen, 2010). Additionally, frameworks have been built to efficiently couple the components of different system models (e.g. Claessens et al., 2009; Schmitz et al., 2016; de Kok et al., 2017). Modellers should consider to test these frameworks, at least as a verification of the currently made assumptions.

Alternatively or additionally, elements of other land-use or urban models can be used to improve CA models. Several examples were already mentioned above, most notably the regional socio-economic macro-scale model of MOLAND and the agent-based CA models of Benenson and Portugali. Also LUTI models, which operate traditionally rather at the regional level, have moved towards agent-based approaches. A good example is UrbanSim (Waddell, 2002), an open-source LUTI modelling platform with four types of agents: households, businesses, governments and developers. A disadvantage of this platform are the many linkages that are needed between the separate components (Hagen-Zanker, 2012). Even the system dynamics models of trade centres are becoming agent-based (Rand, 2012), but here the linkages between the micro-level and the macro-level have still to be better developed. Ideally, urban models should be as detailed as needed at each scale, and integrate all these scales. The variable grid structure for CA models (Andersson et al., 2002a, 2002b; White, 2006a) and the adaptive zoning algorithm for LUTI (or other vector-based) models (Hagen-Zanker & Jin, 2012, 2013) make this integration of all scales possible, which will be further discussed in the next chapters.

To conclude, models are a simplified representation of reality and should stay simple and fast enough so that a large number of runs is possible to calibrate the model parameters, preferably also for large regions. The notion of multifunctionality, some heterogeneity between actors or regions, and a high resolution are however essential to keep the complexity of reality in the model. This thesis focuses on enhancing a CA model making maximum use of the simple spatial CA structure, but incorporating multifunctionality at different scales: effects of regional economic processes and the transportation network are present within the CA model itself.

3. The activity-based cellular automata model

This chapter is partially based on:

1. sections 1 and 2 of: Crols, T., White, R., Uljee, I., Engelen, G., Poelmans, L., & Canters, F. (2015). A travel time-based variable grid approach for an activity-based cellular automata model. *International Journal of Geographical Information Science*, 29(10), 1757–1781.

2. the section ‘Robustness analysis’ of: Crols, T., White, R., Uljee, I., Engelen, G., Canters, F., & Poelmans, L. (2012b). An activity-based CA model for Flanders, Belgium. Robustness analysis and improved distance computation in a variable grid representation. In: N. N. Pinto, J. Dourado, & A. Natálio (Eds.), *Proceedings of CAMUSS, the International Symposium on Cellular Automata Modeling for Urban and Spatial Systems*, November 8-10, 2012, Porto, Portugal, (pp. 195–207). Coimbra: Department of Civil Engineering of the University of Coimbra.

3.1 Introduction

Most urban models, including cellular automata (CA) models, focus on the rather abstract and categorical concept of land-use change. Each modelling unit is in one dominant land-use state at each time step in the simulation interval. However, in the real world, even at the typical scale of land-use models (100 m – 1 km), many urban zones are characterised by mixed land uses and activities (e.g. residential and commercial functions). In regions with extensive urban sprawl, like Flanders, Belgium, where ribbon development is common, urban and non-urban functions also become mixed, turning agricultural and natural areas into highly fragmented landscapes (Poelmans & Van Rompaey, 2009).

Agent-based models capture the interaction between different types of activities in a direct way, but need a huge amount of data to realistically represent the agents and their spatial behaviour (Parker et al., 2003). They are also computationally very heavy, especially for running simulations on

large study areas, which can only be solved by aggregating agents into ‘super-individuals’ and resorting to parallel computing (Parry & Bithell, 2012). A more straightforward solution for this is to introduce activities and their interactions into the simple but efficient grid structure of a CA model. Then the density of all activities that are present in each cell of the model provides an appropriate description of the complexity of land use in both dense urban areas and in regions with urban sprawl.

Roger White (2006a) was the first to propose a theoretical prototype to systematically include activities in a CA model: both land uses and their associated activities were represented and forecasted for each individual cell. White et al. (2011) and Van Vliet et al. (2012) defined a model with one activity type per cell, consistent with the dominant land use. Next, White et al. (2012) introduced a multiple-activity CA model that can really deal with multifunctional land use, where each cell has values for all activity types considered in the model. Initial results obtained with this model seemed promising, yet some challenges had to be resolved.

In this chapter, both the CA model of White et al. (1997), also known as the MOLAND model, and the activity-based CA model of White et al. (2012) are introduced. Then, a number of exploratory sensitivity tests are shown that were performed during the initial stage of this study to analyse the main drawbacks of the model of White et al. (2012) that could be addressed in this research.

3.2 Key elements of the MOLAND model

Urban CA models rely in the first place on distance-dependent interactions among land uses to predict the future development of urban regions. These are captured in the neighbourhood effect. The neighbourhood effect is

calculated by means of influence functions that define the effect of each cell in the neighbourhood on the cell for which the neighbourhood effect is being calculated — i.e. the focal cell — where that effect depends on both the distance and the state of the cell. For example, people do not want to live next to factories, but need a job in their larger neighbourhood: as such, industrial land use has a negative effect on residential land use in the immediate vicinity but a positive effect further away. Most significant effects, however, are distance-decay effects that are largely positive nearby and smaller further away: e.g. it is good to have shops nearby your house, preferably as close as possible. The contributions of all cells in the neighbourhood are summed to get the neighbourhood effect on the focal cell for a particular land use (state) of that cell. A neighbourhood effect is calculated for each possible state that the focal cell could have. Traditionally, neighbourhoods are small and only capture short-distance effects up to a maximum of 1 km, but this is clearly not large enough for a lot of interactions (e.g. people needing a job). In the MOLAND model this problem was tackled by coupling the CA component to a regional gravity-based spatial interaction model of the economy and population (White & Engelen, 1997, 2000; White et al., 1997; Engelen et al., 2007).

The MOLAND model makes a distinction between active, passive and static cell states. Active land uses change directly as a result of CA dynamics as forced by exogenous demands: these are land uses such as residential, industrial or commercial. Passive land uses, such as various agricultural or natural land-cover states, change as a consequence of the dynamics of the active land uses — they are taken over or are abandoned by active land uses. Static land uses, like water and parks, cannot change state and are not subject to the CA dynamics, though they may affect it.

For each time step (usually one year) a transition potential V_{Ui} is computed for an active land use U on a cell i , mostly with the following equation:

$$V_{Ui} = r Z_{Ui} S_{Ui} X_{Ui} N_{Ui} \quad (3.1)$$

where r is a random perturbation term, Z_{Ui} is the zoning status for land use U on cell i , X_{Ui} is a measure of accessibility to the transport network for land use U on cell i , S_{Ui} is the suitability of cell i for land use U , and N_{Ui} is the neighbourhood effect. Alternatively, some applications of the model replace ZS in (3.1) by a weighted sum of Z and S . If repulsion dominates the neighbourhood effect, which rarely happens, so that N_{Ui} is negative, then:

$$V_{Ui} = r (2 - Z_{Ui} S_{Ui} X_{Ui}) N_{Ui} \quad (3.2)$$

because the potential should even be more negative for small values of Z , S and X than for values close to 1.

The random perturbation, which is necessary to account for the possible differences in actor behaviour, is drawn from a highly skewed distribution:

$$r = 1 + [-\ln(rand)]^\alpha \quad (3.3)$$

where $\alpha > 0$ is a parameter that controls the skewed distribution, and $rand$ is a uniform random variable with $0 < rand < 1$. Most perturbations are very small, i.e. r is still close to 1. Some cells, however, really get larger potentials than they would have without the perturbation, especially for large values of α . A fixed seed is used in the random generator to make runs recomputable.

The zoning status is a vector of binary legal restriction factors for active land uses for each time step. Alternatively, it is possible to have a gradually increasing zoning status between two specific time steps. This option allows to deal with the anticipation of reality on the legal status. The suitability is a vector of physical and environmental constraints for active and passive land uses (e.g. slope, soil type, inundation risk, air traffic noise...) with $0 \leq S_{Ui} \leq 1$, determined with expert knowledge in a GIS.

The accessibility $X_{U,i,s}$ to a specific transportation network s (e.g. roads) is computed as:

$$X_{U,i,s} = a_{s,U} / (a_{s,U} + d_{i,s}) \quad (3.4)$$

where $a_{s,U} > 0$ is a distance-decay parameter of the importance of transportation network s for active land use U (actually the half-distance of the decay function), and d is the shortest Euclidean distance between cell i and an element of network s . The accessibility X_{Ui} to all transportation networks is then defined as:

$$X_{Ui} = \frac{1 - \prod_s (1 - w_{s,U} X_{U,i,s})}{1 - \prod_s (1 - w_{s,U})} \quad (3.5)$$

where $0 \leq w_{s,U} \leq 1$ is a parameter that defines the relative importance of transportation network s for land use U (Gobin et al., 2009).

The cell neighbourhood is circular and has a radius of eight cells. All cells fall into thirty discrete distance zones. The neighbourhood effect for land use U at cell i is then:

$$N_{Ui} = \sum_J \sum_d W_{JU,d} C_{J,d} \quad (3.6)$$

where $W_{JU,d}$ is the influence weight of land use J on land use U in distance zone d , and $C_{J,d}$ is the total number of cells with land use J in distance zone d .

For each cell, the potentials for all land uses are ranked, and then all cells are ranked by their highest potential. The transition rule of the CA model is then to give each cell the land use for which it has the highest potential until there is no more demand for that land use, starting with cells with the highest rank. Thanks to influence weights at zero distance (i.e. inertia factors) and high potentials for the already present land use, most cells will not change state. Theoretically, however, all non-static land-use cells could change state at each time step. A full description and discussion of the CA model can be found in White et al. (1997, 2015).

The process described above determines the local change while the quantity of change and regional growth of population and employment are handled by a regional, gravity-based macro model. In early versions the macro-scale model was based on an economic input-output model and a simple demographic model for the whole study area (White & Engelen, 1997). In most later applications a model of regional competitiveness is used, inspired by Warntz (1965) and Batty (1976). Briefly summarised, the attractiveness of a region i to host a certain activity K (population or economic activity in sector K) depends on the total of all already present activities in the region and the average potential for activity K . Activity can migrate from a region j to i depending on the amount of K already present in j , the attractiveness for K in i relative to the total attractiveness of all regions, the distance between j and i , and the inertia of activity K , i.e. the

propensity not to migrate. Extra activity K is attributed to a region depending on its size and relative attractiveness for K . Changes in the amount of activity K are translated into a demand for the corresponding land use on the basis of the current mean density of cells with activity K in the region, as modified by the amount of land still available after claims by competing activities. A detailed description of this regional activity model can be found in Engelen et al. (2007) or White et al. (2015). A disadvantage of the approach is that a large number of parameters have to be calibrated, both inside the regional model as well as in the links with the CA microsimulation model.

3.3 Extending the neighbourhood: computations in a variable grid structure

The coupling of a macro model with a CA model had several disadvantages: additional parameters are needed to link the models, the regions may be large but each region is represented by a single point, typically the centroid, not all levels of spatial interaction are represented, and finally it is still impossible to have several activities in a single CA cell (White et al., 2012). A solution for including multi-scale interactions in the model would be to change the topology of the CA model, which has an influence on its dynamics (Baetens et al., 2013), so that long-distance effects can be efficiently included.

Andersson et al. (2002a, 2002b) introduced a variable grid representation of the cell neighbourhood in order to make it computationally feasible to expand the neighbourhood to include the entire modelled area, so that all cells could have an effect on the possible land use of each individual cell. White (2006a) extended the variable grid approach by including

activities in it. For the purpose of calculating the neighbourhood effect, cells, and their associated land uses and activities, are aggregated into increasingly larger supercells the greater the distance from the focal cell. The expansion factor is three, which gives a set of nested Moore neighbourhoods. In other words, supercells consist of 3^{2L} unit cells and have a resolution of 3^L times the resolution of the modelling grid, where L , an integer, is the level number of the variable grid (Figure 3.1). This approach keeps calculations fast and simple. In the cell neighbourhood, only the cells belonging to the immediate Moore neighbourhood ($L = 0$) have the unit cell resolution R — i.e. the cell size of the CA grid. The Euclidean distance d_{ij} between the centroid of a focal unit cell i and the centroid of one of its variable grid neighbour cells j can only have specific discrete rook or bishop values since the resolution of variable grid cells increases by a factor of 3 for each level. Therefore d_{ij} can only be R , $\sqrt{2} R$, $3R$, $3\sqrt{2} R$, etc.

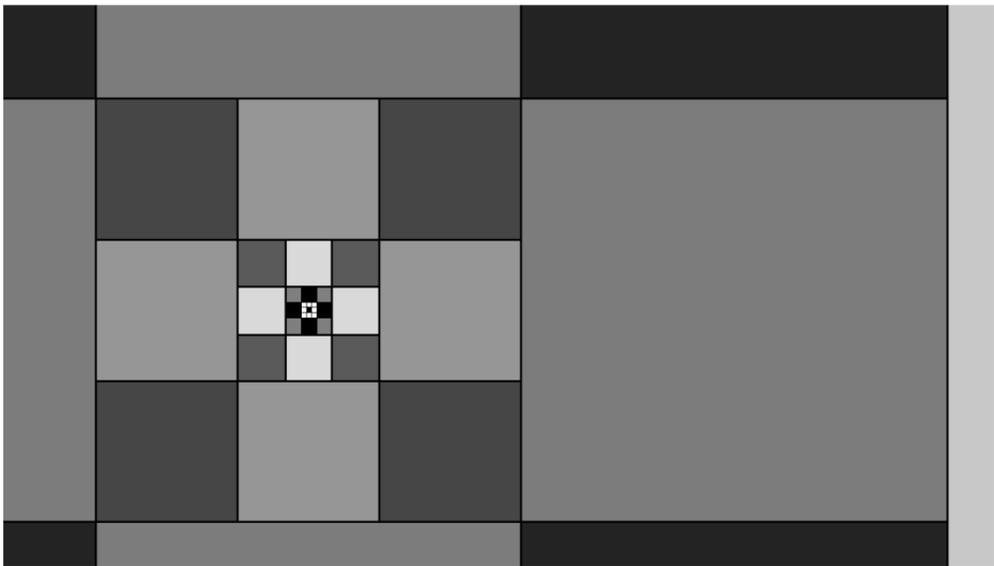


Figure 3.1 Structure of the variable grid.

These concepts were further developed by Van Vliet et al. (2009), who discuss the use of a variable grid approach for land-use change only, and by White et al. (2011, 2012) who developed the full activity-based model.

3.4 The multiple activity cellular automata model

The modelling approach proposed by White et al. (2012) constitutes a variable grid multiple activity-based cellular automaton. As in the MOLAND model, it makes use of active, passive and static land uses. However, in contrast to MOLAND, land-use demands for each active category are needed as input data since there is not longer a regional macro-model to compute these endogenously. The demand comes from historical land-use maps in a historical calibration. In simulations towards the future it can be the result of another land-use change model in business-as-usual runs, or it can be defined in policy scenarios.

By definition, there exists a one-to-one relationship in the model between active land-use types and activity types. In an active land-use cell, primary activities A_P are defined as those associated with the dominant land use U_P (e.g. population in residential land use). However, each cell can also have non-zero values for activities that are primarily associated with other land uses (e.g. employment in residential land use). Such activity is referred to as secondary activity. Passive and static land-use categories have no associated primary activity, but may host secondary activities; for example, agriculture cells may have a resident population. If an active land use cannot be conceptually linked to an activity (e.g. when natural land uses are forced by the model to grow), then it is possible to use the land use directly as the activity: the activity value is set to one when the land use is present in the

cell, otherwise it is zero. We call this an “area-based active land use”.

Several factors are incorporated in the model to determine future activity values, with the neighbourhood effect being the most important one. Within the variable grid approach this effect contains the influence on a cell of all activities throughout the entire study area. The weights W of the influence functions (Figure 3.2) are expressed in the model as functions of log-base 3 cell distances L_{ij} :

$$L_{ij} = \log_3(d_{ij}/R) \quad (3.7)$$

$$W_{JK,d_{ij}} = f_{JK}(L_{ij}) \quad (3.8)$$

where $W_{JK,d_{ij}}$ is the weight given by the influence function f_{JK} for the influence of activity J on activity K at distance d_{ij} , and R is the resolution of the CA grid. The possible rook or bishop values of L_{ij} are then 0, 0.315, 1, 1.315, etc. Inertia factors (weights at zero distance) are not part of the neighbourhood effect in the variable grid but are directly added to the land-use potential (see below) since $\log_3(0)$ cannot be defined.

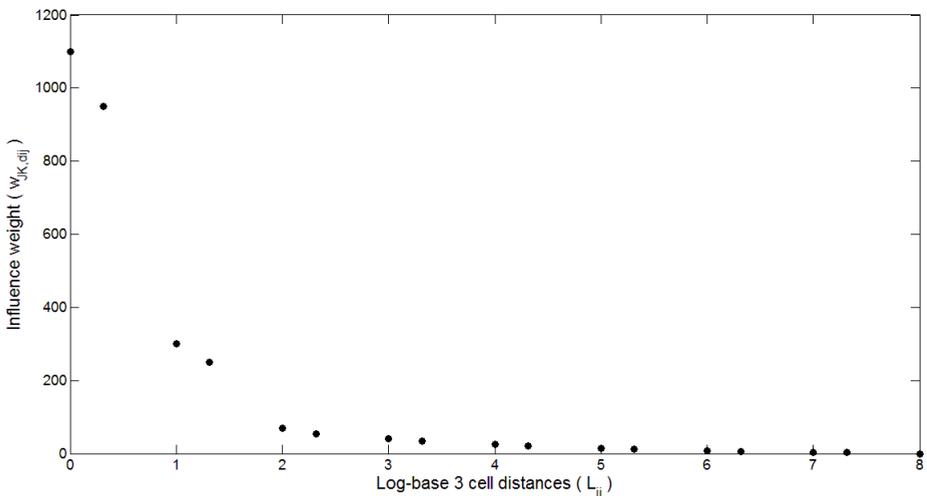


Figure 3.2 Example of a neighbourhood influence function for the variable grid activity based CA model.

The activity potential V_{Ki} for an activity K on a cell i is computed with equation (3.1) of the MOLAND model where a land use U is substituted by an activity K . Negative potential values are by definition put to zero in the ACA model. This is necessary since relative (changes in) potential value(s) are used in equations to update activities, as defined below. Hence, equation (3.2) is not used in the ACA model. All the factors of equation (3.1) are computed as in the MOLAND model, except the neighbourhood effect N_{Ki} on cell i for activity K which is now:

$$N_{Ki} = \sum_J \sum_j W_{JK,dij} (T_{Jj} / T_J) \quad (3.9)$$

where T_{Jj} is the total activity J on cell j , and T_J is the total activity J in the study area. All values of T_J at all time steps have to be provided as input data. These can be census data or future population projections of an external demographic model. Relative activity values are used in this equation since different activities and land uses might have different units. Only a limited set of W_{JK} have non-zero values since not all activities influence each other significantly. Moreover, runtime can strongly increase for a larger number of modelled influences.

Next, the land-use transition potential VT_{Ki} for the associated active land use U_K on cell i is calculated as:

$$VT_{Ki} = D_{Ki} (V_{Ki})^{m_K} + (I_{Ki})^\rho \quad (3.10)$$

where D_{Ki} is a factor representing diseconomies of agglomeration, accounting for the effect of congestion and high land prices on location decisions, caused by a large concentration of population, and m_K is a parameter to be calibrated for each activity K with $0 < m_K \leq 1$. Activities K

with high values of m_K will more easily occupy new locations in (mostly central) parts of the study area in which there is strong competition between land uses. I_{Ki} is the inertia value for activity K in cell i , which is the neighbourhood effect (equation (3.9)) at zero distance, representing the tendency of land uses and activities to remain fixed because of relocation costs, while ρ , the secondary activity power, is a parameter ($0 \leq \rho \leq 1$) to decrease the effect of inertia for secondary activities. For primary activity, i.e. if the cell already has land use U_K , then $\rho = 1$. The inertia of activities not associated to the current land use will then be lower. During the calibration research of this thesis, experiments with activity-dependent parameters ρ_K only had limited success. Therefore, unless indicated otherwise, ρ is the same for all activities. After the computation of all land-use transition potentials, the same land-use transition rule is used as in the MOLAND model (see section 3.2).

The diseconomies of agglomeration D_{Ki} represent the effect that in regions with a high demand and competition for land, it becomes more difficult to convert cells into another land use, mainly due to higher land prices. However, once converted, high values of activity will be present thanks to the high activity potentials. In the model, the diseconomies are a function of the relative neighbourhood effect for population at the regional scale in cell i . If a threshold, determining which cells are affected by the diseconomy effect and controlled by a parameter $\varepsilon > 0$, is exceeded, then exponents $\lambda_K \geq 0$ determine the importance of diseconomies specific to activity K :

$$N_{crit} = \varepsilon \overline{N_{pop}} \quad (3.11)$$

$$D_{Ki} = \frac{1}{(\max[\kappa_K, (1 + N_{i,pop} - N_{crit})])^{\lambda_K}} \quad (3.12)$$

where $\overline{N_{pop}}$ is the mean neighbourhood effect on population (optionally only the effect of population on population) updated after each time step, and $N_{i,pop}$ is the neighbourhood effect on population in cell i for levels L of the variable grid not smaller than a level threshold b ($L \geq b$) since the diseconomy effect operates at the regional scale. The parameters κ_K can control the upper limits of D_{Ki} and even introduce economies ($D_{Ki} > 1$) instead of diseconomies ($0 < D_{Ki} < 1$) when $\kappa_K < 1$. However, only experiments with $\kappa_K = 1$ for all K , i.e. excluding economies, have been successful.

The input scenario defines the total amount T_K of activity K to be located in the study area at each time step. A parameter Q_K determines the proportion of T_K to be distributed as primary activity A_K on the associated land use U_K :

$$A_K = Q_K T_K \quad (3.13)$$

The remainder of T_K is distributed to cells of the other land uses as secondary activity B_K on the basis of compatibility factors c_{KJ} representing the compatibility of activity K with land use U_J :

$$B_K = (1 - Q_K) T_K \quad (3.14)$$

$$B_{KJ} = B_K [(c_{KJ} n_J) / \sum_{(K' \neq K)} (c_{KK'} n_{K'})] \quad (3.15)$$

where B_{KJ} is the total activity K to be located in cells with land use U_J , and n_J is the number of cells with land use U_J .

The updated activity values of all cells are determined in four steps: the first two steps allocate activity in proportion to relative potential values, the last two are rescaling operations to keep all activity totals and proportions consistent with the input scenario. Firstly, the allocation of activity in cells with a changed land-use state is in proportion to the relative value of the activity potential V_{Ki} within all cells of this land-use state. Respectively, for cells i with a changed land-use state to U_K , and cells j with a changed land-use state to U_J :

$$T'_{Ki} = A'_{Ki} = A_K [(V_{Ki})^{\gamma_K} / \sum_{(i \in U_K)} ((V_{Ki})^{\gamma_K})] \quad (3.16)$$

$$T'_{Kj} = B'_{Kj} = B_{KJ} [(V_{Kj})^{\gamma_K} / \sum_{(j \in U_J)} ((V_{Kj})^{\gamma_K})] \quad (3.17)$$

where A'_{Ki} is an intermediate estimate of primary activity K in cell i , B'_{Kj} is an intermediate estimate of secondary activity K in cell j , and $\gamma_K > 0$ is a parameter to be calibrated for each activity. High values of γ_K will immediately put a large amount of activity K in the changed cells. This is useful in regions where new built-up cells have a higher activity density than most existing cells. Since high secondary activity values in changed cells can be an undesired by-product of equation (3.17), experiments were done in this thesis in which secondary activity powers ρ_K are used to lower the effect of γ_K :

$$T'_{Kj} = B'_{Kj} = B_{KJ} [(V_{Kj})^{\rho_K \gamma_K} / \sum_{(j \in U_J)} ((V_{Kj})^{\rho_K \gamma_K})] \quad (3.18)$$

The intermediate estimates T'_{Ki} of all cells i whose land-use state did not change, are equal to T_{Ki} at the previous time step. Next, White et al. (2012)

updated the activity in all cells with relative changes in the activity potential:

$$T''_{Ki} = \left[\frac{(V_{Ki})^{m_K} / \sum_i ((V_{Ki})^{m_K})}{(V_{(t-1)Ki})^{m_K} / \sum_i ((V_{(t-1)Ki})^{m_K})} \right]^{\tau_K} T'_{Ki} \quad (3.19)$$

where $V_{(t-1)Ki}$ is the activity potential for activity K in cell i at the previous time step, $0 \leq \tau_K \leq 1$ is a parameter to be calibrated for each activity, and m_K is the same parameter as in equation (3.10). This equation, that assumes that changes in activity density can be directly related to changes in activity potential, is also used in the model applications of this thesis up to chapter 4. The calibration research (chapter 6) resulted in two newly proposed alternatives for equation (3.19) to allow for more densification in large cities, since fast-growing dense areas turned out to be underestimated with equation (3.19). The first option only updates T'_{Ki} in cells with a high neighbourhood effect for population: if $N_{i,pop} > \zeta_K \overline{N_{pop}}$, then:

$$T''_{Ki} = (T'_{Ki})^{\chi_K} \quad (3.20)$$

where $\zeta_K > 1$ (a threshold) and $\chi_K > 1$ (the densification exponent) are parameters to be calibrated for each activity K . This option is especially useful to have a much larger increase of population in neighbourhoods in which population is already highly present since it assumes a direct link between current activity values and activity growth. A disadvantage is that there is no specific adjustment for all cells in which the neighbourhood effect for population is smaller than the threshold. Alternatively, the activity can be adjusted in proportion to the relative neighbourhood effect for

population, making activity growth directly dependent on the most important component of the activity potential:

$$T''_{Ki} = [(N_{i,pop} / \overline{N_{pop}})^{\tau_K}] T'_{Ki} \quad (3.21)$$

This equation creates a more coherent growth pattern since it is applied to all cells, but τ_K has to be carefully calibrated to avoid an exponential increase of activity in a few centres at the expense of all other locations.

No matter which alternative is chosen, more or less activity can be present in the study area than demanded by the scenario. Therefore, all activity values are rescaled a first time:

$$T'''_{Ki} = [T_K / \sum_i (T''_{Ki})] T''_{Ki} \quad (3.22)$$

Finally, all activity values are again adjusted to get the demanded proportions of primary and secondary activity in the study area. All primary activity is first rescaled and all excess or deficit primary activity is then distributed as secondary activity:

$$T_{Ki} = A_{Ki} = (Q_K / Q'''_K) A'''_{Ki} \quad (3.23)$$

$$T_{Kj} = B_{Kj} = B'''_{Kj} + (B'''_{Kj} / B'''_K) (A'''_K - A_K) \quad (3.24)$$

where i is a cell with land use U_K , j is a cell with land use U_J , Q'''_K is the proportion of activity K that is primary activity after executing equation (3.22), and A'''_K and B'''_K are the total primary and secondary activity K after executing (3.22).

After each time step, most parameters (m_K , ε , λ_K , Q_K , c_{KJ} and χ_K) can be updated. In a historical calibration, the values of Q_K and c_{KJ} (how much of each activity is relatively present in each land use) can be extracted out of

the initial and final ground truth land-use and activity maps. Interpolated values can be used for intermediate time steps. Of all other parameters, only the densification exponent χ_K turned out to improve the results when updated in time in exploratory model runs.

3.5 Robustness analysis: Application to the Antwerp area

During the initial stage of the thesis research, the model of White et al. (2012) was tested in an application to the surroundings of Antwerp, Belgium (the Antwerp province and the districts of Dendermonde and Sint-Niklaas) to define possible shortcomings in the model to be addressed in the research. Therefore, an exploratory robustness analysis of the land-use demand, the model parameters and neighbourhood rules, the accessibility equations and the random perturbation was carried out.

The land-use map of the study area was based on the VITO 10 m land-use map of Flanders for the year 2010 (Van Esch et al., 2011). It was aggregated to a 100 m resolution and the number of land uses was highly reduced to 2 active land uses with associated activities (residential associated to population, and industrial and commercial associated to employment), 2 area-based active land uses (protected nature and agriculture) and a few static categories that were combined for visualisation (Figure 3.3).

The goal of a robustness analysis is to find out whether the model behaves as expected under extreme circumstances. The parameter values, influence rules and land-use demands were not calibrated but iteratively changed to test the model behaviour. Four extreme scenarios had interesting results from a modeller's perspective because they helped to identify the behaviour of the model and its strengths and weaknesses in a first, though

simplified, application to the Flemish context. The scenarios were: (1) a strong decrease of built-up land-use categories, followed by a strong increase (or vice versa), (2) atypical neighbourhood rules for residential activities, (3) major influence of railways on the accessibility for residential activity, and (4) a very strong random perturbation. All scenarios were run from 2010 until 2025 or 2040, depending on whether a longer run still made a big difference or not.

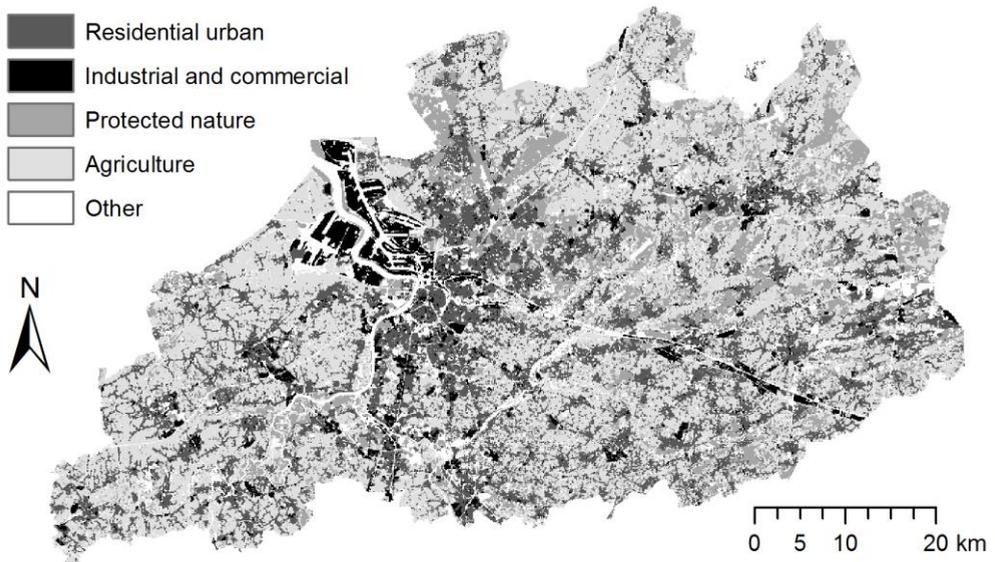


Figure 3.3 Actual land use around Antwerp in 2010.

When the built-up land use strongly decreases (scenario 1), only the centres of the cities remain as residential urban fabric (Figure 3.4a) because higher population densities are present there. However, forecasting industrial and commercial land use based on its associated activity (number of employees) resulted for this scenario in the disappearing of important economic zones. In the large port area of Antwerp for example (in the northwest of the study area), the number of workers per cell is quite limited,

since the employment map used in this initial, exploratory phase of the PhD research attributed an equally low employment density to all cells of the port. Hence, industrial and commercial land use had a small inertia value in the port area. In the case of a shrinking industrial sector (even for a small decrease), the model tended to eliminate port cells prior to other industrial cells, thus ignoring the huge economic importance of the Antwerp port. When the amount of built-up area increased again, the port returned to its original location because the zoning map prevented the area from becoming residential. The residential urban fabric grew around larger towns since they have larger potential values. Smaller villages near the edges of the study area, however, did not return (Figure 3.4b), which indicated that the model had some difficulties to attribute activity to empty cells.

In scenario 2 atypical neighbourhood rules for residential activities were tested. A negative weight for residential activity at medium distance from residential areas was assumed. After an unstable period lasting some years, this resulted after 15 years into a spatial configuration consisting of new urbanised areas, interspaced by a given distance (Figure 3.5).

In the third scenario, raising the influence of the accessibility parameter determining the impact of railways on residential activity did not have a strong effect on the results obtained. With values of neighbourhood rules and diseconomies of agglomeration, both based on the work of White et al. (2012), the interaction rules and even more the diseconomies seemed to dominate the accessibility parameters. Only in the case of lower diseconomies of agglomeration, lower inertia values for all land uses (especially natural categories) and in the absence of a zoning map, growth of new residential urban fabric was a little bit stronger near towns situated in a buffer zone around railways (Figure 3.6). In general however, we could

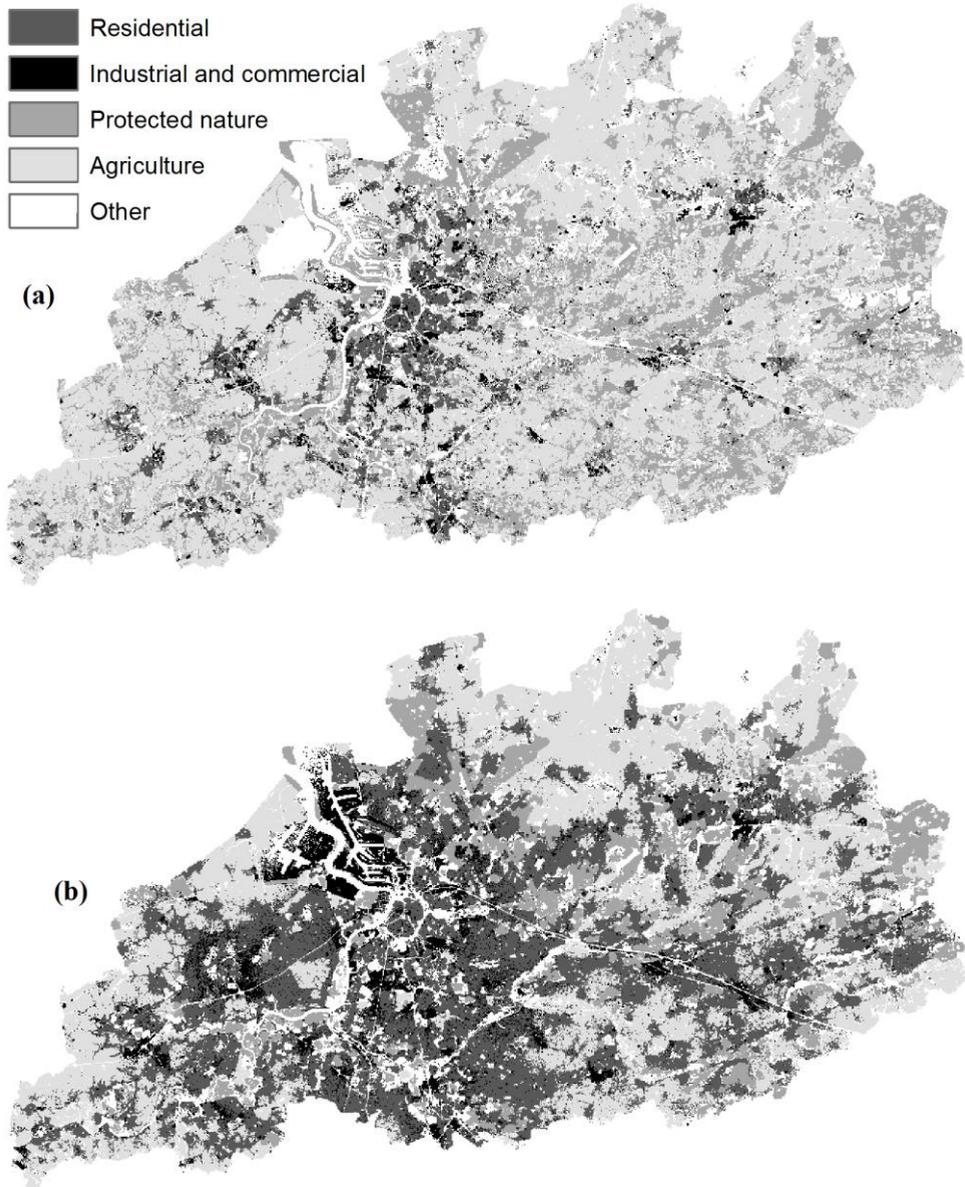


Figure 3.4 Simulated land use around Antwerp in 2025 for (a) an extreme decrease of urban areas, followed by (b) an extreme increase in 2040.

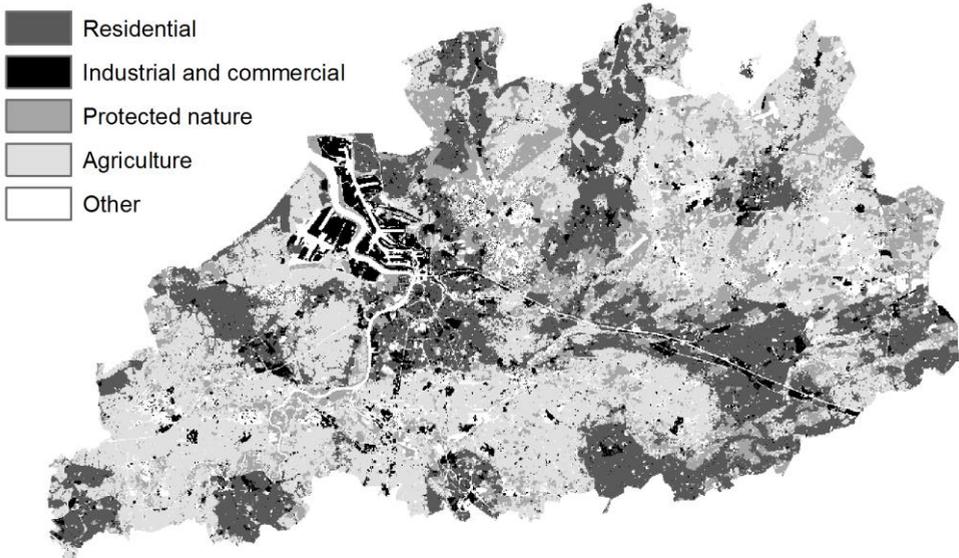


Figure 3.5 Simulated land use around Antwerp in 2025 if residential activity repels itself at medium distances.

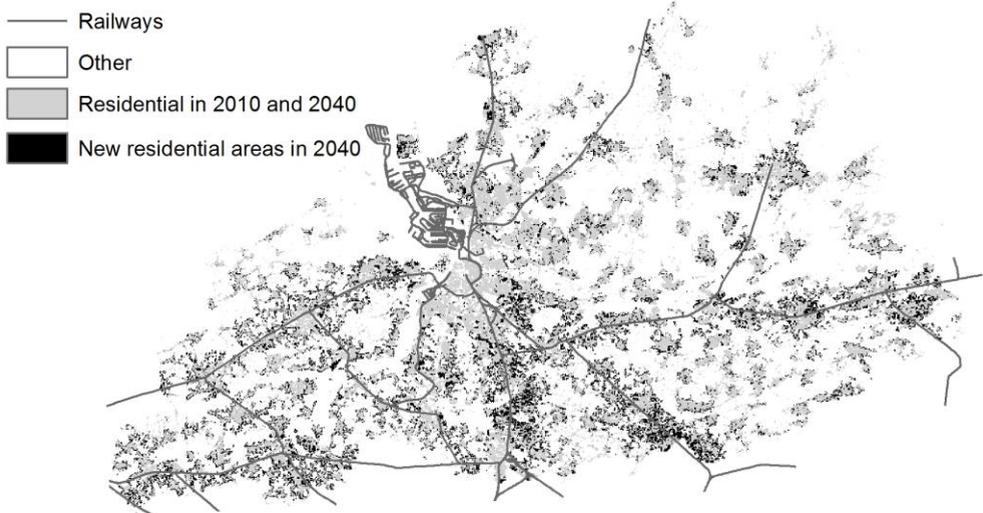


Figure 3.6 Simulated residential land-use change around Antwerp in 2040 with a high railway accessibility parameter, a zero value for other accessibility parameters, low inertia and diseconomies of scale, and no zoning status.

conclude that accessibility parameters in the early-stage model seemed to have an insufficient effect on model behaviour. This was one of the arguments for including network accessibility directly into the neighbourhood rules, as will be discussed in the next chapter.

Finally, we tested the impact of a high random perturbation effect in scenario 4. Extreme random perturbation resulted in numerical instability (when $\alpha > 55$ in equation (3.3)) as some potentials get extremely high values, taking away all activity from all other cells. For values just below this critical value, new residential urban cells are mainly found scattered in agricultural zones, and also in a concentric zone at a certain distance around Antwerp due to the diseconomies of agglomeration effect (Figure 3.7). As in scenario 1, the port area disappears because the employment activity values are low.

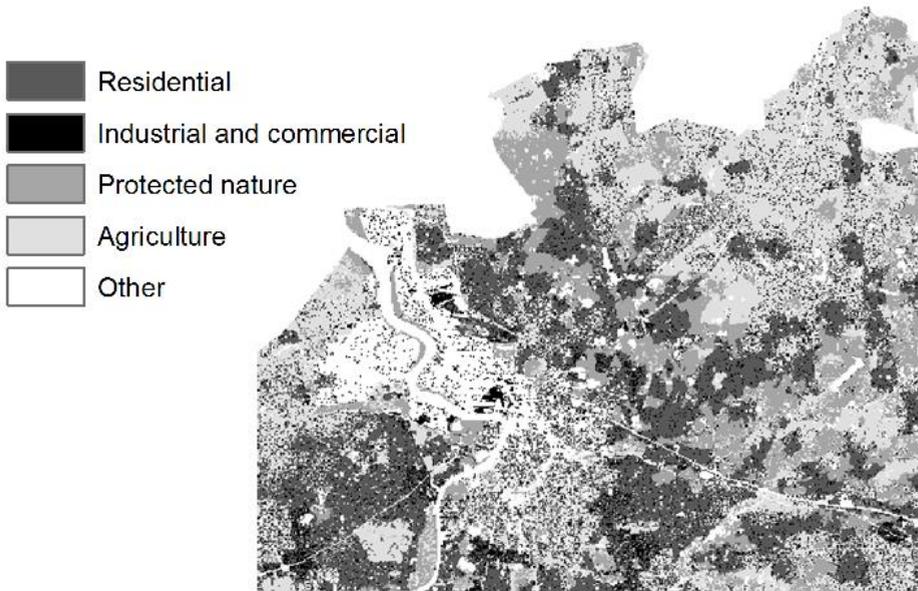


Figure 3.7 Simulated land use north of Antwerp in 2040 with a high random perturbation.

3.6 Discussion and conclusions

The activity-based CA model of White et al. (2012) was without any doubt a big step forward towards the creation of a multifunctional CA-based land-use model, without having to couple different models. Rather than focusing on land-use change only, several types of activities are directly simulated at high resolution, and each cell can host several types of activities no matter which land use it has. Although the simulation of activities at the level of individual cells involves a rather high uncertainty, White et al. (2012) showed that activity totals at the level of small administrative units in the Greater Dublin Region were clearly better simulated than when indirectly computed with the MOLAND model. Many environmental or economic applications are better served by spatially detailed forecasting of population and employment figures rather than only having an idea of the future presence and location of fairly abstract and aggregated land-use types. Crols et al. (2012a) mentioned four interesting applications of detailed forecasts of population and employment: (1) defining the location of new power plants in regions exhibiting population and/or economic growth, (2) assessing the generation of waste products by individuals, (3) defining exposure to health risks for an increasing number of individuals living in highly polluted areas, and (4) predicting the amount of people living close to major roads, a category of residents that is expected to increase by 21% in Flanders in 2030 according to the Flemish Environment Outlook 2030 (Van Steertegem et al., 2009). Obviously, one may envision many other interesting applications of high-resolution population and employment forecasts.

Apart from spatially explicit forecasting of population and employment at cell level, the model proposed by White et al. (2012) has

some other advantages in comparison with its predecessors. Fewer parameters need to be calibrated than in the MOLAND model since there is no longer a need for a separate regional model dealing with long-distance competition between activities. The neighbourhood effect of the CA model is consistently used to capture both regional and local evolution of spatial patterns. The variable grid neighbourhood ensures fast computation enabling the model to be run many times in a calibration procedure. Finally, the activity-based CA model accounts for the negative effects of high land prices and traffic congestion by including diseconomies of agglomeration. The diseconomies have an influence on the land-use transition potentials (see equation 3.10). Hence it will be more difficult to convert non-urban cells into an urban category if activity potentials get too high. But once the cell is converted, it can immediately obtain a high activity density.

The model also has a number of shortcomings, mainly as a result of the direct conceptual relation between land uses and associated primary activities. In contrast with agent-based models and some LUTI models, there is no interaction between different population groups since the activity-based CA model only considers one type of residential activity. On the other hand, the spatial resolution is much more detailed, and socio-demographic information is not easily available at such high resolution. Next, only one type of economic activity can be represented per economic land use. Because of data availability at such high resolution, only employment seems to be a possible choice, yet annual revenue or profit could be much more important to characterise some economic sectors (e.g. ports). In the absence of such data, this problem was solved in later chapters by using a new land-use map in which four economic land uses (industry, wholesale and logistics, retail, services) and four associated employment

values were defined. The port area consists in that map of different economic land uses and the static land use infrastructure, allowing to allocate employment to more specific plots. Finally, although there are fewer parameters to calibrate than in the MOLAND model, the equations to update activities in time needed to be further investigated, since initial runs of the model calibration applied to the whole study area indicated that the model could not deal with high densification of particular urban areas (see chapter 6).

Initial runs with the model in the test area around Antwerp in the early stage of this research revealed one particular issue with the model that had to be resolved. While in the original CA model, neighbourhood effects are considered at the local level only, in a variable grid approach the distance between a focal cell and the centre of a variable grid neighbourhood cell can be extremely long (more than 100 km in the application to Belgium or Flanders). Defining neighbourhood rules based on Euclidean distance in this case seems to result in urban land-use growth patterns that are often too concentric around existing urban agglomerations. Moreover, as mentioned above, even when giving extreme values to the accessibility X in equation (3.1), this seemed to have a limited effect on the results. Therefore, we decided to modify the model by calculating travel times through a multimodal transportation network between the focal cells and their neighbours, and define the influence rules based on travel time instead of Euclidean distance, as will be explained in detail in the next chapter.

4. Integrating travel time into variable grid neighbourhood rules

This chapter (except the appendix) is based on sections 1 and 3-7 of: Crols, T., White, R., Uljee, I., Engelen, G., Poelmans, L., & Canters, F. (2015). A travel time-based variable grid approach for an activity-based cellular automata model. *International Journal of Geographical Information Science*, 29(10), 1757–1781.

4.1 Introduction

One striking simplification in the variable grid models proposed to date (discussed in chapter 3) is that simple Euclidean distances are used in defining the influence rules of a neighbourhood with a size of sometimes several hundred kilometres instead of travel times generated by a transport model component. Many studies have shown that transport and land use are strongly interrelated systems, but the nature of the relation is often debated. Both the road system and the public transport system influence and are influenced by the land-use characteristics of urban areas and their related activity patterns. The standard work of Newman and Kenworthy (1989) suggested a negative relationship between population density and energy use in transportation. Although this work has been criticised as being simplistic (van de Coevering & Schwanen, 2006), other authors have also found statistical relationships between urban structure, activities, access to public transport and commuting patterns (e.g. Schwanen & Mokhtarian, 2005; Sohn, 2005; Næss, 2010; Fuglsang et al., 2011). Attitudes (e.g. Handy et al., 2005) as well as socio-economic characteristics of a country or region, such as income and fuel prices (Giuliano & Dargay, 2006), have also been found to be important. Hansen (2009) used a raster-based approach to show

that many new residential and industrial areas in Northern Denmark have high accessibility to existing towns, and for industry also to motorway exits. In regions with much sprawl, however, it is almost impossible to build an efficient public transport network that can compete with private car usage; nevertheless, public transport can influence dynamics within and between city centres (Camagni et al., 2002). The reality is without any doubt complex, and to some degree specific to countries or regions.

Over the past decades, a considerable number of land use and transportation interaction (LUTI) models have been developed (Chang, 2006; Iacono et al., 2008; Vorel, 2015). According to Wegener (2004), most of these models describe the link between slowly changing systems (land use, networks and buildings), fast changing systems (activities) and immediate changes (transport as such). Hagen-Zanker (2012) compared the standard CA model of White et al. (1997) with well-known examples of LUTI models — in particular MEPLAN (see e.g. Echenique, 2004) and UrbanSim (Waddell, 2002). He concluded that although these three modelling strategies are fundamentally and computationally different, all of them may lead to similar results if their weaknesses are overcome. Unlike LUTI models, most CA models do not have an intrinsic economic or transport component. An activity-based model includes economic activities, but should still be improved by taking transport into account.

Some LUTI models are linked to a stand-alone transport model, others have their own transport subsystem. Aljoufie et al. (2013) coupled a CA-based land-use model, developed with the Metronamica framework and thus based on the work of White et al. (1997), to a stand-alone transport model for the city of Jeddah. The resulting generalised cost of the transport model is used as an input for the accessibility component of the land-use

model. Nevertheless, the accessibility component does not directly influence the distance between cells in the influence rules. Blečić et al. (2011) incorporated distance effects directly into their CA model. They represented the short-distance CA neighbourhood effect and a long-distance accessibility effect separately in what they call a ‘wave model’. Two different signals of land-use influences are propagated between cells to simulate vicinity and accessibility. Vicinity is here represented by a small-neighbourhood Euclidean influence function. Accessibility is modelled as an influence signal that gets weaker with increasing distance.

In this chapter, we use the model of White et al. (2012) to test how travel times can be integrated into the variable grid neighbourhood rules (except for the close neighbourhood) and show how this impacts modelling results. We define different transport network scenarios (a ‘Euclidean scenario’ and four ‘network scenarios’) to compute these travel times with a road-based or a multimodal network. The methodology is applied to the case of Belgium. The resulting model is inspired by the LUTI models discussed above, but it is not a LUTI model itself since the land-use simulations are not coupled back to updates in the transport network.

4.2 Travel time computations within the variable grid CA

White et al. (2012) observed that the calibrated neighbourhood influence rules in their activity-based variable grid CA model divide naturally into two parts: an inner zone with a radius of approximately 1 km, where the influence weights decrease rapidly with increasing distance, and the rest of the study area, where the weights decrease slowly as a function of distance. This can be seen in Figure 3.2. Therefore, in incorporating network-based distances in the activity-based CA model, we decided to deal with the

immediate neighbourhood of a location and more distant areas in different ways. In our approach, the classic Euclidean concept of vicinity still holds for distances within the range of a local CA neighbourhood of approximately 1 km radius, as used in the MOLAND modelling framework (Engelen et al., 2007). For long-distance interactions, however, we introduce time distances through the network. To make this dual approach compatible with the variable grid, the point at which the transition between Euclidean and network distances occurs needs to be between two levels of the variable grid. The level at which the network distances are introduced will be referred to as the *network grid level* or level L_{NG} .

This approach, tailored to the specific requirements of the variable grid CA, is more effective than the alternative of coupling the variable grid CA to a transportation model to provide the required travel time distances. It can offer higher resolution than a typical transport model and thus more accurate distances, while requiring fewer distances to be stored in memory. As an internal algorithm it gives fast run times if updates are desired during the simulation, or if the modeller wants to execute a full model calibration, including transport-related parameters.

Nevertheless, calculating travel times and simulating travel behaviour between a number of widely separated points involves a reasonable amount of computation time. For polygon-based modelling approaches, adaptive zoning can be an alternative approach (Hagen-Zanker & Jin, 2012, 2013). This technique groups polygons systematically and thereby develops a new zone map for each origin with small zones nearby and larger zones further away, based on the idea that exact locations are less important for long-distance interactions. Essentially, the same reasoning lies at the basis of the variable grid approach. Some studies have even

investigated whether network distances, represented by a weighted l_p norm, are just mathematical functions of Euclidean distances (e.g. Berens, 1988; Brimberg et al., 2007). Although these studies led to interesting results for some regions, especially for cities with rectangular road patterns, we believe that these results are not relevant to our model, which must be applicable to a wide variety of network configurations.

Therefore, we define a fixed grid zone system that consists of supercells of a specific level, L_{NG} , but unlike the template of supercells in the variable grid itself, it is not displaced as we move from one unit cell to another. We call this the *network grid*, or NG, and its cells the *network grid cells*, or NG cells. It has resolution $R_{NG} = 3^{L_{NG}} R$, where R is the resolution of the unit cell grid. Note that the resolution R_{NG} is equal to the size of the immediate neighbourhood in which the Euclidean distance calculation is applied.

The network time distance between a unit ($L = 0$) cell within a fixed NG cell and the centroid of a distant supercell ($L \geq L_{NG}$) is then taken to be the network distance between the centroid of the NG cell containing the unit cell and the centroid of the NG cell containing the centroid of the supercell (Figure 4.1). We call the NG cell containing the unit cell the *origin* of the transport computation, and we call the NG cell containing the centroid of the supercell the *destination*. As already indicated, we only have to store network distances between specific needed combinations of NG cells. The only needed ‘destinations’ for an ‘origin’ NG cell are exactly the centroids of all the variable grid neighbour cells of the central unit cell of an NG cell with $L \geq L_{NG}$. Therefore, the results can be stored as an $8 \times (L_{max} - L_{NG} + 1) \times N_{NG}$ matrix, where 8 is the number of possible directions (top, top right...), L_{max} is the highest needed level number of the variable grid, and

N_{NG} is the total number of NG cells. The storage size is significantly smaller than a $N_{\text{NG}} \times N_{\text{NG}}$ matrix, since the value of L_{max} is, even for large regions, often only 6 or 7.

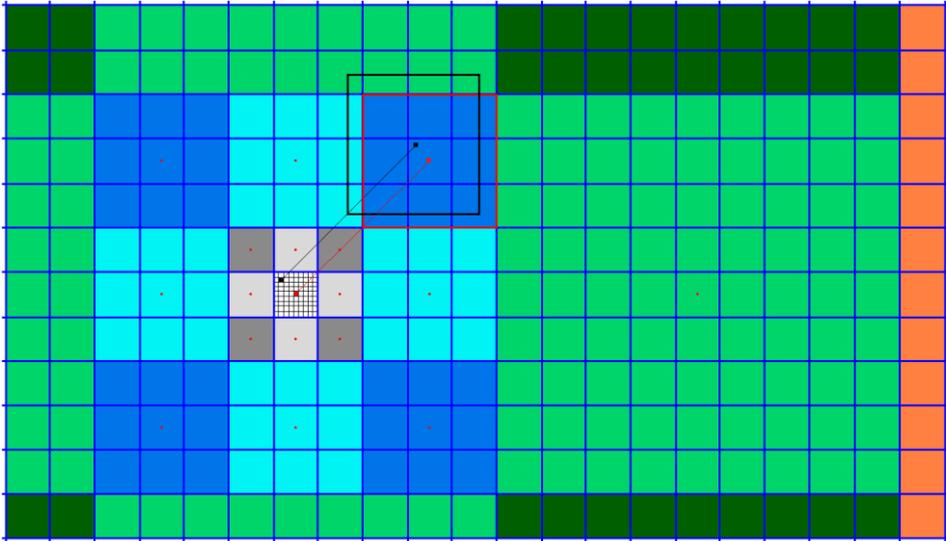


Figure 4.1 Detail of a fixed ‘network grid’ (NG) with resolution 3^{2R} . The model cells with resolution R are shown within the central grey NG cell only. When the distance is needed from a black unit cell to the centroid of its black variable grid neighbour, the distance from the red centroid of the fixed NG cell containing the black unit cell towards the centroid of the red NG cell containing the centroid of the black variable grid neighbour is used.

Some supercells will lie partly outside the area being modelled. Since it does not make sense to calculate distances to points outside the study area, an algorithm was developed to determine the centroid of that part of the supercell that is within the study area. The NG cell containing that centroid is then used to determine the network distance.

For the sake of generality and realism the network used to calculate travel times can be a multimodal one. The one used here has two components: the road network, and public transport — specifically rail in the current application (Figure 4.2). Both components can be combined to

generate weighted time distances, or one specific component can be selected. If roads only are used, the weights are respectively 1 and 0. For public transport, however, displacements are always multimodal since the road network must in general be used to reach a station. These access times by road are calculated with the road network component. Finally, a Euclidean displacement corresponding to a low speed, which we call the *Euclidean speed*, is added to reach the nearest point in the road network. Note that this Euclidean travel time is treated as part of the network distance. To this are added the Euclidean travel times from the origin unit cell to the centroid of the origin NG cell, and from the centroid of the destination NG cell to the centroid of the destination variable grid supercell. The road network itself has known speeds for all road segments, while the public transport network is characterized by an origin-destination travel time matrix, with the origins and destinations being stations on the network.

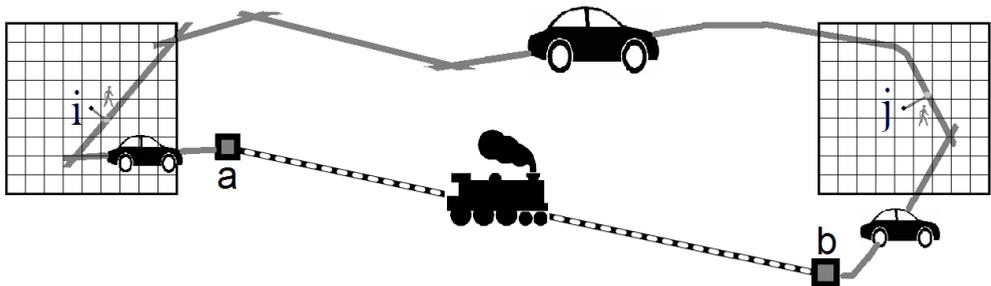


Figure 4.2 The road component and the public transport component can be combined or used separately to generate time distances.

In the road network component, we calculate network travel times with the shortest path algorithm of Dijkstra (1959), applied to all needed origins. We verify whether travelling on a straight line between the two NG centroids at low speed — the ‘Euclidean speed’ — is indeed slower than travelling to and then through the network. This verification is especially

important in cases where the road network consists of major roads only, and the Euclidean distance can be used to represent a more direct route using minor roads that are not included in the network.

The public transport component is mainly designed for rail travel but can also be used for local public transport. In either case, the calculation of effective travel times through the network requires data on (1) average travel time, and (2) frequency of service between each pair of stations on the network. Frequency of service is defined as the total number of possible connections between the stations per hour, either direct or indirect, but earlier departures should not result in later arrivals. The frequency is used to add a penalty to the average travel time, where the penalty is inversely proportional to the frequency. Finally, the road network travel times to reach the stations from the origin and destination centroids in the network grid are added. Thus the multimodal travel time is given by:

$$t_{m,ij} = t_{p,ab} + C/f_{ab} + t_{r,ia} + t_{r,bj} \quad (4.1)$$

where $t_{m,ij}$ is the multimodal transport time between two NG cells i and j , $t_{p,ab}$ is the average public transport travel time between the departure station a and the arrival station b , C is the frequency penalty parameter, f_{ab} is the frequency of service between stations a and b , $t_{r,ia}$ is the road travel time from the origin NG cell i to the departure station a , and $t_{r,bj}$ is the road travel time from the arrival station b to the destination NG cell j .

As already indicated, the travel time distance is a weighted average of the various travel times — specifically road and multimodal in the present case. We include the possibility of giving more weight to the fastest mode with the introduction of a parameter F . If $F > 0$, then it will increase the weight of the fastest mode between two NG cells, and decrease the

weight of the slowest mode. Hence, the parameter introduces the option of developing scenarios where the existence of one fast travel mode between certain locations suffices to make the associated cells relatively more attractive for activities. The weighted network time t_{ij} between NG cells i and j is then calculated as follows:

$$t_{ij} = t_{m,ij} \frac{w_m (t_{r,ij})^F}{w_m (t_{r,ij})^F + w_r (t_{m,ij})^F} + t_{r,ij} \frac{w_r (t_{m,ij})^F}{w_m (t_{r,ij})^F + w_r (t_{m,ij})^F} \quad (4.2)$$

where $t_{m,ij}$ is the multimodal network time between i and j , $t_{r,ij}$ the road network time between i and j , w_m the multimodal weight, and w_r the road weight (with $w_r = 1 - w_m$). The public transport component is fully optional: by choosing $w_m = 0$, only the road network is used. If $w_r = 0$, then travelling by road towards stations is still possible, but a straight displacement at low ‘Euclidean speed’ is used when no station is located near one of the NG cells. This is of course a rather theoretical option to test the abilities of the model. The maximum road distance to reach a station can be specified, and it can be different for major and minor stations. A nearby major station overrides a minor one, except in a specified smaller zone around the minor station.

Travelling via a multimodal network is often slower than travelling by road, thus t_{ij} is often larger than $t_{r,ij}$, which would have an illogical effect on activity potentials of cells close to stations. Therefore, when w_m and w_r are between 0 and 1, equation (4.2) is only used to calculate the weighted travel time for combinations of NG cells where both cells are close to a station. For all other combinations of cells, the calculated road time is penalised proportionally to the importance of multimodal travelling:

$$t_{ij} = \frac{P}{P - w_m} t_{r,ij} \quad (4.3)$$

where $P \geq 1$ is a parameter that defines how much the road time has to be penalised: the smaller the value of P , the larger the increase of the road time.

The final time distance t_{ij} can subsequently be used as an input to determine the weights in the influence functions of the CA neighbourhood effect, except for the close neighbourhood where distances are still Euclidean. This division between the close and the far neighbourhoods generates a continuity problem. The weights of the influence functions are dependent on log-base 3 cell distances, as specified by equations (3.7) and (3.8). The distance d_{ij} of equation (3.7) is still Euclidean if it is shorter than R_{NG} . For the longer distances, we obviously have to convert the travel time t_{ij} into a relative distance d_{ij} that is comparable with Euclidean distances since we do not want influence functions that are in two different intervals dependent on two different quantities. Because the influence weights are values that have to be calibrated for each possible cell distance, it makes sense to work with a continuous range of cell distances. Consequently, we set the shortest calculated travel time between any unit cell in the modelling area and one of its variable grid neighbours equal to the network grid resolution R_{NG} , since R_{NG} is the upper boundary of the range of Euclidean distances. Next, all the other travel times are scaled proportionally as follows:

$$d_{ij} = t_{ij} (R_{\text{NG}} / t_s) \quad (4.4)$$

where d_{ij} is the relative network distance between NG cells i and j , t_{ij} the calculated network travel time between i and j , R_{NG} the resolution of the network grid NG, and t_s the shortest calculated network travel time in the

study area. Finally, all distances are converted with equation (3.7) to be expressed in logarithmic cell distances L , so that these cell distances can be used to find the corresponding neighbourhood influence weights from equation (3.8). For the short Euclidean distances ($L < L_{NG}$), L can only have Euclidean rook or bishop values (0, 0.315, etc.), while for the longer network distances ($L \geq L_{NG}$), any value of L is possible. An example of a mixed Euclidean distance-decay and time distance-decay influence function can be seen in Figure 4.3.

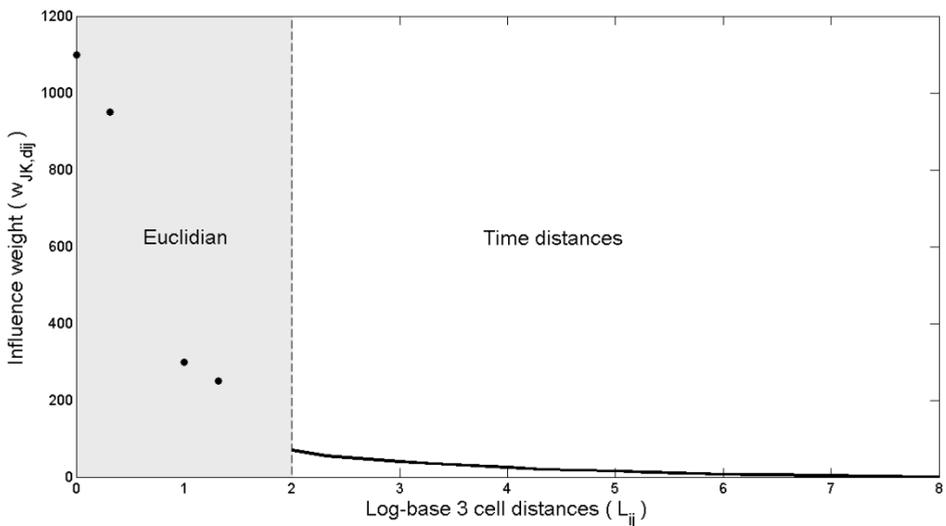


Figure 4.3 Example of a neighbourhood influence function with a Euclidean part for local interactions and a network time-distance part for long-distance interactions. The cell distance axis is made continuous with equation (4.4) and then made logarithmic with equation (3.7).

4.3 Study area and implementation

This chapter is a continuation of the work of White et al. (2012), which focused on the Belgian case. Hence, in this chapter too, we applied and tested the adapted activity-based CA model to Belgium (map: see Figure 4.4). Several studies have discussed the problem of urban sprawl in Belgium (e.g. Poelmans & Van Rompaey, 2009; De Decker, 2011) and the link

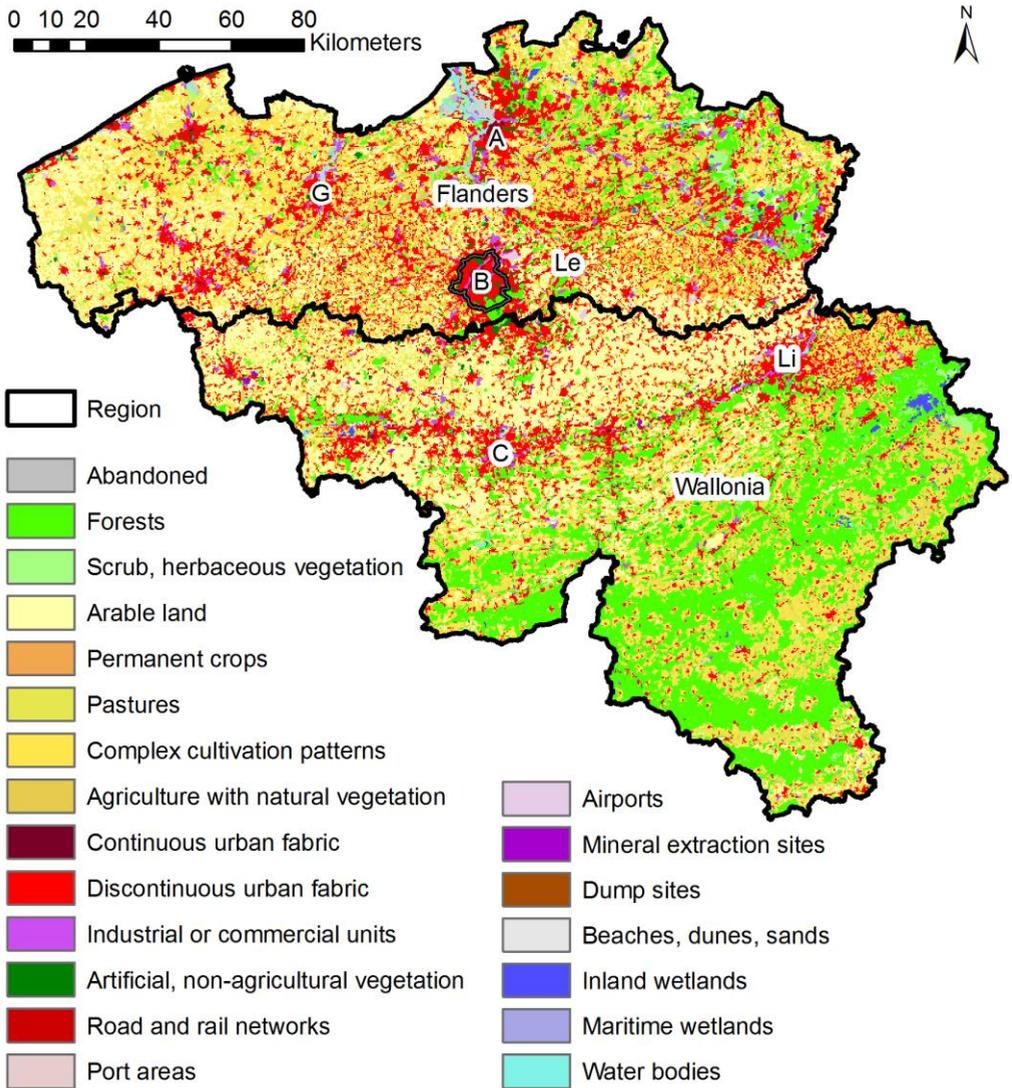


Figure 4.4 CORINE land-use map of Belgium in 2000, with an indication of the regions (Flanders, Wallonia and the Brussels Capital Region (B)) and the largest cities (Brussels (B), Antwerp (A), Ghent (G), Leuven (Le), Liège (Li), and Charleroi (C)).

between urban land use and the transport network (e.g. Vandenbulcke et al., 2009; Boussauw et al., 2012). Political decisions and low land prices have led to large areas of urban sprawl in the northern part of Belgium, and more specifically ribbon development along the roads (Antrop, 2000; De Decker,

2008). Meanwhile, the average commuting distance continues to increase because the population growth in many peri-urban municipalities is higher than the growth of jobs, while the opposite is true for cities (Boussauw et al., 2011).

Studies of both home-to-work travel and general travel in Belgium indicate that individual car travel is far more common than any other transport mode (Thys & Andries, 2011; Vandebulcke et al., 2009). A survey by the Belgian Federal Department of Transport and Mobility indicates that 67% of Belgian employees commute individually by car. Train travel is the second most used mode with 9.5%. It is especially popular for long distances (> 30 km) and for travel to large cities (Thys & Andries, 2011). Thus, a road-based approach seems the most relevant for simulating distance-based interactions in a realistic way, yet train travel towards the most important stations can improve the model significantly. As local public transport, mostly by bus, is less popular except in some cities, we have not yet included it in our analysis. Hence, we compare five different scenarios for determining distance: (1) the Euclidean distance approach of White et al. (2012); (2) road-based network distances; (3) a ‘congestion’ scenario, which uses road-based network distances with congestion effects in the central area of Belgium; (4) a rather theoretical ‘train only’ scenario with multimodal weight $w_m = 1$, which means that Euclidean paths with low speeds are used when train travel is not possible; and (5) a ‘choice’ scenario with road weight $w_r = 0.7$, and multimodal weight $w_m = 0.3$.

Road data come from the NAVSTREETS database for Belgium, the Netherlands and northern France, made available by the Flanders Geographical Information Agency, AGIV. We exclude the least important

local roads (functional class 5) and estimate average speeds on the basis of legal maximum speeds for all road segments as provided in the database (Table 4.1). The ‘Euclidean speed’ is given a low value of 5 km/h. The ‘congestion’ scenario is largely based on a report of the Flemish Traffic Study Centre (*Vlaams Verkeerscentrum*: Hoornaert, 2014). It generally assumes lower speeds in the central area of Flanders between Ghent, Antwerp, Brussels and Leuven on all roads (by downgrading them to a lower road category), and lower motorway segment speeds based on saturation statistics.

Table 4.1 Average modelling speeds for road segments in km/h.

Link type	Legal max. speed (reality)	Average speed (model)
Motorway	120	100
Express road, urban motorway	90-100	85
Major road	90	70
Secondary road	70	55
Local road	50	40
Urban low speed zone	30	30
‘Euclidean speed’	-	5

An origin-destination train time matrix could not be provided by the Belgian National Railway Company (NMBS/SNCB); therefore we reconstructed it from their website for the most important origin-destination pairs. The resulting matrix contains average travel time and frequencies per hour for weekday connections in 2013, between the 18 largest major stations (> 8000 users per weekday in 2009) and all 116 stations (major and minor) with more than 1000 users per weekday in 2009. Travel times and

frequencies are for non-peak hours. Peak-hour connections might be more frequent, but there are also more delays during the peak for which we cannot account anyway. The data can be used to start model runs from the year 2000, since there were no major changes in the train schedule during the period 2000-2013. All missing links are handled as origin-destination pairs where travel by train is not possible. Since train commuting is mainly towards the big cities, we believe that this limited matrix covers the vast majority of the important train connections for a land-use model. After initial tests, and following observations by Thys and Andries (2011), we used a maximum distance of 10 km for the access and egress legs by car to and from train stations in multimodal travelling. A major station overrides a minor one except in the NG cell of the minor station itself.

Land-use data come from the Corine land-use / land-cover data set for 2000, which was aggregated to a 300 m grid (Figure 4.4). Activity maps for population and employment were reused from White et al. (2012). There are three urban categories in the Corine map that were associated with primary activities: the discontinuous urban fabric (associated with population), the continuous urban fabric (associated with ‘urban’ employment: wholesale, retail, hotels and catering, and finance), and industrial or commercial units (associated with all other employment). As our main aim is to assess the impact of different transport scenarios on the model outcome, and compare the results that are obtained with a simple Euclidean scenario, we ran simulations from 2000 to 2060 with the same parameters, rules, land-use area changes, and population and employment growth (Table 4.2) as those used in White et al. (2012), even though that paper indicated that this growth is somewhat excessive, especially for the discontinuous urban land use. However, a large land-use growth helps to

visualise transport scenario differences. Population and employment predictions were extrapolated from data of the Federal Planning Bureau.

Table 4.2 Land-use and activity growth scenarios, and parameter values of the application to Belgium that were copied from White et al. (2012).

Active land use or activity	Value in 2000	Value in 2060
Discontinuous urban fabric (cells)	56,051	101,741
Continuous urban fabric (cells)	527	570
Industrial or commercial units (cells)	5540	6000
Population	10,251,249	12,662,761
‘Urban’ employment	855,822	891,061
Other employment	3,140,781	3,850,000

Since the modelling resolution is 300 m, we could have used $L_{NG} = 1$ so that the boundary between Euclidean and network travel would be 900 m, which is close to the ‘ideal’ boundary of 1 km. However, in order to keep the computation time reasonable for our large study area, we chose $L_{NG} = 2$, which means that $R_{NG} = 2700$ m. With these settings and with a 64-bit version of the software, the calculation of all needed network distances takes +/- 10 min on an Intel® Core™ i5-2520M CPU. Using another distance algorithm could still reduce the run time. The network distance calculations are normally only performed once; if the network or link speeds do not change between simulations the calculations do not need to be repeated. The land use and activity calculations take approximately 2 min for 15 years of simulation without updating the network distances. The program uses about 500 MB of RAM memory.

Initial tests were done to determine suitable values for the parameters C , F and P in equations (4.1), (4.2) and (4.3). We chose $C = 20$

min, $F = 5$ and $P = 2$, yet these choices proved to have only a limited influence on the model results. Only for extreme values, or when the parameters are omitted ($C = 0$ min, $F = 0$, $P \rightarrow +\infty$) do the results differ significantly: yet, even then, differences are still slightly smaller than between the transport scenarios (at most, about 500 cells have a different land use; most population differences are below 1 person/km²).

4.4 Results

Land use and activities were modelled for the five transport scenarios for 2000 to 2060. To make the comparison of the resulting land-use and activity patterns of the alternative transport scenarios easier, we defined an accessibility index to produce maps of the effort needed to travel from each network grid cell towards the centroids of its variable grid neighbours in a specific transport scenario. For the network scenarios, we first calculate two sums: (1) the sum of all possible network times t_{ij} from an origin NG cell i to all the NG cells j containing the centroids of its variable grid neighbours, and (2) the sum of the associated Euclidean distances. We then define the accessibility to be the ratio of these network and Euclidean distance sums. The very long distances towards the high-level variable grid neighbours ($L \geq 5$) were not included in the index to avoid boundary effects. Moreover, the neighbourhood effect weights for these levels are very small. To arrive at a standardised index, we replace the smallest value in the study area by 1: this is the most accessible NG cell. All larger values (less accessible NG cells) are scaled proportionally. The results for the four network scenarios can be seen in Figure 4.5.

In the ‘road-based’ scenario, the north of Belgium has the best accessibility values, especially in the large cities and along the motorways

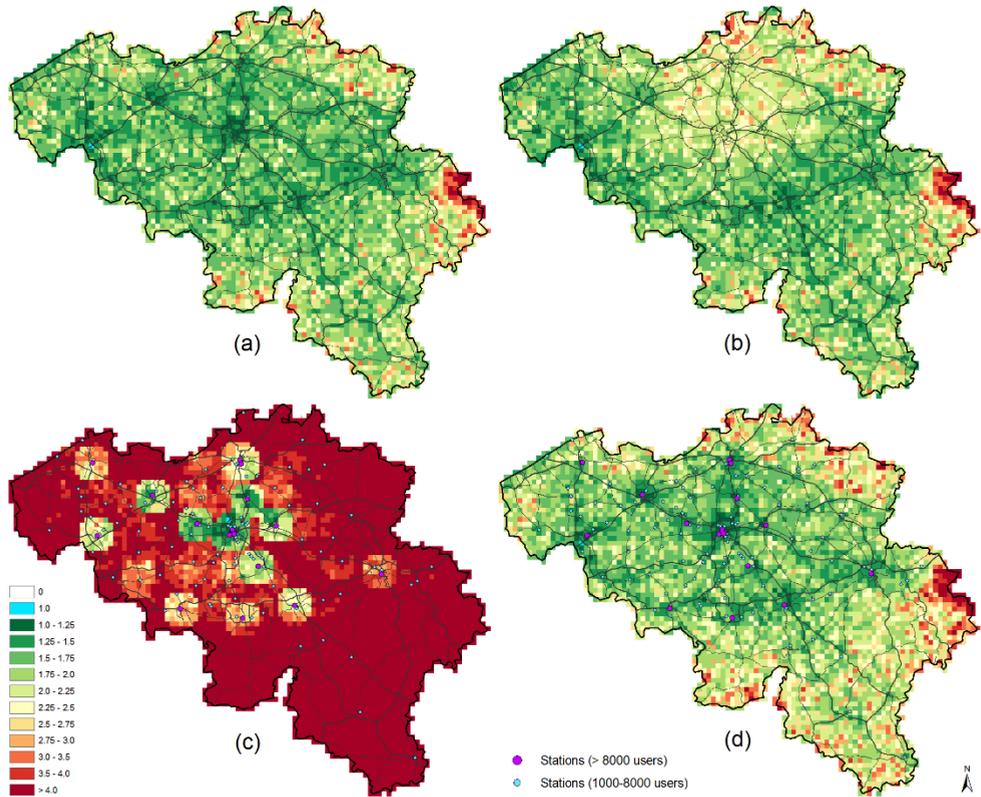


Figure 4.5 Accessibility index for different transport scenarios in Belgium: (a) the standard ‘road-based’ scenario, (b) the ‘congestion’ scenario, (c) the ‘train only’ scenario, and (d) the ‘choice’ scenario.

(Figure 4.5a). The ‘congestion’ scenario produces a rather different pattern, with lower accessibility values in the central part of Flanders and around Brussels (Figure 4.5b); the best values are now situated along the motorways west and south of this region. Especially the areas in between the major cities have a lower accessibility. The ‘train only’ scenario has a very distinct accessibility map with the best accessibility values around the most important stations, especially around Brussels (Figure 4.5c). The accessibility map of the multimodal scenario has a fairly similar pattern to

that of the ‘road-based’ scenario, but logically, values are clearly worse where there is no major station nearby (Figure 4.5d).

In general, the predicted land-use change patterns between 2000 and 2060 are rather similar for all the network scenarios. The discontinuous residential area grows significantly in the central area of Flanders between Brussels, Antwerp and Ghent, as well as in the southwest of Flanders. This is shown for the ‘road-based’ scenario in Figure 4.6. The differences between the ‘road-based’ and most other scenarios are distinct but less pronounced than the overall growth pattern in all scenarios (Figures 4.7, 4.9, 4.11). The differences between the scenarios stand out more strongly in the activity values, especially in the population difference maps (Figures 4.8, 4.10, 4.12).

In the road network-based model the urban growth in the central areas is slightly higher in 2060 than in the Euclidean model and vice versa for the peripheral areas (Figure 4.7). Most interestingly, some of the more accessible cells in the peripheral areas also become residential. Although the general pattern in the population difference map is harder to discern, the population values are somewhat higher next to major roads as well as in the centres of the major cities (Figure 4.8).

The residential land-use difference map between the scenarios without and with congestion clearly shows that the model is sensitive to network speed differences (Figure 4.9). The population growth is also smaller in the suburban regions of central Belgium in this ‘congestion’ scenario (Figure 4.10), while the population in the main cities not really gets affected by the congestion to reach the city centres. The ‘train only’ scenario leads to some remarkable but logical differences, with much more residential growth in central areas that are more accessible to rail stations

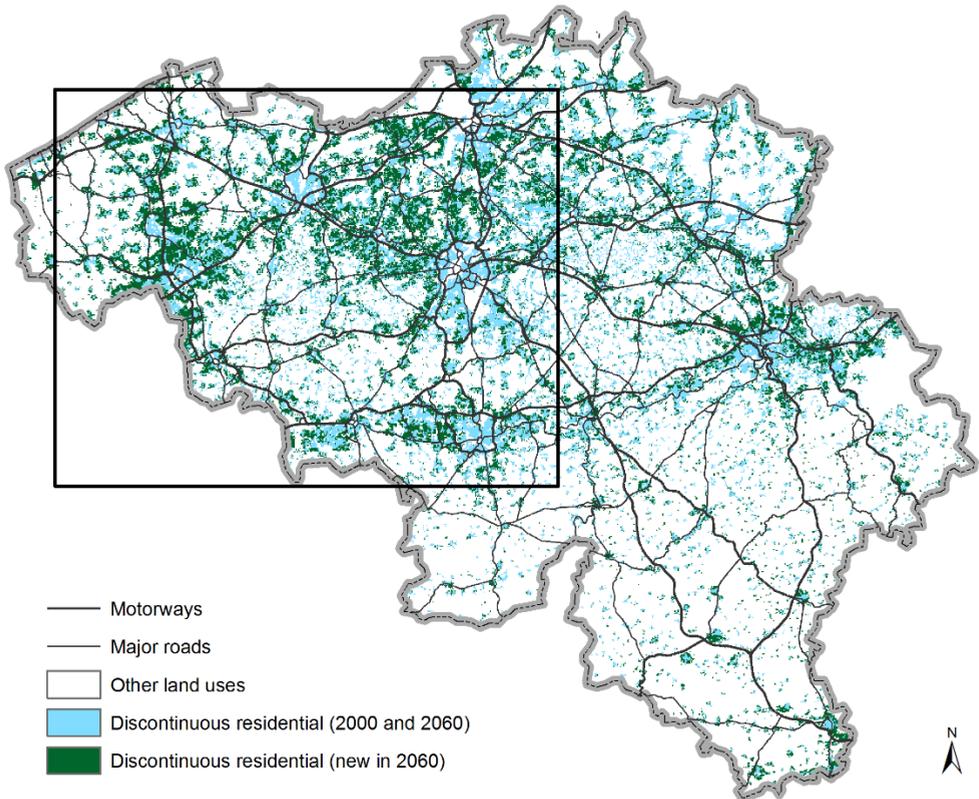


Figure 4.6 New discontinuous residential development in Belgium in the road-based scenario between 2000 and 2060. The rectangle shows the extent of the zooms in the next figures.

(Figure 4.11). The differences in population compared with the ‘road-based’ scenario show a reasonable pattern with more population not only in the ‘train only’ scenario in big cities but also around smaller stations with good connections to big cities. Rural areas, as well as some areas near smaller stations lacking good connections with the centre of Belgium (e.g. in West Flanders), gain less population (Figure 4.12). The ‘choice’ scenario produces growth patterns similar to those of the ‘road-based’ scenario. Only some slight differences can be noted, such as somewhat higher population values in big cities.

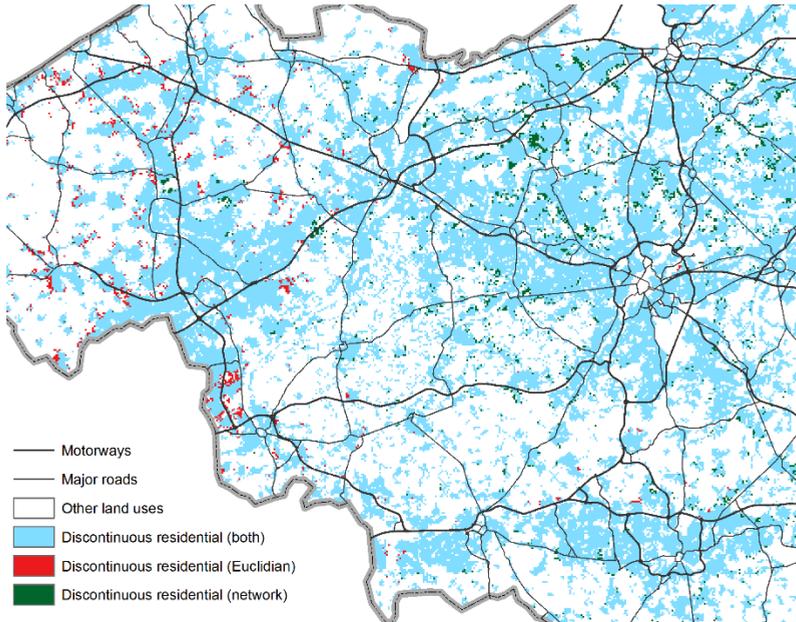


Figure 4.7 Differences in residential development between the Euclidean and the ‘road-based’ scenarios in 2060.

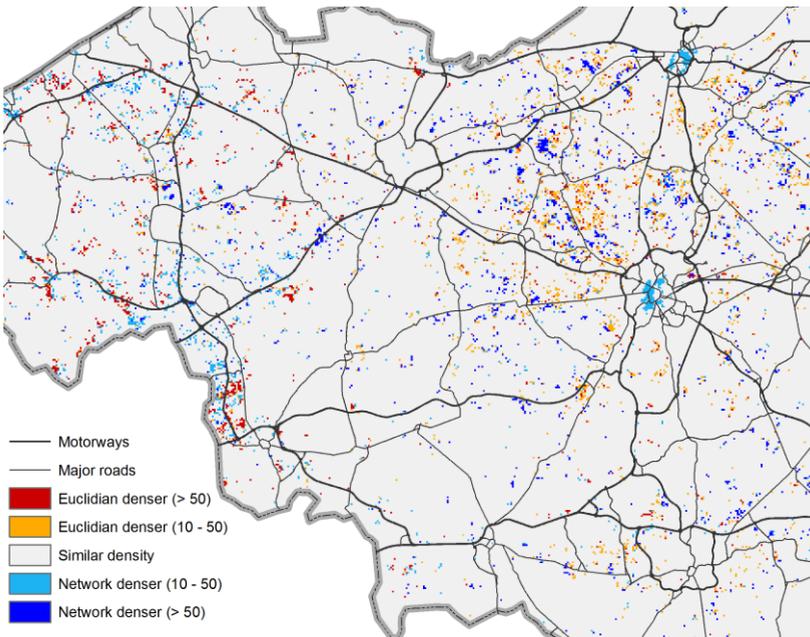


Figure 4.8 Differences in population values (people per cell) between the Euclidean and the ‘road-based’ scenarios in 2060.

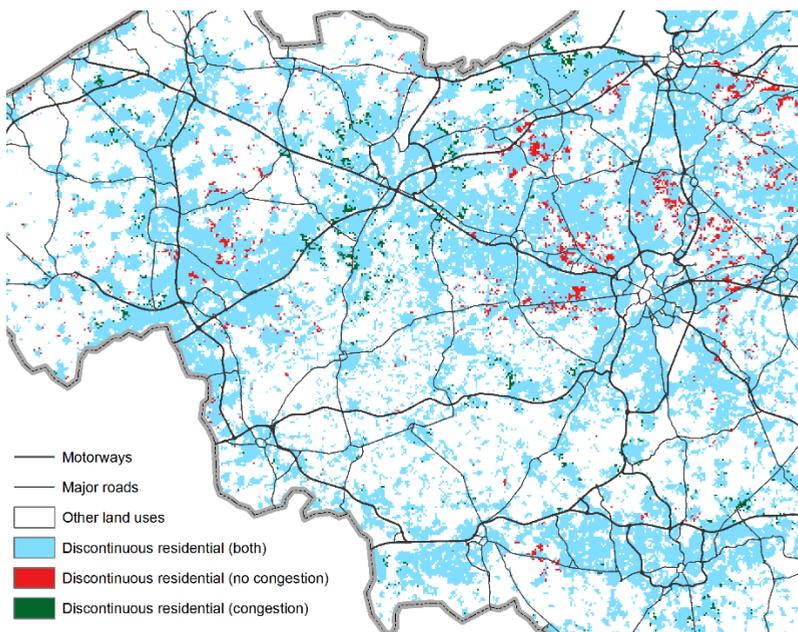


Figure 4.9 Differences in residential development between the ‘road-based’ and the ‘congestion’ scenarios in 2060.

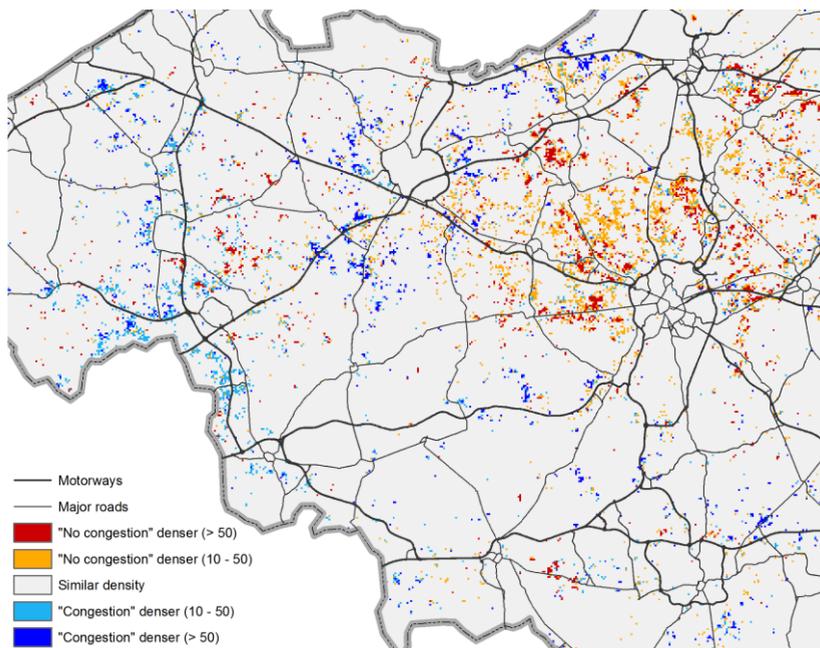


Figure 4.10 Differences in population values (people per cell) between the ‘road-based’ and the ‘congestion’ scenarios in 2060.

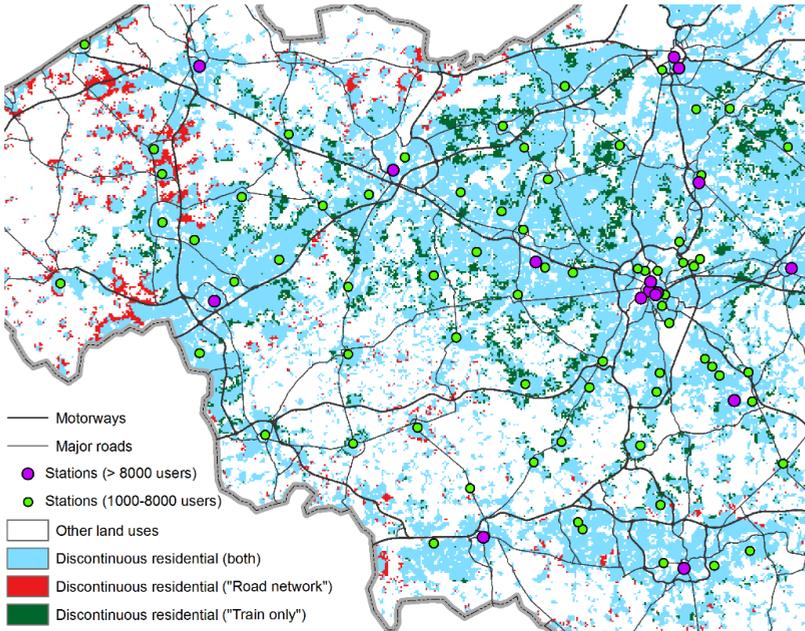


Figure 4.11 Differences in residential development between the 'road-based' and the 'train only' scenarios in 2060.

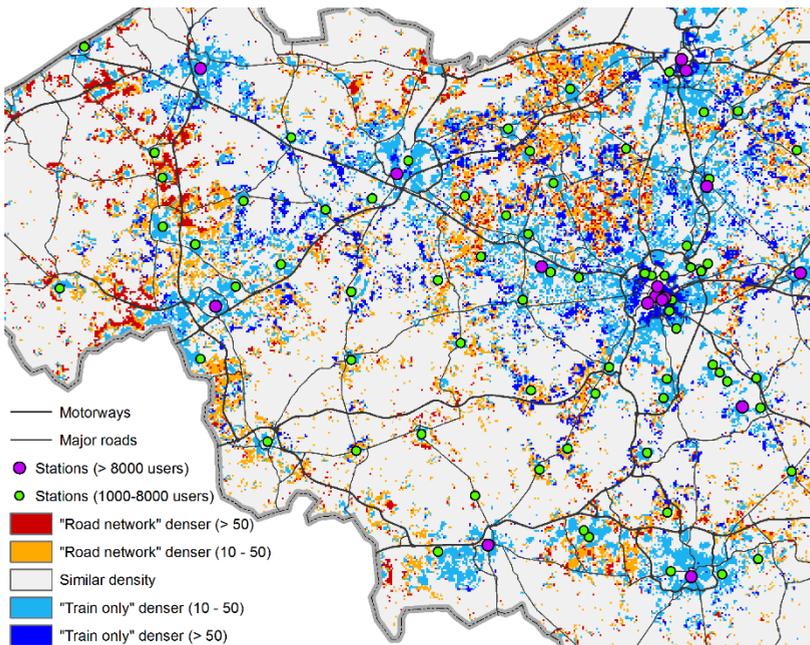


Figure 4.12 Differences in population values (people per cell) between the 'road-based' and the 'train only' scenarios in 2060.

4.5 Discussion

The transport component of the activity based CA model discussed in this chapter is a useful extension of the original version of White et al. (2012) for studying the influence of existing and evolving transport networks on land-use dynamics. The only accessibility measure used in the original model, X_{Ki} in equation (3.1), was based on a weighted Euclidean distance to reach the network from a cell, and thus represented only a cell's accessibility to the transport network, rather than its general accessibility to the region. The new model individually computes the anisotropic accessibility of a unit cell towards its supercells and saves it in a structure adapted to the variable grid. The structure of the distance grid is more independent of urban development than a population-based zone system, as often used in transport models (e.g. the model of the Flemish Traffic Study Centre). This is an advantage in less-populated areas, which are important for future development, but can be a disadvantage in current urban areas. Nevertheless, more detailed zones are possible in our model with a lower network grid level L_{NG} . Moreover, we define useful parameters to enhance an easy definition of long-term transport scenarios with different modes within an activity and land-use modelling context.

The transport scenarios indeed generate different results, both in terms of land use and activity patterns. The 'congestion' and the 'train only' scenarios lead to pronounced differences compared with the 'road-based' scenario. In places where the value of the accessibility index is clearly higher for the 'congestion' scenario (worse accessibility), less land is converted to urban land uses, and the population growth is significantly smaller (Table 4.3). Cells that have slightly worse accessibility in the central areas still seem to attract more people and urban land use. This is logical as

the population and land-use growth were fixed, and other mechanisms in the model normally produce the largest growth in those central areas.

Table 4.3 Comparison in 2060, in terms of discontinuous residential land use and population, between the ‘road-based’ scenario (‘R’) and the ‘congestion’ scenario (‘C’). The classes are defined by relative accessibility index values (division of ‘congestion’ and ‘road-based’ index values).

Relative accessibility index (C / R)	Total area (km ²)	Population density difference ((C - R)/km ²)	Discontinuous residential in 2060 (cells)		
			R only	Both scenarios	C only
< 0.97	9493	3.11	192	12911	222
0.97 - 1	6731	9.80	89	19349	516
1 - 1.15	8746	8.03	572	36577	1276
> 1.15	5692	- 29.11	1806	30245	645

The same conclusions largely hold for the ‘train only’ scenario in comparison with the ‘road-based’ scenario (Table 4.4). The best locations in the ‘train only’ scenario are those close to large stations, and those are often already built-up. New space for urban development is limited. Nevertheless, the model predicts substantial densification, since almost 100,000 extra people (or 160 per km²) are allocated to these areas, compared to the ‘road-based’ scenario. The regions around these areas also gain extra residential land use and population, while there is clearly less growth in remote areas far from stations.

Table 4.4 Comparison in 2060, in terms of discontinuous residential land use and population, between the ‘road-based’ scenario (‘R’) and the ‘train only’ scenario (‘T’). The classes are defined by relative accessibility index values (division of ‘train only’ and ‘road-based’ index values).

Relative accessibility index (T / R)	Total area (km ²)	Population density difference ((T – R)/km ²)	Discontinuous residential in 2060 (cells)		
			R only	Both scenarios	T only
< 1	597	160.62	5	4140	384
1 – 2	5774	53.96	551	24242	2370
2 – 3	14408	- 11.13	4325	41484	5163
> 3	9881	- 25.01	3962	23032	926

Obviously, the ‘choice’ scenario is, by definition, much more realistic than the theoretical ‘train only’ scenario, yet its results do not differ much from those of the ‘road-based’ scenario (Table 4.5). This is clearly caused by the values of the modal weights w in the ‘choice’ scenario, but we chose these values because it is reasonable that the road network plays the most important role in allocating new land uses, especially in the many suburban villages and ribbon developments of Belgium. In comparison with the ‘road-based’ scenario there is a slightly higher population growth near the most important stations. On the contrary, residential land-use growth is slightly greater near the edge of these cities in less accessible cells of this scenario. This may seem contra-intuitive but it is caused by the attraction effects of the model since there is no space for urban growth very close to major stations.

Table 4.5 Comparison in 2060, in terms of discontinuous residential land use and population, between the ‘road-based’ scenario (‘R’) and the ‘choice’ scenario (‘Ch’). The classes are defined by relative accessibility index values (division of ‘choice’ and ‘road-based’ index values).

Relative accessibility index (Ch / R)	Total area (km ²)	Population density difference ((Ch – R)/km ²)	Discontinuous residential in 2060 (cells)		
			R only	Both scenarios	Ch only
< 1	2012	4.80	200	12352	8
1 – 1.04	4241	0.98	289	20446	89
1.04 – 1.08	8984	- 1.02	303	37163	376
> 1.08	15424	- 0.30	100	30888	419

Although the variable grid provides a good computational framework for activity predictions in large study areas, network distance calculations can be further improved. Firstly, in the approach described in section 4.2, distances are measured between centroids of NG cells, of which one may represent the centroid of a larger variable grid supercell. The latter might not be representative of the location of the activities within the supercell. Hence, the logical next step is to represent each supercell by a centre that reflects the actual distribution of activity within the supercell. This approach has been implemented in later versions of the model, used in chapters 6 and 7 (see appendix 4.A1 for the methodology). Secondly, the access and egress legs in the multimodal computations can be made more realistic, since in home-to-work travel there is often a very short leg at the work end of the displacement (Thys & Andries, 2011). Because of the symmetry of the variable grid, such an asymmetric approach has not yet been implemented.

Extra details could also be included in an improved transport-based model, such as different modal weights w for different regions of the study area, or transport networks evolving over time. The functionality to introduce changes to the network at specified time steps during a simulation is already implemented, but has not been used due to a lack of data. Since Belgium already has a dense network, there are likely to be few additions to the network, but a growing population and increased congestion could reduce travel speeds in the future, and road characteristics could change. Congestion scenarios of transport models can therefore provide useful input for the model. It would be an interesting exercise to couple our activity based CA model directly to a transport model in order to get simultaneous predictions of activities, land uses and transport for the future. On the other hand, we fear that a direct coupling would increase the computation time drastically.

Finally, in a more complete and general sensitivity analysis it would be interesting to examine how parameters, rules and functions of the activity based CA could be adapted, added or removed to improve the model. For instance, the accessibility component of the earlier CA models using Euclidean distances seems to have less influence in the network-based implementation, and so it might be redundant. Additionally, it would be interesting to use a generalised cost measure of distance. Such a measure would enable the model to be used to investigate the effects of a subsidised, and hence cheaper, public transport in comparison with congestion pricing schemes for the road network.

4.6 Conclusions

The activity-based CA model of White et al. (2012) was a big step forward in comparison with earlier CA models, since every cell is modelled as a truly multifunctional environment where people live and work. In this chapter, we developed an activity-based CA model with travel time-based interaction rules for long-distance interactions, and simple Euclidean distance-based rules for the local vicinity interactions. The network-based rule sets for the model are clearly more realistic and provide the possibility to test different transport scenarios. The impacts of road congestion and public transport usage on land-use and activity futures can be evaluated with the proposed approach, and possibly new spatial indicators could be derived to clearly display these impacts.

Nevertheless, the version of the model discussed in this chapter still had some shortcomings. Hence, the variable grid was adapted by defining activity centres as the representative locations of large variable grid cells (see appendix 4.A1). Furthermore, more critical assessment, calibration and sensitivity analysis were needed to validate the current assumptions and equations of the model and to optimise parameter values and rule sets. We therefore continued this work with a historical calibration exercise to validate and improve the model (see chapters 5 and 6).

Appendix 4.A1 Activity-based centres of variable grid supercells

The standard variable grid uses distances between the geometric centres of supercells, and therefore does not measure distances towards the actual locations of activities. This can be a problem when most activities are located in one part of the supercell.

Centres of supercells in the standard transport-based variable grid already differ slightly from the Euclidean variable grid when the supercells are partially outside the study area. For these supercells the centre of gravity within the study area is calculated and used as the centre (see section 4.2). At a later stage, we tested some possibilities to move the centre of a supercell towards a position that is representative for the local concentration of activities. Currently, only population is taken into account. A combination of population and economic activity could still be implemented in the future. The activity-based centres are only calculated for supercells that are larger than or equal to the size of the network grid (NG).

We tested two different methods to determine activity-based centres for supercells (Figure 4.A1). As a first possibility, we calculated the centre of gravity of the activity within each supercell. This is first done for all the NG cells. Next, we calculate the centre of gravity for each supercell of the next higher level using the centres of the nine cells of the level below, with the activity totals of these subcells used as weights. The procedure is repeated for each successively higher level.

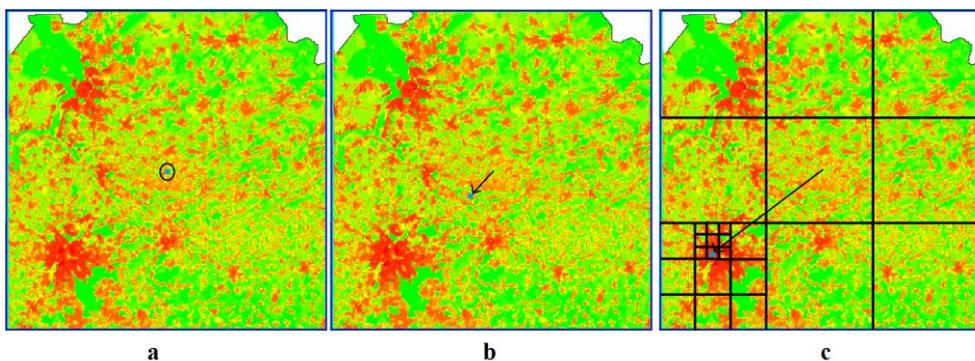


Figure 4.A1 Location of the cell centre of a supercell: (a) geometric centre, (b) centre of gravity, and (c) hierarchic centre.

Alternatively, centres of gravity can be determined hierarchically. In this method we also start with the NG cells, and search for the subcell within a subcell (and so on until the subcell at the level of unit cells is determined...) with the highest activity. The centre of the found unit cell will be the centre of the NG cell. Next, a similar approach is used for larger levels: we search for the subcell of the level below with the highest total activity of all nine subcells. The already determined centre of this subcell is then chosen to be the centre of the supercell. This second approach produces more substantial differences with the standard model results. Two reasons can explain these differences. Firstly, with this approach we create an activity influence model that is almost vector-based instead of raster-based since overlapping supercells have very often exactly the same centres (centres of big cities). Secondly, this approach is very sensitive to border effects. When one moves through the grid and looks at supercells of a specific size, a city can first be entirely within one subcell, next it can be located within two subcells, and finally again be entirely within one subcell. This influences the centre choice in a negative way. Therefore, we ultimately decided to adopt the methodology with the centre of gravity as the standard approach. However, this method also has a disadvantage: the centre of gravity will often be located in an intermediate position where there is not much activity, and where there are maybe even no transport possibilities nearby. This was of course already a problem in the standard variable grid with geometric centres.

5. Datedating high-resolution population density maps using sealed surface cover time series

This chapter (except its appendices) is based on: Crols, T., Vanderhaegen, S., Canters, F., Engelen, G., Poelmans, L., Uljee, I., & White, R. (2017). Datedating high-resolution population density maps using sealed surface cover time series. *Landscape and Urban Planning*, 160, 96–106.

5.1 Introduction

Several land-use change models have been developed over the years to simulate the spatial evolution of the urban sprawl phenomenon (Haase & Schwarz, 2009; Poelmans & Van Rompaey, 2010). One particularly successful type of highly-detailed raster based models are cellular automata (CA) (Santé et al., 2010). Thanks to their high resolution and regular structure, they fully integrate spatial input data and are computationally efficient for large regions. The proposed activity-based CA model of White et al. (2012) and Crols et al. (2015) integrates the simulation of land use, population and employment within the CA structure itself. This implies that for calibration historical land-use maps, as well as detailed maps of population and employment for the past are required. The development of a time series of regional land-use maps is a time consuming task, and existing series are often produced with inconsistent methodologies (van der Kwast et al., 2009; Van de Voorde et al., 2016). Remote sensing imagery can be a useful data source to build consistent land-use time series for recent decades. Spectral unmixing of medium-resolution images, available since the 70s, enables the mapping of sealed surfaces, which are an indicator of urbanised land uses (Wu, 2004; Van de Voorde et al., 2009; Weng, 2012;

Vanderhaegen et al., 2015). Van de Voorde et al. (2011) used a spectral unmixing technique to produce sealed surface maps from Landsat TM/ETM+ data, and to use these maps for distinguishing urban from non-urban land use, and even residential from non-residential urban land uses by computing spatial metrics on pixel-based sealed surface fractions for each urban block. A thematically more detailed mapping of land use though, as is often required for historical calibration of land-use models, cannot be achieved with this approach.

A possibility to obtain historical land-use maps with a higher number of urban classes is to start with present land-use information, and then move progressively back in time by defining a downdating procedure, using remote sensing data available for previous time steps to downdate the most recent map. Using this approach, Fricke and Wolff (2002) succeeded in building a land-use series for the agglomeration of Brussels ranging from 1955 until 1997. All land-use maps they produced involved manual interpretation of orthophotos or aerial photos and topographical maps. For the 1997 map, which formed the basis for downdating, extra information out of satellite remote sensing and a GIS database was used. The procedure involves much manual work, and hence is not feasible for application to larger regions. Nevertheless, downdating a present high-resolution land-use map using remote sensing data available for the past remains an interesting idea, if image interpretation for previous time steps can be automated. Time series of sealed surface maps obtained through spectral unmixing seem promising for developing such a strategy.

In a similar way, sealed surface time-series seem also an interesting data source for downdating population maps. High-resolution population density maps are mostly constructed using dasymetric mapping approaches,

which model population distribution within census units based on spatially detailed ancillary data explaining differences in population density (Wu et al., 2005). Often use is made of multiple regression analysis, with population as the dependent variable and land-use data as one of the independent variables (e.g. Langford et al., 1991; Eicher & Brewer, 2001; Mennis, 2003; Gallego et al., 2011). Stevens et al. (2015) proposed nonparametric modelling to include a large set of land cover and other spatial variables. Several publications confirm that population density within a certain land use and census unit is related to sealed surface cover (Lu et al., 2006; Wu & Murray, 2007; Zandbergen & Ignizio, 2010; Batista e Silva et al., 2013a). If available, information on the building type in parcels (Xie, 2006; Jia et al., 2014) or in land-use polygons (Goerlich & Cantarino, 2013) can improve the results, and can be combined with land cover data such as sealed surface cover (Jia & Goughan, 2016). Recently, the number of address points within a land use and census unit was suggested as the most accurate factor to compute high-resolution population density maps (Tapp, 2010; Cockx & Canters, 2015). Alternatively, nighttime mobile phone data can be used to downscale population maps (Deville et al., 2014) or to update existing outdated population maps (Douglass et al., 2015). Unfortunately, address points and mobile data are generally not available for the past. Hence, in the absence of address data, sealed surface cover can be considered the best option for estimating population density (Zandbergen, 2011).

In this chapter, we define a workflow for downdating both land use and activity maps using time series of sealed surface maps for the Flanders-Brussels region, to be used as an input for the historical calibration of the activity-based land-use model. We focus on downdating of population,

which plays an important role in the dynamics of the model and has proven useful for documenting the phenomenon of urban sprawl. Datedating of land-use data using sealed surface cover as ancillary data has proven to be a complex and partially ad hoc process; a description of the technique can be found in appendix 5.A1. Datedating of employment is difficult due to a lack of spatially detailed data that can be related to employment as there is no direct link with sealed surface cover. The simple solution therefore was to keep the current spatial pattern, but limited to the employment land-use categories of the obtained land-use maps of the past: a brief discussion can be found in appendix 5.A2.

Starting from a recent high-resolution population map, and using remote sensing derived sealed surface data and statistical population data for previous time steps, in this chapter we demonstrate the population datedating method proposed. The analysis focuses on the years 1986, 2001, and 2013. As such, the development of the population time series, in combination with the sealed surface data for each time step enables us to analyse in detail the last stage of the urban flight of the late 20th century in the Flanders-Brussels region, the 21st century revival of the regional cities, and the still declining population density of new built-up areas around these cities.

5.2 Methods

5.2.1 Study area and land use

As discussed in chapter 1, the northern half of Belgium is a region with a long history of urban sprawl, caused by transportation and housing policies going back to the 19th century. Tax policies in Belgium have always promoted home ownership. In 2013, 70% of the Flemish population

lived in a unit they owned, and 85% of these units were single-family dwellings (Winters et al., 2015). From 1970 till 2000 there was still an ongoing ‘urban flight’ in Belgian cities with a declining urban population, while the rate of suburban population growth did not fall (Hertogen, 2013). At first sight, the Flemish cities and Brussels seem to have regained their attractiveness after the year 2000 due to improved spatial policies and urban revitalisation projects, yet the total picture is more complex since the main reason for urban population growth is external migration (De Decker, 2011).

Flanders has a relatively high population considering its size (6,381,859 inhabitants in 2013, or 472 per km²). The Brussels Capital Region (BCR), a territory of only 161.38 km², forms an enclave in Flanders (Figure 5.1). With a population of 1,154,635 inhabitants in 2013 (data of the Belgian Federal Department of Economics), it has a density of 7155 inhabitants/km². Just outside the BCR, the population density rapidly falls: in comparison with other large European capitals, Brussels has one of the highest densities in the core but one of the lowest in the zone between 8 and 15 km from the city centre (European Commission, 2014).

According to the 10 m resolution land-use map of the Flemish Institute for Technological Research (VITO) of 2013 used in this study (Figure 5.1, Tables 5.1 and 5.2), the 13,522 km² of the Flemish Region consist of 27.6% of built-up land uses, while the Brussels Capital Region contains 69.3% built-up land uses. Not all sources agree on the amount of built-up surfaces. The Cadastre of Belgium reports only 20.1% of built-up registered parcels for Flanders and 58.6% for the BCR in 2013 (FOD Economie, 2015) (Table 5.2). The big differences can be attributed to other land-use definitions: in the Cadastre, infrastructure and unbuilt parts of industrial and commercial terrains are regarded as unbuilt.

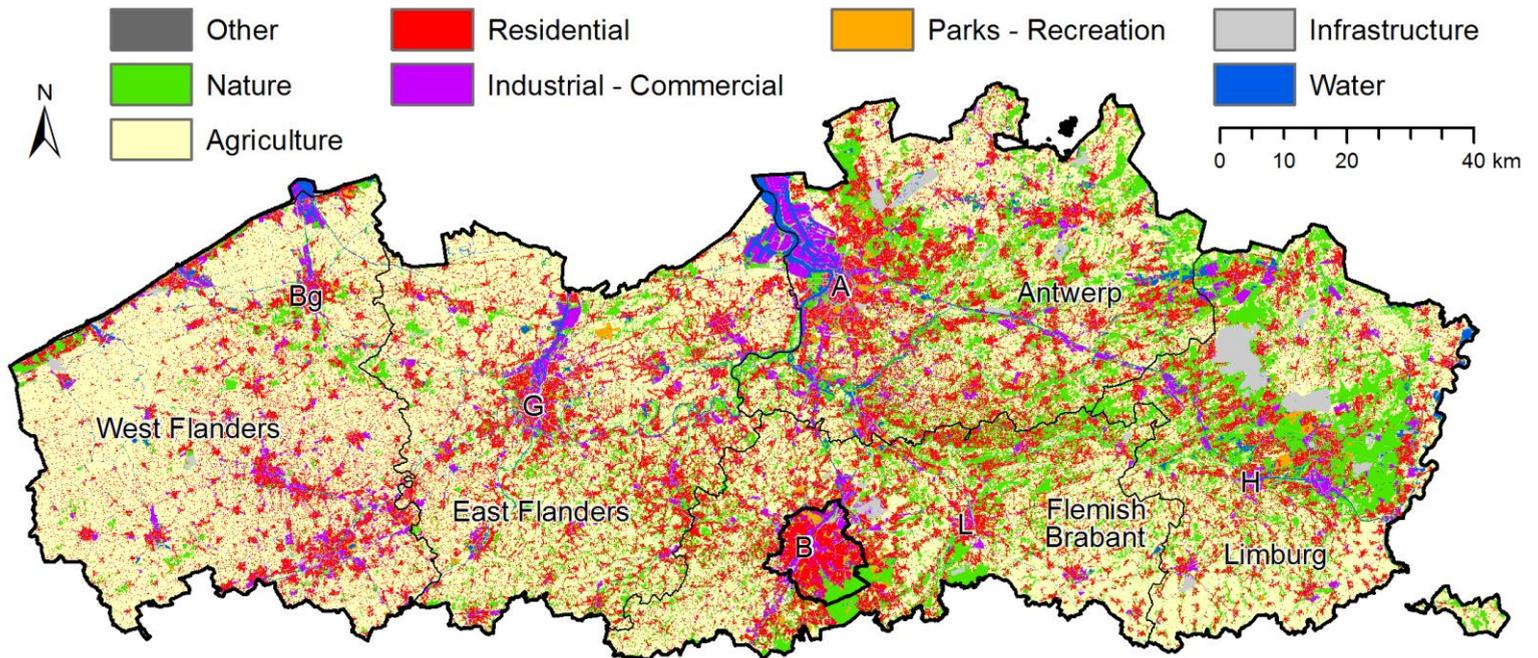


Figure 5.1 Land-use map of Flanders and Brussels in 2013 (10 m resolution map developed by VITO aggregated to a 100 m resolution), with an indication of both regions (thick black lines; Brussels Capital Region (B)), the Flemish provinces (thin black lines) and their capitals (Bruges (Bg), Ghent (G), Antwerp (A), Leuven (L) and Hasselt (H)).

Table 5.1 Proportions of land uses (%) in Flanders and the Brussels Capital Region according to the 2013 land-use map of the Flemish Institute for Technological Research (VITO). Residential, industrial and commercial, and infrastructure are built-up land uses. Natural land covers in military terrains were moved from infrastructure to nature to produce this table.

Land use	Flanders	Brussels Capital Region
Residential	15.8	66.1
Industrial and commercial	5.4	3.2
Infrastructure	6.4	14.9
Nature	17.8	3.9
Agriculture	48.8	9.7
Parks and recreation	2.2	1.1
Water	2.7	0.9
Other	0.9	

Table 5.2 Proportion of built-up land use / land cover (%) in Flanders and the Brussels Capital Region for 1986, 2001, and 2013 according to the 2013 land-use map of VITO, the Belgian Cadastre, and the sealed surface time series ($SSF \times \text{pixel area}$) used in this study.

LU/LC data	Flanders			Brussels Capital Region		
	1986	2001	2013	1986	2001	2013
LU map of VITO	-	-	27.6	-	-	69.3
LU in the Cadastre	12.7	17.8	20.1	50.9	55.8	58.6
Sealed surface data	13.5	17.2	19.1	45.2	48.3	51.3

5.2.2 Sealed surface fraction mapping

This study uses sealed surface fraction estimates for Flanders and Brussels for the years 1987 (as a proxy for 1986), 2001 and 2013, produced from a time series of medium-resolution Landsat imagery at 30 m resolution by

applying a sub-pixel mapping approach, as described in Vanderhaegen et al. (2015). In this approach, the sealed surface fraction within each pixel that is part of the urban area is assumed to be the complement of the vegetation fraction. Therefore, the first step in the mapping is to determine an urban mask for each date delineating the urbanised area. The mask for 2013 was defined based on large-scale reference data sets of Flanders (GRB Vlaanderen) and Brussels (UrbIS), combined with the 10 m resolution land-use map of VITO. The masks for previous years were considered to be subsets of the 2013 mask. The Normalised Difference Vegetation Index (NDVI: see e.g. Mather, 2004) was used to exclude pixels fully covered by vegetation for each year. Based on four high-resolution IKONOS images (4 m resolution) covering part of the study area, a detailed mapping of vegetation cover was accomplished. Based on this detailed mapping, the vegetation fraction of each 30 m Landsat pixel located within the IKONOS coverage was determined. Next, a linear regression model was built with known vegetation fraction, derived from the IKONOS images as the dependent variable, and the Landsat NDVI value as independent variable:

$$VF_i = a_0 + a_N \cdot NDVI_i + \varepsilon_i \quad (5.1)$$

with VF_i the reference vegetation fraction in pixel i , a_0 the model's intercept, a_N the regression coefficient for the NDVI, $NDVI_i$ the NDVI value of pixel i , and ε_i the residual for pixel i . Multiple linear regression with Landsat spectral band values as extra independent variables did not improve the model's outcome, hence the model with NDVI as the sole variable was applied to estimate the vegetation fraction for each Landsat pixel within the urban mask. Finally, as indicated above, the sealed surface fraction SSF was defined as the complement of VF :

$$SSF_i = 1 - VF_i \quad (5.2)$$

To produce a temporally consistent *SSF* series and correct for errors due to uncertainty in the estimation procedure, the *SSF* value of a pixel for a later time step was defined to be minimally the value of the previous step. As such, the sealed surface cover is assumed to only increase over time, which can be expected to be true for almost the entire study area. For a full description and discussion of the technique used, the reader is referred to Vanderhaegen and Canters (2016).

5.2.3 High-resolution mapping of population density for 2013

For each first day of the year, there exist population data made available by the Federal Department of Economics at the municipality level, and for certain years also at the more detailed level of statistical sectors. Flanders and the BCR include 327 municipalities and 9906 statistical sectors. We obtained data per statistical sector for 1981, 1991, 2001 and 2011, and made an interpolation for 1986 and an extrapolation for 2013; in both the latter cases we rescaled the statistical sector estimates so that the totals were correct for municipalities.

For the purpose of downdating, we required a detailed population density raster for 2013. The Belgian Federal Department of the Interior provided population data at 100 m resolution for 2013 which we resampled to a 30 m raster to enable computations with the sealed surface series. However, the original population raster had a number of drawbacks. The raw population data were extracted from a database of the Belgian National Population Register (Rijksregister) which contains the addresses of all inhabitants. Unfortunately, these addresses were transformed into coordinates with an unknown interpolation method. Especially in rural areas

we noticed a mismatch between the built-up area in the land-use map and the sealed surface maps on the one hand, and the population map on the other. Furthermore, the totals were slightly different from those of the Federal Department of Economics.

Hence, we decided to evenly redistribute all the population living in cells without buildings, i.e. cells without an urban land use or without sealed surfaces, over nearby cells having a ‘habitable’ urban land use with sealed surface cover. At 10 m resolution, residential and industrial cells, commerce and services, port areas, and military buildings were considered to be ‘habitable’, while cells labelled as infrastructure and all rural and natural land uses were not. Thus all 30 m pixels containing at least one ‘habitable’ cell in the 10 m land-use map, and having $SSF_i > 0$, were defined to be ‘habitable’. The redistribution was done in 5 x 5 focal neighbourhoods around the wrongly located 30 m ‘inhabited’ cells, and repeated with remaining erroneous locations in 11 x 11 and 17 x 17 focal neighbourhoods. The remaining very small amount of population not having ‘habitable’ cells within a 17 x 17 focal neighbourhood was considered to be totally wrong, and was therefore excluded. In the end, we performed a rescaling operation

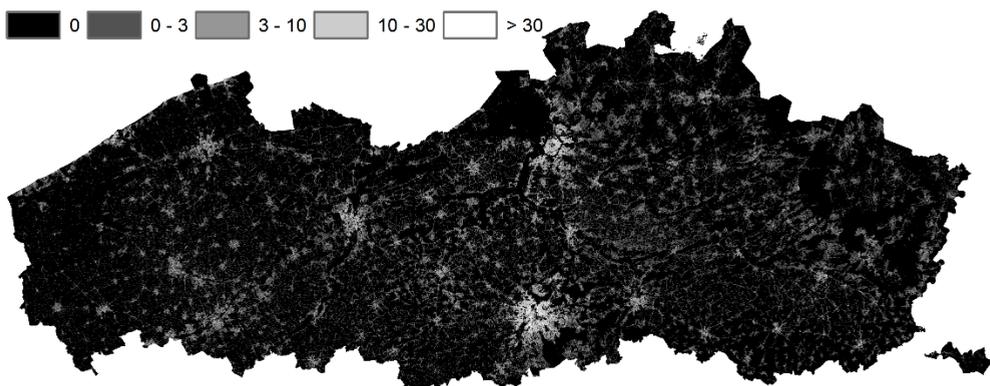


Figure 5.2 Population density map for 2013 (inhabitants per 30 m pixel).

to equate the statistical sector totals to those obtained based on the Federal Department of Economics data. The rescaling was proportional to the population already present in a cell. The resulting map is shown in Figure 5.2.

5.2.4 DOWNDATING OF THE POPULATION DENSITY MAP

To produce a population density map for 1986 and 2001, we developed a downdating strategy using the population raster constructed for 2013, as well as the sealed surface fractions of 1987 and 2001 and the statistical data for 1986 and 2001. The results of the downdating for 2001 served as input data for the 1986 estimates.

In the proposed strategy, cells that are not part of the urban mask in a previous time step do not receive any population. For cells with changing *SSF* values we hypothesise a local relationship between the evolution of the population and sealed surface cover. We do so by assuming that the local building style imposes a more or less constant population density per unit of built-up area. Therefore, increases in sealed surface cover in urban cells between an earlier and a later date can be associated with increases in population:

$$\bar{P}_{i,t-1} = P_{i,t} \times SSF_{i,t-1} / SSF_{i,t} \quad (5.3)$$

with $\bar{P}_{i,t-1}$ the estimated population in pixel *i* at an earlier time step *t* – 1, $P_{i,t}$ the population in pixel *i* at step *t*, and $SSF_{i,t-1}$ and $SSF_{i,t}$ the sealed surface fraction in pixel *i* at time steps *t* – 1 and *t*. The fact that ribbon development next to existing roads is one of the most common mechanisms of built-up area growth in Flanders supports the hypothesis that *SSF* increase goes along with local population growth, with little changes in population density

per unit of built-up area. As such, the proposed procedure seems a valid option to create rasters with spatial resolution higher than the level of statistical sectors for past decades. On the other hand, the approach does not account for large-scale major evolutions such as population decline or re-urbanisation of an entire urban core. Furthermore, enlargements of a building are not necessarily related to an increase in the number of inhabitants. As such, a rescaling of local population estimates obtained is required, accounting for these effects. The rescaling ensures correct totals at the statistical sector level:

$$P_{i,t-1} = \bar{P}_{i,t-1} \times P_{s,t-1} / \sum_{j=1}^n \bar{P}_{j,t-1} \quad (5.4)$$

with $P_{i,t-1}$ the rescaled population in pixel i at time step $t - 1$, $P_{s,t-1}$ the total population in statistical sector s , j a pixel in sector s , and n the number of pixels in sector s . The applied correction factor $P_{s,t-1} / \sum_{j=1}^n \bar{P}_{j,t-1}$ for every statistical sector serves both as a validation of the constant density hypothesis and, in case of a rejection of this hypothesis in a statistical sector, as an indicator of density changes. If the correction factor for a statistical sector is close to 1, then use of equation (5.3) to estimate the population seems justified. A correction factor not close to 1 indicates that population density trends in the statistical sector are not in line with the extension of the urban tissue. In that case, the factor gives an indication of population density changes in relation to the sealed surface fraction evolution within the sector. The inverse of the correction factor can be defined as a densification index $D_{s,t-1,t}$ for a statistical sector s between time steps $t - 1$ and t :

$$D_{s,t-1,t} = \frac{\sum_{j=1}^n \bar{P}_{j,t-1}}{P_{s,t-1}} \quad (5.5)$$

For sectors with an increasing population, D_s serves as a detection method for urban densification (if $D_s > 1$) or deteriorative sprawl (if $D_s < 1$). The index is not suited though to analyse causes for a shrinking population since there is clearly no direct link between population decrease and sealed surface growth.

5.2.5 Uncertainty propagation analysis

Because sealed surface mapping from remotely sensed data inevitably introduces uncertainty in the estimation of sealed surface fractions at pixel level (Cockx et al., 2014) and may be considered the most important source of error in the downdating method proposed in this study, we analysed how this uncertainty affects the computation of the historical population maps and the densification indices through Monte Carlo uncertainty propagation analysis.

The sealed surface fraction estimates obtained in this study were validated for a set of Landsat pixels, located in parts of the study area covered by the IKONOS images used for calibrating the sub-pixel mapping model (see section 5.2.2). For a set of validation pixels, not used in the calibration, and for which it was verified that sealed surface cover did not change between 1986 and 2013, the error in sealed surface fraction estimates was calculated, using the average sealed surface fraction in the underlying IKONOS pixels as ground truth. Validation pixels were selected with stratified random sampling: each of ten quantiles of all possible sealed surface fraction values are equally represented in the validation set. Next, random errors were drawn for all Landsat pixels from a multivariate normal

distribution defined by the mean error vector and the variance-covariance error matrix, calculated from the errors observed at each time step. This way, error perturbed versions of the sealed surface maps for each of the three time steps could be produced, accounting for temporal correlation in the errors. The error simulation process was repeated 100 times, producing 100 sealed surface maps for all time steps. These maps were used to compute 100 population maps of 2001 and 1986 and 100 densification index maps for the periods 1986-2001 and 2001-2013. As such, we can analyse the uncertainty in the resulting population and densification maps.

5.3 Results

5.3.1 Population trends between 1986 and 2013

Sealed surface cover increased in Flanders and the Brussels Capital Region (BCR) from 1911km² in 1987 to 2687 km² in 2013 (Vanderhaegen & Canters, 2016), which corresponds to an overall growth of 41%. The proportions of sealed surface cover in each time step largely match the proportions of built-up land use reported by the Belgian Cadastre (Table 5.2). The average daily growth of sealed surfaces was 10 ha between 1987 and 2001, and 6 ha between 2001 and 2013. During the same period, the population evolved from 6,652,730 in 1986 to 7,536,494 in 2013, or a 13% overall growth (Table 5.3). The average annual population growth increased from 19,246 during the 1986–2001 period to 49,589 during the 2001–2013 period. In other words, while the population growth rate has risen significantly, the sealed surface growth rate has declined.

Table 5.3 Population of Flanders and the Brussels Capital Region for 1986, 2001, and 2013.

Year	Flanders	Brussels Capital Region	Total
1986	5,676,194	976,536	6,652,730
2001	5,967,946	973,475	6,941,421
2013	6,381,859	1,154,635	7,536,494

To show and analyse the trends observed, we have aggregated our 30 m population maps obtained for the three dates to a 300 m resolution. Between 1986 and 2001 the population decreased in almost all urban centres, Antwerp, Ghent and Bruges being the best examples, as well as in some rural areas (Figure 5.3a). In the BCR the trend was more mixed, with some neighbourhoods increasing and others decreasing in population, resulting in almost no total population change over the whole period (Table 5.3). Suburban areas attracted the most population, especially around Antwerp and in the north of Limburg province. Another growth zone was the Belgian coast, situated at the northwest edge of West Flanders province. Between 2001 and 2013 population in all cities increased, especially in Brussels, Antwerp and Ghent (Figure 5.3b). At the same time, the population increase was still considerable in suburban areas, as well as along the Belgian coast. Decreases in population were mostly limited to rural areas, with some neighbourhoods at the edges of cities as exceptions.

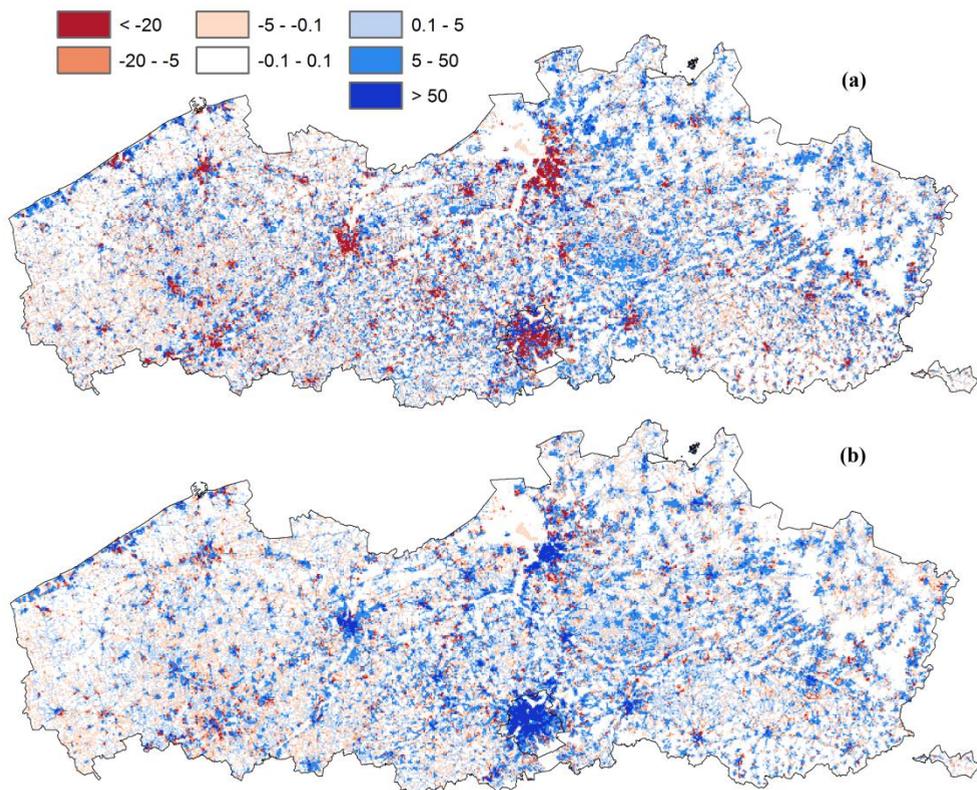


Figure 5.3 Population difference maps (aggregated to a 300 m resolution) for the periods (a) 1986 – 2001, and (b) 2001 – 2013.

5.3.2 Population change versus changes in built-up area

We also aggregated the results obtained for the densification index of equation (5.5) to a 300 m resolution, and excluded all 300 m cells with a population below 10 in the start year of a time period to avoid outliers, as well as all cells with a population decline during that time period. Before 2001 the trend towards sprawl is clear in almost the whole of Flanders (Figure 5.4a): index values below 1 indicate that sealed surfaces grew more strongly than the population. The lower the population density of a cell, the more obvious is the trend (Table 5.4).

Table 5.4 Average (Avg) densification index values and standard deviations (SD) for different categories of population per 300 m cell in the population density map for 2013. Cells with a population below 10 were excluded as outliers.

Population per 300 m cell (2013)	Densification index D_s for cells with growing population			
	1986 – 2001		2001 – 2013	
	Avg	SD	Avg	SD
10 – 30	0.866	0.187	1.006	0.139
30 – 100	0.905	0.181	1.022	0.170
100 – 300	0.948	0.201	1.043	0.170
> 300	0.990	0.152	1.098	0.169

There are only a few exceptions, with the Belgian coast standing out. After 2001 some regional differences appear (Figure 5.4b). In general, the densification index is higher for cells with a higher population density (Table 5.4). In most cities, the densification process brought extra population to areas where sealed surface fractions were already high. In many suburban zones, the increase of sealed surfaces was proportional to the population growth. However, in some rural areas with an increasing population, and especially in suburban areas north of Antwerp and in the centre of Limburg province, there was still a stronger increase of the built-up area than of the population. This continuing process of sprawl is also observed at the level of municipalities in the north and east of Flanders (Figure 5.5). Not all municipalities covering urban core areas have the highest values for the densification index, since they often are much larger than only the urban core. Some smaller municipalities in the ‘Flemish Diamond’, the central region between the cities of Brussels, Ghent, Antwerp and Leuven, have equally high index values.

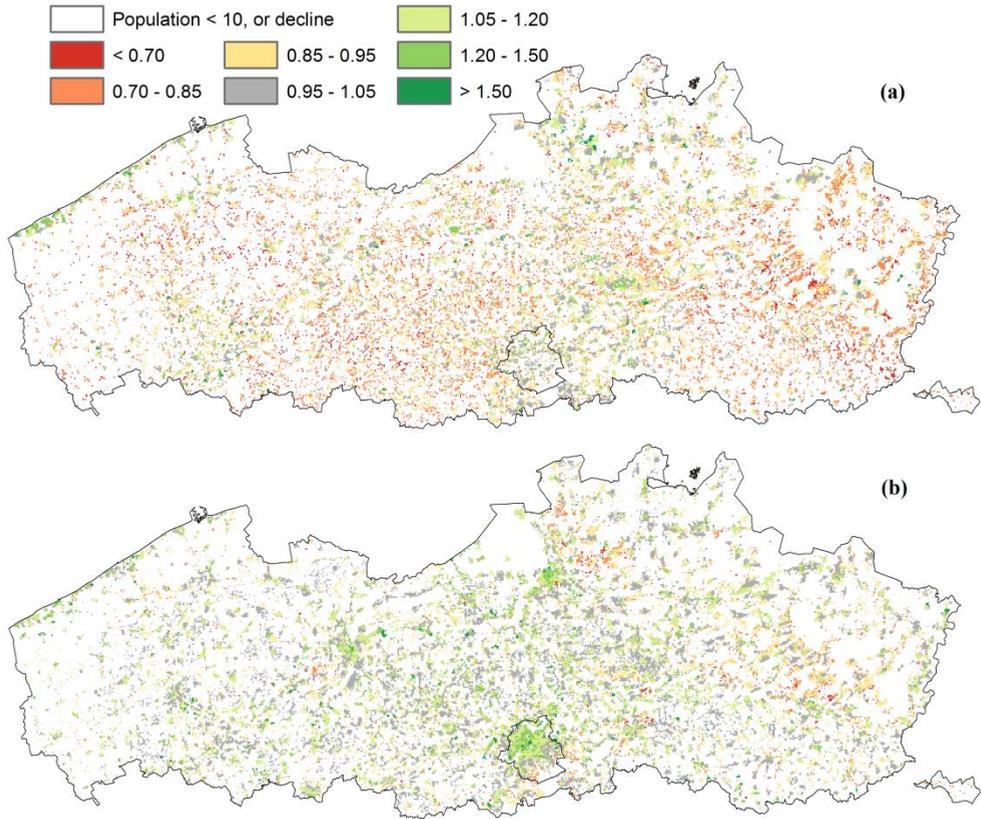


Figure 5.4 Densification index maps (300 m resolution) for cells with a growing population and a minimum of 10 inhabitants for the periods (a) 1986 – 2001, and (b) 2001 – 2013.

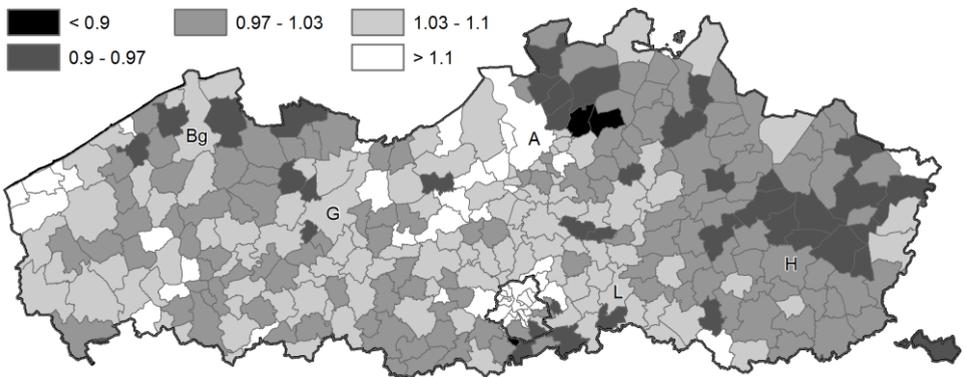


Figure 5.5 Average densification index values per municipality for the period 2001 – 2013. The Flemish provincial capitals are indicated: Bruges (Bg), Ghent (G), Antwerp (A), Leuven (L) and Hasselt (H).

5.3.3 Uncertainty propagation analysis

Since error in sealed surface maps can be substantial, we wanted to investigate if this causes a high uncertainty in the resulting population and densification maps. There is no significant bias in sealed surface error and the errors for each time step prove to be normally distributed, so 100 perturbed sealed surface maps for each time step could be produced using a multivariate normal error distribution. Like in the original procedure, the maps obtained for the three time steps in each of the 100 simulations were made temporally consistent. The standard deviation of the 100 resulting sealed surface fractions, averaged over all cells, which provides a measure of the magnitude of the uncertainty, is for all time steps around 15% (Table 5.5). Errors are lower in urban areas, and more specifically in the three largest cities (Brussels, Antwerp and Ghent). While these errors are significant, they seem to level each other out when sealed surface ratios are used together with other available information to produce population maps. The average standard deviation in the population maps (aggregated to the 300 m resolution used in the population and densification analysis) is only around 1 person per cell (Table 5.5). Absolute uncertainty values are higher in dense urban areas since the population per cell is higher and the sealed surface differences in time are mostly limited. However, the standard deviations are still relatively small with values of around 10 persons per 300 m cell (Figure 5.6), while most of these cells are inhabited by 1000 to 2000 persons. The densification indices have a limited uncertainty too (Table 5.5) in comparison with the range of observed values in Figure 5.4. There is only very little spatial variation in densification uncertainty, except in a few cells with changed urban functions, like old industrial terrains that were reconverted into dense residential areas.

Table 5.5 Standard deviation of sealed surface fraction (30 m cells), population estimate (inhabitants per 300 m cell) and resulting densification index (300 m cells), averaged over all cells, in a Monte Carlo uncertainty propagation analysis.

	1986	1986-2001	2001	2001-2013	2013
Sealed surface fraction	0.1560		0.1592		0.1574
Population	1.4635		1.0296		---
Densification index		0.01655		0.01589	

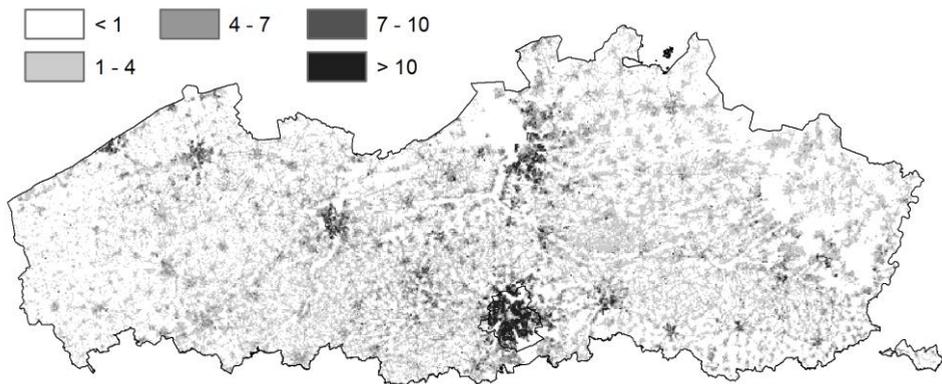


Figure 5.6 Standard deviation of the estimated number of inhabitants in a Monte Carlo analysis at 300 m resolution for 1986.

5.4 Discussion

The origins of urban sprawl and ribbon development in Flanders go back to the 19th century (De Decker, 2008; Verbeek et al., 2014). The combination of population growth with the construction of a dense road network in the second half of the 20th century have caused even more widespread sprawl

with large areas of ribbon development between villages. This phenomenon can be seen in the land-use map (Figure 5.1), but is even more striking when analysing detailed population maps as produced in this study (Figure 5.2). A real urban flight during the last decades of the 20th century towards suburban areas caused a huge intake of natural and agricultural land. Sealed surfaces grew more than would be expected from population trends, as can be inferred from the low densification factor values in almost all areas with a population increase in the period 1986–2001. An exception is the Belgian coastline, which has been almost entirely filled up with high rise apartment buildings for decades. These buildings were originally intended for recreational stays but have gradually been more and more used as permanent residences, especially by the growing number of older people (Schillebeeckx et al., 2013). Moreover, urban growth is not possible anywhere near the coast since the dunes are protected by a “Dune Decree”. One of the causes of the population decline per unit of built-up area in the rest of Flanders and the BCR is the shrinking household size. In Belgium an average household had 2.73 members in 1981, 2.52 in 1991, and 2.31 in 2008 (Deboosere et al., 2009; FOD Economie, 2011). 30% of all households in Flanders and 50% in the BCR consisted of singles in 2008.

The massive sprawl has a long history, and the proposed federal and later regional policies have not always been effective. The first Belgian law on spatial planning was only established in 1962, and as a result 48 subregional zoning plans (“Gewestplannen”) were defined in the 1970s and 1980s (Albrechts, 1999). Many differences exist between the different subregional plans, but generally, rather large zones were still defined as possible construction areas. During the development phase of the plans, local authorities served the interests of local landowners and developers by

increasing the number of housing construction zones. Exceptions could also be approved even after the development of the plans (Albrechts, 1999; Halleux et al., 2012, Boussauw & Boelens, 2015). Likewise, ribbon development continued to occur along roads where building was still allowed (Verbeek et al., 2014). Later, spatial planning policy was regionalised and Flanders developed a general Spatial Structure Plan (Ruimtelijk Structuurplan Vlaanderen, RSV) in 1997. To reduce future urban sprawl, the plan adopted a few clear concepts from the rigid Dutch spatial planning process, like ‘deconcentrated clustering’: i.e. clustering but at different places in Flanders (Albrechts, 1999). Urban areas were demarcated that should contain most additional homes, but a revised version of the RSV specified that only 54% of the new housing needed to be within the urban boundaries (Voets et al., 2010; Boussauw et al., 2013). As a consequence, even in this plan residential development is still allowed in a large number of non-urban parcels (De Decker, 2008).

Since 2001, the urban population has significantly increased. Household sizes are larger again in cities, especially in the BCR (FOD Economie, 2011). Migration is an important factor in this new urban revival since cities serve as an arrival station for newcomers (Schillebeeckx et al., 2013). In our analysis we see indeed that migrant neighbourhoods in Brussels and Antwerp have the largest densification index values. However, we also clearly observe population growth and a densification increase in almost all other urban neighbourhoods, including those of smaller towns and in centres of suburban municipalities, especially in the ‘Flemish Diamond’. On the other hand, most suburban areas in the north and east of Flanders and some rural areas still suffer from sprawl.

The beginning of a trend towards densification seems to be there, but many authors believe that spatial planning in Flanders still has to improve. The RSV is in the process of being replaced by the Spatial Policy Plan for Flanders (Beleidsplan Ruimte Vlaanderen, BRV), yet Boussauw and Boelens (2015) criticise the content of the green paper of the plan. Among other things, they indicate the persistence of exceptions for large projects, the lack of quantification and measurable goals, the lack of cooperation between the Belgian regions, and more specifically the restrictions on the growth of the Brussels urban area in the fringe governed by the Flanders region encapsulating the city. Due to regional politics, the municipalities of the fringe do not want to accommodate more inhabitants of the BCR who are mainly non-Dutch speaking. Nevertheless, the recent strong growth of the population in Brussels caused a rise of real-estate market prices which is leading to new suburbanisation further away from Brussels, in both Flanders and Wallonia (Boussauw et al., 2013). Leaving the city is often not the preferred choice of young adults in Belgium who desire to acquire a home, but the prices force them to do so. Living in less attractive urban neighbourhoods or renting instead of buying can be intermediate solutions or alternatives (Slegers et al., 2012). Within the cities, there is also an increasing segregation between ethnic groups. The proportion of migrants living in often poorer neighbourhoods seems to be an important negative attractiveness factor for those who can afford to live elsewhere (Schuermans et al., 2015).

An enhanced spatial planning policy in Belgium clearly needs quantifiable goals to prevent the growing population from moving into the rural zones, and rather to house it in available space in the existing urban centres. Both long-term land use modelling and specific case studies

assessing the impact of different planning scenarios can help to reach these goals. Job accessibility should again become a more important factor in residential location choices in order to decrease the average commuting distance. The recently published white paper of the BRV (Ruimte Vlaanderen, 2016, 2017) seems to be more ambitious and has set up clear, quantifiable guidelines on which factors need to be taken into account to reduce the area of the present zones for urban expansion. Such spatial visions can be examined with the activity-based land-use model (see chapter 7) which needs to be fed with reliable population and employment data to allow for a proper calibration. The approach proposed in this study for producing detailed historical population maps may help to do so. Between 1986 and 2001 the relation between population and sealed surface growth was not constant: for a 27% increase of sealed surface cover in Flanders, there was only a 5% increase of the population. This means that the computed population density map for 1986 may be less reliable as an input for historical calibration of the activity-based land-use model for Flanders. From 2001 densification factor values are close to 1 for the less populated cells, which suggests that the hypothesis of equation (5.3) is valid for this period (Table 5.4). A better input population raster for the most recent year (2013) would help to reduce some of the errors observed, especially in rural neighbourhoods. Within urban cores, and especially those of the largest cities, a densification trend is observed between 2001 and 2013, yet one should keep in mind that it is more difficult to examine the ratio between sealed surfaces and population in an environment that is already largely built-up. The detailed population predictions within city centres for 2001 are therefore assumed to be less reliable than the local predictions in rural areas. Since the prime purpose of an urban growth model remains the prediction of

the future extension of urban land uses, the limitations of the proposed downdating approach seem acceptable for historic calibration purposes.

In summary, downdating a spatially detailed population density map has the advantage of using today's high resolution information for predicting population densities in the past, but involves the risk that errors will increase when going further back in time. To build scenarios for Flanders' and Brussels' future, a calibration based on the most recent period in the past (2001 – 2013 in this study) should give the best results since this period includes the start of the re-urbanisation process and the hypothesis of a local relationship between population density and sealed surface cover seems to hold rather well for this time frame. Downdating of population based on time series of sealed surface maps may be useful for the calibration of the activity-based CA model proposed in this research and was also the incentive for developing the mapping approach proposed in this chapter.

5.5 Conclusions

This chapter presents an approach for downdating a detailed population density map for a recent date, using a time series of sealed surface data for the past, obtained through remote sensing. The computed historical population density maps revealed spatial differences and changing trends in the relation between population and sealed surface cover in Flanders and Brussels. In the last decades of the 20th century the region experienced an urban population decline and a massive increase of urban sprawl. Since 2001 the urban population has been increasing again, primarily due to immigration, but in the centres of smaller towns also as a result of the advent of an overall spatial structure planning policy that was missing before. Nevertheless, the built-up area used per inhabitant is still increasing

in some rural areas and several suburban municipalities in the north and east of Flanders.

Land-use models can help spatial planners to define objectives for the spatial allocation of the still expected population growth in the future. These land-use models should incorporate population, employment and transport data at a detailed scale level. Datedating of recent population data based on an existing high-resolution population raster and sealed surface fraction estimates of the past, as proposed in this study, may be useful for producing historical data needed for the calibration of such models.

Appendix 5.A1 Datedating of land-use maps

Historical land-use maps are necessary to feed the historical calibration of the activity-based CA model. Methodologies that try to derive urban land-use categories directly from remote sensing images often have poor quality and low consistency between years, especially when simple pixel-based classification algorithms are used (Grinblat et al., 2016). Therefore, we used a workflow similar to the one applied above to the population map: VITO's land-use map for Flanders is datedated with the help of the sealed surface maps of the past. The validation of the approach is more complicated than for activities. We have municipal land-use totals of the Cadastre for each year, but the Cadastre uses another definition of built-up land uses. Therefore, we decided to do a visual validation with a sample of aerial photographs. The photographs were re-used from the PhD research of Lien Poelmans who developed land-use maps for Flanders for the years 1988 and 2000 (Poelmans & Van Rompaey, 2009, 2010), two dates which are close to 1986/1987 and 2001.

An automated workflow was programmed in ArcPy, the Python-based scripting language of the ArcGIS software. A parcel map of the whole study area has to be provided. The user can define a whole set of thresholds on values of sealed surface proportions in parcels in the present and past maps, and in difference maps between the present and the past. The key concept in the proposed approach is the computation of the percentage of changed cells per parcel. The simplest definition of changed cells is to consider all cells with sealed surfaces in the present map but without any sealed surfaces in the past. Alternatively, cells with a “small” amount of sealed surface in the past which is considerably “lower” than in the present (two thresholds to be calibrated) can be added when the above-mentioned percentage is computed. This addition helps to deal with small errors in the sealed surface maps, and more generally with land-use changes from agricultural to urban categories since agricultural land uses could have had a small portion of sealed surface in the past. Note that built-up plots in agricultural lands are in principle classified as residential land use in the land-use map of the present that we used. Next, different thresholds of the minimum percentage of changed cells can be specified for different urban land uses. When the threshold is exceeded in a parcel, then the parcel is considered to be unbuilt in the past.

After the previous parcel-based step, some isolated pixels are still marked as built-up because of inconsistencies between the parcel map and the land-use map. Therefore, a filter is applied to all cells adjacent to a minimum number of already removed cells and a maximum number of still built-up cells of the same land-use category. Initially, the filter was only applied to cells nearby roads since most of the isolated pixels are next to

roads. However, this extra condition was dropped in order to provide a solution for all isolated cells.

Finally, the land-use states in the past of the changed parcels have to be determined. Agriculture and nature are the two rather general vacant unbuilt classes in the land-use map. Remote sensing products do not provide a solution here since some agricultural and natural land uses are difficult to distinguish in remote sensing images at the level of individual pixels as both have a land cover of vegetation. Hence, we decided to use a simple context-based approach: the vacant unbuilt category that is the most present today in a radius of 1 km on the land-use map is chosen as the most likely category of the parcel in the past.

The visual validation with aerial photographs generally had very good results. However, in rural areas of the West Flanders province, and especially in 1986, small agricultural built-up plots seem to be underestimated in the downdated map. Due to these errors and since the land-use growth and population growth trends are rather different in the periods 1986-2001 and 2001-2013, the 1986 land-use map was finally not used in the calibration discussed in the next chapter.

Appendix 5.A2 Datedating of employment maps

Employment values (number of jobs) in June 2011 and coordinates of large company establishments (32.6 % of total employment) were provided by the Belgian Federal Ministry of Mobility, and total number of jobs per municipality per economic sector (NACE code) in 1986, 2001 and 2013 by the Belgian Federal Ministry of Economics. The 10 m resolution land-use map of VITO groups the NACE code sectors into 21 classes. Employment in the primary sector (agriculture) was not considered as agriculture is

simulated as an area-based land use. Total employment in 2013 was computed for each of the 21 sectors per municipality, and then in a first stage evenly redistributed over all 10 m cells of the economic sector within the municipality.

This distribution was then updated with the data of the large company establishments, even though there is a time difference between both data sets of 1.5 years. Total employment of the establishment was redistributed over the parcel containing the point location of the establishment. Finally, each 10 m cell gets the maximum value of the simple first-stage map and the large companies map, and this resulting map is rescaled to get correct municipal totals for the 21 employment sectors in 2013. For model calibration and simulation, the maps were generalised to a 100 m resolution and the 21 sectors were grouped into 4 more general economic sectors: industry (including energy, mining, waste and water treatment), wholesale and logistics, retail, and services.

The construction of the employment maps of 1986 and 2001 starts from the 2013 maps of the 4 general sectors. Since (changes in) employment cannot be related to (changes in) sealed surface fractions, the employment values per economic sector were simply rescaled within the already available cells of the economic sector on the land-use map of 1986 and 2001 to have the correct totals per municipality in 1986 and 2001.

6. Semi-automated calibration

An alternative version of this chapter will be submitted as: Crols, T., van der Meulen, M., White, R., Poelmans, L., Uljee, I., Engelen, G., & Canters, F. Semi-automated calibration of an activity-based cellular automata model. *Environment and Planning B*, special issue on Cellular Automata Modeling for Urban and Spatial Systems.

6.1 Introduction

Cellular automata (CA) models and other land-use change models can be calibrated with expert-based knowledge, statistical analysis, manual trial-and-error calibration, or automated calibration (van Vliet et al., 2016). Modellers have experimented with a wide range of automated calibration techniques. Specific algorithms were implemented to improve CA influence functions, based on the iterative modification of neighbourhood evaluation functions to improve simulation results (Straatman et al., 2004; van Vliet et al., 2013b; Liao et al., 2016). Others applied more general optimisation algorithms: some examples are brute-force calibration (Silva & Clarke, 2002), artificial neural networks (Li & Yeh, 2002; Pijanowski et al., 2002; Almeida et al., 2008; Basse et al., 2014), ant colony optimisation (Liu et al., 2007), particle swarm optimisation (Feng et al., 2011; Liao et al., 2014), support vector machines (Yang et al., 2008), and genetic algorithms (Colonna et al., 1998; Goldstein, 2004; Li et al., 2008, 2013; Shan et al., 2008). Artificial neural networks (ANN) are amongst the most powerful optimisation routines, but they operate as a black box and involve a certain risk of over-calibration towards the end map of the calibration period. Genetic algorithms (GA) are powerful too, allow for more flexibility since

error functions can be easily defined, and are as such easier to adopt in a semi-automated approach.

Fully automated calibration algorithms for land-use models can have two important disadvantages. Firstly, the algorithms can get stuck in local minima instead of reaching a global minimum. Secondly, some mathematically optimal solutions can involve unrealistic relationships. The model is then over-calibrated, as the optimisation routine returns a combination of parameter values that are illogical in terms of the actual processes that determine land use. Nevertheless, they may lead to a low value of the error function. If the model with these parameter values is run into the future, the resulting patterns are likely to be seriously in error. The calibration of the activity-based CA model has initially been done in a manual fashion, applying trial-and-error to avoid the problems listed above. However, since the model not only involves the weights of the neighbourhood functions but also a number of activity relocation parameters, manual calibration gets overly complicated and slow. Therefore, it seemed interesting to test if we can define a semi-automated workflow to obtain realistic parameters and neighbourhood rules for the simulation of activities and land uses.

To compare model results with ground truth land-use maps an error function has to be defined. Some functions involve averaging of local errors while others look at differences in the general pattern obtained. Local differences in raster maps are often analysed with the Kappa statistic, a cellular accuracy metric with a correction for the agreement as expected by chance (Cohen, 1960; Monserud & Leemans, 1992). Fuzzy Kappa is a variant in which near-hits in location or category are treated as more or less correctly classified cells (Hagen, 2003; Hagen-Zanker, 2009). Recently, Van

Vliet et al. (2011, 2013a) presented an alternative with (Fuzzy) Kappa Simulation in which only the error of those cells whose category changed in comparison with the start map is considered. An alternative to the Kappa statistic is to use multiple-resolution error functions (Costanza, 1989; Pontius et al., 2004), which summarise the errors at different scales. Although a numerical Kappa exists for continuous data (van Vliet et al., 2010), the multiple-resolution functions are easier to use with a model that combines continuous and categorical data (Pontius et al., 2008), as is the case in the activity-based CA model.

Multiple-resolution functions analyse both local and regional errors, but not the differences in the shape and the pattern of patches. The general land-use pattern can be explored with spatial metrics or fractal analysis (Engelen & White, 2007). Spatial metrics are indices that quantify the structure and patterns of the landscape. They have their foundations in landscape ecology but have become widely used in land-use and urban modelling (Herold et al., 2003, 2005; Huang et al., 2007; Liu et al., 2010a; Van de Voorde et al., 2016). Metrics can be calculated at the patch, class and landscape level or within a moving window. Patches are defined as homogenous regions for a specific landscape property, such as specific land-use types. Class metrics represent the spatial distribution and pattern of all patches belonging to a single land-use type, whereas landscape metrics represent the spatial pattern of the entire landscape, considering all land-use types simultaneously (McGarigal et al., 2002). Unfortunately, no standard set of best-suited spatial metrics for model calibration and validation is available. Fractal analysis is another possibility to evaluate the global map similarity, e.g. by plotting a graph of the cluster size-frequency relationship (White & Engelen, 1993; White, 2006b). This relationship should remain

more or less constant over time, even for long-term predictions. Both spatial metrics and fractal analysis apply to categorical maps only, and were therefore in this study only considered to validate land-use change but not to calibrate the model.

In the following sections, we present a semi-automated approach to calibrate the activity-based CA model with a genetic algorithm and multiple-resolution error functions, applied to the Greater Dublin Region in Ireland for the period 1990–2000, and to Flanders and the Brussels Capital Region in Belgium for the period 2001–2013. The proposed calibration methodology allows us to compare the different versions of the model presented in this thesis. Other than evaluating the Euclidean and the network-based model versions, we also used the proposed calibration framework to assess small changes in the definition of some model equations. Two types of independent validation of the results for the residential land-use class are proposed: (1) class-level spatial metrics to compare the final simulation map with an observed map; and (2) cluster size-frequency dimension graphs to check if the fractal dimensions of the residential class remain constant when the simulation is extended to a long period.

6.2 Methods

6.2.1 The activity-based CA model

A full description of the activity-based CA model without and with network-based rules can be found in chapters 3 and 4. In this chapter we discuss a calibration of both model versions, and we also evaluate the different equations to update activity densities that were proposed in chapter 3 (equations (3.19), (3.20) and (3.21)). Equations (3.20) and (3.21) increase

the direct impact of the relative size of the neighbourhood effect on population, the most important factor in cities to determine the location of more rapidly densifying areas. The first alternative uses a densification exponent, while the second alternative makes activity growth directly dependent on the relative neighbourhood effect on population.

6.2.2 Semi-automated calibration with a genetic algorithm

Genetic algorithms (GA) are optimisation routines (Holland, 1975; Goldberg, 1989) that are well suited to handle the non-linear relationships of CA land-use change models. Originating in genetic and evolutionary theory, a number of individuals (possible solutions) each endowed with a different chromosome set (parameters set) is evaluated with a fitness function (mostly an inverse error function). To iteratively optimise the fitness function, new generations of individuals will have genes based on the fittest individuals of the previous generation, but slightly altered by mutation and crossover processes, while many less fit individuals will have been eliminated.

For the activity-based CA calibration, we used the freely available GAUL library (Adcock, 2005), which we extended with a GA based on NSGA-II (Deb et al., 2002). The NSGA-II algorithm has previously been successfully applied in spatial land-use optimisation problems (Datta et al., 2007; Cao et al., 2011). We replaced the NSGA-II ranking method by a novel ranking method, the ‘push-point method’, to keep as much diversity as possible within the population. First, a fitness threshold (the push point) is computed, which is exceeded by 95% of the population (‘group 1’ individuals). Then, mutation and crossover are executed to double the population. All individuals better than the push point are ranked in group 1, while the other individuals are ranked in ‘group 2’. Finally, the NSGA-II

elimination is executed to halve the population, but a maximal diversity is retained within the group 1 individuals.

The chromosome set of the GA includes two types of parameters to be optimised. The first parameter type are the weights of the neighbourhood influence functions. These are the inertia values I_K of equation (3.10) and all weights $W_{JK,dij}$ of equation (3.9). We simplified the functions by representing the weights at all possible distances by only seven parameters per influence of activity J on activity K (Figure 6.1). The first parameter $p_{K,1}$ is the weight at zero distance, i.e. the inertia value I_K , which is zero for $J \neq K$. A different distance-decay behaviour of the weights for immediate neighbourhood effects (roughly 1 km) and long-distance effects (> 1 km) was observed by White et al. (2012) and in the results of the network-based model in chapter 4. In the network-based model, travel time-decay functions are used for the long-distance part of the function. We define $p_{JK,2}$, $p_{JK,3}$ and $p_{JK,4}$ as the weights of the three most nearby rook variable grid neighbour influences (at log-base 3 cell distances 0, 1, and 2). These represent the effect that activities have on each other in the immediate neighbourhood. The three weights are interpolated to get the values of bishop variable grid neighbours (at log-base 3 cell distances 0.315 and 1.315). A distance ($p_{JK,5}$) and a weight ($p_{JK,6}$) define a turning point between short-distance and long-distance influences. Finally, a decay exponent ($p_{JK,7}$) determines an inverse power function that is used to compute all weights of influences that activities have on each other at distances larger than the turning point:

$$W_{JK} = \frac{p_{JK,6}}{(p_{JK,5})^{-p_{JK,7}}} L^{-p_{JK,7}} \quad (6.1)$$

with W_{JK} the influence weight of activity J on activity K , and L the log-base 3 cell distance. The seven parameters have to be optimised for a (preferably limited) chosen set of activity combinations, which are considered to capture the essential influence effects. Next, the second parameter type of the calibration are the most important parameters that control the different potentials, the diseconomies of scale effect and the changes in density of activities (equations (3.10) to (3.12) and (3.16) to (3.21)). There are two general parameters (ρ and ε) and four activity-specific parameters (λ_K , m_K , γ_K , and τ_K or ζ_K) that can be optimised. The role of these parameters in the model is discussed in chapter 3.

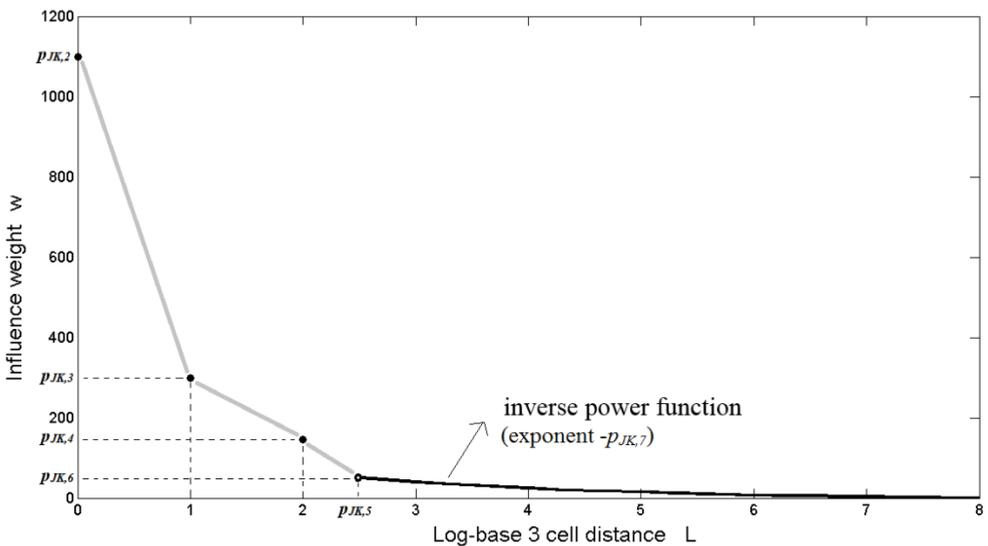


Figure 6.1 An influence function of activity J on activity K , defined by its inertia $p_{JK,1}$ (not in the figure) and 6 other parameters ($p_{JK,2} - p_{JK,7}$).

We defined a semi-automated calibration strategy that includes expert-based knowledge to avoid over-calibration. The modeller chooses a subset of influence rules and activity parameters to be calibrated. For each input value, a minimum, a maximum, and an initial best guess, have to be

defined. The first generation of the GA always contains an individual which has all best guesses as its values. For the other individuals, all values are randomly drawn with an equal chance to be between the minimum and the best guess or between the best guess and the maximum.

After running the model, the fitness of each individual is evaluated with the following function:

$$F = 1 / (1 + E) \quad (6.2)$$

with F the fitness of the individual, and E a multiple-resolution error function that maps the difference between the simulated and observed land uses and activities.

6.2.3 Multiple-resolution error functions

The quantities of all activities and land-use categories are always exogenously defined for all time steps in the ACA model. As such, errors due to quantity are impossible, hence, only the errors due to location have to be quantified. We chose to systematically account for near-hits by computing differences at several resolutions and in all possible overlapping windows. The most straightforward way to implement this methodology in the activity-based CA model is by recycling the concept of the variable grid and use its different resolutions (or ‘levels’ L) as the different window sizes. The smallest level ($L = 0$) are the individual cells of the CA model (with resolution R), and then the resolution increases by a factor of 3 (and the area by a factor of 9) with every level. Therefore, we decided to calculate weighted sums of land-use category root mean square errors (RMSE) per level of the variable grid, and activity RMSE per level of the variable grid. Individual errors of all categories at all resolutions are normalised to allow

for the definition of a general error function that combines the different errors. This is necessary since activities can have different units, while land-use categories are just present or not in a cell at the model resolution.

The error $E_{i,L}$ due to deviations in land-use category i at level L is then:

$$E_{i,L} = \sqrt{\frac{\sum_{w=1}^{N_w} (\hat{n}_{i,w} - n_{i,w})^2}{N_w}} \frac{1}{S_w \bar{n}_{i,0}} \quad (6.3)$$

with w a window with resolution $3^L R$ and area $S_w = 3^{2L} R$, N_w the total number of overlapping windows, $\hat{n}_{i,w}$ the number of simulated cells of land-use category i in map window w , $n_{i,w}$ the number of observed cells of i in w , and $\bar{n}_{i,0}$ the proportion of cells at the model resolution ($L = 0$) having land use i .

The error $E_{k,L}$ due to deviations in activity k at level L is then:

$$E_{k,L} = \sqrt{\frac{\sum_{w=1}^{N_w} (\hat{A}_{k,w} - A_{k,w})^2}{N_w}} \frac{1}{S_w \bar{A}_{k,0}} \quad (6.4)$$

with $\hat{A}_{k,w}$ the total of simulated activity k in map window w , $A_{k,w}$ the total of observed activity k in w , and $\bar{A}_{k,0}$ the average value of activity k of all cells at the model resolution.

Functions with mean absolute errors (MAE) were initially also tested, but finally we chose the above mentioned equations with RMSE. Additionally, errors averaged over administrative units, instead of the multiple-resolution approach, were implemented too. As such, the

calibration results will approach correct regional totals. Specifically, we used the coefficient of variation of the RMSE (CV RMSE) of regional totals of activity or the number of land-use cells of a category. The error R_i due to deviations in land-use category i in all regions is then:

$$R_i = \sqrt{\frac{\sum_{r=1}^{N_r} (\hat{n}_{i,r} - n_{i,r})^2}{N_r}} \frac{1}{\bar{n}_{i,R}} \quad (6.5)$$

with r a region, N_r the total number of regions, $\hat{n}_{i,r}$ the number of simulated cells of land-use category i in region r , $n_{i,r}$ the number of observed cells of i in r , and $\bar{n}_{i,R}$ the average number of cells per region having land use i .

The error R_k due to deviations in activity k in all regions is then:

$$R_k = \sqrt{\frac{\sum_{r=1}^{N_r} (\hat{A}_{k,r} - A_{k,r})^2}{N_r}} \frac{1}{\bar{A}_{k,R}} \quad (6.6)$$

with $\hat{A}_{k,r}$ the total of simulated activity k in region r , $A_{k,r}$ the total of observed activity k in r , and $\bar{A}_{k,R}$ the average value of activity k per region.

A combination of optimising on multiple windows and on regions turned out to be particularly successful, although experiments in which only the windows or the regions were used also led to interesting observations. The total error E is a weighted average of the categorical multiple-resolution errors E_i and E_k at all levels L and the categorical regional errors R_i and R_k :

$$E = \frac{\sum_{L=0}^{L_{max}} (\sum_{i=1}^F (g_{i,L} E_{i,L}) + \sum_{k=1}^T (g_{k,L} E_{k,L})) + \sum_{i=1}^F (h_i R_i) + \sum_{k=1}^T (h_k R_k)}{\sum_{L=0}^{L_{max}} (\sum_{i=1}^F g_{i,L} + \sum_{k=1}^T g_{k,L}) + \sum_{i=1}^F h_i + \sum_{k=1}^T h_k} \quad (6.7)$$

with F the number of active land-use functions, T the number of different activities, L_{max} the highest level of the variable grid, $g_{i,L}$ the weights of the

multiple-resolution land-use category errors $E_{i,L}$, $g_{k,L}$ the weights of the multiple-resolution activity errors $E_{k,L}$, h_i the weights of the regional land-use category errors R_i , and h_k the weights of the regional activity errors R_k .

Finally, an option in which the land-use errors $E_{i,L}$ and R_i only handle cells with a changed land use in comparison with the start map, inspired by Kappa Simulation of Van Vliet et al. (2011), was initially developed, but did not lead to good results due to the fairly ubiquitous growth patterns present in the activity-based model.

6.2.4 Study area and data

The above presented methodology was applied to two study areas for which we have good data and which experienced a rather different development history during the time intervals considered. The Greater Dublin Region (GDR) had a strong, monocentric urban growth and population growth pattern in the 1990s, while Flanders had a scattered polycentric urban land-use growth pattern in the 2000s accompanied by a densification of the population in the largest city centres. The urban growth in Dublin should be easier to calibrate than the complex trends of the Flemish application. For Flanders though, we can fully test the abilities of the model since we have employment data and spatially more detailed population data, which are both missing for Dublin.

6.2.4.1 Greater Dublin Region

Economic and population growth in Dublin and its metropolitan region were high and also much stronger than in the rest of Ireland (Williams & Shiels, 2002) in the 1990s and early 2000s. As a consequence, urban sprawl was rapid: ‘edge cities’ developed next to highways, containing both new residential and economic activity. The Greater Dublin Region (GDR) has

been extensively studied in the Urban Environment Project of the Urban Institute Ireland and the Joint Research Centre of the European Union. It is the best developed example of the MOLAND model (Engelen et al., 2007) and was one of two study areas of the activity-based CA model of White et al. (2012).

For this study, we re-used the available CORINE land-use maps (1990, 2000, 2006) and population data of the earlier studies. We worked with 20 land-use categories at an aggregated resolution of 200 m: one active land-use with activity (residential), four area-based economic active land uses (industry, commerce, services and ports), seven passive land uses (three agricultural, three natural and one abandoned category) and eight static land uses (Figure 6.2). The 1990 and 2000 maps were used in the semi-automated calibration while the 2006 map was used for validation. Population data of EDs (629 small areas in the GDR) of the censuses of 1986, 1991, 1996, 2002 and 2006 were interpolated to obtain population data for the same dates as for the land-use maps. Population raster maps were constructed with the initialisation module of the activity-based CA model, as discussed by White et al. (2012).

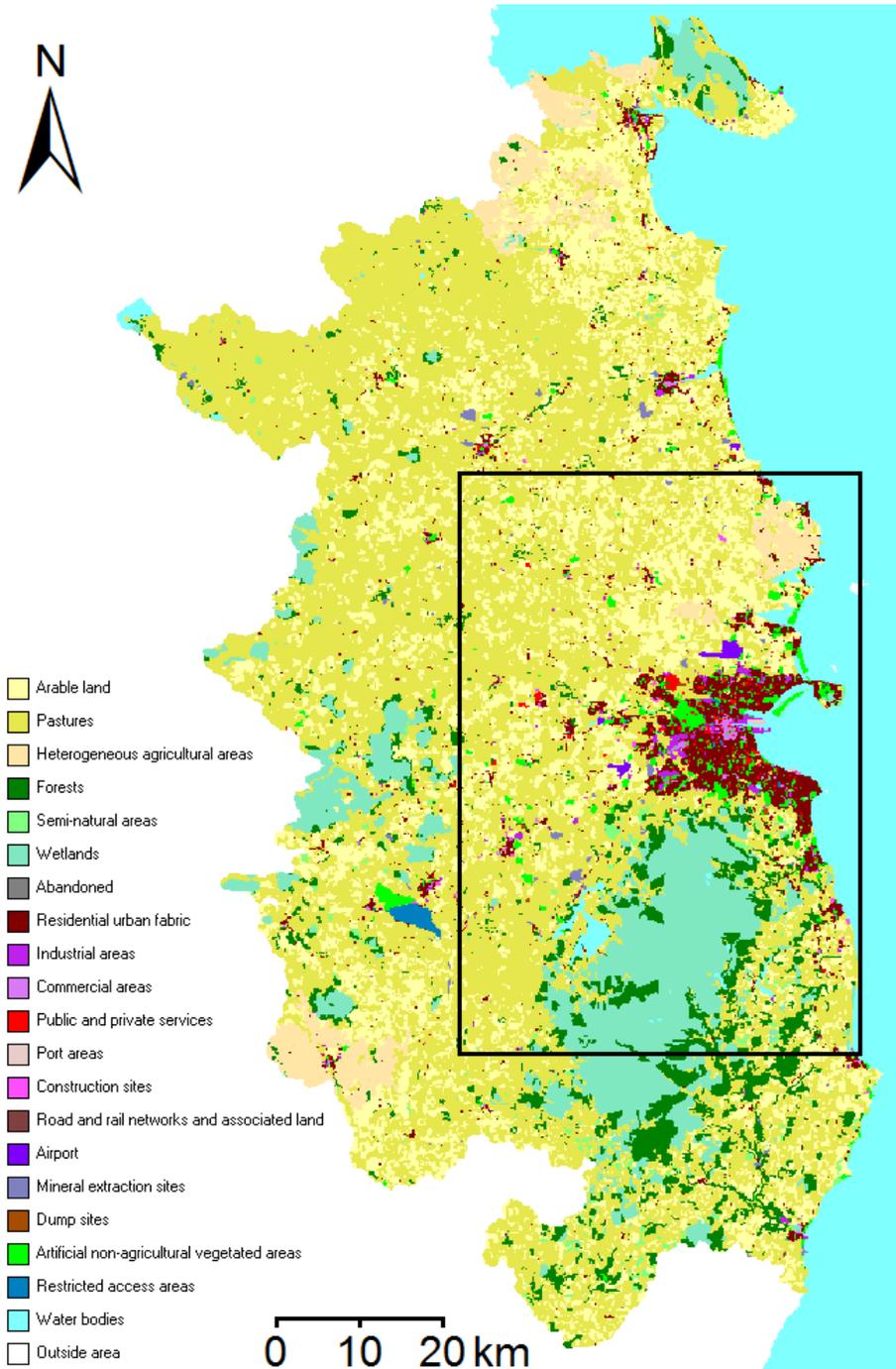


Figure 6.2 CORINE land-use map of the Greater Dublin region in 1990, aggregated to a 200 m resolution. The rectangle shows the extent of the zooms in Figures 6.4 and 6.5.

6.2.4.2 Flanders and the Brussels Capital Region

The northern half of Belgium, consisting of Flanders and the Brussels Capital Region (BCR), is a densely populated region with a long history of increasing urban sprawl (De Decker, 2008; Poelmans & Van Rompaey, 2009). Recently, however, densification of urban centres has started to emerge, and the rate of new sprawl in rural areas seems to be declining, as discussed in the previous chapter.

To feed our calibration routine, we used land-use and activity maps of Flanders and the BCR of the years 2001 and 2013 at 100 m resolution. The 2013 land-use map (original resolution 10 m) was provided by the Flemish Institute for Technological Research (VITO) (Poelmans et al., 2016). We reduced the number of land-use categories to 13 to make the calibration process faster and easier. The resulting map has five active land uses with activities (residential and four economic land uses), one area-based active land use (agriculture), and two passive land uses (nature and ‘other’, a rest category), while five other categories are kept static (Figure 6.3). A discussion of the population and employment maps used, and how the maps of 2001 were downdated, starting from the maps of 2013, can be found in chapter 5.

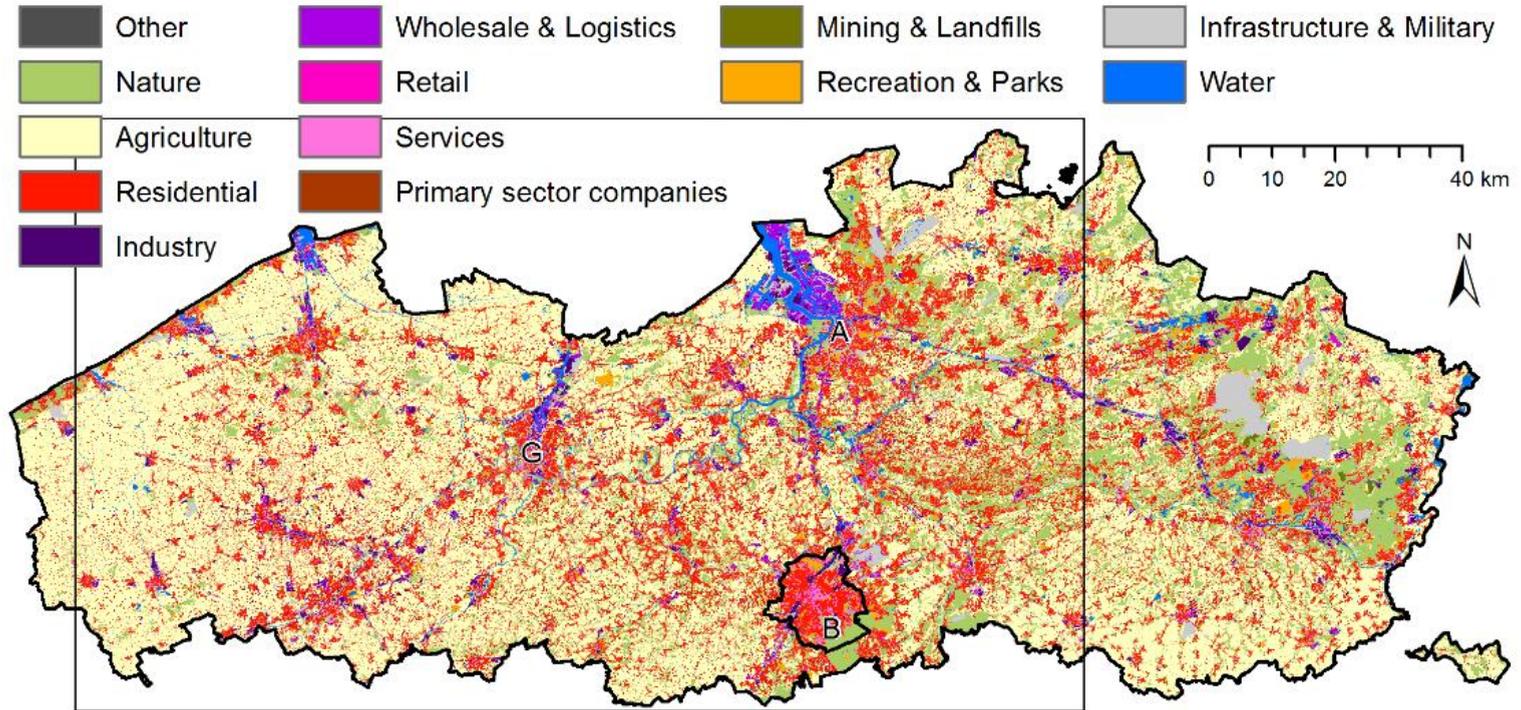


Figure 6.3 VITO land-use map of Flanders and the Brussels Capital Region in 2013, aggregated to a 100 m resolution, with an indication of the three largest cities (Brussels (B), Antwerp (A) and Ghent (G)). Regional borders in thick black lines. The rectangle shows the extent of the zooms in Figures 6.12 and 6.14.

6.2.5 Experimental design of the calibration

Different optimisation experiments were chosen to test if the semi-automated calibration routine is able to find lower errors for the results of the activity-based CA model when applied to the Greater Dublin Region (time period 1990–2000) and to Flanders and Brussels (time period 2001–2013), in comparison with manual calibration, and if the routine can be applied to compare different versions of the model. Not all possible influence rules had to be tested: to avoid a possible impact of non-causal relationships on the outcome of the calibration process, we carefully selected the dominant influences that are known to cause change, based on earlier work with both the MOLAND model and the ACA model. We focused on positive influences (pull forces), mainly between the different urban categories: see also Table 6.5 and its footnote in the Results section for Flanders. Negative influence (push) weights are difficult to calibrate and to extend into the future, and their behaviour can generally be captured by a positive effect in other neighbourhood functions. For the Dublin Region, we re-used all influence rules of White et al. (2012) except the influence of population and residential land use on itself. We did a straightforward automated optimisation of this influence rule and of the activity-based parameters for the different versions of the model and densification equations (see below) and with two different weighted error functions. For Flanders, we fully tested the semi-automated approach by alternately selecting larger and smaller subsets of influence weights and parameters to be optimised in order to progressively reach globally optimised results, capturing all important influences between land uses and activities. The range of the parameter values, defined by the minimum and maximum value

that needed to be specified, were relatively large during initial experiments and were gradually made more specific too.

Depending on the size of the subset of weights and parameters to be optimised, GA experiments were set up with a population of 200 to 2000 individuals, 20 to 50 generations, and 1 to 4 clean starts with different random seeds to initialise the GA. For computational reasons, we calculated the errors in a variable grid with a base grid resolution that is three times larger than the model resolution (i.e. level 1 instead of level 0, or 300 m in Flanders and 600 m in Dublin), and we did not calculate the errors of the highest level which has cells almost as large as the study area. Hence, overlapping windows at all resolutions are also separated by a distance of three times the model resolution. Next to multiple-resolution errors, we also computed errors with the Greater Dublin counties and the Flemish districts (or “arrondissements”) as spatial units, and we computed a weighted sum of errors in multiple windows and administrative regions. After initial experiments, the weights of the total error function (E in equation (6.5)) were chosen with more emphasis on population and residential land use, and mostly with lower weights for the lowest levels (see Table 6.1 in the Results Section for Dublin and Tables 6.A1 and 6.A2 in Appendix 6.A1 for Flanders).

The general methodology can be used to simultaneously optimise the parameters and rules of all activities and land uses by minimising the error of all activities and active land-use totals per window size. Nevertheless, experiments in which only specific errors are optimised are often very useful: e.g. by only looking at population or residential land use. For Flanders there are more activities and urban land uses in the model, thus we experimented with optimising all or a subset of them. We decided to focus

in some experiments on only optimising errors of population (and residential land use) since this is the key activity to simulate in the model and since the population data have a better spatial accuracy than the employment data.

The full experimental design was applied to several versions of the activity-based CA model, including the original version of White et al. (2012) with Euclidean distances only, and the update proposed in chapter 4 with travel time-based influence rules, both with the original and the alternative activity density equations – equations (3.19), (3.20) and (3.21) – as described in chapter 3. Due to the extensive nature of this experiment, we will focus in the results section on the most significant results and findings and will not give a complete overview of all experiments.

6.2.6 Validation of the global pattern of the residential land-use class

A validation of the results can be done with the same error functions or metrics used in the calibration, but preferably for a separate validation period. For the Belgian study area, however, maps were only available for two dates. Therefore, we did both the calibration and validation of this study area on the calibration time period (2001 – 2013). For the calibration, the error is represented by multiple-resolution and/or regional RMSE error functions, while spatial metrics were used as an independent validation of the results. For Dublin, we used the same procedure for the calibration period (1990 – 2000), but since data are available for 2006 we also validated the calibration with runs to that date.

We used a number of class-level spatial metrics to analyse the resulting pattern of the residential urban fabric. All metrics were calculated using the Fragstats 3.3 software (McGarigal et al., 2002). Spatial metrics

can be classified according to the aspects of the landscape pattern that they quantify. In this study, we used four area/density/edge metrics (number of patches (NP), patch density (PD), largest patch index (LPI) and edge density (ED)), one shape metric (shape index distribution with area-weighted mean (SHAPE_AM)), one isolation/proximity metric (proximity index distribution with area-weighted mean (PROX_AM)), and one contagion/interspersion metric (effective mesh size (MESH)).

Another useful validation approach is to set up long-term runs and test if the fractal properties of the results are constant in time: e.g. by plotting the cluster size-frequency dimension graph. The logarithm of the number of patches of urban land (or specifically residential land) of a specific size is in almost any region a decreasing linear function of the logarithm of the size of the patches (see also section 2.1). The test is a necessary (but not a sufficient) condition to indicate that the model is not over-calibrated since a model that is successfully calibrated according to this criterion will continue to produce realistic urban forms, cluster distributions, and fractal properties when run far into the future.

6.3 Results

6.3.1 Greater Dublin Region

6.3.1.1 Semi-automated calibration

We used the proposed error function (equation (6.7)) with two different sets of weights to calibrate the application to the Greater Dublin Region (Table 6.1). We first tried to improve the regional error of the residential land use and the population as in the previous manual calibrations of White et al. (2012), though now both without and with network distance-based influence rules. Next, we experimented with a weighted combination of regional and

multiple resolution errors. The weight of residential land use was set higher than that of population in these runs since the first experiments resulted in good population patterns but weak urban land-use growth predictions. The higher resolution errors ($L > 2$) were omitted since the regional totals were already acceptable after the first experiments. With this second set of weights, we also tested if the new density equations (3.20) and (3.21), which were initially developed for the Flanders case, would have a positive effect on the results of the Dublin simulation runs. Tables 6.2 and 6.3 contain the settings of the different runs, their calibration error function weight type, and the different computed errors (of population in Table 6.2, of residential land use in Table 6.3). White et al. (2012) had already concluded that their results for population outperformed those of the MOLAND model and a constant share null model, but still we managed to lower the regional population error R_{pop} from 0.009318 (best run presented in White et al. (2012)) to 0.006633 (run D3, best run of this study).

Table 6.1 Definition of calibration error function weight types for the Greater Dublin Region. The regional error weight h and the multiple resolution error weights g_L are given per activity or land-use category.

Type	Activity K or LU i	h	g_1	g_2
1	$K = \text{Population}$	0.5	-	-
	$i = \text{Residential}$	0.5	-	-
2	$K = \text{Population}$	0.4	0.0005	0.05
	$i = \text{Residential}$	0.6	0.001	0.25

The northern and central neighbourhoods of the city of Dublin largely had a declining population in the 1990s while population grew in the

Table 6.2 Errors in population in 2000 of different calibration runs for the Greater Dublin Region with indication of the model version ('Netw': network-based influence rules or not), the number of the density equation ('Dens'), the error function weight type ('Type') according to Table 6.1, the regional error R , and the errors E_L at different resolutions. Underlined values are optimised together with the underlined values in Table 6.3 for residential land use.

Run	Netw	Dens	Type	R	E_1	E_2	E_3	E_4	E_5	E_6
D1	no	(3.19)	1	<u>0.008612</u>	1.8496	1.2164	0.7200	0.3528	0.0862	0.0102
D2	yes	(3.19)	1	<u>0.007703</u>	1.9124	1.3033	0.7707	0.3621	0.0870	0.0104
D3	no	(3.19)	2	<u>0.006633</u>	<u>1.9118</u>	<u>1.2355</u>	0.7322	0.3560	0.0867	0.0102
D4	yes	(3.19)	2	<u>0.009218</u>	<u>1.9942</u>	<u>1.2946</u>	0.7700	0.3735	0.0871	0.0101
D5	yes	(3.20)	2	<u>0.010348</u>	<u>1.9688</u>	<u>1.2701</u>	0.7638	0.3736	0.0862	0.0100
D6	yes	(3.21)	2	<u>0.014403</u>	<u>1.9169</u>	<u>1.2857</u>	0.7987	0.3769	0.0838	0.0096

Table 6.3 Errors in the number of residential cells in 2000 of different calibration runs for the Greater Dublin Region with indication of the model version ('Netw': network-based influence rules or not), the number of the density equation ('Dens'), the error function weight type ('Type') according to Table 6.1, the regional error R , and the errors E_L at different resolutions. Underlined values are optimised together with the underlined values in Table 6.2 for population.

Run	Netw	Dens	Type	R	E_1	E_2	E_3	E_4	E_5	E_6
D1	no	(3.19)	1	<u>0.008712</u>	1.3357	0.6572	0.2679	0.0846	0.0153	0.0024
D2	yes	(3.19)	1	<u>0.012329</u>	1.3655	0.6988	0.2915	0.0997	0.0224	0.0033
D3	no	(3.19)	2	<u>0.012623</u>	<u>1.3072</u>	<u>0.6172</u>	0.2306	0.0725	0.0151	0.0022
D4	yes	(3.19)	2	<u>0.010891</u>	<u>1.3084</u>	<u>0.6183</u>	0.2203	0.0614	0.0115	0.0021
D5	yes	(3.20)	2	<u>0.014208</u>	<u>1.3084</u>	<u>0.6133</u>	0.2171	0.0629	0.0128	0.0025
D6	yes	(3.21)	2	<u>0.013960</u>	<u>1.2941</u>	<u>0.6006</u>	0.2234	0.0674	0.0134	0.0026

southern part of the city (actual difference map: see Figure 6.4a). All simulations have difficulties to generate this spatial pattern: the population and residential expansion are overestimated in the north of Dublin and underestimated in the south (Figures 6.4b-c and 6.5). The best simulations without network distances (runs D1 and D3) generally have lower errors for population than those with network distances (runs D2 and D4), although run D2 has a lower regional error than run D1 (Table 6.2). If we compare both runs with low regional population errors (D3 and D2), then run D2 has a stronger overestimation of the population in central Dublin (population decline was much stronger in reality) and in some new residential cells in the north (Figures 6.4b). Run D3 is more balanced but has a slightly too small increase of population at the edge of the suburban towns (Figure 6.4c). Using density equation (3.20) (run D5) does not lead to big changes in Dublin in comparison with equation (3.19) (run D4). Equation (3.21) generates an even stronger central population growth in run D6 which is not in line with the observations in Dublin. On the other hand, the errors of population at large resolutions (E_5 and E_6) are slightly lower because population increase in smaller remote towns is better predicted with this equation (Table 6.2). The residential land-use development in the year 2000 is best predicted by run D4 (Table 6.3, Figure 6.5b), which has rather low error values at all resolutions and a low regional error too. When only the regional error is optimised (run D2), some suburban towns near Dublin (especially to the north) grow too fast (Figure 6.5a). The run with activity density equation (3.21) has the least errors close to Dublin and therefore slightly lower errors at the local resolutions (E_1 and E_2 , Table 6.3). On the other hand, there is an overestimation of scattered residential growth in rural areas (Figure 6.5c) leading to a larger regional error.

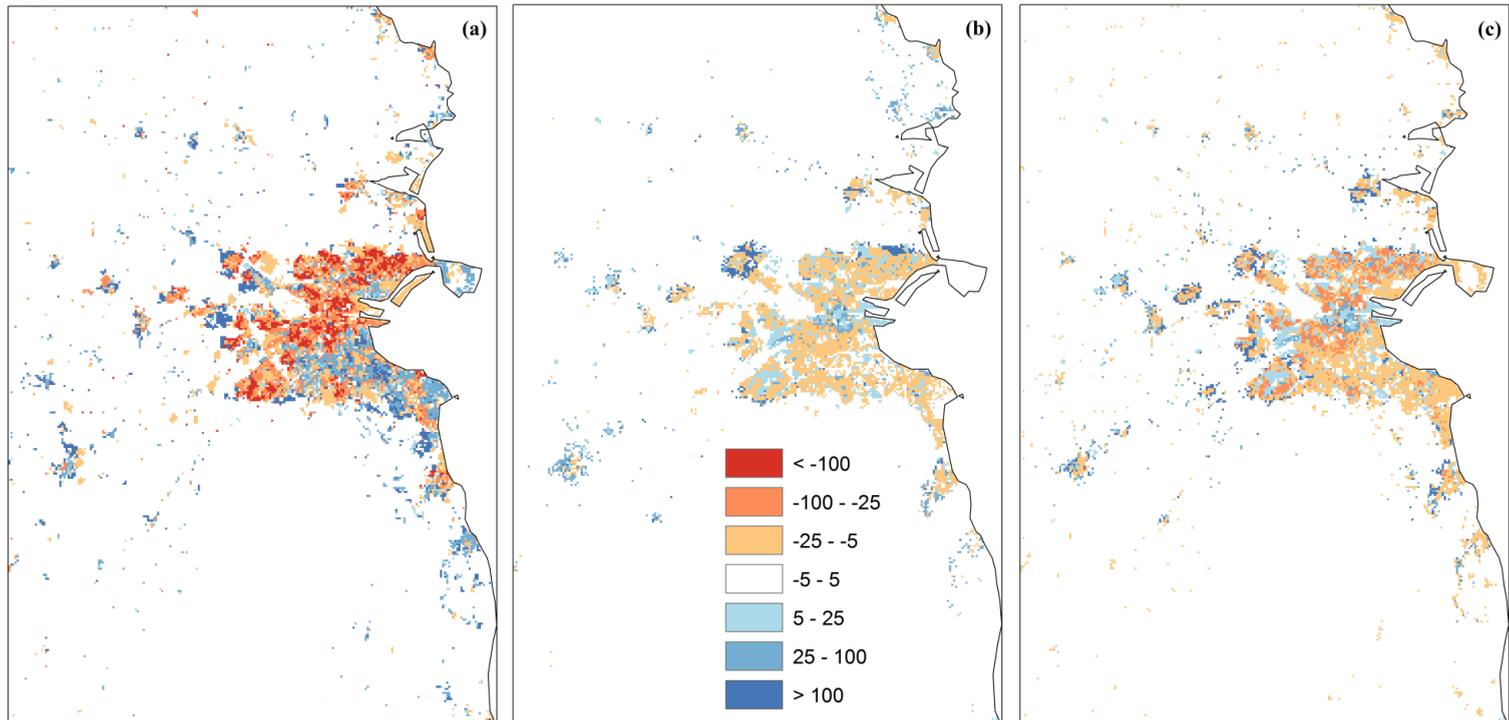


Figure 6.4 Evolution of population (per 200 m cell) between 1990 and 2000 around the city of Dublin in (a) the actual maps, (b) simulation D3 (no network distances, error weight type 2), and (c) simulation D2 (network distances, error weight type 1).

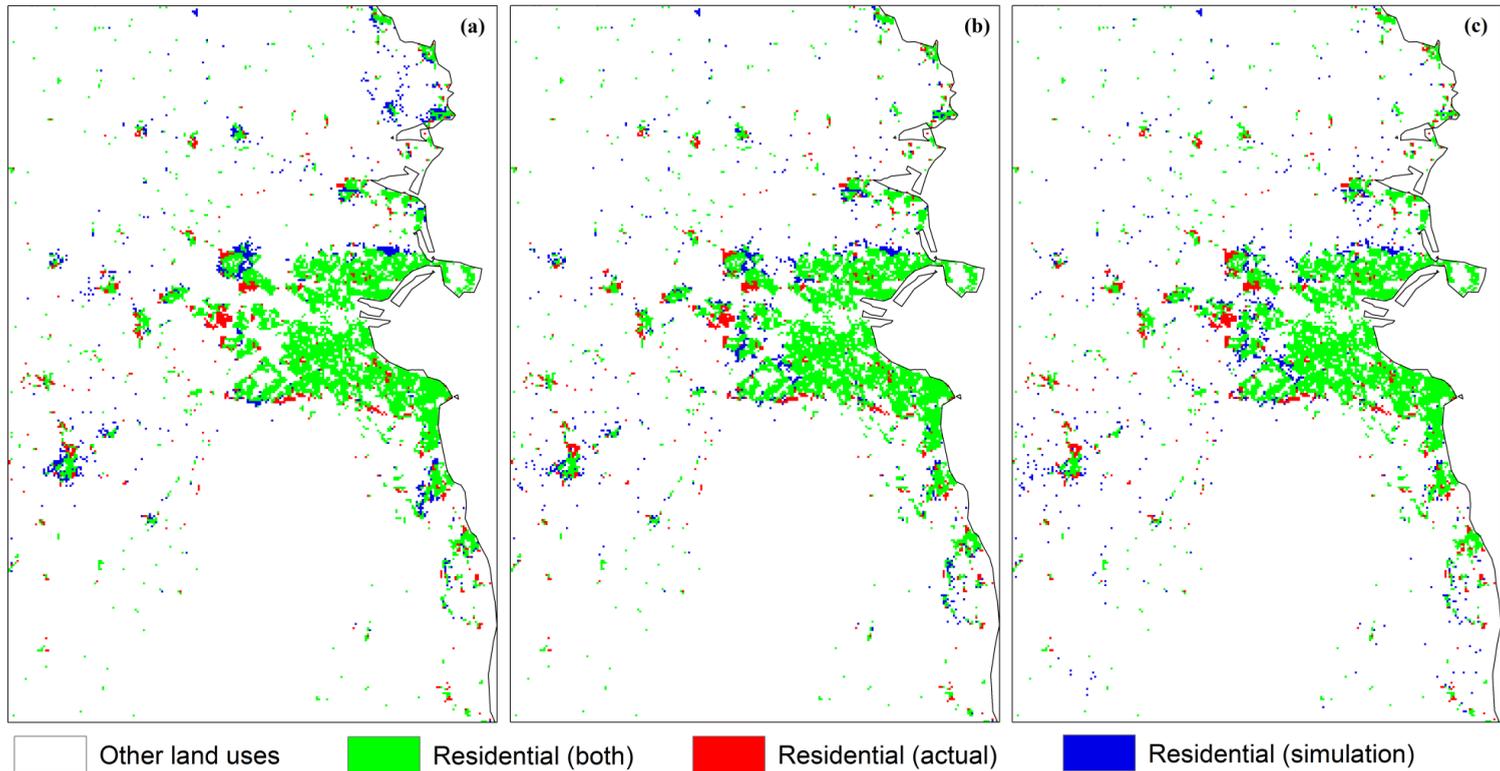


Figure 6.5 Difference in residential development around the city of Dublin in the year 2000 between the actual CORINE land-use map and three simulations with network distances: (a) simulation D2 (error weight type 1), (b) simulation D4 (error weight type 2), and (c) simulation D6 (error weight type 2 with alternative density equation (3.21)).

6.3.1.2 Validation of the residential land-use pattern

We computed the different spatial metrics for the actual maps of 1990, 2000 and 2006 and for simulation results of 2000 and 2006 (Table 6.4). Most simulation maps of 2000 have spatial metric values which evolve in a similar way as the actual metric evolution after 1990. The results for run D2 (with network distances and regional optimisation) were not the best for the calibration period: this is again clear in the values of LPI, PROX_AM and MESH for 2000 which are closer to those of the initial map of 1990 than to those of the final map of 2000. In 2006 this run has many spatial metric values close to the actual values. However, this is because the excess in residential cells to the north of Dublin expands, so the location of the urban growth is again not optimal. The runs without network distances (D1 and D3) create an exaggeratedly large central cluster around Dublin in 2006, as can be observed in their too small number of patches (NP) and too high largest patch index (LPI). Run D6 continues to produce scattered growth, leading to a very large number of patches and a large edge density (ED).

Runs over a longer period of time should still generate realistic map patterns. We did simulations until 2050 in which we assume that the strong urban growth of the calibration and validation period will continue. As such, unrealistically large residential areas are created, but these areas should still have the same fractal dimensions. The resulting patterns of runs without and with network distances are completely different (Figure 6.6). The simulations without network turn Dublin into a large megalopolis, while those with network create extreme linear development along the existing major roads and motorways extending from Dublin. Run D6 is similar to run D4 but again with a higher proportion of scatter. Note though that simulations with network distances could result in compact growth too for

other settings of the accessibility parameter X of equation (3.1), which was not included in the semi-automated calibration.

Table 6.4 Spatial metrics of the residential land-use pattern of the actual land-use maps ('Act') and different calibration runs for the Greater Dublin Region.

Year	Map	NP	PD	LPI	ED	SHAPE_ AM	PROX_ AM	MESH
1990	Act	989	0.0580	0.4455	1.2565	4.8493	97.4334	43.8745
2000	Act	1028	0.0603	0.5278	1.4122	5.0934	86.6512	58.7787
	D1	1014	0.0594	0.5440	1.3852	5.4694	77.5298	62.0843
	D2	1036	0.0607	0.4483	1.4068	4.7820	98.7016	47.4601
	D3	1065	0.0624	0.5522	1.4324	5.6299	84.9746	64.9556
	D4	1066	0.0625	0.5587	1.4204	5.6837	85.5851	65.6850
	D5	1081	0.0634	0.5557	1.4324	5.6299	84.9746	64.9556
	D6	1324	0.0776	0.5604	1.5494	5.7719	99.6421	65.3907
2006	Act	1211	0.0710	0.5548	1.6964	4.8154	81.9937	66.4065
	D1	1027	0.0602	0.6563	1.5613	5.9268	88.4964	88.4596
	D2	1262	0.0740	0.5111	1.7405	5.0707	92.9476	60.2707
	D3	1070	0.0627	0.6567	1.5991	5.9387	102.880	89.3277
	D4	1427	0.0836	0.5862	1.7913	5.3386	93.6385	73.8663
	D5	1531	0.0897	0.5838	1.8354	5.2965	92.7182	72.9127
	D6	1987	0.1165	0.5883	2.0701	5.4332	132.061	72.8468

We plotted the cluster size-frequency dimension graphs of residential land use of the actual 1990 map and three simulation results for 2050 (Figure 6.7). The logarithm of the number of residential clusters is clearly a decreasing linear function of the logarithm of the size of these clusters, when the smallest frequencies are omitted. In all results, the largest cluster gets much larger. The slope and intercept of the run without network

distances (D3) corresponds much better to the original map than those of the other runs. In run D4, the relationship is a little bit less linear than in the other graphs confirmed by a lower R^2 . As such, the extreme linear development of run D4 is not really rejected by this test, but nevertheless, it seems to be the weakest simulation for 2050. The large number of small residential plots in run D6 makes the slope of the linear relationship steeper, but R^2 is larger than for the other runs.

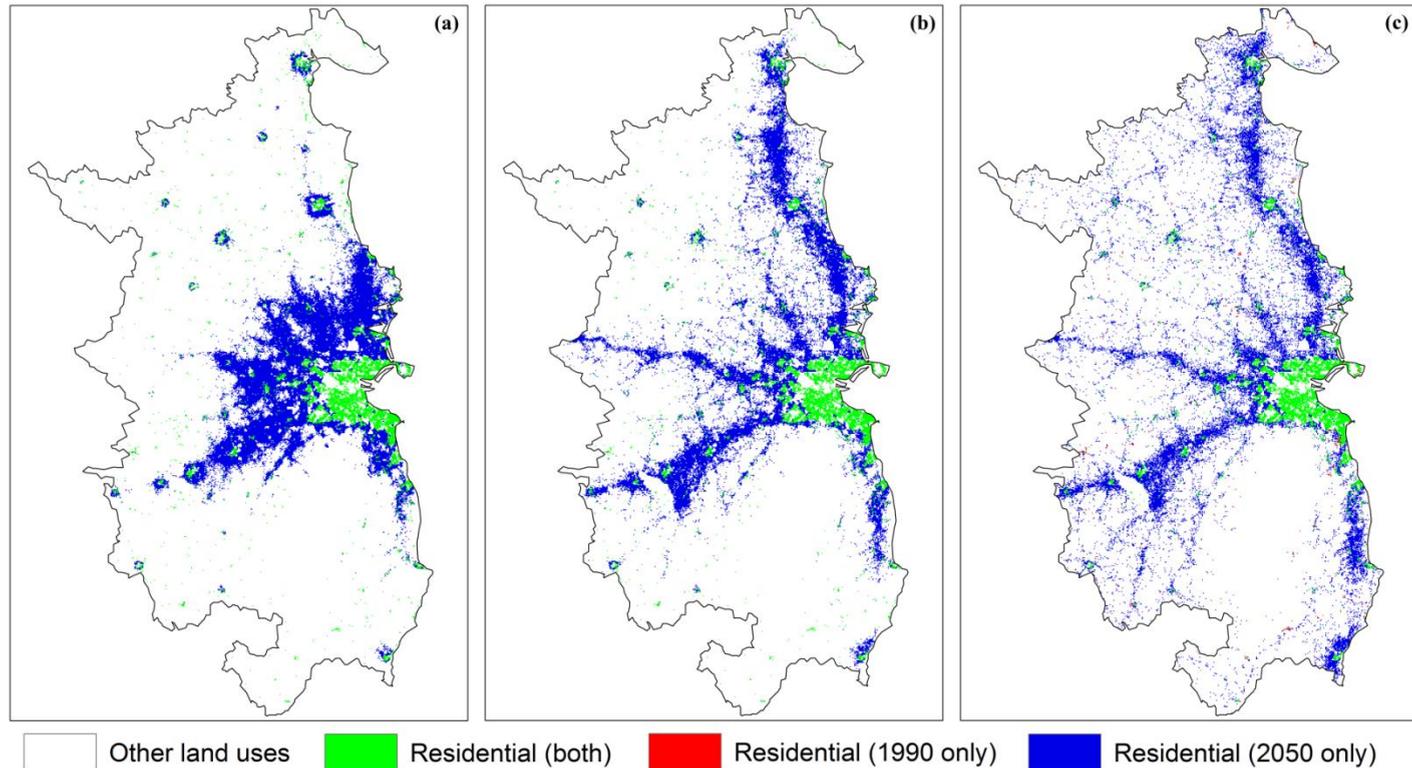


Figure 6.6 Difference in residential development in the Greater Dublin Region between 1990 and 2050 in three different simulations with error weight type 2: (a) run D3 (no network distances), (b) run D4 (network distances), and (c) run D6 (network distances with alternative density equation (3.21)).

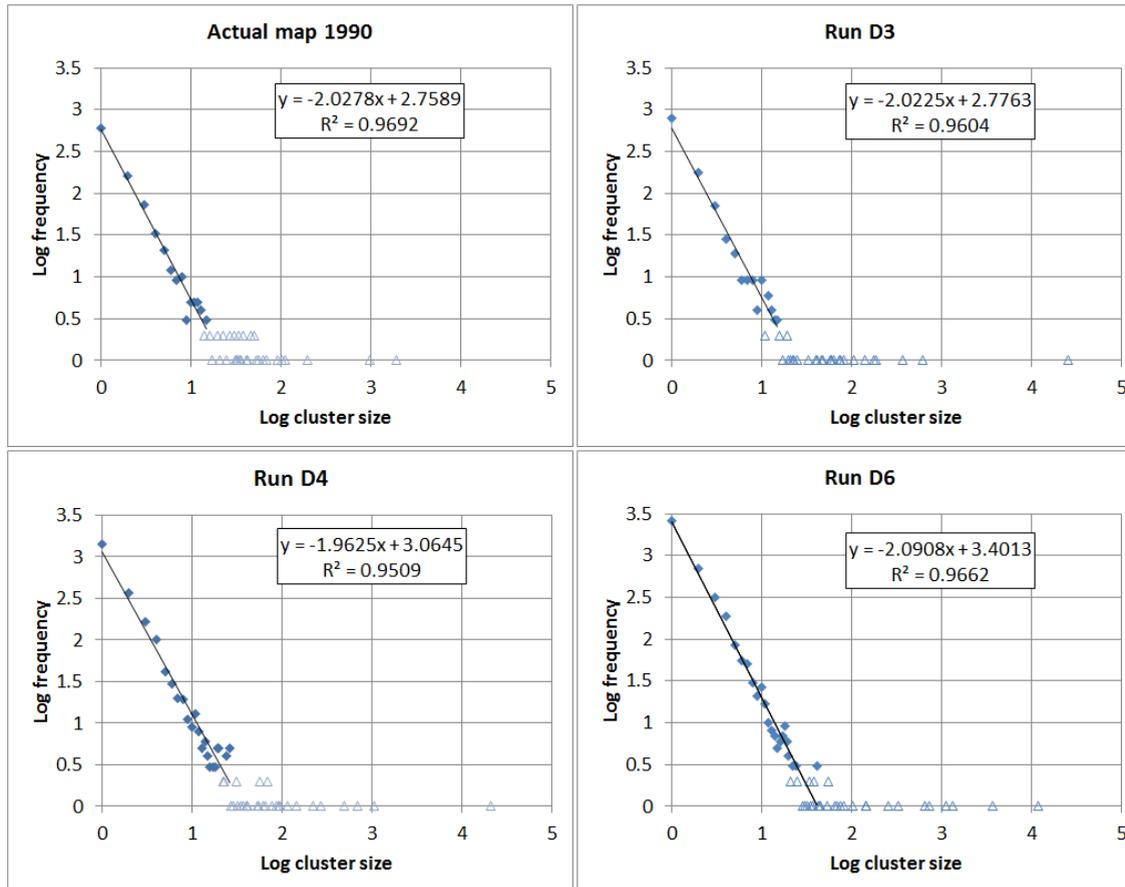


Figure 6.7 Cluster size-frequency graphs of residential land use of the GDR for the actual map (1990) and three simulations (runs D3, D4 and D6). The open triangles are cluster sizes with a low frequency not used to compute trend lines.

6.3.2 Flanders and the Brussels Capital Region

6.3.2.1 Semi-automated calibration

The residential land use and the population of Flanders and the BCR have rather different growth patterns in reality between 2001 and 2013 (Figures 6.8 and 6.9). The population mainly increases in larger cities while most new residential cells can be found in suburban and rural areas. Given this distinct nature of the urban growth and the large number of parameters and influence functions of the activity-based CA model, a manual calibration did not lead to satisfying results. This was the main reason to test the

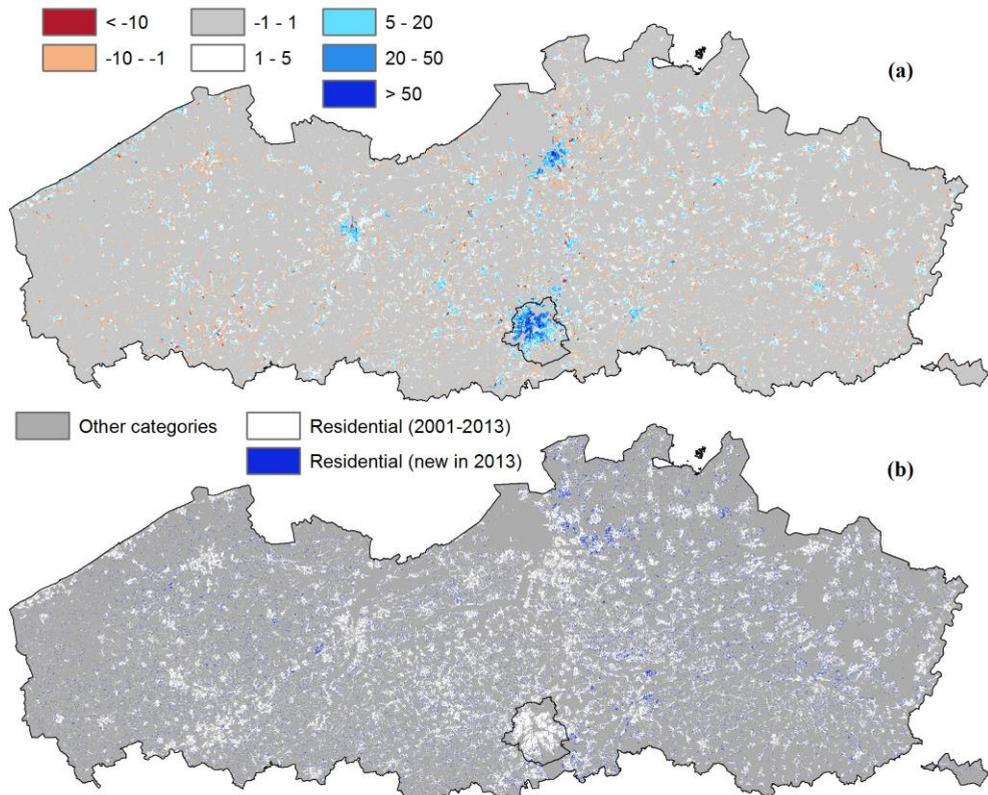


Figure 6.8 Actual urban development in Flanders and the Brussels Capital Region between 2001 and 2013: (a) evolution of the population per ha, and (b) development of new residential cells.

semi-automated calibration approach. The initial manual calibration also made clear that population of the largest cities, and especially Brussels, was systematically underestimated when population density equation (3.19) was used (4 to 5 % underestimation of the actual population in Brussels in 2013, i.e. +/- 30 % underestimation of the actual growth of the population between 2001 and 2013).

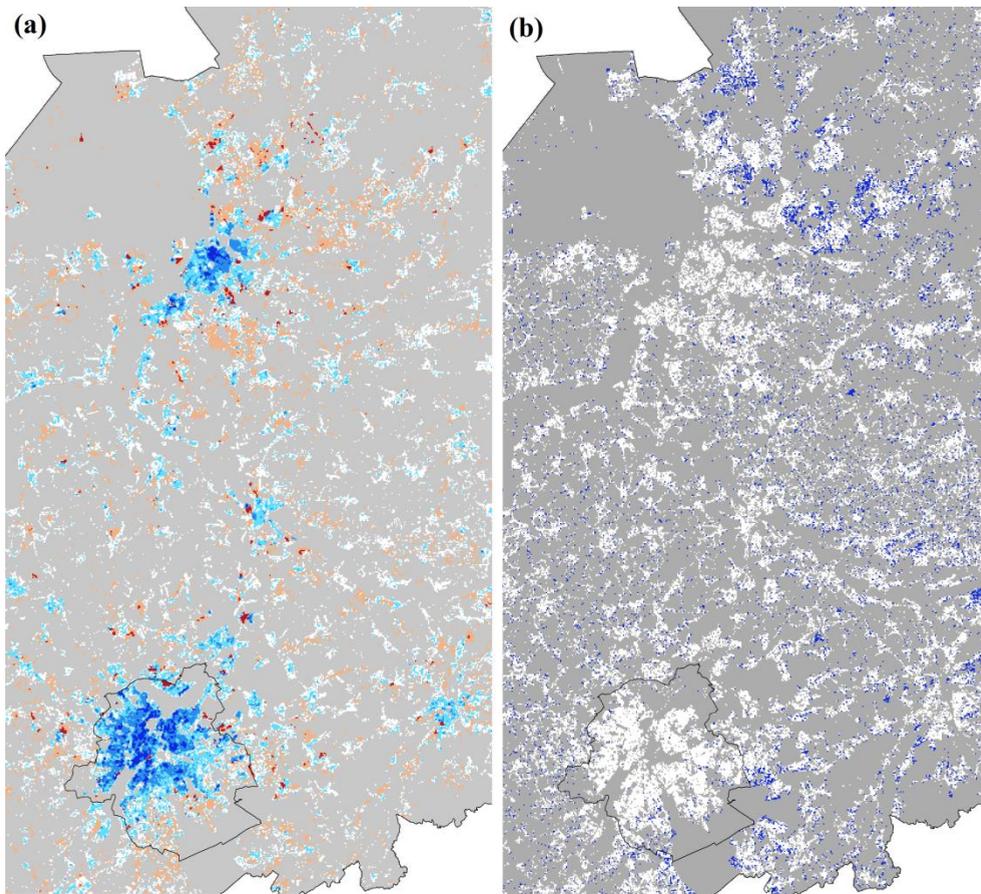


Figure 6.9 Actual urban development in the Antwerp-Brussels area between 2001 and 2013: (a) evolution of the population per ha, and (b) development of new residential cells. Legend as in Figure 6.8.

In the semi-automated approach, we tested a large number of calibration and model settings which were iteratively upgraded. Therefore, all results discussed in this section were obtained in several steps and cannot be as straightforwardly interpreted as the results obtained for the Dublin region. Apart from initial tests, 43 optimisation experiments were carried out, mostly with network-based distances but some without, with one of the alternative population density equations, with different error function weights, and with different subsets of influence weights and parameters to be optimised. In the discussion of the results, we focus on ten interesting simulation runs from a modeller's perspective, having all low errors. They were all identified as the best result of a particular optimisation experiment. Most of these have for all or some activities the lowest regional or multiple-resolution errors, but some results are also discussed because they seem good at first sight but are actually over-calibrated. Other experiments led to larger errors or were not significantly different from the results presented. The settings of the ten runs can be found in Table 6.5 and their errors in population and in the number of residential land-use cells can be found in Tables 6.6 and 6.7. A more comprehensive set of tables with calibration error weights of the ten runs, chosen parameters to optimise and the errors in all activities and active land uses of the according results can be found in Appendix 6.A1.

Generally, all simulations result in good regional population predictions, but they also have the same shortcomings at a more local scale. Population increase is underestimated in urban neighbourhoods that already have the highest densities (e.g. the west of Brussels and the east of the Antwerp city centre) and is overestimated in lower-density neighbourhoods of urban cores (e.g. the southeast of Brussels). The pattern of the residential

Table 6.5 Overview of calibration runs of the Belgian application with indication of the model version ('Netw': network-based influence rules or not), the number of the density equation ('Dens'), the error function weight type ('Type') according to Tables 6.A1 and 6.A2, and the influence weights ($W_{i,j}$ is the influence of category i on category j) and parameters (p_K is a parameter for activity K) which were optimised by the calibration algorithm. Abbreviations are used for population (pop) and employment (ret = retail, ser = services).

Run	Netw	Dens	Type	Optimised influence weights	Optimised parameters
F1	no	(3.20)	1		ρ (activity-specific
F2	yes	(3.20)	1	$W_{\text{pop,pop}} ; W_{\text{ret,pop}} ;$	ρ_K in F5), $\varepsilon, \lambda_{\text{pop}},$
F3	yes	(3.21)	1	$W_{\text{ser,pop}} ; W_{\text{pop,ret}} ;$	$m_{\text{pop}}, \gamma_{\text{pop}}, \zeta_{\text{pop}}$ or $\tau_{\text{pop}},$
F4	yes	(3.20)	2	$W_{\text{ret,ret}} ;$	$\lambda_{\text{ret}}, m_{\text{ret}}, \gamma_{\text{ret}}, \lambda_{\text{ser}},$
F5	yes	(3.21)	3	$W_{\text{pop,ser}} ; W_{\text{ser,ser}}$	$m_{\text{ser}}, \gamma_{\text{ser}}$
F6	yes	(3.20)	1	all non-zero	ρ (activity-specific
F7	yes	(3.21)	4	influence rules (18 functions) ¹	ρ_K in F7), $\varepsilon, \text{all } \lambda, \text{all } m, \text{all } \gamma, \zeta_{\text{pop}}$ or τ_{pop}
F8	yes	(3.20)	5	$W_{\text{pop,pop}} ; W_{\text{ret,pop}} ; W_{\text{ser,pop}} ; W_{\text{pop,ret}} ; W_{\text{pop,ser}}$	$\rho, \varepsilon, \lambda_{\text{pop}}, m_{\text{pop}}, \gamma_{\text{pop}}, \zeta_{\text{pop}}$
F9	yes	(3.20)	6	all $W_{K,K}$	$\rho, \varepsilon, \text{all } \lambda, \text{all } m, \text{all } \gamma, \zeta_{\text{pop}}$
F10	yes	(3.20)	7	all $W_{K,K} ; \text{all } W_{\text{pop},K} ; W_{\text{ret,pop}} ; W_{\text{ser,pop}}$	$\rho, \varepsilon, \text{all } \lambda, \text{all } m$

land-use growth is too clustered in rural areas where most new cells are popping up adjacent to other urban cells. The proportion of new cells in the actual 2013 map not being adjacent to existing urban land is considerable, although it should be noted that scattered residential land in the West Flanders province might be slightly underestimated in the 2001 map as discussed in section 5.A1 due to errors in the downdating technique used to

¹ Activity influences: (1) of population, industry, retail, services, nature and recreation and parks on population; (2) of population, industry and water on industry; (3) of population, water, industry and wholesale and logistics on wholesale and logistics; (4) of population, retail and services on retail; and (5) of population and services on services.

Table 6.6 Errors in population in 2013 of different calibration runs for Flanders and the BCR with indication of the model version ((‘Netw’: network-based influence rules or not), the number of the density equation (‘Dens’), the regional error R and the errors E_L at different resolutions. Underlined values are optimised together with the underlined values in Table 6.7 for residential land use, and in some cases with an optimisation of employment and economic land uses (see Tables 6.A3 and 6A4).

Run	R	E_1	E_2	E_3	E_4	E_5	E_6
F1	0.0158	<u>0.4294</u>	<u>0.2420</u>	<u>0.1265</u>	<u>0.0471</u>	<u>0.0131</u>	<u>0.0040</u>
F2	0.0133	<u>0.4431</u>	<u>0.2495</u>	<u>0.1273</u>	<u>0.0443</u>	<u>0.0114</u>	<u>0.0030</u>
F3	0.0344	<u>0.4725</u>	<u>0.2768</u>	<u>0.1634</u>	<u>0.0813</u>	<u>0.0260</u>	<u>0.0061</u>
F4	0.0141	<u>0.4208</u>	<u>0.2329</u>	<u>0.1172</u>	<u>0.0405</u>	<u>0.0107</u>	<u>0.0026</u>
F5	0.0278	<u>0.3665</u>	<u>0.2117</u>	<u>0.1132</u>	<u>0.0518</u>	<u>0.0187</u>	<u>0.0056</u>
F6	0.0144	<u>0.4275</u>	<u>0.2363</u>	<u>0.1187</u>	<u>0.0412</u>	<u>0.0121</u>	<u>0.0039</u>
F7	0.0225	<u>0.4044</u>	<u>0.2304</u>	<u>0.1216</u>	<u>0.0500</u>	<u>0.0162</u>	<u>0.0030</u>
F8	<u>0.0116</u>	<u>0.4585</u>	<u>0.2585</u>	<u>0.1271</u>	0.0427	0.0121	0.0034
F9	<u>0.0139</u>	<u>0.4291</u>	<u>0.2439</u>	<u>0.1201</u>	0.0421	0.0126	0.0039
F10	<u>0.0168</u>	0.4546	0.2545	0.1239	0.0440	0.0146	0.0051

make this map. Next, residential land-use growth is underestimated in a few suburban low-density residential areas with large gardens (e.g. northeast of Antwerp).

Except for the issues discussed above, run F2, an optimisation result with network distances and activity density equation (3.20), has low population errors in most cells (Figure 6.10a) and no significant over- or underestimation of residential land-use in any part of the study area (Figure 6.10b). The run also has a good balance between errors in different activity (Table 6.8) and land-use (Table 6.9) categories at the level of individual districts. Figure 6.11 is a map of the districts. The relative error in population of all districts is below 5% and mostly below 2%. This error is

Table 6.7 Errors in the number of residential cells in 2013 of different calibration runs for Flanders and the BCR with indication of the model version (‘Netw’: network-based influence rules or not), the number of the density equation (‘Dens’), the regional error R and the errors E_L at different resolutions. Underlined values are optimised together with the underlined values in Table 6.6 for population, and in some cases with an optimisation of employment and economic land uses (see Tables 6.A3 and 6A4).

Run	R	E_1	E_2	E_3	E_4	E_5	E_6
F1	0.0371	<u>0.4600</u>	<u>0.2031</u>	<u>0.0980</u>	<u>0.0520</u>	<u>0.0269</u>	<u>0.0077</u>
F2	0.0405	<u>0.4633</u>	<u>0.2060</u>	<u>0.1016</u>	<u>0.0555</u>	<u>0.0296</u>	<u>0.0086</u>
F3	0.0422	<u>0.4654</u>	<u>0.2091</u>	<u>0.1046</u>	<u>0.0573</u>	<u>0.0298</u>	<u>0.0089</u>
F4	0.0462	<u>0.4729</u>	<u>0.2197</u>	<u>0.1095</u>	<u>0.0596</u>	<u>0.0315</u>	<u>0.0091</u>
F5	0.2108	0.6857	0.4279	0.2878	0.2076	0.1401	0.0581
F6	0.0414	<u>0.4638</u>	<u>0.2072</u>	<u>0.1023</u>	<u>0.0561</u>	<u>0.0305</u>	<u>0.0092</u>
F7	0.0538	<u>0.4505</u>	<u>0.2036</u>	<u>0.1100</u>	<u>0.0686</u>	<u>0.0406</u>	<u>0.0158</u>
F8	<u>0.0338</u>	<u>0.4681</u>	<u>0.2088</u>	<u>0.1037</u>	0.0552	0.0253	0.0067
F9	<u>0.0378</u>	<u>0.4695</u>	<u>0.2037</u>	<u>0.0930</u>	0.0481	0.0257	0.0077
F10	<u>0.0331</u>	0.4713	0.2102	0.0947	0.0451	0.0225	0.0063

not much higher in the other runs and rural districts with little population have the largest relative error. The error in the number of residential land-use cells ranges up to about 7%. The different employment and economic land use categories generally have larger errors than population and the residential land use in all runs. Most districts have relative errors below 10% in run F2 for economic activities and land uses, but some outliers are present in regions with little or declining employment in a particular sector. Other runs have often more relative errors of the economic sectors above 10%, especially when all or some sectors did not get any weight in the calibration.

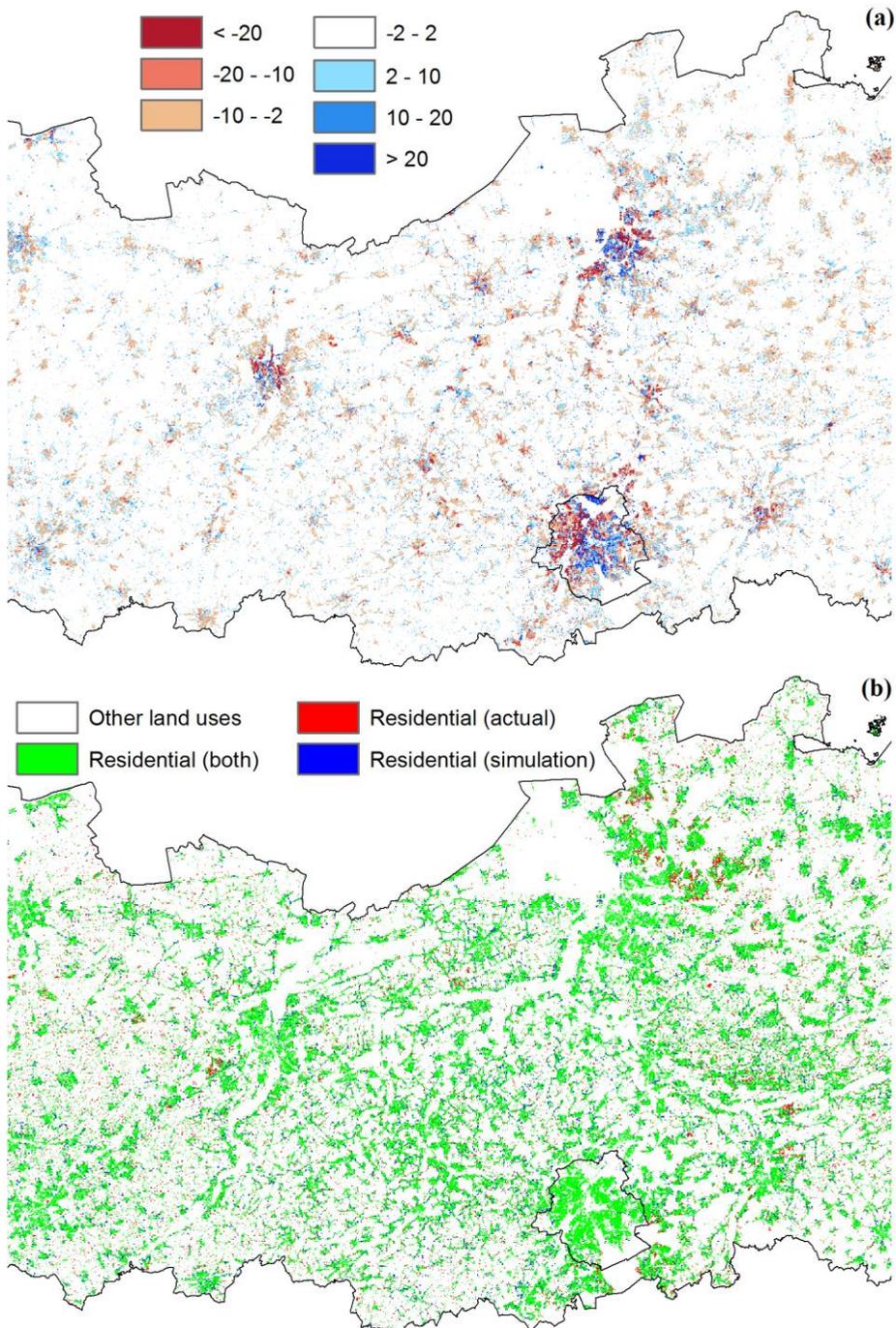


Figure 6.10 Difference in (a) population and (b) residential development in central Flanders and the Brussels Capital Region in the year 2013 between the actual maps and simulation F2.

Table 6.8 Relative error (%) in activity K per district in 2013 of calibration run F2 for Flanders and the BCR. Abbreviations are used for population (pop) and economic sectors (ind = industry, whl = wholesale and logistics, ret = retail, ser = services).

District	$K = \text{pop}$	$K = \text{ind}$	$K = \text{whl}$	$K = \text{ret}$	$K = \text{ser}$
Antwerpen	0.48	7.30	-2.32	0.84	-1.71
Eeklo	2.66	-19.12	27.14	8.32	7.20
Oostende	-2.18	-13.88	5.56	-2.59	7.95
Brugge	2.26	-5.48	12.87	-3.46	1.19
Maaseik	-2.74	0.75	5.67	12.98	4.83
Veurne	-1.15	-1.15	113.45	-8.82	10.94
Mechelen	-1.09	-6.85	-8.75	-0.13	-5.67
Sint-Niklaas	-1.75	-12.09	-15.47	6.68	2.29
Dendermonde	0.90	-5.86	34.88	9.91	3.35
Gent	-0.61	-5.99	1.75	-0.07	-6.55
Diksmuide	0.92	-9.56	3.13	3.24	10.18
Hasselt	-0.89	10.97	-10.74	-2.45	-0.51
Tielt	4.16	-3.87	14.93	7.62	0.70
Aalst	-1.17	7.24	-21.31	9.23	5.50
Roeselare	0.36	-3.95	3.02	2.39	-10.48
Leuven	0.62	8.70	-1.45	7.59	-0.21
Tongeren	-1.21	-4.97	-3.20	1.23	9.40
Kortrijk	2.81	5.11	-8.28	5.59	0.85
Oudenaarde	1.98	15.98	57.48	3.07	3.47
Ieper	3.98	-14.08	7.86	0.48	7.05
Turnhout	-2.52	-6.53	1.51	4.76	-2.78
Halle-Vilvoorde	-0.21	3.33	-4.28	3.45	5.98
Brussels Capital Region	0.18	3.35	5.07	-12.45	-0.09

Table 6.9 Relative error (%) in the number of cells of active land-use category i per district in 2013 of calibration run F2 for Flanders and the BCR. Abbreviations are used for residential land use (res) and economic sectors (ind = industry, whl = wholesale and logistics, ret = retail, ser = services).

District	$i = \text{res}$	$i = \text{ind}$	$i = \text{whl}$	$i = \text{ret}$	$i = \text{ser}$
Antwerpen	-4.64	2.53	0.85	5.92	-0.39
Eeklo	-2.42	-4.00	7.29	1.47	6.16
Oostende	-2.37	-3.67	-2.92	-15.89	-0.35
Brugge	-4.77	-0.59	-1.06	7.45	1.56
Maaseik	0.97	-0.88	-7.22	-19.59	-4.49
Veurne	-2.78	2.19	2.08	4.44	1.73
Mechelen	1.25	-0.65	1.90	0.35	-2.34
Sint-Niklaas	4.47	-2.14	-0.41	1.30	-8.35
Dendermonde	3.27	-1.32	11.46	3.38	4.75
Gent	0.70	-0.26	1.26	8.70	0.47
Diksmuide	-7.41	-11.30	-8.94	-1.72	-2.70
Hasselt	1.82	2.64	-5.54	-1.48	0.77
Tielt	-6.86	-2.52	-5.54	0.00	-3.77
Aalst	4.83	2.34	1.52	8.54	2.16
Roeselare	2.00	0.48	-7.51	3.05	0.78
Leuven	-0.94	0.74	-1.14	-7.91	5.43
Tongeren	4.51	-3.52	-6.13	-5.33	-4.73
Kortrijk	5.67	1.24	-3.12	7.67	5.24
Oudenaarde	-1.36	3.27	-3.39	5.88	-0.82
Ieper	-5.58	-5.53	-0.43	-2.26	-25.93
Turnhout	-4.09	0.24	0.42	-6.34	-6.38
Halle-Vilvoorde	4.03	2.84	7.87	7.73	4.54
Brussels Capital Region	2.07	-0.70	19.05	-1.24	2.92



Figure 6.11 The Flemish districts and the Brussels Capital Region (“Brussel”).

The difference in population and residential land use between run F2 and run F1, which is the result of an optimisation with the same settings but without network distances, is actually limited. The main difference is a slightly different shape of the densification area of the large cities, the extent in which equation (3.20) assures a strong increase of the population density, determined by the value of the threshold ζ_{pop} . This area is less circular and somewhat following the roads in run F2. The regional error of population and the errors of larger scales (E_4 to E_6) are lower for run F2 than for run F1, but the errors of local scales (E_1 to E_3) are somewhat higher (Table 6.6). Using network distances seems to improve growth patterns at the larger scale and make them more complex like in reality, but finding the exact right location of growth might be more difficult.

The two new population density equations both improve the simulation of high-density neighbourhoods and have a clear positive impact on the regional population error, they have important disadvantages too. Equation (3.20) allocates extra activity to cells with a population potential higher than a given threshold. In most results the threshold is only exceeded in the cores of the largest cities. Hence, too many other town cores have a

decreasing and as such underestimated population because of a strong rescaling effect to get the correct population totals (e.g. run F4, see Figure 6.12a). Nevertheless, the errors in population of run F4 at the larger scales (E_4 to E_6) are the lowest of all runs. Alternatively, equation (3.21) directly uses a power function of the population potential to amplify population increase. While the best calibration results for population in city cores in general (including smaller towns) can be obtained with this equation, the power function is difficult to calibrate and easily leads to unstable behaviour and exaggerated residential growth in some regions. A good example is the over-calibrated run F5: the simulated pattern of population change in the cities (Figure 6.12b) highly matches the actual pattern (Figure 6.8a) and as such, errors in population at local resolutions (E_1 , E_2 and E_3) are the lowest of all runs (Table 6.6). Unfortunately, too many new residential cells appear in the rural parts of West Flanders province where the population actually decreases. The residential land use errors of this run are extremely high (Table 6.7). Other runs with equation (3.21) have other issues with exaggerated urban growth or population increase: run F7 results in several suburban neighbourhoods between Brussels and Antwerp becoming too large with respect to their actual size (Figure 6.13a) and population densities. Scattered residential growth in rural regions is in almost all runs too limited (Figure 6.13b): even if the regional errors of residential land use are low as in run F10, the new residential cells are closer to the existing major towns than in reality.

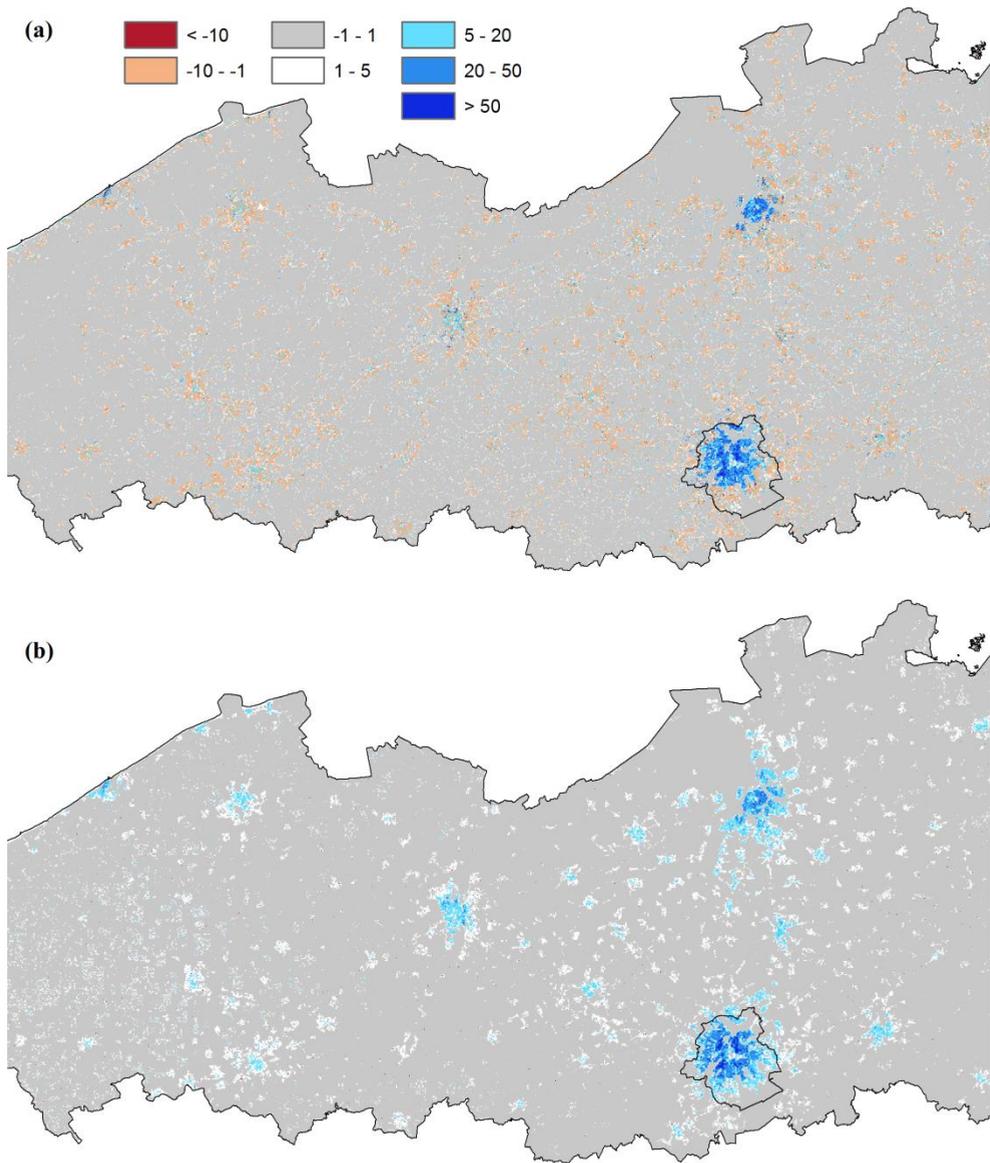


Figure 6.12 Evolution of population per ha between 2001 and 2013 in western and central Flanders and the Brussels Capital Region in (a) simulation F4 (with density equation (3.20)), and (b) simulation F5 (with density equation (3.21)).

Retail employment in Brussels is underestimated in most runs in comparison with the strong actual increase (+ 29 %). Employment in industry and in wholesale and logistics both have large regional and especially local errors (Table 6.A3). The local errors are due to plant shutdowns and decreases in logistics employment near Brussels Airport after 2001. Such local economic changes are beyond the scope of the model and over-calibration of this phenomena should not be encouraged, especially not if this would increase errors in other categories. Relative errors in the number of cells are the largest in the retail sector (Table 6.A4), but this is the active land use with the smallest amount of cells and only a very limited, local growth. Therefore, the exact location of the growth is particularly difficult to simulate. Employment in retail and services generally has more acceptable errors (Table 6.A3). The still relatively high local errors are caused by larger absolute errors in the Brussels Capital Region which can partially be explained by the poorer quality of the input data in this part of the study area: employment density is high in Brussels but data about the location of most companies was not available. Therefore, all urban cells get rather high background values of employment in the 2001 input maps (actual employment in services: Figure 6.14a). Simulated changes in service economy employment in Brussels are in all runs rather little per cell (e.g. run F3: Figure 6.14c) in comparison with the observed changes (Figure 6.14b). Other local errors can be attributed to the large value of the compatibility coefficient in the model of service employment in residential land use which leads to an overestimation of employment in residential areas.

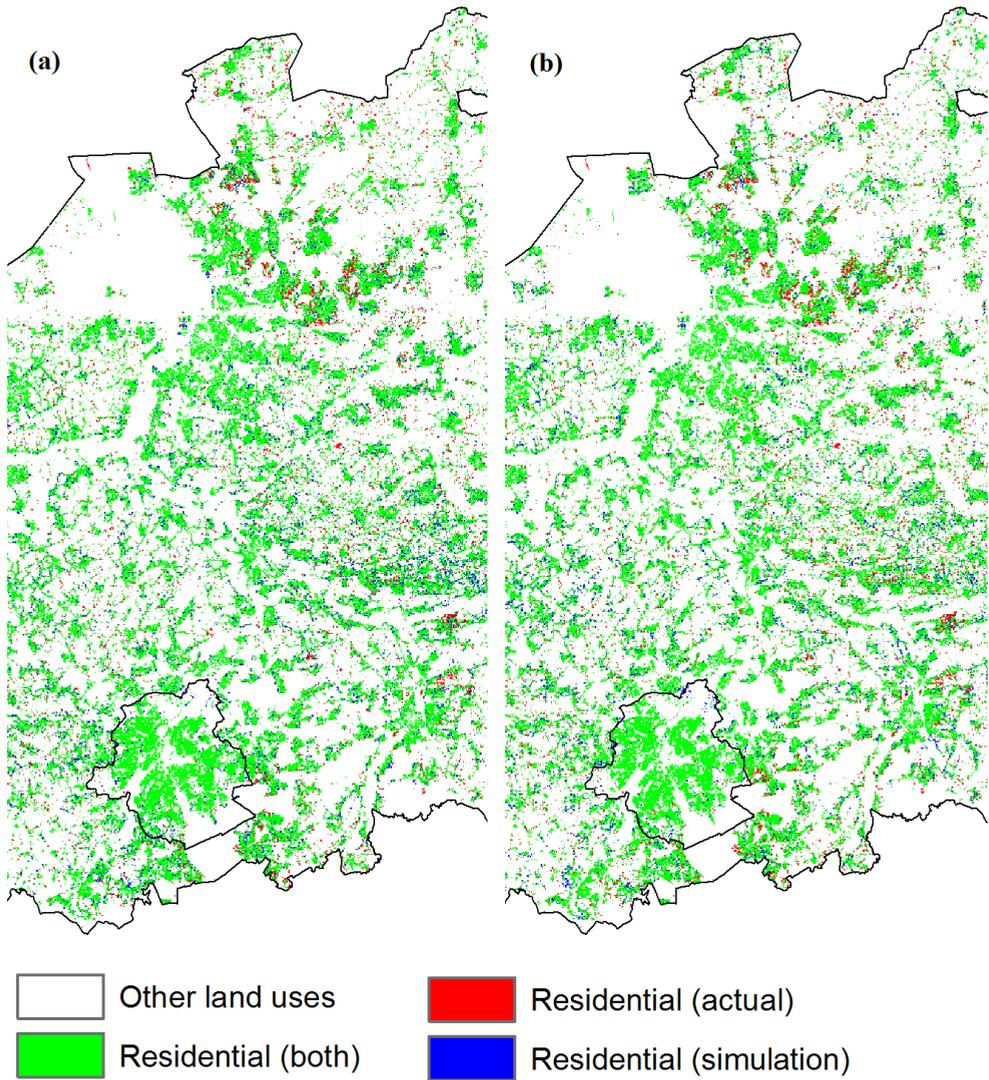


Figure 6.13 Difference in residential development in the Antwerp-Brussels area in the year 2013 between the actual land-use map and (a) simulation F7 (multiple resolution optimisation, density equation (3.21)), and (b) simulation F10 (regional optimisation, density equation (3.20)).

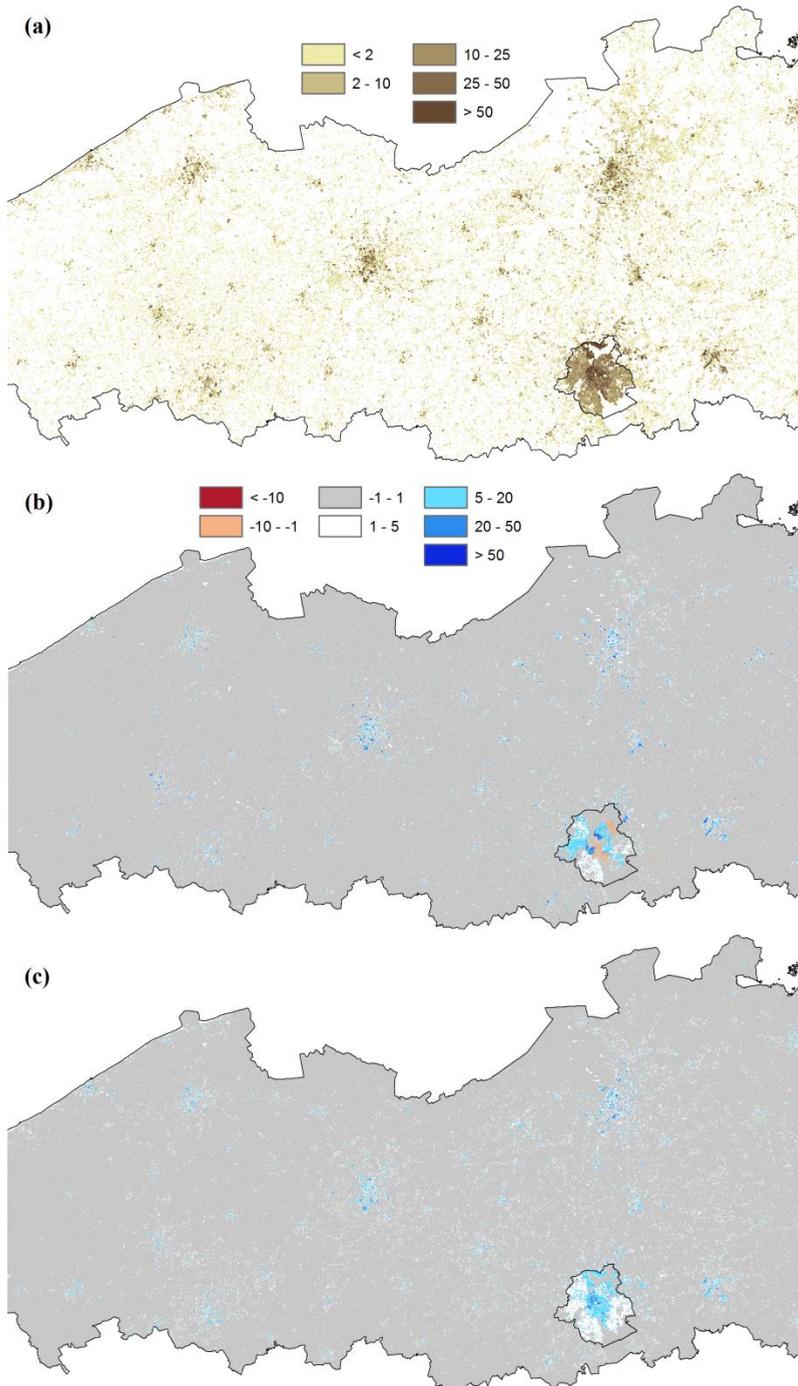


Figure 6.14 Employment in the services sector in western and central Flanders and the Brussels Capital Region: (a) actual map in 2001; (b) actual change in employment between 2001 and 2013; (c) simulated change in run F3.

6.3.2.2 Validation of the residential land-use pattern

The spatial metrics of all simulations indicate a poor match between the actual and simulated patterns of residential land use in 2013 (Table 6.10). The area/density/edge metrics of the actual and simulated maps even evolve in the opposite direction between 2001 and 2013. Residential patches become too large or merge overly easily, leading to a low number of patches and an extremely low patch density. The shape and proximity metrics have acceptable values.

Like for the Dublin region, we performed long-term simulations by simply extending business-as-usual growth until 2060. The predicted amount of residential urban fabric is again exaggerated but it should have the same fractal dimension as in the original map. Most simulations for 2060 result in a uniform growth of urban sprawl across Flanders (e.g. run F6, see Figure 6.15a). In some but not all simulations with density equation (3.21) some suburban neighbourhoods evolve into new large contiguous urban areas (e.g. run F7, see Figure 6.15b). The pattern of the population change towards 2060 is difficult to validate but it shows the same behaviour in most simulations: the population often decreases in all smaller town centres, while high growth rates are observed in larger cities, and, population is allocated to the large amount of new residential cells.

Table 6.10 Spatial metrics of the residential land-use pattern of the actual land-use maps ('Act01' for 2001 and 'Act13' for 2013) and different calibration runs for the Belgian application.

Map	NP	PD	LPI	ED	SHAPE	PROX	MESH
					_AM	_AM	
Act01	26367	1.9141	0.2626	29.3509	4.5845	75.4347	44.3390
Act13	28298	2.0543	0.3191	33.6032	5.1084	104.0999	69.9386
F1	24099	1.1053	0.2042	18.8694	5.2029	112.6321	53.1286
F2	23750	1.0892	0.2059	18.7320	5.3251	127.2193	58.2310
F3	23473	1.0765	0.1912	18.6818	5.3330	118.7476	55.7984
F4	23298	1.0685	0.1681	18.5748	5.1628	112.6916	50.5588
F5	25450	1.1672	0.1659	20.7179	4.4460	72.4822	32.5665
F6	23652	1.0848	0.2061	18.6650	5.2642	119.9040	55.3176
F7	23982	1.0999	0.2295	18.8046	5.5092	167.9303	66.4630
F8	24315	1.1152	0.1962	18.9388	5.1964	110.4928	52.6811
F9	24124	1.1064	0.1923	19.0733	5.0052	106.5843	46.7523
F10	24475	1.1225	0.2168	19.3620	5.1729	112.1334	50.9154

The cluster size-frequency graphs of Belgium are typically not entirely linear because widespread ribbon development results in fewer small clusters. This can be seen in Figure 6.16 as a curve in the upper-left corner of the graph, both for the actual map of 2001 and three simulations. The apparently bad simulation F7 has indeed a graph with a slope and an intersect that are most different to the actual initial distribution. The other results have slightly less steep slopes but the distribution of the clusters is in all runs very comparable to the original map. As mentioned before, good results for the cluster size-frequency analysis are necessary but do not prove

that calibrations that are successful by this criterion are satisfactory in other respects.

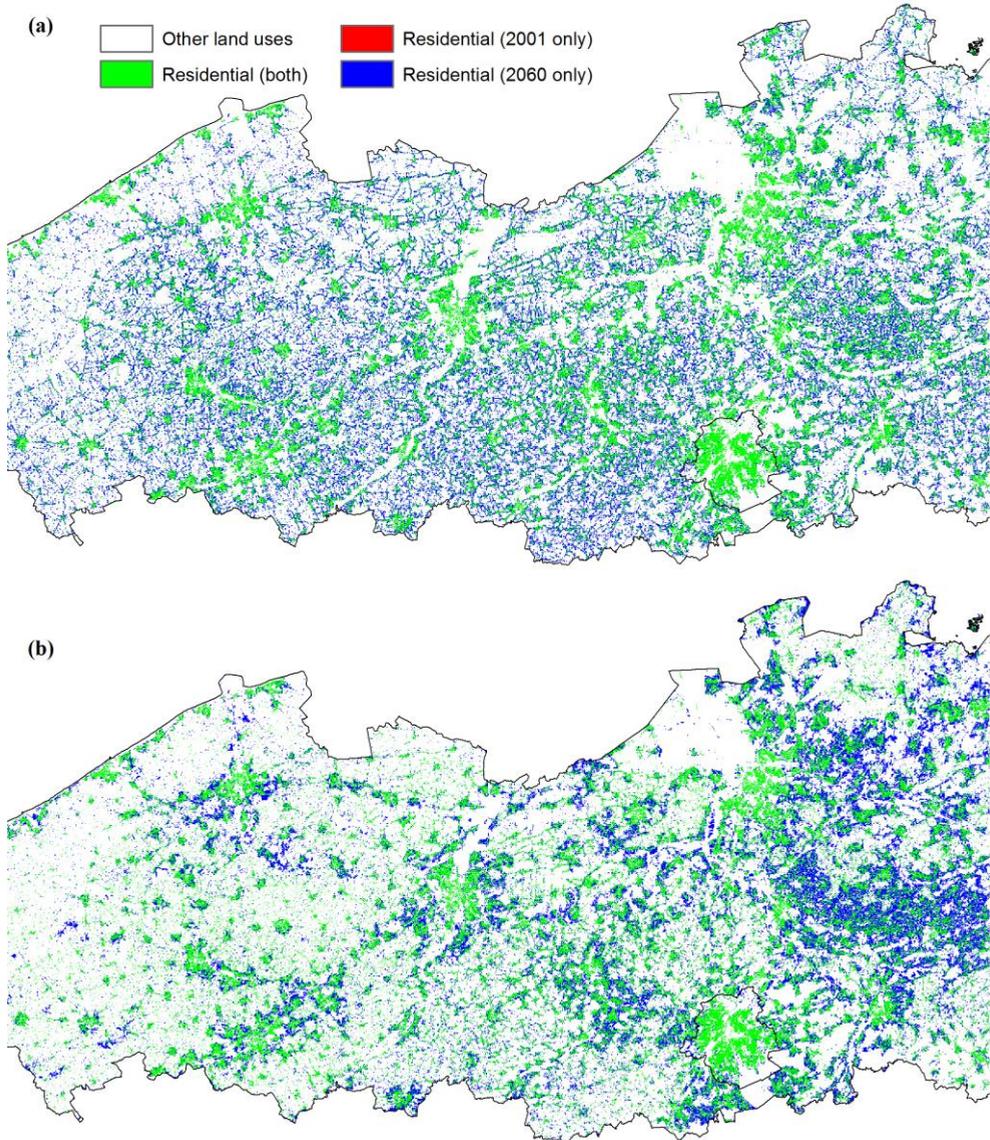


Figure 6.15 Difference in residential development in western and central Flanders and the Brussels Capital Region between 2001 and 2060 in (a) simulation F6 and (b) simulation F7.

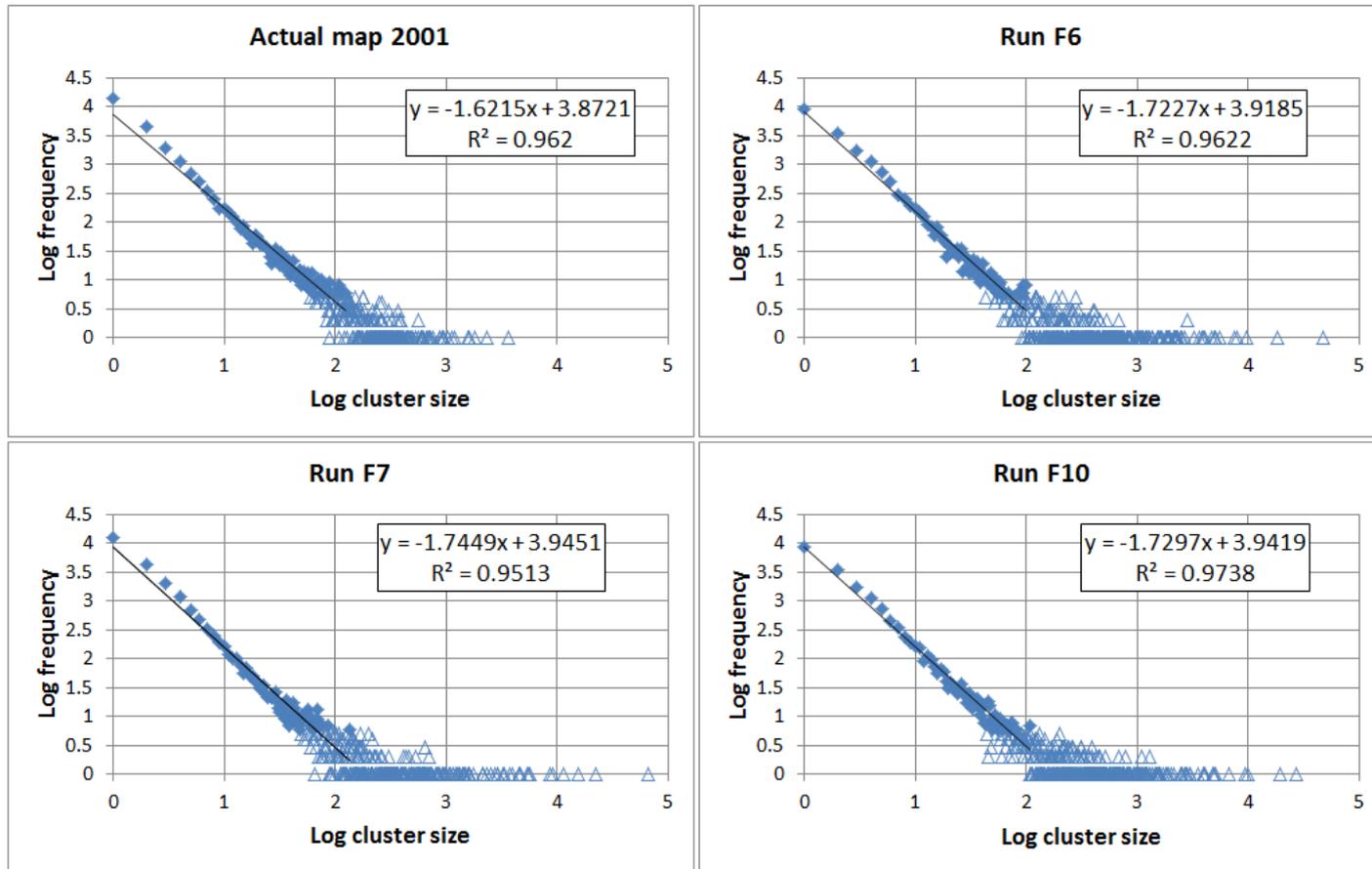


Figure 6.16 Cluster size-frequency graphs of residential land use of Flanders and the BCR for the actual map (2001) and three simulations (runs F6, F7 and F10). The open triangles are cluster sizes with a low frequency not used to compute trend lines.

6.4 Discussion

The activity-based cellular automata model clearly has a number of advantages in comparison with more traditional land-use change models. It directly predicts the pattern of several activities and is as such able to provide the user with multifunctional urban growth simulations. Population is the main driving factor of the model and is more correctly simulated than employment. Employment is more difficult to simulate anyway and depends on a number of economic trends with which the model does not directly deal. The model is complex with many parameters that were not all taken into account during the calibration. Weights and settings that determine the random factor r and the accessibility X in the activity potential (equation (3.1)) and some settings of the diseconomies of scale D (in equation (3.10)) were still determined with expert-based knowledge to avoid that the automated calibration would become messy.

The model can both handle monocentric and polycentric growth. The calibration is more difficult though when the residential land-use and population trends have a clearly different spatial pattern, as is the case in the north of Belgium. A majority of new residential land-use plots are located in suburban and rural areas. On the other hand, the population mainly increases in town centres, and especially in high-density neighbourhoods of the largest cities (Figure 6.8). The model, however, is built upon the hypothesis that the same influence function weights can be used to model both a land use and its associated activity. The two solutions proposed in this thesis to improve population density predictions of large cities in Flanders seem to be a step into the right direction when compared to the density change algorithm of White et al. (2012), although both have a limited negative

impact on either population or land-use growth in smaller cities and rural areas. The errors of population and residential land use in the best runs are certainly not excessive, as discussed in the results section, while the population of Brussels was clearly underestimated without the proposed updates.

Within cities though, the model is not sufficiently able to distinguish different types of neighbourhoods with specific population change rates — probably because it omits economic and demographic factors. Different types of population, based on e.g. household composition, income data, and/or nationalities, would allow to have growth patterns that are more adapted to neighbourhood characteristics. However, these data are not available at the desired resolution and the basic concepts of the model would need to be largely changed since active land uses are now associated to one single activity. It is maybe also not a goal in the first place of this type of model to provide that kind of detail, but rather to simulate the multifunctional long-term evolution of a larger region. The model cannot deal with all little details of the complexity of the planning process anyway. Even though those details are not present, planners can get insights into the current macro-scale behaviour of the system from the calibrated parameter value results obtained. Two parameters for population have significantly lower values in northern Belgium as opposed to Dublin. Firstly, the parameter m of equation (3.10) is lower than 0.1 in most results, which means that the diseconomies of scale (negative effects) and inertia (status quo) determine much more the future residential land use in Belgium than spatial attraction factors. Secondly, the parameter ε of equation (3.11) is low too in most results, indicating that diseconomies of scale are not only present in the most populated areas.

The semi-automated calibration method discussed is an interesting, novel approach with the potential of being used to calibrate other spatial models too. Yet, the method will not always lead to a simple, straightforward calibration process. A whole list of possible calibrations can be quickly generated with user-defined constraints, which is the main advantage, but this also still requires the modeller to analyse and choose among the results. The enhanced genetic algorithm creates a diverse solution space. The best model version is not necessarily the one for which the chosen error function has the lowest value, but rather the one that can produce many realistic futures. Brown et al. (2005) proved that this path dependence is an essential property of urban simulation models. Calibrations that look perfect at first sight might lead to very erroneous predictions of future patterns. Therefore, different solutions with close-to-optimal errors should be examined and validated. The best results are obtained by gradually putting more constraints on the parameter ranges of the model. Thus, experiments were carried out with different sets of weights for the error functions: residential land-use and population growth should be the main drivers of the calibration, and the weights of the local levels should be lower than those of the regional levels unless regional errors are combined with multiple resolution errors in the error function. High weights for the most local levels of the multiple-resolution errors lead to over-calibration. Indeed, it is not the most important goal of this type of simulation to predict as many as possible correct cells and activity values of individual cells, but to generate realistic growth patterns at the scale of neighbourhoods, cities and regions. Nevertheless, strong local errors should be avoided and extremely high activity values in even a minority of the cells are not desirable since they may lead to model instability by introducing

high potentials in the next time steps. The second proposed new density equation (3.21) easily creates such outliers (a few cells having an extremely high population) when the exponent is only slightly increased. This is one of the reasons why RMSE-based error functions seem to perform better than MAE-based error functions: the RMSE is more sensitive to outliers (Pontius et al., 2008).

The land-use category area demand and the compatibilities between activities and land uses in the model, that determine total activity values within each land-use category, are predefined, which was the only possible working assumption in the absence of a linked regional demographic and economic model. An evolution in time can be specified and we used the real evolution of the compatibilities for the calibration period. Yet, this model assumption introduces rather artificial rescaling equations, and does not meet user expectations that an urban growth model should endogenously translate an increasing population potential into a total urban land-use demand. Only scenarios with predefined future land-use category cell totals can be directly analysed with the model, as was done in the next chapter. For future model versions we should consider redefining these compatibilities. Separate activity and land-use influence functions could be tested but would make the model more complex. An updated macro model could be another solution, possibly with variable units at different scales instead of, or combined with administrative units. However, the number of parameters would then further increase and links would again be needed with the cellular model.

Travel time-based influence rules, as defined in chapter 4, add realism to the long-distance relationships considered in the model. Differences between results with and without network distances in regions

like Flanders with a dense network are limited but have a somewhat less simple circular pattern around the cities, as could be expected. We did not include public transport in the calibration runs discussed in this chapter and we always used Euclidean distances for local influences below level 2 of the variable grid. Updates in time of the road network were provided for the calibration period, but are not available for the long-term predictions since the model is not coupled to a transportation model that could provide future travel times. As such, locations in the study area which are well connected to centres of activity in the present might get overly high potentials and activity values in the future. The linear coastal development in the network-based Dublin simulations for 2050 is excessive and would be less pronounced if the model would consider that new roads are built in new urban areas. A strong reduction of the impact of the accessibility X would be an alternative solution. As indicated before, the density of roads in Flanders is already much higher today, and therefore the results for Flanders do not seem to be as much affected by this problem.

Different comparison methods for calibration and/or validation could still be tested. Firstly, we considered fuzziness in location with the multiple resolution approach but not fuzziness in category. For example, if an actual retail cell would get the category ‘services’ in the simulation, this could be considered as a result not differing too much from reality (value between 0 and 1 in the validation) since both are economic land uses, while the category ‘nature’ would still be totally wrong (value 0 in the validation). Secondly, we used pattern-based techniques (specifically spatial metrics) for validation but not for calibration. It should however be noted that calibrating with spatial metrics is difficult too: it depends on the choice of the metrics, and should best be combined with local or focal error functions

(Hagen-Zanker & Martens, 2008; van der Kwast et al., 2009; Van de Voorde et al., 2016). Moreover, spatial metrics are typically applied to compare categorical maps, while we wanted to systematically compare differences both in quantity (for activities) and category (for land uses). Finally, if input maps were available for more time steps, we could test the calibration framework more thoroughly. Calibrations with more than two time steps are likely to generate better results (Blečić et al., 2015).

6.5 Conclusions

Automated calibration of a land-use change model is faster and more comprehensive than manual calibration but it can result in over-calibrated or meaningless parameter sets. In this chapter we proposed a semi-automated approach that allows modellers to use their expert knowledge to determine the possible range of the considered parameters. An adapted genetic algorithm generates diverse sets of parameter values with optimised scores of a weighted error function. The vast number of close-to-optimal results still needs to be examined with different validation techniques though and does not directly indicate the most optimal set of parameter values.

Urban growth can be analysed more multifunctionally with the activity-based CA model in comparison with models that only simulate land-use categories. The prediction of population at the regional and at the city scale outperforms results of the MOLAND model for the Greater Dublin Region. Travel-time based CA influence rules enable to include long-distance effects in a realistic way, although in the experiments that were performed the method proved less suitable for running future simulations in regions with a less dense road network, since the CA model is not yet coupled to a transportation model that updates the network.

The results of the calibration reveal that some of the model hypotheses and equations could still be improved. The model can have difficulties in properly capturing the evolution of both an activity and its associated land use. Two new density equations were assessed to enable an increased densification of city centres. The population change patterns in the largest cities are better predicted in the Flemish application with both these equations than with the original density equation, although other inconvenient model behaviour is also introduced. The first proposed equation – (3.20) – somewhat underestimates population growth in smaller city centres, while the other equation – (3.21) – is very sensitive to small parameter changes and can have a negative effect on suburban or rural residential land-use growth. Since the best choice of the activity density equation does not seem to be universal (e.g. the original equation works better in the Greater Dublin Region), future research still has to come up with better ways to simultaneously deal with land uses and activities regardless of the region studied. As for now and because of its good results in highly densifying areas, equation (3.20) was preferred to run model simulations of a business-as-usual and a more sustainable land-use scenario for the region of Flanders with the model. These two scenarios are presented in the next chapter.

Appendix 6.A1 Additional tables of the results for Flanders and the BCR

The weights of the error functions that were used in the final calibration stage of the ten representative simulation runs can be found in Tables 6.A1 (multiple resolution errors only) and 6.A2 (including regional errors). Table 6.A3 contains the different computed errors in activity, and Table 6.A4 the

different computed errors in the number of cells of active land-use categories.

Table 6.A1 Definition of calibration error function weight types for Flanders and the BCR without regional errors. The multiple resolution error weights g_L are given per activity K or land-use category i . Abbreviations are used for population (pop), residential land use (res) and economic sectors (ind = industry, whl = wholesale and logistics, ret = retail, ser = services).

Type	Activity K or LU i	g_1	g_2	g_3	g_4	g_5	g_6
1	$K = \text{pop}$	0.02	0.02	0.05	0.05	0.03	0.02
	$K = \text{ind}$	0.0005	0.0005	0.001	0.001	0.001	0.001
	$K = \text{whl}$	0.0005	0.0005	0.001	0.001	0.001	0.001
	$K = \text{ret}$	0.002	0.002	0.005	0.005	0.002	0.001
	$K = \text{ser}$	0.01	0.01	0.04	0.04	0.02	0.01
	$i = \text{res}$	0.01	0.01	0.04	0.04	0.02	0.02
	$i = \text{ind}$	0.002	0.002	0.004	0.004	0.004	0.004
	$i = \text{whl}$	0.002	0.002	0.004	0.004	0.004	0.004
	$i = \text{ret}$	0.002	0.002	0.004	0.004	0.002	0.002
	$i = \text{ser}$	0.005	0.005	0.02	0.02	0.01	0.01
2	$K = \text{pop}$	0.01	0.05	0.1	0.1	0.1	0.1
	$i = \text{res}$	0.01	0.01	0.04	0.04	0.02	0.02
3	$K = \text{pop}$	0.01	0.05	0.1	0.1	0.1	0.1
4	$K = \text{pop}$	0.01	0.05	0.1	0.1	0.1	0.1
	$K = \text{ind}$	0.0002	0.001	0.002	0.002	0.002	0.002
	$K = \text{whl}$	0.0001	0.0005	0.001	0.001	0.001	0.001
	$K = \text{ret}$	0.001	0.005	0.01	0.01	0.01	0.01
	$K = \text{ser}$	0.001	0.01	0.025	0.025	0.025	0.025
	$i = \text{res}$	0.0075	0.0375	0.075	0.075	0.075	0.075
	$i = \text{ind}$	0.0004	0.002	0.0075	0.0075	0.0075	0.0075
	$i = \text{whl}$	0.0004	0.002	0.0075	0.0075	0.0075	0.0075
	$i = \text{ret}$	0.0001	0.0005	0.001	0.003	0.003	0.003
	$i = \text{ser}$	0.0005	0.0025	0.01	0.01	0.01	0.01

Table 6.A2 Definition of calibration error function weight types for Flanders and the BCR with regional errors. The regional error weight h and the multiple resolution error weights g_L are given per activity K or land-use category i . Abbreviations are used for population (pop), residential land use (res) and economic sectors (ind = industry, whl = wholesale and logistics, ret = retail, ser = services).

Type	Activity K or LU i	h	g_1	g_2	g_3
5	$K = \text{pop}$	0.5	0.01	0.05	0.1
	$i = \text{res}$	0.5	0.01	0.05	0.1
6	$K = \text{pop}$	1.0	0.01	0.075	0.125
	$K = \text{ind}$	0.001	-	-	-
	$K = \text{whl}$	0.001	-	-	-
	$K = \text{ret}$	0.001	0.0001	0.0005	0.001
	$K = \text{ser}$	0.1	0.001	0.005	0.01
	$i = \text{res}$	0.5	0.01	0.075	0.125
	$i = \text{ind}$	0.001	-	-	-
	$i = \text{whl}$	0.001	-	-	-
	$i = \text{ret}$	0.001	0.0001	0.0005	0.001
	$i = \text{ser}$	0.05	0.001	0.005	0.01
7	$K = \text{pop}$	1.0	-	-	-
	$K = \text{ind}$	0.1	-	-	-
	$K = \text{whl}$	0.1	-	-	-
	$K = \text{ret}$	0.25	-	-	-
	$K = \text{ser}$	0.25	-	-	-
	$i = \text{res}$	0.95	-	-	-
	$i = \text{ind}$	0.05	-	-	-
	$i = \text{whl}$	0.05	-	-	-
	$i = \text{ret}$	0.125	-	-	-
	$i = \text{ser}$	0.125	-	-	-

Table 6.A3 Errors in activity K in 2013 of different calibration runs for Flanders and the BCR: the regional error R and the errors E_L at different resolutions. Underlined values are optimised together with the underlined values in Table 6.A4 for the active land uses. Abbreviations are used for population (pop) and economic sectors (ind = industry, whl = wholesale and logistics, ret = retail, ser = services).

Run	K	R	E_1	E_2	E_3	E_4	E_5	E_6
F1	pop	0.0158	<u>0.4294</u>	<u>0.2420</u>	<u>0.1265</u>	<u>0.0471</u>	<u>0.0131</u>	<u>0.0040</u>
	ind	0.0853	<u>1.5878</u>	<u>0.8131</u>	<u>0.4196</u>	<u>0.1898</u>	<u>0.0614</u>	<u>0.0130</u>
	whl	0.0865	<u>3.9686</u>	<u>2.1044</u>	<u>1.1101</u>	<u>0.5056</u>	<u>0.1324</u>	<u>0.0202</u>
	ret	0.0721	<u>1.0030</u>	<u>0.5338</u>	<u>0.2870</u>	<u>0.1322</u>	<u>0.0461</u>	<u>0.0092</u>
	ser	0.0412	<u>1.1520</u>	<u>0.6470</u>	<u>0.2900</u>	<u>0.1042</u>	<u>0.0281</u>	<u>0.0052</u>
F2	pop	0.0133	<u>0.4431</u>	<u>0.2495</u>	<u>0.1273</u>	<u>0.0443</u>	<u>0.0114</u>	<u>0.0030</u>
	ind	0.0851	<u>1.6085</u>	<u>0.8170</u>	<u>0.4176</u>	<u>0.1874</u>	<u>0.0605</u>	<u>0.0130</u>
	whl	0.0882	<u>3.9636</u>	<u>2.1009</u>	<u>1.1120</u>	<u>0.5090</u>	<u>0.1361</u>	<u>0.0226</u>
	ret	0.1105	<u>1.0011</u>	<u>0.5695</u>	<u>0.3359</u>	<u>0.1844</u>	<u>0.0669</u>	<u>0.0134</u>
	ser	0.0397	<u>1.1538</u>	<u>0.6411</u>	<u>0.2861</u>	<u>0.1032</u>	<u>0.0282</u>	<u>0.0051</u>
F3	pop	0.0344	<u>0.4725</u>	<u>0.2768</u>	<u>0.1634</u>	<u>0.0813</u>	<u>0.0260</u>	<u>0.0061</u>
	ind	0.0758	<u>1.6511</u>	<u>0.8187</u>	<u>0.4039</u>	<u>0.1765</u>	<u>0.0595</u>	<u>0.0129</u>
	whl	0.0886	<u>3.9638</u>	<u>2.1039</u>	<u>1.1139</u>	<u>0.5084</u>	<u>0.1362</u>	<u>0.0224</u>
	ret	0.0806	<u>1.0929</u>	<u>0.5660</u>	<u>0.2952</u>	<u>0.1436</u>	<u>0.0524</u>	<u>0.0103</u>
	ser	0.0353	<u>1.1544</u>	<u>0.6396</u>	<u>0.2864</u>	<u>0.1019</u>	<u>0.0272</u>	<u>0.0048</u>
F4	pop	0.0141	<u>0.4208</u>	<u>0.2329</u>	<u>0.1172</u>	<u>0.0405</u>	<u>0.0107</u>	<u>0.0026</u>
	ind	0.0764	1.5487	0.7885	0.3928	0.1716	0.0601	0.0145
	whl	0.0861	3.9552	2.0951	1.1058	0.5037	0.1338	0.0218
	ret	0.1436	1.9714	1.0155	0.5500	0.2493	0.0863	0.0230
	ser	0.0714	1.2338	0.7043	0.3612	0.1700	0.0580	0.0114
F5	pop	0.0278	<u>0.3665</u>	<u>0.2117</u>	<u>0.1132</u>	<u>0.0518</u>	<u>0.0187</u>	<u>0.0056</u>
	ind	0.0890	1.5896	0.8344	0.4372	0.1983	0.0637	0.0127
	whl	0.1017	4.1000	2.1844	1.1651	0.5341	0.1419	0.0227
	ret	0.0583	1.2169	0.6016	0.3061	0.1220	0.0351	0.0066
	ser	0.0644	1.1857	0.6626	0.3274	0.1539	0.0513	0.0101

Table 6.A3 (continued)

Run	<i>K</i>	<i>R</i>	<i>E</i> ₁	<i>E</i> ₂	<i>E</i> ₃	<i>E</i> ₄	<i>E</i> ₅	<i>E</i> ₆
F6	pop	0.0144	<u>0.4275</u>	<u>0.2363</u>	<u>0.1187</u>	<u>0.0412</u>	<u>0.0121</u>	<u>0.0039</u>
	ind	0.0859	<u>1.5173</u>	<u>0.7940</u>	<u>0.4126</u>	<u>0.1873</u>	<u>0.0609</u>	<u>0.0130</u>
	whl	0.1366	<u>4.1911</u>	<u>2.2630</u>	<u>1.2132</u>	<u>0.5670</u>	<u>0.1618</u>	<u>0.0272</u>
	ret	0.0768	<u>1.0430</u>	<u>0.5606</u>	<u>0.3059</u>	<u>0.1400</u>	<u>0.0498</u>	<u>0.0101</u>
	ser	0.0458	<u>1.1525</u>	<u>0.6417</u>	<u>0.2902</u>	<u>0.1066</u>	<u>0.0289</u>	<u>0.0050</u>
F7	pop	0.0225	<u>0.4044</u>	<u>0.2304</u>	<u>0.1216</u>	<u>0.0500</u>	<u>0.0162</u>	<u>0.0030</u>
	ind	0.0943	<u>1.7888</u>	<u>0.8794</u>	<u>0.4472</u>	<u>0.2039</u>	<u>0.0663</u>	<u>0.0136</u>
	whl	0.1482	<u>4.4242</u>	<u>2.4146</u>	<u>1.3007</u>	<u>0.6070</u>	<u>0.1748</u>	<u>0.0292</u>
	ret	0.0724	<u>1.0044</u>	<u>0.5375</u>	<u>0.2869</u>	<u>0.1312</u>	<u>0.0482</u>	<u>0.0106</u>
	ser	0.0395	<u>1.1624</u>	<u>0.6380</u>	<u>0.2864</u>	<u>0.1050</u>	<u>0.0300</u>	<u>0.0074</u>
F8	pop	<u>0.0116</u>	<u>0.4585</u>	<u>0.2585</u>	<u>0.1271</u>	0.0427	0.0121	0.0034
	ind	0.0716	1.4966	0.8006	0.4216	0.1893	0.0618	0.0151
	whl	0.1800	4.4747	2.2508	1.2116	0.5843	0.1918	0.0431
	ret	0.0668	1.4755	0.6482	0.3045	0.1308	0.0453	0.0091
	ser	0.0730	1.1867	0.6630	0.3094	0.1276	0.0429	0.0064
F9	pop	<u>0.0139</u>	<u>0.4291</u>	<u>0.2439</u>	<u>0.1201</u>	0.0421	0.0126	0.0039
	ind	<u>0.0986</u>	2.0971	1.1431	0.5882	0.2605	0.0816	0.0173
	whl	<u>0.1340</u>	4.2023	2.4036	1.3470	0.6274	0.1778	0.0277
	ret	<u>0.1138</u>	<u>0.9927</u>	<u>0.5666</u>	<u>0.3405</u>	0.1887	0.0633	0.0111
	ser	<u>0.0409</u>	<u>1.1520</u>	<u>0.6459</u>	<u>0.2920</u>	0.1094	0.0323	0.0056
F10	pop	<u>0.0168</u>	0.4546	0.2545	0.1239	0.0440	0.0146	0.0051
	ind	<u>0.0643</u>	1.5055	0.7790	0.3947	0.1713	0.0528	0.0101
	whl	<u>0.0883</u>	3.9601	2.0996	1.1119	0.5094	0.1366	0.0231
	ret	<u>0.0278</u>	6.7406	2.2508	0.7193	0.1932	0.0364	0.0055
	ser	<u>0.0381</u>	1.1688	0.6472	0.2960	0.1127	0.0331	0.0058

Table 6.A4 Errors in the number of cells of active land-use category i in 2013 of different calibration runs for Flanders and the BCR: the regional error R and the errors E_L at different resolutions. Underlined values are optimised together with the underlined values in Table 6.A3 for the activities. Abbreviations are used for residential land use (res) and economic sectors (ind = industry, whl = wholesale and logistics, ret = retail, ser = services).

Run	i	R	E_1	E_2	E_3	E_4	E_5	E_6
F1	res	0.0371	<u>0.4600</u>	<u>0.2031</u>	<u>0.0980</u>	<u>0.0520</u>	<u>0.0269</u>	<u>0.0077</u>
	ind	0.0208	<u>0.9501</u>	<u>0.3672</u>	<u>0.1328</u>	<u>0.0471</u>	<u>0.0162</u>	<u>0.0035</u>
	whl	0.0524	<u>1.1299</u>	<u>0.4231</u>	<u>0.1674</u>	<u>0.0796</u>	<u>0.0426</u>	<u>0.0155</u>
	ret	0.0959	<u>3.2991</u>	<u>1.4684</u>	<u>0.6035</u>	<u>0.2074</u>	<u>0.0637</u>	<u>0.0215</u>
	ser	0.0408	<u>1.7377</u>	<u>0.6918</u>	<u>0.2550</u>	<u>0.0901</u>	<u>0.0291</u>	<u>0.0070</u>
F2	res	0.0405	<u>0.4633</u>	<u>0.2060</u>	<u>0.1016</u>	<u>0.0555</u>	<u>0.0296</u>	<u>0.0086</u>
	ind	0.0221	<u>0.9514</u>	<u>0.3688</u>	<u>0.1336</u>	<u>0.0478</u>	<u>0.0166</u>	<u>0.0036</u>
	whl	0.0392	<u>1.1187</u>	<u>0.4132</u>	<u>0.1570</u>	<u>0.0670</u>	<u>0.0307</u>	<u>0.0105</u>
	ret	0.0996	<u>3.2939</u>	<u>1.4615</u>	<u>0.6006</u>	<u>0.2064</u>	<u>0.0661</u>	<u>0.0231</u>
	ser	0.0481	<u>1.7304</u>	<u>0.6889</u>	<u>0.2637</u>	<u>0.1050</u>	<u>0.0391</u>	<u>0.0108</u>
F3	res	0.0422	<u>0.4654</u>	<u>0.2091</u>	<u>0.1046</u>	<u>0.0573</u>	<u>0.0298</u>	<u>0.0089</u>
	ind	0.0622	<u>1.0340</u>	<u>0.4236</u>	<u>0.1819</u>	<u>0.0900</u>	<u>0.0449</u>	<u>0.0153</u>
	whl	0.0407	<u>1.1195</u>	<u>0.4141</u>	<u>0.1579</u>	<u>0.0680</u>	<u>0.0315</u>	<u>0.0108</u>
	ret	0.1439	<u>3.5064</u>	<u>1.6235</u>	<u>0.6781</u>	<u>0.2488</u>	<u>0.0916</u>	<u>0.0316</u>
	ser	0.0491	<u>1.7622</u>	<u>0.7117</u>	<u>0.2743</u>	<u>0.1075</u>	<u>0.0374</u>	<u>0.0100</u>
F4	res	0.0462	<u>0.4729</u>	<u>0.2197</u>	<u>0.1095</u>	<u>0.0596</u>	<u>0.0315</u>	<u>0.0091</u>
	ind	0.0656	1.0407	0.4271	0.1848	0.0931	0.0497	0.0179
	whl	0.0392	1.1312	0.4196	0.1606	0.0680	0.0306	0.0105
	ret	0.4565	3.7495	1.9440	1.0931	0.6758	0.3419	0.1001
	ser	0.0689	1.9131	0.8296	0.3527	0.1445	0.0468	0.0116
F5	res	0.2108	0.6857	0.4279	0.2878	0.2076	0.1401	0.0581
	ind	0.0783	1.1348	0.4990	0.2236	0.1029	0.0530	0.0195
	whl	0.0389	1.1291	0.4179	0.1590	0.0671	0.0307	0.0108
	ret	0.3478	3.4929	1.6414	0.7951	0.4450	0.2613	0.0882
	ser	0.0542	1.7708	0.7219	0.2822	0.1154	0.0432	0.0114

Table 6.A4 (continued)

Run	<i>i</i>	<i>R</i>	<i>E</i> ₁	<i>E</i> ₂	<i>E</i> ₃	<i>E</i> ₄	<i>E</i> ₅	<i>E</i> ₆
F6	res	0.0414	<u>0.4638</u>	<u>0.2072</u>	<u>0.1023</u>	<u>0.0561</u>	<u>0.0305</u>	<u>0.0092</u>
	ind	0.0180	<u>0.9626</u>	<u>0.3808</u>	<u>0.1376</u>	<u>0.0464</u>	<u>0.0148</u>	<u>0.0034</u>
	whl	0.0315	<u>1.1262</u>	<u>0.4580</u>	<u>0.1948</u>	<u>0.0863</u>	<u>0.0277</u>	<u>0.0076</u>
	ret	0.0928	<u>3.3264</u>	<u>1.4731</u>	<u>0.6066</u>	<u>0.2092</u>	<u>0.0644</u>	<u>0.0189</u>
	ser	0.0443	<u>1.7485</u>	<u>0.6958</u>	<u>0.2609</u>	<u>0.0986</u>	<u>0.0349</u>	<u>0.0093</u>
F7	res	0.0538	<u>0.4505</u>	<u>0.2036</u>	<u>0.1100</u>	<u>0.0686</u>	<u>0.0406</u>	<u>0.0158</u>
	ind	0.0221	<u>0.9530</u>	<u>0.3707</u>	<u>0.1348</u>	<u>0.0481</u>	<u>0.0169</u>	<u>0.0037</u>
	whl	0.0270	<u>1.0903</u>	<u>0.4347</u>	<u>0.1789</u>	<u>0.0787</u>	<u>0.0237</u>	<u>0.0053</u>
	ret	0.1125	<u>3.2952</u>	<u>1.4660</u>	<u>0.6053</u>	<u>0.2123</u>	<u>0.0721</u>	<u>0.0240</u>
	ser	0.0615	<u>1.8347</u>	<u>0.7506</u>	<u>0.2976</u>	<u>0.1260</u>	<u>0.0504</u>	<u>0.0138</u>
F8	res	<u>0.0338</u>	<u>0.4681</u>	<u>0.2088</u>	<u>0.1037</u>	0.0552	0.0253	0.0067
	ind	0.2333	1.8435	1.2236	0.7945	0.4358	0.1725	0.0484
	whl	0.2470	2.8700	1.6418	1.0287	0.5829	0.2282	0.0482
	ret	0.0964	3.8076	1.7281	0.7107	0.2428	0.0714	0.0209
	ser	0.1224	2.0227	0.8375	0.3504	0.1851	0.0896	0.0193
F9	res	<u>0.0378</u>	<u>0.4695</u>	<u>0.2037</u>	<u>0.0930</u>	0.0481	0.0257	0.0077
	ind	<u>0.0426</u>	1.2796	0.6760	0.3030	0.1152	0.0371	0.0098
	whl	<u>0.1116</u>	1.9182	1.1945	0.6823	0.3037	0.0973	0.0222
	ret	<u>0.2204</u>	<u>3.5709</u>	<u>1.6540</u>	<u>0.7109</u>	0.2982	0.1523	0.0581
	ser	<u>0.0477</u>	<u>1.7482</u>	<u>0.6984</u>	<u>0.2613</u>	0.0971	0.0346	0.0094
F10	res	<u>0.0331</u>	0.4713	0.2102	0.0947	0.0451	0.0225	0.0063
	ind	<u>0.0352</u>	1.0245	0.4052	0.1489	0.0568	0.0213	0.0056
	whl	<u>0.0399</u>	1.1341	0.4199	0.1612	0.0692	0.0312	0.0108
	ret	<u>0.1054</u>	3.3457	1.4896	0.6081	0.2092	0.0671	0.0230
	ser	<u>0.0414</u>	1.7811	0.7282	0.2756	0.1034	0.0322	0.0074

7. Simulation of future land use in Flanders and Brussels

7.1 Introduction

As discussed in earlier chapters, Flanders is a good example of a region characterised by strong, uncontrolled growth of sealed surfaces during the last decades, leading to excessive sprawl and ribbon development. Currently, 14% of Flanders is occupied by sealed surfaces (Ruimte Vlaanderen, 2017). Urban sprawl is one of the triggering drivers of climate change due to an increase in energy consumption and travel distance, leading to high transport emissions (Newman & Kenworthy, 1989; Camagni et al., 2002; Bart, 2010). Short-term extra emission reduction (up to 2030) is needed if global warming has to be kept under 2 °C (Riahi et al., 2015; Rogelj et al., 2016) as proposed by the Paris climate agreement (UNFCCC, 2015). The Flemish government recently presented a plan at its Flemish Climate Summit Meeting in December 2016 to reduce emissions by taking measures in all policy domains (Department LNE, 2016). The widespread Flemish urban sprawl has a large impact on total emissions since natural areas shrink and home-to-work travel distances keep on growing. Hence, the climate plan includes the ambition that spatial development should be *land-take neutral* by 2040. Such an environmentally friendly spatial planning policy is also in line with European policy guidelines aimed at a decrease in land take to preserve open space, ecosystem services and natural soil processes (European Commission, 2016). A land-take neutral future was therefore also proposed in the recent white paper of the Spatial Policy

Plan for Flanders (*Witboek Beleidsplan Ruimte Vlaanderen*, BRV white paper) (Ruimte Vlaanderen, 2016, 2017).

Land take is defined as an increase in settlement area. In turn, settlement areas can be defined as the collection of land dominated by human activity and containing man-made structures. It is not a synonym of built-up land or sealed surface cover: it includes all land having sealed surfaces (built-up plots and infrastructure), but also parks, recreational land, mining and dump sites, etc. (Prokop et al., 2011). Only nature, agriculture and water are not settlement areas. Currently, settlement areas cover 33 % of the total Flemish area. While daily land take in Flanders was around 6 ha in 2013 (Poelmans & Engelen, 2017) (on a total area of 13,522 km²), the BRV white paper specifies that it has to decrease to 3 ha by 2025 and to a net zero growth by 2040.

Different measures are proposed in the BRV white paper to optimise the location of the still possible land take in the near future. New residential development will be strictly limited to areas in or adjacent to existing cores of cities and villages with a high public transport accessibility and with a high proximity to a sufficiently wide variety of retail and public services (from here on called “services”), as defined in a study of Verachtert et al. (2016). An earlier study of Arts et al. (2014) had already figured out that the future Flemish population could be hosted in the current urban areas and a small proportion (5 – 30 %, depending on densities) of the designated zones for new residential land with good accessibility, as long as the population density of the built-up area would highly increase. Existing residential areas (especially within cores, and to a lesser extent outside core areas) should densify if they have a high proximity to services and well-connected transit nodes of public transport. The BRV white paper suggests four strategies for

densification of the existing built-up area in Flanders: intensification of the use of space (e.g. higher or denser building structures, underground parkings), mixed use (multifunctional plots and neighbourhoods), re-use of built-up areas (e.g. the redevelopment of former industrial land), and temporary use (e.g. pop-up stores, green spaces on yet-to-be-developed business terrains) (Ruimte Vlaanderen, 2017). Inserting the adopted spatial development goals into a dynamic land-use model, such as the activity-based CA model proposed in this research, could lead to interesting insights for Flemish policy makers. By feeding the model with alternative input parameters, the effects of different spatial development policies can be simulated and analysed.

In general, land-use development scenarios can be based on demographic, economic and environmental projections or they can be specific what-if analyses. Projections can be business-as-usual (BAU) scenarios or can be based on possible alternative political developments. An example of this last category are the “world views” developed by the Dutch Bureau for Economic Policy Analysis (CPB) (CPB et al., 2006). In Flanders, De Kok et al. (2012) used the MOLAND model (White et al., 1997) to simulate future land use for BAU trends and for the different world views which were adapted to the Flemish context. As an alternative to projections, proposed policy changes can be directly analysed in what-if scenarios. In this chapter we will focus on a land-take neutral (LTN) what-if scenario, based on the guidelines of the BRV white paper. We will start from a weighted indicator of sustainable land-use development, measuring accessibility to public transport and proximity to services and will use this indicator to define zoning and suitability maps. These maps will then be used as input for the activity-based CA model. In other words, we will run

the model to find the most likely locations of future land take given the indicator values. The results will be compared to a BAU scenario.

The next section gives an overview of the model version and the land-use and activity data used in this chapter, and a detailed description of the business-as-usual (BAU) and the land-take neutral (LTN) scenario. Next, the results of both scenarios are presented. The discussion and conclusion sections focuses on possible implications of the results for policy makers.

7.2 Methods

7.2.1 The activity-based CA model and input data

A detailed description of the activity-based CA model can be found in chapters 3 and 4. Equation (3.20) was used in this chapter to update activity densities in time since the best results of strong urban densification without negative impacts on land-use growth were obtained with this equation in chapter 6. Simulations were run in annual time steps from 2013 until 2050.

Although Flanders is the principal study area in this chapter, we modelled the Brussels Capital Region (BCR) too since this region is an enclave in Flanders, and the future of the Flemish spatial development can hardly be modelled without taking Belgium's largest city into account. We used the land-use, population and employment maps of 2013 that were discussed in chapter 5 and section 6.2.4.2 as initial maps. The land-use map of 2013 is shown in Figure 6.3. In contrast with the historical calibration discussed in chapter 6, agriculture is now considered a passive category and nature is an area-based active land use. The growth of nature was in both of our scenarios based on the BAU land-use change scenario (2010 – 2030) of the Nature Report 2009 of the Research Institute for Nature and Forest

(INBO) (Van Reeth & Van Daele, 2009), extrapolated to 2050 by De Kok et al. (2012). The future total population and working age population of Flanders and Brussels until 2050 was provided per district by the Federal Planning Bureau (FPB) (FPB & ADS, 2016). The FPB computes employment predictions per economic sector per district until 2030 in its PLANET model (FPB & FOD Mobiliteit, 2015). We extrapolated the employment (number of jobs) per sector until 2050 based on (1) the working age population trend per year and (2) the average evolution of each economic sector in years before 2030, both for the whole study area. Since both before and after 2030 some years have decreasing working age population, and employment trends are clearly different if so, trends of employment per economic sector before 2030 were separately computed for years with an increasing or decreasing working age population.

7.2.2 The business-as-usual (BAU) scenario

De Kok et al. (2012) applied the MOLAND model to Flanders to implement a business-as-usual (BAU) scenario. In that model, the local CA component is coupled to a regional gravity-based spatial interaction model of the economy and population, and the necessary land-use growth to host the future population and economic activities is determined endogenously (see chapter 3). The activity-based model needs predetermined land-use growth values as an input. Therefore, we have used the BAU land-use growth trends of De Kok et al. (2012) for all active land-use categories (i.e. the relative growth of each category for each time step), but starting with the new land-use map of Poelmans et al. (2016) for the year 2013. Zoning, suitability and accessibility were copied from the original study. Protected agricultural and natural areas and zones with a high risk of flooding were

excluded from the urban zoning and suitability maps. An overview of all protected areas considered can be found in Appendix 7.A1. The ACA model allows the definition of phase-in zoning maps: some cells are immediately available to be transformed into a specific land-use category, while others only become available at later dates. Urban development in Flanders has often been allowed outside designated areas (De Decker, 2008, 2011) (see the discussion in section 5.4). Therefore, zoning maps in the BAU scenario for the initial time step only allow building in designated areas by the Spatial Planning Department (cfr. the “Spatial Planning Accountancy” (*Ruimteboekhouding*)), but gradually allow building elsewhere (mainly in agricultural land), as long as it is not in the protected areas.

The influence rules and density parameters of population were calibrated in such a way that the population in the largest cities keeps on increasing, even if many suburban neighbourhoods are still expanding. This is consistent with trends since 2001 showing that most centres of cities and even smaller towns have begun to densify, as was demonstrated in the historical population analysis of chapter 5. The recent population projections of the FPB until 2050 indicate that this trend will continue since large increases in population are predicted for districts containing large cities. Therefore, we calibrated the BAU scenario in order to keep the differences between modelled population per district in 2050 and the FPB projections as small as possible. In the simulated BAU scenario, daily urban land take gently increases from 6 to 7 ha per day between 2013 and 2050, i.e. a conversion of 89,369 ha of open space during the total time period (6.6% of the total Flemish area). The results of the BAU scenario are used as a reference to which the LTN scenario can be compared.

7.2.3 The land-take neutral (LTN) scenario

The land-take neutral scenario was defined together with policy makers of ‘Ruimte Vlaanderen’, the Flemish Department of Spatial Planning. The decreasing daily land take called for in the BRV white paper (6 ha in 2013, 3 ha by 2025, 0 ha by 2040) was converted into an exponential decay function of annual land take, which is actually close to a linear decreasing function (the exponent is -1.2). The total possible land take until 2040 is then 25,056 ha (1.8% of the total Flemish area). A comparison of daily land take in both the BAU and LTN scenarios can be seen in Figure 7.1. The proportion of the LTN land take attributed to each particular built-up land-use category is for every year the same as the proportion of the BAU land take attributed to that category in the BAU scenario. After 2040, the urban evolution is really ‘land-take neutral’ since any new urban cell has to be compensated by another urban cell disappearing. As a consequence, a strong densification of the existing residential land will be required in the LTN scenario, assuming similar population and employment growth as in the BAU scenario.

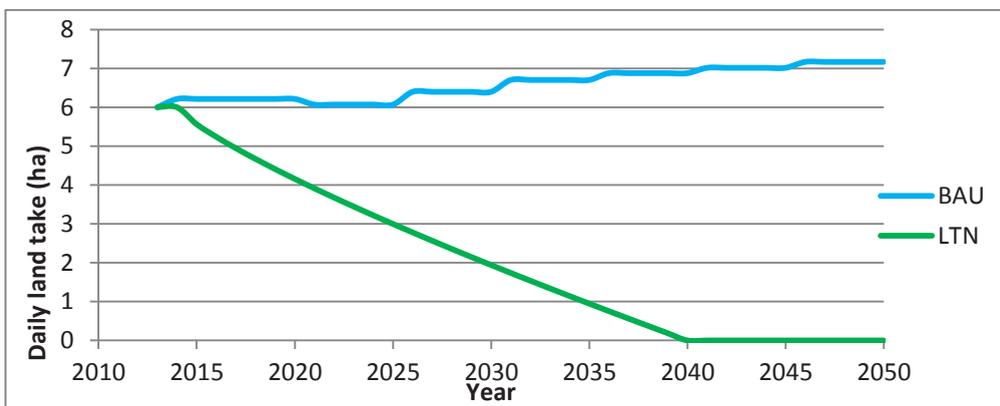


Figure 7.1 Daily urban land take in Flanders and the BCR in both scenarios.

The optimal locations for residential growth and densification in the LTN scenario were defined based on a study by Verachtert et al. (2016) who computed for each location in Flanders the accessibility by bicycle and on foot to attractive transit nodes of public transport, and the service proximity. Four categorical scores (poor, average, good or excellent) are possible for both criteria, leading to a matrix of 16 categories (Figure 7.2). The public transport modes considered are railways, tram lines, metro lines and major bus lines. Public transport accessibility was computed as a weighted value representing distance to public transport nodes via bicycle routes and foot paths, frequency of service and connection possibilities. Services are categorised as basic services, regional services and metropolitan services.

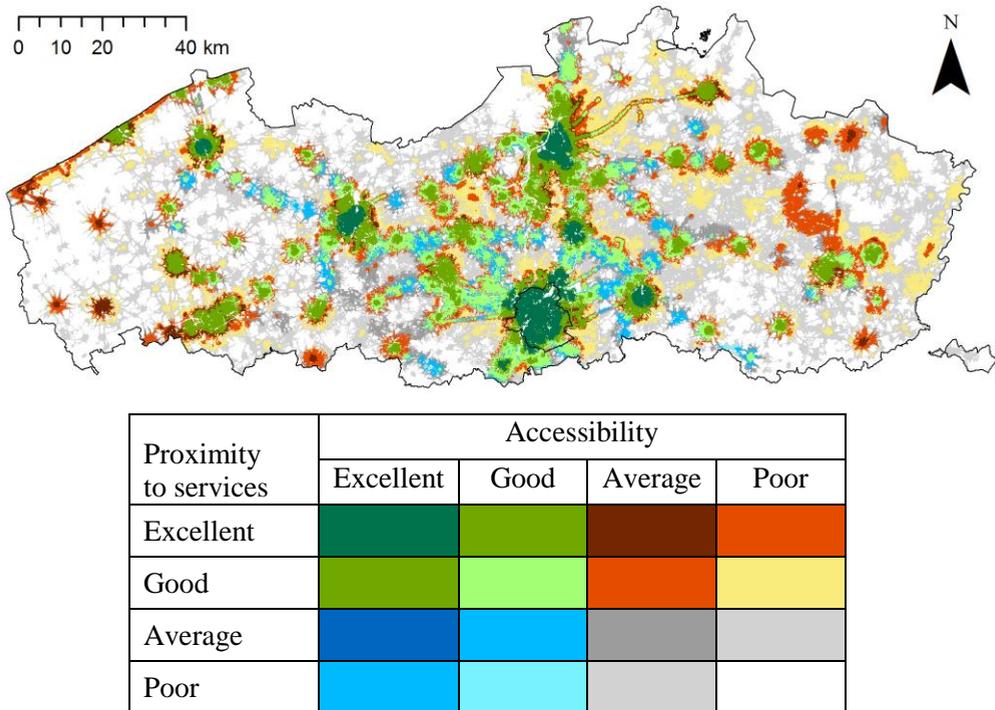


Figure 7.2 Categorical scores per ha in Flanders and the BCR for public transport accessibility and proximity to services (based on data of Verachtert et al., 2016).

Some basic services have to be accessible on foot, all other services by bicycle. The different service categories define the possible maximal travel time on foot (4 km/h) or by bicycle (15 km/h) to the service. A weighted score is awarded to service accessibility depending on the travel time and the service category. Distance decay functions, based on a study of Martínez and Viegas (2013), were used to increasingly reduce accessibility and service proximity weights the further away from the transport node or the service. More details are available in Verachtert et al. (in preparation).

Zoning and suitability values for the land-use categories ‘Residential’, ‘Retail’ and ‘Services’ in the land-take neutral scenario of this study were determined by the scores of Verachtert et al. (2016). It is important that zoning maps in the LTN scenario be as strict as possible. If actual development turns out to be worse than what was planned, then at least the simulated LTN land-use maps will show how undesired urban growth could have been avoided. Compliant with the spatial development goals of the BRV, urban growth was limited to land with a good or excellent score for at least one of both criteria, and with a location in or adjacent to an urban core. Phase-in zoning maps were also defined for the LTN scenario. As such, cells in or adjacent to urban cores with good scores on both criteria can be converted into residential, retail or services from the start, but cells with only one good score can increasingly be converted too. Cells in protected natural and agricultural areas and flood-risk zones were again excluded (see Appendix 7.A1). The total vacant or natural land that possibly can be converted into residential, retail or services is then 35,200 ha, which proved more than sufficient for the required land take (see above).

The suitability maps were created in different steps. A suitability score (Table 7.1) was first defined for each category of the accessibility-

proximity matrix. Cells outside the present cores of cities and villages get only half of the value. Cores were determined with an index of building density. Next, the suitability was increased or decreased depending on the potential to install district heating, according to a study of Renders et al. (2015). Table 7.2 gives an overview of the district heating correction factors. Cells i which are outside the defined areas but already have land use U in the initial map get by default $S_{U,i} = 0.1$ and $Z_{U,i} = 1$. Suitabilities of cells in protected natural and agricultural areas were set to zero in the same way as done for the zoning maps. The resulting suitability values can be seen in Figure 7.3.

Table 7.1 Suitability values for each possible combination of accessibility and service proximity scores of Verachtert et al. (2016).

		Accessibility			
		Excellent	Good	Average	Poor
Proximity to services	Excellent	1.0	0.8	0.4	0.3
	Good	0.8	0.6	0.3	0.2
	Average	0.4	0.3	0.1	0.05
	Poor	0.3	0.2	0.05	0.025

Table 7.2 Suitability correction factors for six different categorical potential scores to install district heating in Flanders, according to a study of Renders et al. (2015).

District heating potential						
Score	---	--	-	+	++	+++
Correction factor	0.85	0.9	0.95	1.05	1.1	1.15

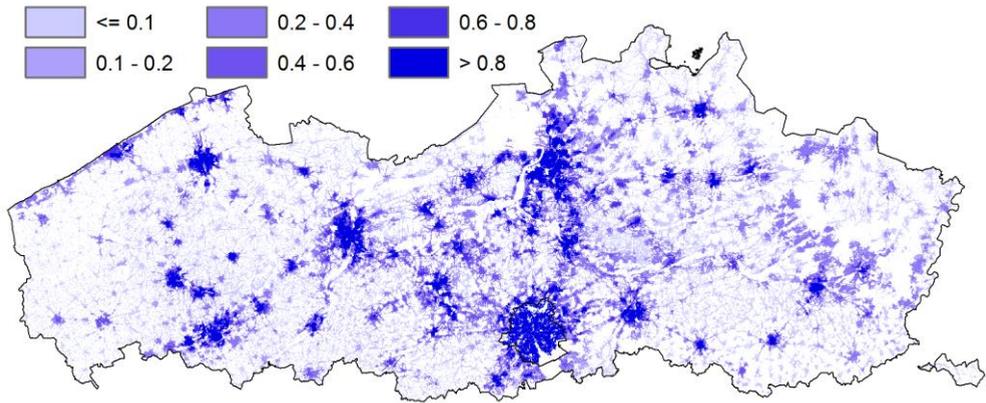


Figure 7.3 Suitability values for the land-use categories residential, retail and services in the LTN scenario.

Zoning and physical suitability for nature, industry, wholesale and logistics were copied from the BAU scenario but with the exclusion of the LTN zoning for residential land use, retail and services. The accessibility parameter X of equation (3.1) was set to 1 by default in the LTN scenario for residential, retail and services since both the network-based influence rules and especially the suitability, which is now rather defined as an accessibility parameter, already deal with proximity to roads and public transport. The accessibility parameters of industry and wholesale and logistics were copied from the BAU scenario in order to have large potentials for these categories nearby motorway junctions, major roads, railways and navigable waterways.

The LTN scenario was simulated with a higher value of γ_{POP} of equations (3.16) and (3.17) to attribute more population to new residential cells than in the BAU scenario (Table 7.3), since this is also a goal of the BRV spatial policy. Additionally, the distance-decay curves of the influences of population and of the employment categories on themselves were made steeper, in order to have a higher clustering effect of activities. The relative value of these spatial attraction factors and of the zoning and

suitability was made higher by increasing m_{POP} of equation (3.10), and by lowering the diseconomy of scale effect (lower λ_{POP} of equation (3.12)) and limiting this effect to the large cities only (higher value of the threshold ε in equation (3.11)). The densification exponent χ of equation (3.20), which makes already highly populated areas even denser, linearly decreases per year from its initial value in 2013 to a final value in 2030 and remains constant thereafter. This choice was made to ensure a large enough population growth in the cities in the beginning of the simulation, since the model would otherwise underestimate initial urban densification. The decrease over time is necessary to avoid an exponential growth effect. A lower value of the threshold ζ_{POP} was chosen in the LTN scenario to have high densification in smaller towns too.

Table 7.3 Parameters of population in the two scenarios.

	BAU	LTN
ρ	0.04	0.04
ε	1.25	4.0
λ_{POP}	0.21	0.10
m_{POP}	0.19	0.40
γ_{POP}	1.14	1.20
ζ_{POP}	5.11	2.00
$\chi_{POP,2013}$	1.005	1.004
$\chi_{POP,2030}$	1.0003	1.0003

7.3 Results

The population in the BAU scenario was calibrated to have values close to the projection of the FPB in 2050. The final parameter values (Table 7.3) and influence rules lead to low relative differences between the BAU population and the FPB projection in Brussels and the central Flemish districts with a large population (Antwerp, Ghent, Leuven, Mechelen) (Figure 7.4b). In some districts, however, the model simulations substantially differ from the projection, no matter which influence rules or population density parameters were chosen (within reasonable bounds of the parameters as determined in the historical calibration of chapter 6). The CV RMSE of population is 0.0497, which is considerably higher than the values obtained in the historical calibration (Table 6.6). The simulation length is of course also substantially longer in the scenario runs (37 years) than in the calibration runs (only 12 years). The FPB predicts a population decline in some more rural districts (Bruges, Ieper, Veurne, Tongeren, Maaseik) (Figure 7.4a), which leads to an overestimation of these regions in the model since a general population decline in larger parts of the study area is not very likely in the model due to the way the population density equations have been defined. The population strongly increases in the suburban Aalst district between Brussels and Ghent in the FPB projection (+ 24 % by 2050) even though there is no larger city with high densities. The model does not seem to be able to reproduce that trend no matter which parameter values are used.

The calibration of the LTN scenario was based on the BRV white paper policy and not on the FPB projection. Differences in regional population growth can therefore to some extent be interpreted as consequences of the policy. Some districts with a strong suburban growth

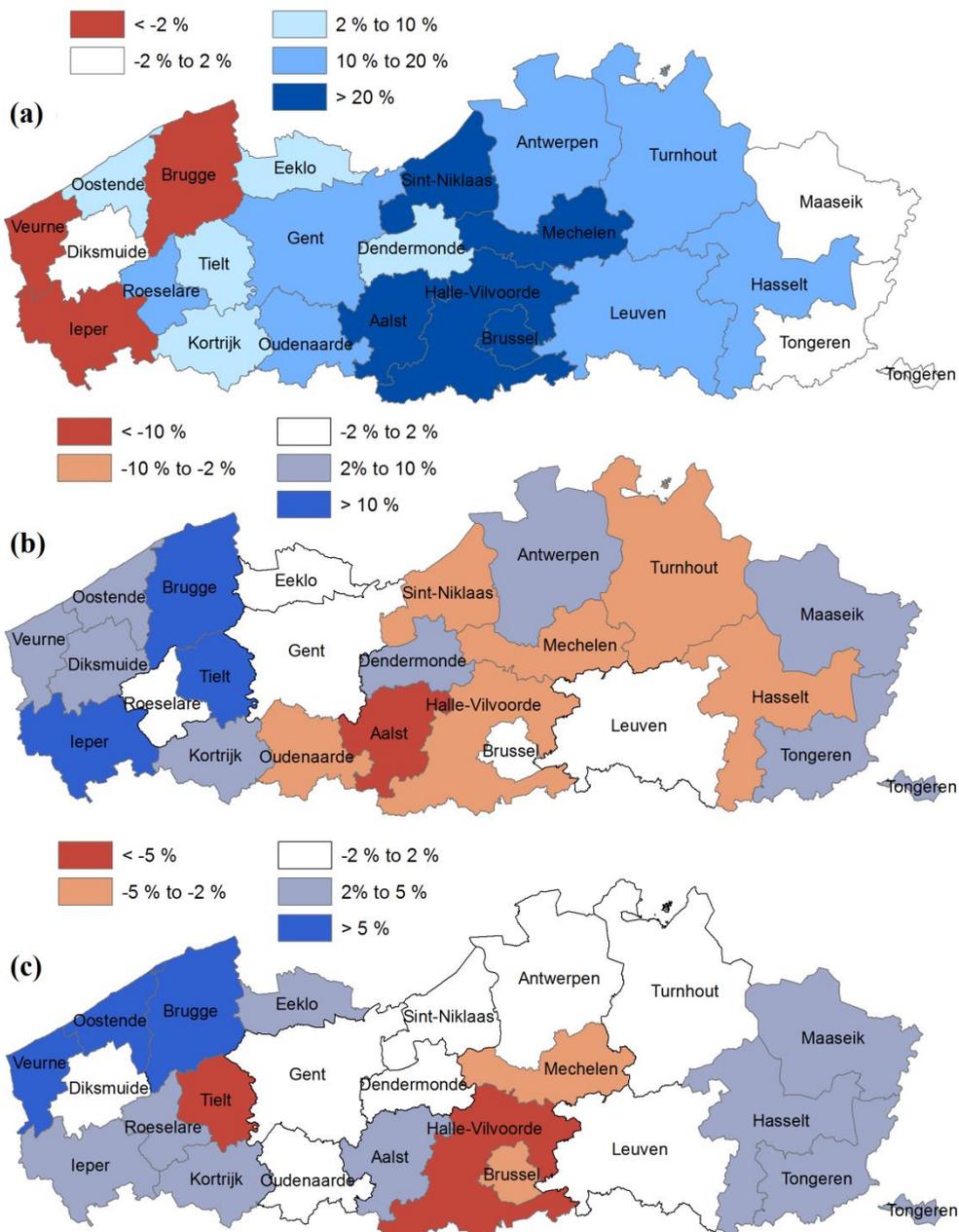


Figure 7.4 Relative difference in population (%) per district (a) between 2013 and the projection for 2050 of the Federal Planning Bureau; and in 2050 between (b) the projection of the Federal Planning Bureau and the BAU scenario, and (c) the BAU and the LTN scenario.

in the BAU scenario (Tielt, Mechelen, Halle-Vilvoorde, Oudenaarde, Dendermonde) get less population in the LTN scenario (Figure 7.4c). Also Brussels gets less population, while many rural districts get more in the LTN scenario. The scenario specifies that new residential cells get a higher density. However, since Brussels, as well as core areas of other larger cities, have little space left for new cells, the result is that the population increase in these areas is slightly lower than in the BAU scenario. On the other hand, existing residential neighbourhoods in smaller (rural) towns get more population in the LTN scenario because the higher total of centrally located new residential cells in these towns serves as an attraction pole for population growth in all nearby cells. This explains the higher regional totals in rural regions.

There is a clear difference between the BAU and the LTN in the location of new residential cells (Figure 7.5). The large majority of the BAU growth consists of new suburban neighbourhoods and ribbon development. As many as 90 % of these new cells are located outside the LTN zoning for new residential areas (Figure 7.6a). After Verachtert et al. (2016), we grouped the 16 accessibility-proximity matrix categories for analysis into four main categories: the A category (the green area in Figure 7.2) has good scores for both, the B category (blue) has good scores for public transport, the C category (orange/brown) for services, and the D category (grey) does not have any good scores. In the LTN scenario, the zoning maps force the new residential cells to be located near urban centres (Figure 7.6b). 69 % of the new cells are located in the A zone. However, due to the addition of secondary inertia (inertia of an activity outside its associated land use) in the land-use potential equation (3.10), non-residential cells with a large enough population can still be converted to residential land use. Therefore, 3.5 % of

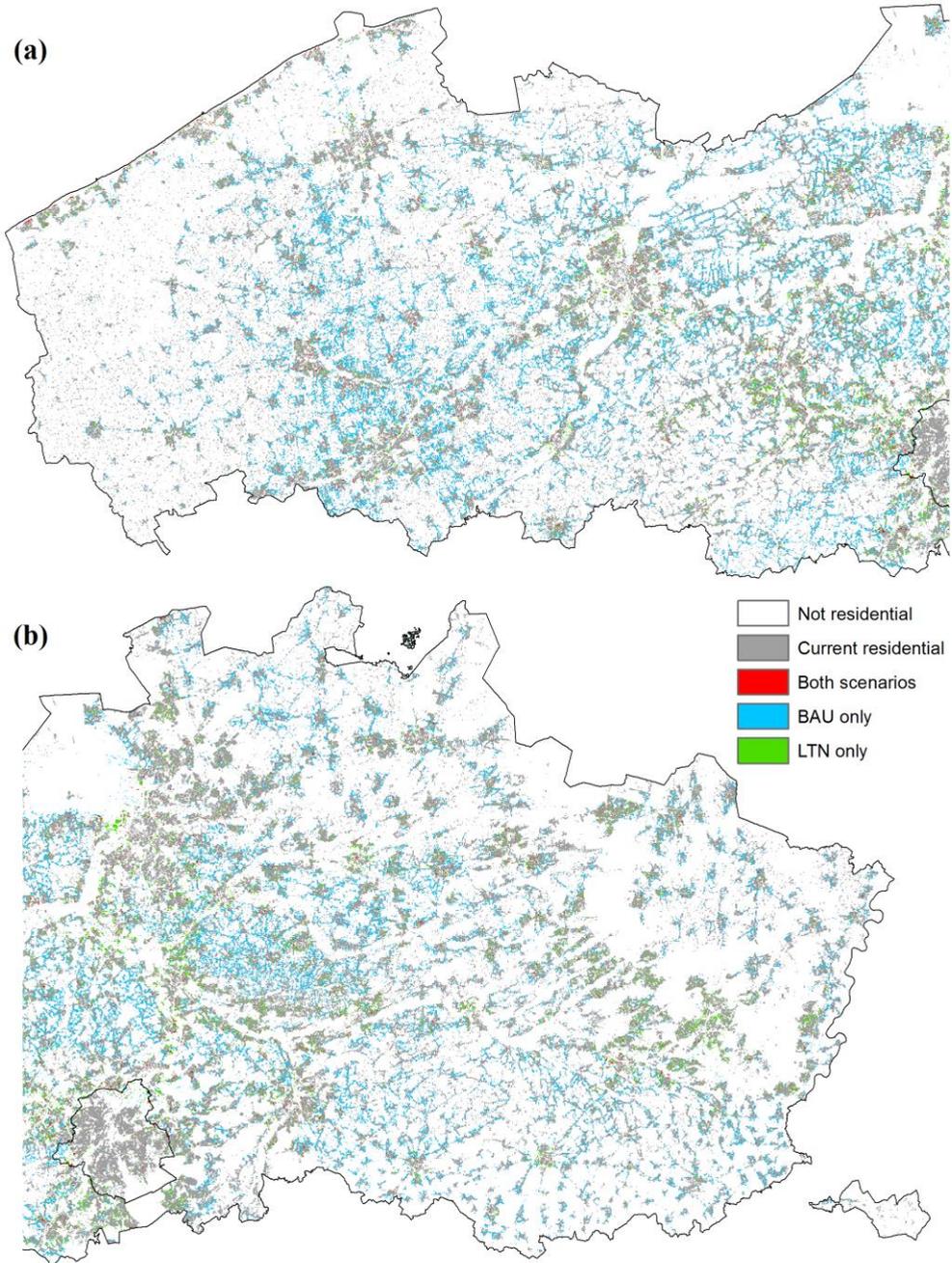


Figure 7.5 Current and new residential development in Flanders and the BCR between 2013 and 2050 in both scenarios in (a) the western and (b) the eastern half of the study area.

the new residential cells in the LTN scenario turn out to be located outside the predefined zoning. Note though that ‘non-residential cells with a large population’ in the initial map can be considered as multifunctional cells containing residential land use. Until a few years after 2020, there is little difference in the loss of open space (agriculture, nature, water) between the scenarios (Figure 7.7) because the daily land take only gradually decreases in the LTN scenario (Figure 7.1). After 2025 the BAU land take continues, resulting in an extra loss of open space of more than 60,000 ha in comparison with the LTN scenario. Since natural land uses need to grow in both scenarios, the loss of land is mainly agricultural.

Population growth in larger cities is rather similar in both scenarios, though slightly stronger in the BAU scenario (Figure 7.8). The recent trend of strong urban densification is indeed also in line with the LTN scenario and therefore differences between both simulations in Brussels, Antwerp and Ghent are limited. All other cities and smaller towns clearly have a stronger population increase in the LTN scenario. The population slightly decreases in a lot of existing suburban areas in the BAU scenario, while all new residential cells get low densities. New residential neighbourhoods in the LTN scenario get a population density that is on average almost double of the BAU density (Table 7.4), although 21 inhabitants per ha (i.e. 59 inhabitants with an average household density in 2013) that were put forward by policy makers in the Spatial Structure Plan for Flanders (RSV) (Vlaamse Overheid, 2008). The current activity-based model does not allow the user to impose minimum densities for new residential cells. A higher value of γ_{pop} in equation (3.16) does increase the average density of the new cells, but would also lead to more unrealistic outliers with extremely high densities. The density of

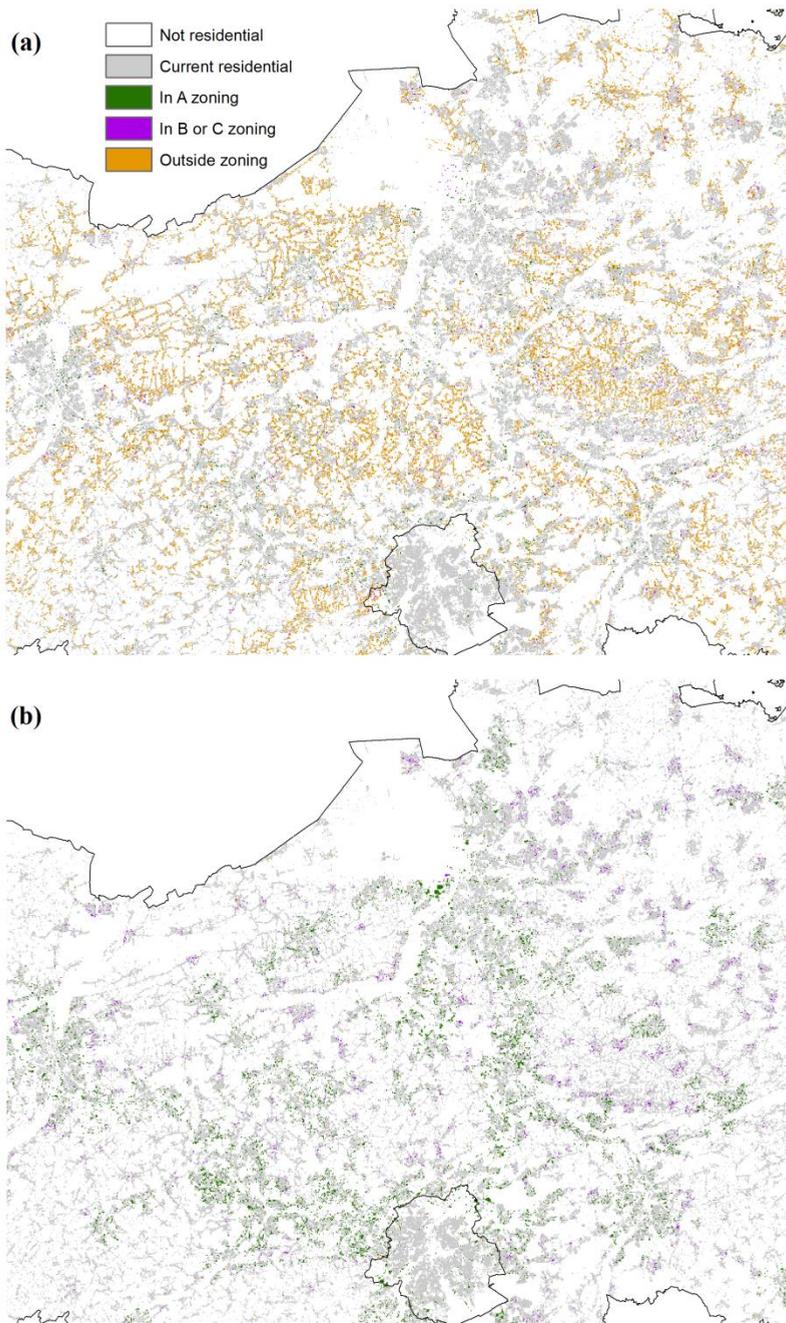


Figure 7.6 Current and new residential development in central Flanders and the BCR between 2013 and 2050 in (a) the BAU scenario and (b) the LTN scenario. With an indication if the new cells are located inside the LTN zoning and if so, in which of the main categories of the accessibility-proximity matrix.

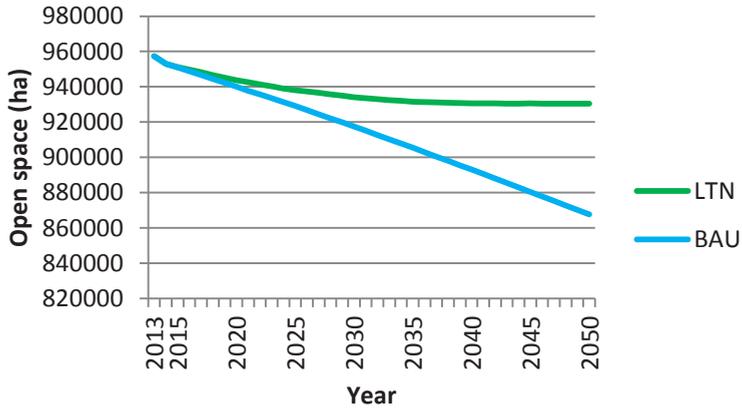


Figure 7.7 Decline of open space in Flanders and the BCR between 2013 and 2050 in the two scenarios.

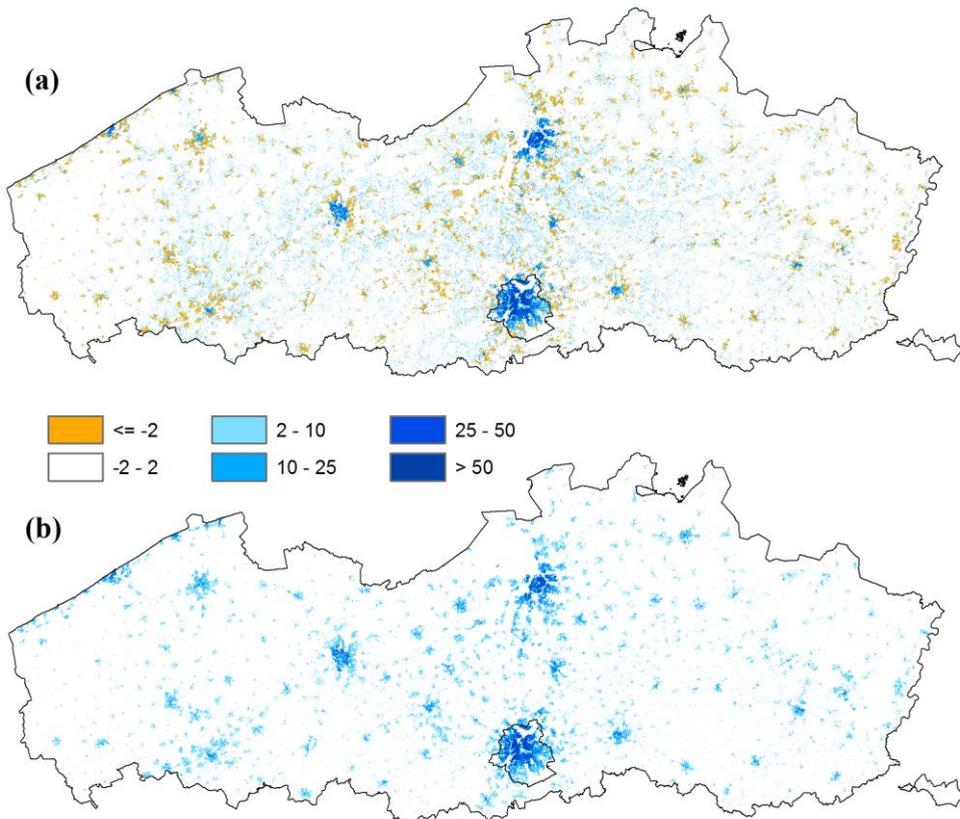


Figure 7.8 Change in population per 100 m cell in Flanders and the BCR between 2013 and 2050 in (a) the BAU scenario and (b) the LTN scenario. Increasing population in blue, decreasing population in orange.

Table 7.4 Comparison of population and population density between actual values in 2013 and simulated values in both scenarios in 2050 in non-residential cells, residential cells already existing in 2013, and new residential cells in 2050.

Cell type	2013			2050	
	Total population	Pop. density per cell (ha)		Total population	Pop. density per cell (ha)
Non-residential	2248223	1.99	BAU	2590136	2.46
			LTN	2536034	2.29
Current residential	5284744	21.47	BAU	5229347	21.34
			LTN	5699710	23.17
New residential			BAU	861069	11.47
			LTN	445462	21.47

current residential areas remains 21 inhabitants per ha in the BAU scenario and increases to 23 in the LTN scenario.

In the LTN scenario, the population should increase the most in areas with high scores in the matrix of public transport accessibility and service proximity. Tables 7.5 and 7.6 show the population in both scenarios in the four main categories of the matrix. In the initial map (2013), 59.1 % of the population lives in the A zone. Until 2025, the population strongly increases in the A zone in the LTN scenario, and even stronger in the BAU scenario. The high density demand for new residential cells is especially clear in the C zone in which population increases in the LTN scenario, while it decreases in the BAU scenario. The B zone is limited in area and population since most neighbourhoods with good public transport accessibility have a good service level too. After 2025, the relative proportion of population in the A zone starts to decline in the BAU scenario while it still grows slowly in the LTN scenario up to 61.7 %. As centrally located land still available for suburban growth increasingly gets filled up in the BAU scenario, more remote locations start to grow and finally, the D

zone gets almost 300,000 extra inhabitants in 2050 in comparison with 2013. Although the population still slightly increases in the D zone in the LTN scenario too (67,000 extra inhabitants), its relative proportion decreases from 15.6 % to 14.3 %.

Table 7.5 Total population and proportion of population (%) in the BAU scenario in the four generalised categories of the matrix of Verachtert et al. (2016).

Matrix score	2013	2025	2040	2050
A	4449997 (59.1 %)	4946754 (61.8 %)	5161723 (61.0 %)	5211808 (60.0 %)
B	105789 (1.4 %)	105961 (1.3 %)	117976 (1.4 %)	125905 (1.4 %)
C	1801481 (23.9 %)	1763282 (22.0 %)	1846274 (21.8 %)	1884892 (21.7 %)
D	1175424 (15.6 %)	1194602 (14.9 %)	1331421 (15.7 %)	1461224 (16.8 %)

Table 7.6 Total population and proportion of population (%) in the LTN scenario in the four generalised categories of the matrix of Verachtert et al. (2016).

Matrix score	2013	2025	2040	2050
A	4449997 (59.1 %)	4904050 (61.2 %)	5204189 (61.5 %)	5359599 (61.7 %)
B	105789 (1.4 %)	104220 (1.3 %)	108672 (1.3 %)	110784 (1.3 %)
C	1801481 (23.9 %)	1836121 (22.9 %)	1925421 (22.8 %)	1971110 (22.7 %)
D	1175424 (15.6 %)	1166307 (14.6 %)	1219213 (14.4 %)	1242437 (14.3 %)

Living in a central location, preferably close to work, is only one goal of the Flemish land-use policy. The residential neighbourhoods of the future should also provide the population with a high quality standard of living. Therefore, as many people as possible should have access to urban

green at walking or cycling distance. To assess this condition, we computed the accessibility of the population to different functional types of green as defined in a study of Van Herzele and Wiedemann (2003), based on the Flemish Environment Report 2000 (MIRA-S 2000). Five categories are defined depending on minimum size of the green area, with an indication of the maximum distance residents are willing to cover to reach each type of green (Table 7.7). In this study, we considered ‘nature’ and ‘recreation and parks’ as equally important green areas. The accessibility was computed with motorways, dual carriageways, railways and navigable waterways as barriers. With these definitions, 74 % of the population currently has access to neighbourhood green space at a maximum distance of 400 m (Table 7.8). It should be noted that since the resolution of the land-use map and the minimum surface of neighbourhood green are equal (100 m), not all patches of neighbourhood green might be detected by the modelling. Of all other categories, city green space (at least 60 ha within 3.2 km) is the most reachable (80 %). In the BAU scenario, access to neighbourhood green space slightly decreases to 72.5 % by 2050. Quarter and district green space remain stable, and access to city green space and urban forest clearly improves due to the growth of nature. The accessibility to urban green in the LTN scenario (Table 7.9), however, is generally much worse than in the BAU scenario. By 2050, only 64 % of the population still has at least 1 ha of urban green space in its immediate neighbourhood. Quarter and district green space have also lower accessibility values in the LTN scenario. Only for city green and urban forest the situation improves over time. In 2050 urban forests can be accessed by 76 % of the population in the LTN scenario due to the formation of larger natural clusters and the protection of open space in general.

Table 7.7 Definition of different functional types of green space that can be reached by the population, based on a study of Van Herzele and Wiedemann (2003).

Functional type of green space	Maximum distance from home (m)	Minimum surface (ha)
Neighbourhood green	400	1
Quarter green	800	10
District green	1600	30
City green	3200	60
Urban forest	5000	200

Table 7.8 Accessibility of the population to different types of green space in the BAU scenario.

Functional type of green	2013	2025	2040	2050
Neighbourhood green	74.08 %	73.22 %	72.88 %	72.54 %
Quarter green	64.03 %	63.91 %	63.99 %	63.66 %
District green	68.74 %	69.41 %	69.70 %	69.62 %
City green	80.32 %	83.40 %	83.88 %	84.26 %
Urban forest	67.67 %	70.13 %	72.01 %	72.41 %

Table 7.9 Accessibility of the population to different types of green space in the LTN scenario.

Functional type of green	2013	2025	2040	2050
Neighbourhood green	74.08 %	66.72 %	64.54 %	64.32 %
Quarter green	64.03 %	59.00 %	59.31 %	59.22 %
District green	68.74 %	67.13 %	67.67 %	67.89 %
City green	80.32 %	81.69 %	83.35 %	83.97 %
Urban forest	67.67 %	70.57 %	75.49 %	75.69 %

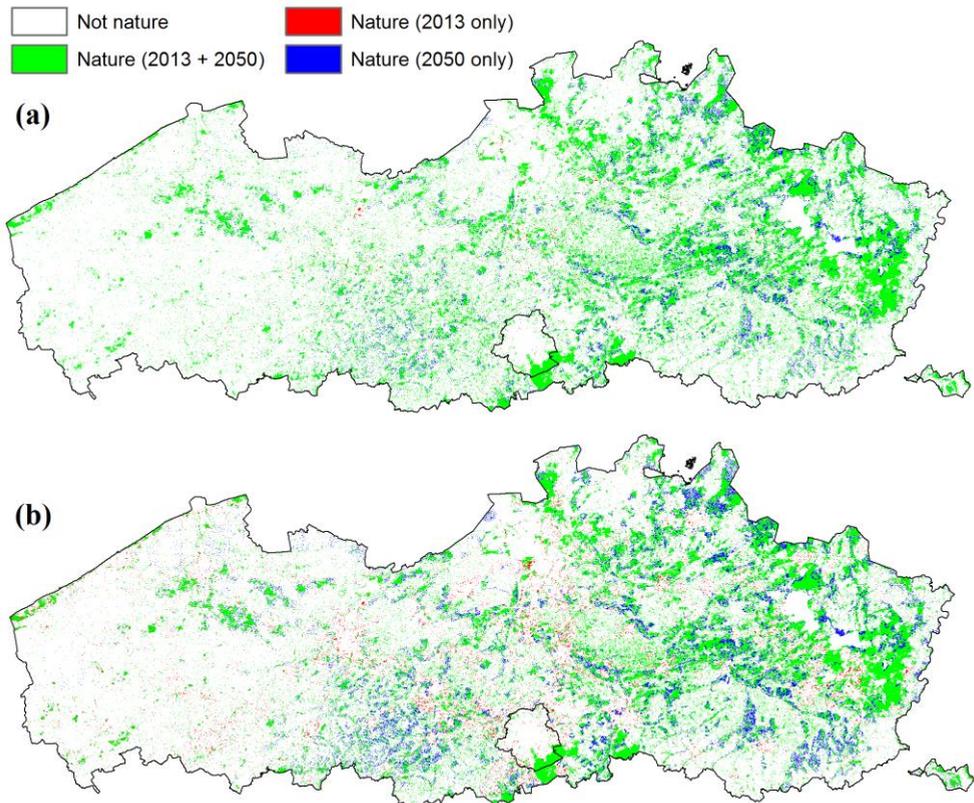


Figure 7.9 Change in natural land use in Flanders and the BCR between 2013 and 2050 in (a) the BAU scenario and (b) the LTN scenario.

The presence of less green space in the LTN towns might be counterintuitive but it is the logical effect of limiting land take to urban cores. No less than 75 % of all cells of the urban extension zones in the LTN scenario have nature as their present land use. Protected natural areas and parks were excluded but small urban green areas were not. If all of these were to be excluded, then the zoning maps of the LTN scenario simply do not provide enough space for the imposed densification in and around core urban areas. Land take should then even be more limited or anyway be allowed further away from present public transport nodes and/or retail and public services. The disappearing urban green space in the LTN scenario

can also be seen in Figure 7.9 which shows the change in nature in both scenarios. While urban green space disappears in the LTN scenario, there is logically a stronger growth of nature elsewhere leading to more and more consistent natural clusters. Both the physical suitability values and the neighbourhood effect of nature (used for the clustering) lead to more nature in regions that already have much nature, which is mainly in the eastern two thirds of Flanders and hardly at all in the West Flanders province.

7.4 Discussion

The simulation of a business-as-usual and a land-take neutral scenario with the activity-based CA model has led to interesting and even unsuspected spatial effects of two possible future policies of urban development in Flanders and the Brussels Capital Region. The striking decline of unprotected urban green in the LTN scenario certainly does not correspond with the vision of policy makers, who are aware of the importance of urban green in future dense cities, as this is also clearly indicated in the BRV white paper. If all vacant land in urban cores would effectively be converted into residential or mixed areas, then urban green should be provided at smaller scales than considered in the model. New residential cells at the 1 ha scale could indeed represent a mix of buildings, green space and water. There may also be other solutions to host the future population in central locations, and maybe the model does not consider all options to do so, since a land-use change model cannot deal with the full complexity of the planning process. More multifunctional or high-rise residential buildings could be alternatives, even though many high buildings may not be desirable in smaller towns or villages. Conversion of industrial land into residential land is theoretically possible in the model, but it is unlikely if a

high enough present employment in industry leads to a large inertia factor for industry. In reality, re-use of land will certainly be necessary to reach the land-take neutral goals. Of course, if the total industrial area would remain constant or still grow, which is assumed in this study, then open space also has to disappear for the creation of new industrial terrains. Ideally, economic land uses would densify too in the future.

Alternatively, improvement of transport infrastructure or extra retail and public service clusters could also be considered in some of the more populated suburban areas. The chosen areas could then be added to the densification zoning and attract extra population. More investment in railways, tram lines or major bus lines would help to achieve climate change and sustainability goals, and would especially be useful in regions with good services but weak accessibility (e.g. the north of the Limburg province). Generally, a large proportion of people working in Flanders still commutes by car (68.5 % in 2014, a number that has remained constant during the last 10 years) (Pauwels & Andries, 2016). Infrastructure should normally be easier to plan for the government than implementing a more stringent land-use policy, which may take many decades to fully accomplish. However, in Flanders, the space to build new public transport infrastructure is also very limited and costs may be high.

Similar studies abroad could also be an inspiration to improve policies. Caparros-Midwood et al. (2015) determined optimal areas for future urban growth in the city of Middlesbrough in the UK. They also considered ways to minimise travel costs and urban sprawl, but additionally prevented development in green spaces and minimised the risk of future heat waves. Fertner et al. (2016) studied recent urban expansion in six European and two American regions and focused on growth management.

Public acquisition of vulnerable open spaces or other types of support (e.g. for sustainable agriculture in valuable agricultural areas) can help to prevent these areas from being taken over by urban sprawl (Bengston et al., 2004; Fertner et al., 2016).

7.5 Conclusions

The activity-based CA model which was further developed in this thesis, turns out to be a particularly useful tool to analyse future spatial land-use and population patterns generated by different policy scenarios. A land-take neutral (LTN) scenario for the region of Flanders, Belgium was compared to a business-as-usual (BAU) scenario. Annual land take in the LTN scenario steadily decreases by 2040 to a zero net growth per year. New residential land plots that still appear before 2040 would get much higher densities than in the BAU scenario, in which low-density urban sprawl and ribbon development still greatly increase. A larger part of the population would live close to good public transportation nodes in the LTN scenario, which could contribute to climate change goals. However, further densification of the present urban cores is difficult without giving up valuable but unprotected urban green areas. Policy makers should therefore be careful when setting up the LTN scenario in order to make sure that land-take goals, possible land-take locations and urban green areas do not conflict with each other.

In any case, both the BAU and the LTN scenario indicate that it could be more difficult than expected to host the future population of Flanders without affecting valuable urban green or open space. However, modelling at a resolution of 1 ha also has its limits. Urban green could be present in future residential cells at smaller scales. Furthermore, local governments can focus on the different solutions indicated in the BRV white

paper to tackle the problems discussed: higher and multifunctional buildings will be necessary, and vacant buildings should be re-used. Finally, a limited and well-chosen number of suburban neighbourhoods could be selected to be provided with better public transport and services, and these areas will then be new attraction poles for densification. Anyway, a good cooperation between all policy levels will be needed to obtain the goals of the BRV white paper.

Appendix 7.A1 Protected natural and agricultural areas

Below, an overview is given of natural and agricultural areas that were excluded from the zoning and suitability maps of built-up categories. Data were provided by different departments of the Flemish Government.

These areas were excluded in all scenarios:

- Special Areas of Conservation according to the EU Habitat Directive (Natura 2000)
- Flemish forest reserves
- Protected maritime dunes by the Dune Decree
- Forests within designated zoned land for ‘Forest’, ‘Nature’ and ‘Other green’ (*Ruimteboekhouding*)
- “Flemish Ecological Network” (*Vlaams Ecologisch Netwerk (VEN)*)
- Listed nature reserves (*Vlaamse en erkende natuurreservaten*)
- Designated anchor places from the Landscape Atlas (*Vastgestelde Landschapsatlas*) and heritage landscapes (*Beschermde cultuurhistorische landschappen*)
- Effective flood-prone areas

These areas were excluded in the BAU scenario only:

- “Confirmed agricultural areas” (*Herbevestigde Agrarische Gebieden* (HAG))

These areas were excluded in the LTN scenario only:

- High risk of landslides, based on the “Flemish ‘Underground’ (Soil) Database” (*Databank Ondergrond Vlaanderen, DOV*)
- Large or valuable spatially and functionally coherent agricultural areas in the “Agricultural Impact Study” (*Landbouwimpactstudie, LIS*)

8. Conclusions and future perspectives

The activity-based CA model, or ACA model, developed by White et al. (2012), was the first CA land-use model to succeed in simulating multiple activities for each cell. A multifunctional model is more suitable to represent the real complexity of cities and regions than a model that only simulates dominant land-use states. In this thesis, several updates of the model, as well as techniques to improve its calibration and validation were presented. The main research objective was to enhance the ACA model influence rules through the integration of network-based travel times, and to validate and improve the activity density equations used in the model. Other objectives of the PhD project included the development of a new approach for producing historical land-use and activity maps, and the development and testing of a semi-automated historical calibration framework. In the last chapter of the thesis, the enhanced model was applied to simulate the future impact of a land-take neutral policy scenario for Flanders, as proposed in the recent white paper of the Spatial Policy Plan for Flanders.

8.1 Network-based influence rules

The results obtained by applying the initial model of White et al. (2012) to Belgium (at a coarse resolution of 300 m) and the Greater Dublin Region were promising, especially with respect to high-resolution population density simulation. However, the early research of this PhD project indicated a few shortcomings of the model. Use of Euclidean distances proved to work well in the 8-cell circular neighbourhood of the MOLAND model, but not in the variable grid neighbourhood where the maximal

distance has a size similar to that of the study area. Hence, network distances were needed to bring the model closer to reality. Computing network distances can be a slow process, so the results had to be stored so that they could be used in different model runs. On the other hand, a full matrix of all network distances needed by the model would again require a large computation time and storage capacity. Since local vicinity interactions could actually remain Euclidean distance-based, a fixed network grid was defined, coinciding with a chosen level of the variable grid: travel time computations were then only necessary between network grid cells and the centres of their variable grid neighbour supercells. Centres of supercells were initially geometric centres but were later on redefined to be the centre of gravity of population. Different centres for different activities could be considered in the future but would be difficult to define for influences of one activity on another.

The travel time-based interaction rules for long-distance interactions add realism to the model and allow for the definition of transport scenarios to evaluate the possible long-term effects of accessibility on land-use and activity changes. Five scenarios were defined: one with Euclidean distances, and four network-based scenarios ('road-based', 'congestion', 'train only' and 'choice'). The general growth pattern of residential land use is rather similar in the different scenarios presented in chapter 4, with residential clusters becoming the largest in the central area and the southwest of Flanders. Still, there are distinct and logical differences between the scenario results of residential land use. Also, in terms of population the differences between the scenarios are more pronounced. Some better accessible cells in all parts of the study area (but more in the central area than elsewhere) become residential and population is higher near major

roads in the ‘road-based’ scenario in comparison with the ‘Euclidean’ scenario. If space is available, extra cells appear near stations in the ‘train only’ scenario and less extra cells appear in areas with congestion in the ‘congestion’ scenario, when both are compared to the ‘road only’ scenario. As such, the model proves to be sensitive to differences in network speed. However, unavailability of space in the best accessible cells and the strong neighbourhood effect of population in dense areas still play an important role too as they push extra cells sometimes a bit further away from the most accessible location. This is actually what we expect: network accessibility is only one of the model’s drivers, and the model has to remain sensitive as well to the other drivers (zoning, suitability, diseconomies of scale, ...). The fact that alternative scenarios do not lead to very strong differences in terms of residential land-use patterns for the Flanders case has a lot to do with the very dense transportation network present in Flanders. Yet the model behaves as expected, and observed differences between the scenarios can be logically explained.

In the current implementation of the model, the public transport component of the multimodal travel time computations is still limited to travelling from or to large railway stations, since it was until recently difficult to obtain a full origin-destination train time matrix of the Belgian railways. This policy has recently changed and Verachtert et al. (2016) already used a full matrix in their study including also metro lines, tram lines and major bus lines. With these data, the results of chapter 4 could be updated. The data were already used to define zoning and suitability maps for the land-take neutral scenario in chapter 7, and therefore show how the ‘train only’ scenario for Flanders could look like with full data. However, the impact of using full data on the ‘choice’ scenario would probably be

limited since only a minority of commuters effectively take public transport to go to their job. A more realistic computation of commuting in the model should also make a difference between the access and egress legs of the multimodal displacements since normally no car can be used at the work end of the trip. Finally, travel time computation could be improved by coupling the ACA model to a transportation model. While this might slow down computation speed, it would make network times dynamic, enabling a more realistic simulation of impacts of future spatial development on accessibility and vice versa.

8.2 Assumptions and equations of the model

The computation of potentials for land-use change and the transition rules to update land-use states in the ACA model were based on the CA component of the MOLAND model and were as such extensively tested and calibrated in earlier research. Yet, the activity density equations defined by White et al. (2012) needed to be assessed in a historical calibration and compared with possible alternatives. The original equation of White et al. (2012) is based on relative changes in activity potential and is therefore the most generically applicable. Nevertheless, Brussels and to a lesser extent the largest Flemish cities recently experienced an increase of the population which is much stronger than the increase in residential land use. As such, when applying the model to Flanders during the initial, manual calibration stage of the project, the original equation systematically underestimated this densification. Two alternatives were therefore proposed in chapter 3 and evaluated in chapter 6: (1) introducing a densification exponent χ to directly update activity values but limited to areas with a high population neighbourhood effect, and (2) update activity values based on the relative

value of the neighbourhood effect. The first alternative does not change activity values in all cells in which the neighbourhood effect threshold is not exceeded; these are then only handled by the rescaling equations. Initial attempts to combine the first alternative equation in high-density areas with the original equation in all other areas did not really solve the problem. The second alternative changes the values of all cells, but its parameter τ has a limited range of values that do not lead to instability. The first alternative seems to be the most appropriate to run scenarios in which densification of larger towns is the main population change. Hence, this equation was used to run the two scenarios discussed in chapter 7. At present, none of all equations can be considered as the best option independent of the study area. In regions where land uses and their associated activities have similar growth patterns, the original equation performs well. In Flanders, however, the alternatives proposed in the PhD project clearly lead to better results for urban population density simulations.

In general, all model components and parameters which are currently included in the model seem to have a logical effect on the results. Since different study areas all have their specific policy context, the key components or parameters that steer the results the most are indeed dependent on the study area. In Flanders, the effect of the diseconomies of scale is large, and therefore the impact of all other components of the land-use transition potential (zoning, suitability, accessibility and the neighbourhood effect) is more limited than in the Greater Dublin Region. Business-as-usual residential growth in Flanders seems to be rather determined by ‘negative’ externalities (land prices and congestion) than by ‘positive’ spatial properties (suitability and proximity).

The compatibility coefficients of activities per land use serve in the first place to get realistic activity proportions in each land use. The coefficients can actually be considered as parameters to couple land-use and activity simulations, just as the MOLAND model needed to couple the regional model (handling population) with the local model (handling land use). An advantage of the definition of the compatibility coefficients is that these can be computed from the actual maps and can be changed in time. This is of course only possible in the historical calibration, and for simulations for the future it would be useful if the model could endogenously come up with changes of the compatibility as activities and land uses evolve. A macro-scale model of demographic and economic effects without regions could be considered as a tool to compute land-use demands, although many extra parameters would again be involved.

8.3 Historical data

Historical calibration of an activity-based CA model requires historical land-use and activity maps. A new methodology, ‘downdating’, was developed to infer high-resolution population maps of the past from current population maps, using ancillary data documenting urban growth over the last decades. Specifically, changes in sealed surface cover obtained from remotely sensed data were analysed to go back in time. A local relationship between population density and sealed surface cover was assumed to estimate population densities for previous time steps.

The methodology provided some interesting results for population in the first place. Local spatial trends of population in Flanders and Brussels between 1986 and 2013, obtained by applying the downdating technique, show an evolution from ever deteriorating urban sprawl in the 1990s

towards more urban densification in more recent years. A densification factor was proposed to visualise these trends for two time periods (1986 – 2001 and 2001 – 2013). The factor shows how the population increases or decreases more strongly than would be expected from the changes in sealed surface fractions. Densification of residential land use is the highest in large cities. Several studies have indicated that this densification is mainly caused by migration, while many other inhabitants still prefer to live in suburban areas or are forced to do so due to rising real-estate market prices (Slegers et al., 2012; Boussauw et al., 2013; Schuermans et al., 2015). The main novelty of our approach is that trends in densification can be monitored at a high resolution. As such, we observed increasing densities in the centres of most smaller towns, but on the other hand, some new suburban areas still show decreasing densities.

Downdating of land use was more an ad-hoc GIS operation, but the results for 2001 were generally good when compared to aerial photographs of the same period. Downdating might be less needed in the future as more and more time series of spatially detailed data will become available. Still, the technique could remain useful for new datasets that do not have a historical record yet.

8.4 Calibration and validation

The calibration approach proposed in this research aims to combine the advantages of both manual and automated calibrations into a semi-automated calibration framework. Modellers can use their domain knowledge in a manual calibration, but the task can easily become a slow, complex and ad-hoc process for models involving a large number of parameters, such as the ACA model. On the other hand, fully automated

calibrations lack domain knowledge: while they will find apparently optimised parameter value sets without the big effort needed in a manual calibration, there is always a risk that the parameter value sets are over-calibrated, do not represent the processes ongoing in reality, and would lead to nonsensical results when the model is further propagated into the future. The semi-automated calibration framework proposed tackles the weaknesses of both approaches by allowing the researcher to iteratively define possible combinations of parameter values. Non-specialists will still need a learning curve to successfully perform a semi-automated calibration, but this is more due to the inherent complexity of land-use change modelling, which is far from being a simple and straightforward process to be automated. The semi-automated methodology proposed in this study certainly has the potential to be more generally applicable than to the activity-based CA model only.

Regional and multiple-resolution error functions can be combined with user-defined weights. While multiple-resolution error functions offer a clear and straightforward approach to validate the results at all spatial scales in a framework that is consistent with the definition of the variable grid, regional error optimisation can be useful since different regions can have different activity evolutions. Some of the best results in chapter 6 were obtained when multiple-resolution errors and regional errors were combined. Focal error computations at all scales give more information on the quality of the obtained maps than global error measures or averaged local error measures, such as Kappa (Simulation), that focus too much on deviations in individual cells.

Future calibration research could (1) include all model parameters since some were not yet considered, (2) not only validate with fuzziness in location but also in category, and (3) compare the activity-based CA model

with MOLAND model results and a null model. Some researchers recently argued that cells with a changed land-use state are the only ones that matter in the calibration and/or validation of land-use change models since nothing happens in most cells (van Vliet et al., 2011; Omrani et al., 2016). This argument appears to be true for validation: e.g. when a simple *percentage correctly classified* (PCC) is computed considering all cells, then this value is misleadingly high due to all unchanged cells. The calibration in this study did not focus on cells with a changed land-use state for two reasons. Firstly, activities are modelled too and these change in all cells and secondly, the potential should also be able to determine which cells do not change state.

Finally, it should be mentioned that the goal of the historical calibration is in the first place to evaluate the model and its equations. The calibration should not necessarily provide parameter values that can directly be used in simulations for the future. An extension of historically calibrated values only determines one possible future, specifically a business-as usual scenario that presumes that the next years still look exactly the same as the previous years. Other possible futures and their impacts can be evaluated with policy scenarios.

8.5 Future urban development in Flanders

The comparison between two scenarios of future activities and land use in Flanders presented in chapter 7 is a good example of how the activity-based model can be applied to give policy makers better insights into the spatial impact of their possible decisions. Specifically, a land-take neutral scenario, building on guidelines of the recent white paper of the Spatial Policy Plan for Flanders, was compared to a business-as-usual scenario. Thanks to the direct simulation of activities, the model could be applied to compute spatial

indicators showing different impacts on the future population (accessibility to urban green, accessibility to public transport, ...). Results of the application show that the land-take neutral scenario preserves much better the remaining clusters of open space in Flanders and would probably encourage inhabitants of new residential neighbourhoods to commute via the public transport network. Yet, it turns out that unprotected urban green space might be threatened by a densification of urban cores. Hence, policy makers should reflect about alternative densification strategies that would not or less affect urban green areas. A non-extensive list of possible options could be: re-use of vacant built-up areas, mixed use, higher buildings, including urban green in built-up plots, densification of economic land uses, new public transport infrastructure and new public service clusters. Future scenario research could follow up changes to the proposed policy.

The research of this PhD project demonstrated that the ACA model is able to generate different simulations of the future depending on the implementation of scenarios that describe possible futures. Although maybe considered as a weakness since there is not one clear model prediction that can be validated, this is actually a model strength as the future is not yet determined. As long as the parameter values are well documented and simulation runs can be repeated, the model can be used to analyse the possible effects of what-if scenarios. Furthermore, since activity maps are computed for all time steps, all kinds of spatial indicators that need high-resolution maps of population or employment could benefit from scenario results of the model, and their dynamic evolution in time could be mapped. A non-extensive list of possible indicators could be: proximity of population to green areas, waste and wastewater production generated by households, population living in flood-prone areas, proximity of population and

employment to public transportation and energy hubs, defining locations for new public transport stations and routes, defining new locations of energy hubs, defining new areas with high potential for employment, etc.

8.6 Where do we go from here?

The activity-based CA model presented in this study is a multifunctional tool, envisioning how urban land-use categories, population and employment can evolve in time. The research presented in this thesis has led to a few improvements of the approach: network-based CA influence rules, population density equations capable to deal with highly densifying areas, a downdating technique to create better historical maps, and a semi-automated calibration strategy to mix the benefits of expert knowledge and computer power. As outlined in this final chapter, some model assumptions and equations could certainly still be further improved, yet, the model has already produced relevant results, as was shown in the application to two policy scenarios in Flanders.

Thanks to the computationally efficient variable grid computations, vicinity and long-distance effects of land uses and activities on each other are modelled together in a consistent framework that can be applied to large regions or even entire countries. In comparison with other types of urban models, as discussed in chapter 2, the activity-based CA model has a better spatial resolution than classic LUTI models, simpler and faster computations than most agent-based models of equally large study areas, and more multifunctionality than classic CA models. Agent-based models are maybe more suitable to explore the detailed demographics of individual cities or neighbourhoods, but it can also be difficult to feed and calibrate

them with the right data. For a larger region, CA modelling still seems to have its benefits as more calibration runs can be done in less time.

The ACA model allows population, economic activities and land-use categories to influence each other in bidirectional dynamic feedbacks. White et al. (2015) state that different coupled components can improve each other and their errors, although one would think at first sight that bidirectional feedbacks involve a larger risk of error propagation. The model assumes that global economic trends and social and cultural values are constant, in other words that the system as such does not dramatically changes its behaviour. Changes in all of these can only be incorporated by recalibrating the model. Forecasts of economic trends can indirectly be incorporated as changes in employment. It is not a goal of this type of research to go much further: a land-use change model is not a full model of how a society can evolve, providing such a model could be defined or calibrated anyway.

At present, only two types of activities have fully been tested: population and employment in different sectors. An effort to include agricultural activity has also been made in another research project (the GroWaDRISK project, funded by the Belgian Science Policy Office). Theoretically, any activity type could be coupled to active land uses, so natural and agricultural activities could be added. Agricultural activity would most likely be different types of yield, while natural activity could be based on ecosystem services (and have a monetary value) or more simply be the proportion of green in each cell if a land-use map at a more detailed resolution is available. Next to employment, other types of economic activity could be tested (e.g. annual revenue) if high-resolution data would be available.

Different types of population depending on demographic characteristics would be interesting to include as different activities as these could improve the modelling of intra-urban population density changes. More information on demographic and economic characteristics of neighbourhoods would allow to define spatially different calibrations of the activity-density equations. Examples of data could be household composition, migration or income data. Demographic neighbourhood properties are not constant in time though and the model would have to become a proper demographic model to keep these data up-to-date in future time steps. Building density data could alternatively be used to define different types of residential land use (very dense, urban, suburban, rural, ...). However, unless the different types would still be coupled to different population types, the one-to-one coupling of activities and active land uses would need to be dropped to implement this option.

Some researchers have proposed the use of specific influence rules for subareas of the study area (e.g. Li et al., 2008). While it seems to make sense to deal with differences between dense and rural regions, this makes the definition of influence rules again complex and the neighbourhood rules per region should then evolve in time so that rural influence rules can become urban in regions that are densifying. Or in other words, a cell should be able to evolve from rule set A to rule set B. Future research could also test if the diseconomies of scale can be made more realistic with spatial information on land prices and congestion. Again, information on the trends of both in relation to the changing land use in their neighbourhood would be necessary in order not to make the diseconomies of agglomeration effect static.

All of the previous suggestions might make the model more complex and less flexible. An alternative future task would be to make the model even more simple. The calibration revealed that many different combinations of parameter values result in similar errors. It would be interesting to assess whether some model components and parameters are redundant and can be eliminated because they are in some way already present in other components or parameters. A simpler model would be easier to understand for policy makers and other non-specialists.

Finally, other projects of the Flemish Institute for Technological Research (VITO) have proven that client-specific applications (with in the case of scenario development policy makers as clients) improve models and calibration techniques, and ensure that the project results are really interesting to the clients or policy makers and not only to scientists. Further developers of the activity-based model should certainly keep in mind that relevant applications of the model could still lead to relevant additions to the model. Involving more researchers in the development of the model (making the model open source or setting up specific collaborations) might be worthwhile to encourage an ongoing scientific discussion about urban modelling in general and, more specifically, the activity-based CA model.

References

- Adcock, S. (2005). The Genetic Algorithm Utility Library (GAUL). Retrieved from <http://gaul.sourceforge.net/index.php> [Accessed 11 May 2016].
- Al-Ahmadi, K., See, L., Heppenstall, A., & Hogg, J. (2009). Calibration of a fuzzy cellular automata model of urban dynamics in Saudi Arabia. *Ecological Complexity*, 6(2), 80–101. doi - 10.1016/j.ecocom.2008.09.004
- Albrechts, L. (1999). Planners as catalysts and initiators of change. The new structure plan for Flanders. *European Planning Studies*, 7(5), 587–603. doi - 10.1080/09654319908720540
- Aljoufie, M., Zuidgeest, M., Brussel, M., van Vliet, J., & van Maarseveen, M. (2013). A cellular automata-based land use and transport interaction model applied to Jeddah, Saudi Arabia. *Landscape and Urban Planning*, 112, 89–99. doi - 10.1016/j.landurbplan.2013.01.003
- Allen, P. (1997). *Cities and Regions as Self-Organizing Systems: Models of Complexity*. Amsterdam: Gordon and Breach.
- Allen, P., & Sanglier, M. (1979). A dynamic model of growth in a central place system. *Geographical Analysis*, 11(3), 256–272. doi - 10.1111/j.1538-4632.1979.tb00693.x
- Allen, P., Sanglier, M., Engelen, G., & Boon, F. (1984). Evolutionary spatial models of urban and regional systems. *Sistemi Urbani*, 1, 3–36.
- Almeida, C. M., Gleriani, J. M., Castejon, E. F., & Soares-Filho, B. S. (2008). Using neural networks and cellular automata for modelling intra-urban land-use dynamics. *International Journal of Geographical Information Science*, 22(9), 943–963. doi - 10.1080/13658810701731168
- Alonso, W. (1964). *Location and Land Use: Toward a General Theory of Land Rent*. Cambridge, MA, USA: Harvard University Press.

An, L. (2012). Modeling human decisions in coupled human and natural systems: review of agent-based models. *Ecological Modelling*, 229, 25–36. doi - 10.1016/j.ecolmodel.2011.07.010

Anas, A., Arnott, R., & Small, K. A. (1998). Urban spatial structure. *Journal of Economic Literature*, 36(3), 1426–1464.

Andersson, C., Lindgren, K., Rasmussen, S., & White, R. (2002a). Urban growth simulation from first principles. *Physical Review E*, 66(2), 1–9. doi - 10.1103/PhysRevE.66.026204

Andersson, C., Rasmussen, S., & White, R. (2002b). Urban settlement transitions. *Environment and Planning B: Planning and Design*, 29(6), 841–865. doi - 10.1068/b12813

Antrop, M. (2000). Changing patterns in the urbanized countryside of Western Europe. *Landscape Ecology*, 15(3), 257–270. doi - 10.1023/A:1008151109252

Antrop, M. (2004). Landscape change and the urbanization process in Europe. *Landscape and Urban Planning*, 67(1-4), 9–26. doi - 10.1016/S0169-2046(03)00026-4

Arts, P., Boussauw, K., & Loris, I. (2014). Scenario's voor woonlocatiebeleid in Vlaanderen: criteria en doorrekening (Scenarios for the residential location policy in Flanders: criteria and computation). *Ruimte & Maatschappij*, 5(4), 8–31.

Baetens, J. M., De Loof, K., & De Baets, B. (2013). Influence of the topology of a cellular automaton on its dynamical properties. *Communications in Nonlinear Science and Numerical Simulation*, 18(3), 651–668. doi - 10.1016/j.cnsns.2012.08.018

Bak, P. (1994). Self-Organized Criticality: A Holistic View of Nature. In G. Cowan, D. Pines, & D. Meltzer (Eds.), *Complexity: Metaphors, Models, and Reality* (pp. 447–496). Reading, MA, USA: Addison-Wesley.

Balmer, M. (2007). *Travel demand modeling for multi-agent transport simulations: Algorithms and systems*. Doctoral dissertation. Zurich: ETH Zurich. doi - 10.3929/ethz-a-005429518

Barredo, J. I., Kasanko, M., McCormick, N., & Lavallo, C. (2003). Modelling dynamic spatial processes: simulation of urban future scenarios through cellular automata. *Landscape and Urban Planning*, 64(3), 145–160. doi - 10.1016/S0169-2046(02)00218-9

Barreira-González, P., Gómez-Delgado, M., & Aguilera-Benavente, F. (2015). From raster to vector cellular automata models: A new approach to simulate urban growth with the help of graph theory. *Computers, Environment and Urban Systems*, 54, 119–131. doi - 10.1016/j.compenvurbsys.2015.07.004

Bart, I. L. (2010). Urban sprawl and climate change: A statistical exploration of cause and effect, with policy options for the EU. *Land Use Policy*, 27(2), 283–292. doi - 10.1016/j.landusepol.2009.03.003

Barthelemy, M. (2016). *The Structure and Dynamics of Cities*. Cambridge, UK: Cambridge University Press.

Basse, R. M., Omrani, H., Charif, O., Gerber, P., & Bódis, K. (2014). Land use changes modelling using advanced methods: Cellular automata and artificial neural networks. The spatial and explicit representation of land cover dynamics at the cross-border region scale. *Applied Geography*, 53, 160–171. doi - 10.1016/j.apgeog.2014.06.016

Batista e Silva, F., Gallego, J., & Lavallo, C. (2013a). A high-resolution population grid map for Europe. *Journal of Maps*, 9(1), 16–28. doi - 10.1080/17445647.2013.764830

Batista e Silva, F., Lavallo, C., Jacobs-Crisioni, C., Barranco, R., Zulian, G., Maes, J., et al. (2013b). *Direct and indirect land use impacts of the EU cohesion policy. Assessment with the land use modelling platform*. Luxembourg: Publication Office of the European Union. doi - 10.2788/60631

Batty, M. (1976). *Urban Modelling: Algorithms, Calibrations, Predictions*. Cambridge, UK: Cambridge University Press.

Batty, M. (2005). *Cities and complexity. Understanding cities with cellular automata, agent-based models, and fractals*. Cambridge, MA, USA: The MIT Press.

- Batty, M., Besussi, E., & Chin, N. (2003). Traffic, urban growth and suburban sprawl. *CASA Working Papers*, 70. London, UK: Centre for Advanced Spatial Analysis, UCL. Retrieved from <http://eprints.ucl.ac.uk/216/>.
- Batty, M., & Longley, P. (1986). The fractal simulation of urban structure. *Environment and Planning A*, 18(9), 1143–1179. doi - 10.1068/a181143
- Batty, M., & Longley, P. (1987). Fractal-based description of urban form. *Environment and Planning B: Planning and Design*, 14(2), 123–134. doi - 10.1068/b140123
- Batty, M., & Xie, Y. (1994). From cells to cities. *Environment and Planning B: Planning and Design*, 21(7), S31-S48. doi - 10.1068/b21S031
- Benenson, I. (1998). Multi-agent simulations of residential dynamics in the city. *Computers, Environment and Urban Systems*, 22(1), 25–42. doi - 10.1016/S0198-9715(98)00017-9
- Benenson, I., Aronovich, S., & Noam, S. (2005). Let's talk objects: generic methodology for urban high-resolution simulation. *Computers, Environment and Urban Systems*, 29(4), 425–453. doi - 10.1016/j.compenvurbsys.2003.11.008
- Benenson, I., Omer, I., & Hatna, E. (2002). Entity-based modeling of urban residential dynamics: the case of Yaffo, Tel Aviv. *Environment and Planning B: Planning and Design*, 29(4), 491–512. doi - 10.1068/b1287
- Bengston, D. N., Fletcher, J. O., & Nelson, K. C. (2004). Public policies for managing urban growth and protecting open space: policy instruments and lessons learned in the United States. *Landscape and Urban Planning*, 69(2), 271-286. doi - 10.1016/j.landurbplan.2003.08.007
- Berens, W. (1988). The suitability of the weighted lp-norm in estimating actual road distances. *European Journal of Operational Research*, 34(1), 39–43. doi - 10.1016/0377-2217(88)90453-5
- Blečić, I., Cecchini, A., & Trunfio, G. A. (2011). Modelling proximal space in urban cellular automata. In B. Murgante, et al. (Eds.), *Proceedings of the*

11th International Conference on Computational Science and Its Applications (ICCSA 2011), Part I, 20–23 June 2011, Santander, Spain, (pp. 477–491).

Blečić, I., Cecchini, A., & Trunfio, G. A. (2015). How much past to see the future: a computational study in calibrating urban cellular automata. *International Journal Of Geographical Information Science*, 29(3), 349–374. doi - 10.1080/13658816.2014.970190

Boussauw, K., Allaert, G., & Witlox, F. (2013). Colouring inside what lines? Interference of the urban growth boundary and the political-administrative border of Brussels. *European Planning Studies*, 21(10), 1509–1527. doi - 10.1080/09654313.2012.722952

Boussauw, K., & Boelens, L. (2015). Fuzzy tales for hard blueprints: the selective coproduction of the Spatial Policy Plan for Flanders, Belgium. *Environment and Planning C: Government and Policy*, 33, 33(6), 1376–1393. doi - 10.1068/c12327

Boussauw, K., Derudder, B., & Witlox, F. (2011). Measuring spatial separation processes through the minimum commute: the case of Flanders. *European Journal of Transport and Infrastructure Research*, 11(1), 42–60. Retrieved from http://www.ejtir.tudelft.nl/issues/2011_01/pdf/2011_01_02.pdf

Boussauw, K., Neutens, T., & Witlox, F. (2012). Relationship between spatial proximity and travel-to-work distance: the effect of the compact city. *Regional Studies*, 46(6), 687–706. doi - 10.1080/00343404.2010.522986

Briassoulis, H. (2000). *Analysis of Land Use Change: Theoretical and Modeling Approaches*. Morgantown, WV, USA: Regional Research Institute, West Virginia University. Available at <http://www.rri.wvu.edu/webbook/briassoulis/contents.htm>

Brimberg, J., Walker, J. H., & Love, R. F. (2007). Estimation of travel distances with the weighted l_p norm: some empirical results. *Journal of Transport Geography*, 15(1), 62–72. doi - 10.1016/j.jtrangeo.2006.01.004

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(2), 153–174. doi - 10.1080/13658810410001713399

Brown, D. G., & Robinson, D. T. (2006). Effects of heterogeneity in residential preferences on an agent-based model of urban sprawl. *Ecology and Society*, 11(1), 46. Retrieved from <http://www.ecologyandsociety.org/vol11/iss1/art46/>

Brueckner, J. K. (2000). Urban sprawl: diagnosis and remedies. *International Regional Science Review*, 23(2), 160-171. doi - 10.1177/016001700761012710

Camagni, R., Gibelli, M. C., & Rigamonti, P. (2002). Urban mobility and urban form: the social and environmental costs of different patterns of urban expansion. *Ecological Economics*, 40(2), 199–216. doi - 10.1016/S0921-8009(01)00254-3

Cao, K., Batty, M., Huang, B., Liu, Y., Yu, L., & Chen, J. (2011). Spatial multi-objective land use optimization: extensions to the non-dominated sorting genetic algorithm-II. *International Journal of Geographical Information Science*, 25(12), 1949–1969. doi - 10.1080/13658816.2011.570269

Caparros-Midwood, D., Barr, S., & Dawson, R. (2015). Optimised spatial planning to meet long term urban sustainability objectives. *Computers, Environment and Urban Systems*, 54, 154–164. doi - 10.1016/j.compenvurbsys.2015.08.003

Cecchini, A., & Rizzi, P. (2001). The reasons why cellular automata are a useful tool in the working-kit for the new millennium urban planner in governing the territory. In *Proceedings from the 7th International Conference on Computers in Urban Planning and Urban Management (CUPUM 2001)*. Honolulu, HI, USA: University of Hawaii.

Chang, J. S. (2006). Models of the relationship between transport and land-use: a review. *Transport Reviews*, 26(3), 325–350. doi - 10.1080/01441640500468432

Chaudhuri, G., & Clarke, K. C. (2013). The SLEUTH land use change model: A review. *International Journal of Environmental Resources Research*, 1(1), 88–104.

Christaller, W. (1966 [1933]). *Central Places in Southern Germany*. Translated by C. Baskin. Eaglewood Cliffs, NJ, USA: Prentice Hall.

Claessens, L., Schoorl, J. M., Verburg, P. H., Geraedts, L., & Veldkamp, A. (2009). Modelling interactions and feedback mechanisms between land use change and landscape processes. *Agriculture, Ecosystems & Environment*, 129(1–3), 157–170. doi - 10.1016/j.agee.2008.08.008

Clarke, K. C., Gazulis, N., Dietzel, C. K., & Goldstein, N. C. (2007). A Decade of SLEUTHing: Lessons Learned from Applications of a Cellular Automaton Land Use Change Model. In P. Fisher (Ed.), *Classics from IJGIS. Twenty Years of the International Journal of Geographical Information Systems and Science* (pp. 413–425). Boca Raton, FL, USA: Taylor and Francis.

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(2), 247–261. doi - 10.1068/b240247

Clarke-Lauer, M. D., & Clarke, K. C. (2011). Evolving simulation modeling: Calibrating SLEUTH using a genetic algorithm. In *Proceedings of the 11th International Conference on GeoComputation*, July 20–22, 2011, London, UK. <http://www.geocomputation.org/2011/papers/clarke-lauer.pdf>

Cleland, J. (2013). World population growth; past, present and future. *Environmental and Resource Economics*, 55(4), 543–554. doi - 10.1007/s10640-013-9675-6

Cockx, K., & Canters, F. (2015). Incorporating spatial non-stationarity to improve dasymetric mapping of population. *Applied Geography*, 63, 220–230. doi - 10.1016/j.apgeog.2015.07.002

Cockx, K., Van de Voorde, T., & Canters, F. (2014). Quantifying uncertainty in remote sensing-based urban land-use mapping. *International*

Journal of Applied Earth Observation and Geoinformation, 31, 154–166.
doi - 10.1016/j.jag.2014.03.016

Coelho, J., & Wilson, A. (1976). The optimum location and size of shopping centres. *Regional Studies*, 10(4), 413–421.
doi - 10.1080/09595237600185441

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46.

Colonna, A., Di Stefano, V., Lombardo, S. T., Papini, L., & Rabino, G. A. (1998). L.A.U.D.E.: Learning Automata for Urban Development Exploration. The case study of Rome urban system. In *Proceedings of the 38th Congress of the European Regional Science Association: Europe Quo Vadis? – Regional Questions at the Turn of the Century*, August 28-September 1, 1998, Vienna, Austria. Retrieved from https://www.econstor.eu/bitstream/10419/113551/1/ERSA1998_302.pdf

Costanza, R. (1989). Model goodness of fit: a multiple resolution procedure. *Ecological Modelling*, 47(3), 199–215.
doi - 10.1016/0304-3800(89)90001-X

Couclelis, H. (1985). Cellular worlds: a framework for modeling micro-macro dynamics. *Environment and Planning A*, 17(5): 585–596.
doi - 10.1068/a170585

Couclelis, H. (1997). From cellular automata to urban models: new principles for model development and implementation. *Environment and Planning B: Planning and Design*, 24(2), 165–174. doi - 10.1068/b240165

Couclelis, H. (2002). Modelling frameworks, paradigms, and approaches. In K. C. Clarke, B. E. Parks, & M. P. Crane (Eds.), *Geographic Information Systems and Environmental Modelling* (pp. 36–50). London: Prentice Hall.

CPB, MNP, & RPB (2006). *Welfare, Prosperity and Quality of the Living Environment*. Retrieved from http://www.welvaartenleefomgeving.nl/context_UK.html [Accessed 23 January 2017].

Crane, R. (1996). The influence of uncertain job location on urban form and the journey to work. *Journal of Urban Economics*, 39(3), 342-356. doi - 10.1006/juec.1996.0018

Crols, T., Vanderhaegen, S., Canters, F., Engelen, G., Poelmans, L., Uljee, I., & White, R. (2017). Dated high-resolution population density maps using sealed surface cover time series. *Landscape and Urban Planning*, 160, 96–106. doi - 10.1016/j.landurbplan.2016.12.009

Crols, T., White, R., Uljee, I., Engelen, G., Canters, F., & Poelmans, L. (2012a). Development of an activity-based cellular automata land-use model: the case of Flanders, Belgium. In R. Seppelt, A. A. Voinov, S. Lange, & D. Bankamp (Eds.), *Managing Resources of a Limited Planet: Pathways and Visions under Uncertainty, Proceedings of the 6th International Congress on Environmental Modelling and Software (iEMSs 2012)*, July 1-5, 2012, Leipzig, Germany, (pp. 2000–2007). http://www.iemss.org/sites/iemss2012//proceedings/F1_0580_Crols_et_al.pdf

Crols, T., White, R., Uljee, I., Engelen, G., Canters, F., & Poelmans, L. (2012b). An activity-based CA model for Flanders, Belgium. Robustness analysis and improved distance computation in a variable grid representation. In: N. N. Pinto, J. Dourado, & A. Natálio (Eds.), *Proceedings of CAMUSS, the International Symposium on Cellular Automata Modeling for Urban and Spatial Systems*, November 8-10, 2012, Porto, Portugal, (pp. 195–207). Coimbra: Department of Civil Engineering of the University of Coimbra.

Crols, T., White, R., Uljee, I., Engelen, G., Poelmans, L., & Canters, F. (2015). A travel time-based variable grid approach for an activity-based cellular automata model. *International Journal of Geographical Information Science*, 29(10), 1757–1781. doi - 10.1080/13658816.2015.1047838

Crooks, A. T., & Heppenstall, A. J. (2012). Introduction to agent-based modelling. In A. J. Heppenstall, A. T. Crooks, L. M. See, & M. Batty (Eds.), *Agent-based models of geographical systems* (pp. 85–105). Dordrecht, Netherlands: Springer.

Datta, D., Deb, K., Fonseca, C. M., Lobo, F., Condado, P., & Seixas, J. (2007). Multi-objective evolutionary algorithm for land-use management

problem. *International Journal of Computational Intelligence Research*, 3(4), 371–384.

De Block, G., & Polasky, J. (2011). Light railways and the rural-urban continuum: technology, space and society in late nineteenth-century Belgium. *Journal of Historical Geography*, 37(3), 312-328.
doi -10.1016/j.jhge.2011.01.003

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2), 182–197. doi - 10.1109/4235.996017

Deboosere, P., Lesthaeghe, R., Surkyn, J., Willaert, D., Boulanger, P.-M., Lambert, A., et al. (2009). Huishoudens en gezinnen in België (Households and families in Belgium). In FOD Economie (Ed.), *Sociaal-Economische Enquête 2001. Monografieën (Socio-Economic Survey 2001. Monographs)*, nr 4/2009. Brussels: Algemene Directie Statistiek en Economische Informatie. Retrieved from http://statbel.fgov.be/nl/binaries/mono_200104_nl%5B1%5D_tcm325-92942.pdf

De Decker, P. (2008). Facets of housing and housing policies in Belgium. *Journal of Housing and the Built Environment*, 23(3), 155–171.
doi - 10.1007/s10901-008-9110-4

De Decker, P. (2011). Understanding housing sprawl: the case of Flanders, Belgium. *Environment and Planning A*, 43(7), 1634–1654.
doi - 10.1068/a43242

de Kok, J.-L., Overloop, S., & Engelen, G. (2017). Screening models for integrated environmental planning – A feasibility study for Flanders. *Futures*, 88, 55–68. doi - 10.1016/j.futures.2017.03.007

de Kok, J.-L., Poelmans, L., Engelen, G., Uljee, I., & Van Esch, L. (2012). Spatial dynamic visualization of long-term scenarios for demographic, social-economic and environmental change in Flanders. In R. Seppelt, A. A. Voinov, S. Lange, & D. Bankamp (Eds.), *Managing Resources of a Limited Planet: Pathways and Visions under Uncertainty, Proceedings of the 6th International Congress on Environmental Modelling and Software (iEMSs 2012)*, July 1-5, 2012, Leipzig, Germany, (pp. 1984–1991).

http://www.iemss.org/sites/iemss2012//proceedings/F1_0400_deKok_et_al.pdf

De Meulder, B., Schreurs, J., Cock, A., & Notteboom, B. (1999). Sleutelen aan het Belgische stadslandschap/Patching up the Belgian urban landscape. *Oase: Journal for Architecture*, 52, 79–113.

de Nijs, T. C. M. (2009). Future Land-use in The Netherlands: Evaluation of the National Spatial Strategy. In S. Geertman, & J. Stillwell (Eds.), *Planning Support Systems. Best Practice and New Methods* (pp. 53–67). Dordrecht, Netherlands: Springer.

Department LNE (2016). *De Vlaamse Klimaattop*. Brussels: Flemish Department of the Environment, Nature and Energy. Retrieved from <http://vlaamseklimaattop.be/> [Accessed 25 January 2017].

Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F. R., Gaughan, A. E., et al. (2014). Dynamic population mapping using mobile phone data. *Proceedings of the National Academy of Sciences*, 111(45), 15888–15893. doi - 10.1073/pnas.1408439111

Dieleman, F., & Wegener, M. (2004). Compact city and urban sprawl. *Built Environment*, 30(4), 308-323. doi - 10.2148/benv.30.4.308.57151

Dietzel, C., & Clarke, K. C. (2007). Toward optimal calibration of the SLEUTH land use change model. *Transactions in GIS*, 11(1), 29–45. doi - 10.1111/j.1467-9671.2007.01031.x

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1), 269–271. doi - 10.1007/BF01386390

Douglass, R. W., Meyer, D. A., Ram, M., Rideout, D., & Song, D. (2015). High resolution population estimates from telecommunications data. *EPJ Data Science*, 4, 4. doi - 10.1140/epjds/s13688-015-0040-6

Echenique, M. (2004). Econometric models of land use and transportation. In D.A. Hensher, et al. (Eds.), *Handbook of Transport Geography and Spatial Systems* (pp. 185–202). Kidlington, UK: Pergamon/Elsevier Science.

Eicher, C. L., & Brewer, C. A. (2001). Dasymetric mapping and areal interpolation: Implementation and evaluation. *Cartography and Geographic Information Science*, 28(2), 125–138. doi - 10.1559/152304001782173727

Engelen, G. (1988). The theory of self-organization and modelling complex urban systems. *European Journal of Operational Research*, 37(1), 42–57. doi - 10.1016/0377-2217(88)90279-2

Engelen, G. (2003). Development of a decision support system for the integrated assessment of policies related to desertification and land degradation in the Mediterranean. In C. Giupponi, & M. Shechter (Eds.), *Climate Change in the Mediterranean: Socio-economic Perspectives of Impacts, Vulnerability and Adaptation* (pp. 159–195). Cheltenham, UK, & Northampton, MA, USA: Edward Elgar Publishing.

Engelen, G., Lavallo, C., Barredo, J. I., van der Meulen, M., & White, R. (2007). The MOLAND modelling framework for urban and regional land use dynamics. In E. Koomen, J. Stillwell, A. Bakema, & H. J. Scholten (Eds.), *Modelling Land-Use Change: Progress and Applications* (pp. 297–320). Dordrecht, Netherlands: Springer.

Engelen, G., & White, R. (2007). Validating and calibrating integrated cellular automata based models of land use change. In S. Albeverio, D. Andrey, P. Giordano, & A. Vancheri (Eds.), *The Dynamics of Complex Urban Systems* (pp. 185–211). Heidelberg: Physica-Verlag. doi - 10.1007/978-3-7908-1937-3_10

Engelen, G., White, R., & de Nijs, T. (2003). Environment Explorer: spatial support system for the integrated assessment of socio-economic and environmental policies in the Netherlands. *Integrated Assessment*, 4(2), 97–105. doi - 10.1076/iaij.4.2.97.16707

Engelen, G., White, R., Uljee, I., & Drazen, P. (1995). Using cellular automata for integrated modelling of socio-environmental systems. *Environmental Monitoring and Assessment*, 34(2), 203–214. doi - 10.1007/BF00546036

European Commission (2014). *Investment for jobs and growth. Promoting development and good governance in EU regions and cities. Sixth report on*

economic, social and territorial cohesion. Luxembourg: Publication Office of the European Union. doi - 10.2776/81072

European Commission (2016). Future brief: no net land take by 2050? *Science for Environment Policy*, 14. Retrieved from <http://www.inspiration-h2020.eu/actualites/no-net-land-take-2050-new-future-brief-published-science-environment-policy> [Accessed 14 April 2017].

European Environment Agency (2016). *Urban sprawl in Europe - Joint EEA-FOEN report*. EEA Report No 11/2016. Luxembourg: Publications Office of the European Union. Retrieved from <http://www.eea.europa.eu/publications/urban-sprawl-in-europe>

Ewing, R. H. (2008). Characteristics, causes, and effects of sprawl: a literature review. In J. M. Marzluff, E. Shulenberger, W. Endlicher, M. Alberti, G. Bradley, C. Ryan, et al. (Eds.), *Urban Ecology. An International Perspective on the Interaction Between Humans and Nature* (pp. 519-535). Springer US. doi - 10.1007/978-0-387-73412-5_34

Feng, Y., Liu, Y., Tong, X., Liu, M., & Deng, S. (2011). Modeling dynamic urban growth using cellular automata and particle swarm optimization rules. *Landscape and Urban Planning*, 102(3), 188–196. doi - 10.1016/j.landurbplan.2011.04.004

Fertner, C., Jørgensen, G., Nielsen, T. A. S., & Nilsson, K. S. B. (2016). Urban sprawl and growth management – drivers, impacts and responses in selected European and US cities. *Future Cities and Environment*, 2, 9. doi - 10.1186/s40984-016-0022-2

FOD Economie (2011). *In the spotlight, 6 December 2011. Small and less small households*. Brussels: Statistics Belgium. Retrieved from http://statbel.fgov.be/en/statistics/organisation/statistics_belgium/dissemination/statbel/in_the_spotlight_archives/in_the_spotlight_2011/20111206_small_and_less_small_households.jsp

FOD Economie (2015). *Bodembezetting volgens het Kadaster (1982-2015) (Land occupation according to the Cadastre (1982-2015))*. Brussels: Federal Department of Economics, SMEs, Retail and Energy. Retrieved from <http://statbel.fgov.be/nl/modules/publications/statistiques/environnement/do>

[wloadbare bestanden/bodembezetting volgens het Kadaster \(1982 - 2015\).jsp](#)

Forrester, J. (1969). *Urban Dynamics*. Cambridge, MA, USA: MIT Press.

FPB, & ADS (2016). *Demografische vooruitzichten 2015-2060. Bevolking, huishoudens en prospectieve sterftequotiënten (Demographic outlook 2015-2060. Population, households and mortality ratios)*. Brussels: Federal Planning Bureau, & Statistics Belgium. Retrieved from http://www.plan.be/admin/uploaded/201603111541350.FOR_POP1560_11178_N.pdf [Accessed 8 March 2017].

FPB, & FOD Mobiliteit (2015). *Vooruitzichten van de transportvraag in België tegen 2030 (Outlook of the transport demand in Belgium until 2030)*. Brussels: Federal Planning Bureau, & Federal Department of Mobility. Retrieved from http://www.plan.be/admin/uploaded/201512081036570.For_Transport_1230_11128_N.pdf [Accessed 9 March 2017].

Frankhauser, P. (1988). Fractal aspects of urban structures. In *Proceedings of the International Symposium of the Sonderforschungsbereich*, 230. Stuttgart: University of Stuttgart.

Frankhauser, P. (1994). *La fractilité des structures urbaines*. Paris: Anthropos.

Fricke, R., & Wolff, E. (2002). The MURBANDY project: development of land use and network databases for the Brussels area (Belgium) using remote sensing and aerial photography. *International Journal of Applied Earth Observation and Geoinformation*, 4(1), 33-50. doi - 10.1016/S0303-2434(02)00010-7

Fuglsang, M., Hansen, H. S., & Münier, B. (2011). Accessibility analysis and modelling in public transport networks – a raster based approach. In B. Murgante, et al. (Eds.), *Proceedings of the 11th International Conference on Computational Science and its Applications (ICCSA 2011)*, Part I, 20–23 June 2011, (pp. 207–224). Santander, Spain. Berlin: Springer-Verlag.

Fujita, M. (1989). *Urban economic theory: land use and city size*. Cambridge, UK: Cambridge University Press.

- Gallego, F. J., Batista, F., Rocha, C., & Mubareka, S. (2011). Disaggregating population density of the European Union with CORINE land cover. *International Journal of Geographical Information Science*, 25(12), 2051–2069. doi - 10.1080/13658816.2011.583653
- Gardner, M. (1970). Mathematical games: The fantastic combinations of John Conway's new solitaire game "life". *Scientific American*, 223(4), 120–123.
- Geertman, S., Hagoort, M., & Ottens, H. (2007). Spatial-temporal specific neighbourhood rules for cellular automata land-use modelling. *International Journal of Geographical Information Science*, 21(5), 547–568. doi - 10.1080/13658810601064892
- Giuliano, G., & Dargay, J. (2006). Car ownership, travel and land use: a comparison of the US and Great Britain. *Transportation Research Part A*, 40, 106–124. doi - 10.1016/j.tra.2005.03.002
- Giuliano, G., Gordon, P., Pan, Q., & Park, J. Y. (2010). Accessibility and residential land values: some tests with new measures. *Urban Studies*, 47(14), 3103–3130. doi - 10.1177/0042098009359949
- Gobin, A., Uljee, I., Van Esch, L., Engelen, G., de Kok, J.-L., Hens, M., et al. (2009). *Landgebruik in Vlaanderen*. Scientific report, MIRA 2009 and NARA 2009, VMM, INBO. R.2009.20, www.milieurapport.be, www.nara.be.
- Goerlich, F. J., & Cantarino, I. (2013). A population density grid for Spain. *International Journal of Geographical Information Science*, 27(12), 2247–2263. doi - 10.1080/13658816.2013.799283
- Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Boston, MA, USA: Addison-Wesley.
- Goldstein, N. C. (2004). Brains versus brawn — Comparative strategies for the calibration of a cellular automata-based urban growth model. In P. M. Atkinson, F. Wu, G. M. Foody, & S. E. Darby (Eds.), *GeoDynamics* (pp. 249–272). Boca Raton, FL, USA: CRC Press.

Grinblat, Y., Gilichinsky, M., & Benenson, I. (2016). Cellular Automata Modeling of Land-Use/Land-Cover Dynamics: Questioning the Reliability of Data Sources and Classification Methods. *Annals of the American Association of Geographers*, 106(6), 1299–1320.
doi - 10.1080/24694452.2016.1213154

Guan, Q., Wang, L., & Clarke, K. C. (2005). An artificial-neural-network-based, constrained CA model for simulating urban growth. *Cartography and Geographic Information Science*, 32(4), 369–380.
doi - 10.1559/152304005775194746

Haase, D., & Schwarz, N. (2009). Simulation models on human-nature interactions in urban landscapes: a review including spatial economics, system dynamics, cellular automata and agent-based approaches. *Living Reviews in Landscape Research*, 3(2). Retrieved from <http://lrlr.landscapeonline.de/Articles/lrlr-2009-2/>.

Hagen, A. (2003). Fuzzy set approach to assessing similarity of categorical maps. *International Journal Of Geographical Information Science*, 17(3), 235–249. doi - 10.1080/13658810210157822

Hagen-Zanker, A. (2009). An improved Fuzzy Kappa statistic that accounts for spatial autocorrelation. *International Journal Of Geographical Information Science*, 23(1), 61–73. doi - 10.1080/13658810802570317

Hagen-Zanker, A. (2012). A comparison of three urban models of land use, transport and activity. In N. N. Pinto, J. Dourado, & A. Natálio (Eds.), *Proceedings of the International Symposium on Cellular Automata Modeling for Urban and Spatial Systems (CAMUSS)*, November 8–10, 2012, Porto, Portugal, (pp. 337–338). Coimbra: Department of Civil Engineering of the University of Coimbra.

Hagen-Zanker, A., & Jin, Y. (2012). A new method of adaptive zoning for spatial interaction models. *Geographical Analysis*, 44(4), 281–301.
doi - 10.1111/j.1538-4632.2012.00855.x

Hagen-Zanker, A., & Jin, Y. (2013). Adaptive zoning for transport mode choice modelling. *Transactions in GIS*, 17(5), 706–723.
doi - 10.1111/j.1467-9671.2012.01372.x

- Hagen-Zanker, A., & Martens, P. (2008). Map comparison methods for comprehensive assessment of geosimulation models. In O. Gervasi, et al. (Eds.), *Proceedings of the International Conference on Computational Science and its Applications – ICCSA 2008*, Part I, 30 June – 3 July 2008, Perugia, Italy, (pp. 194-209). Berlin/Heidelberg: Springer-Verlag.
- Halleux, J.-M., Marcinczak, S., & van der Krabben, E. (2012). The adaptive efficiency of land use planning measured by the control of urban sprawl. The cases of the Netherlands, Belgium and Poland. *Land Use Policy*, 29(4), 887–898. doi - 10.1016/j.landusepol.2012.01.008
- Handy, S., Cao, X., & Mokhtarian, P. (2005). Correlation or causality between the built environment and travel behavior? Evidence from Northern California. *Transportation Research Part D: Transport and Environment*, 10(6), 427–444. doi - 10.1016/j.trd.2005.05.002
- Hansen, H. S. (2009). Analysing the role of accessibility in contemporary urban development. In O. Gervasi, et al. (Eds.), *Proceedings of the International Conference on Computational Science and its Applications – ICCSA 2009*, Part I, 29 June–2 July 2009, Seoul, Korea, (pp. 385–396). Berlin: Springer-Verlag.
- Hansen, H. S. (2010). Modelling the future coastal zone urban development as implied by the IPCC SRES and assessing the impact from sea level rise. *Landscape and Urban Planning*, 98(3–4), 141–149. doi - 10.1016/j.landurbplan.2010.08.018
- Hassan, A. M., & Lee, H. (2015). Toward the sustainable development of urban areas: An overview of global trends in trials and policies. *Land Use Policy*, 48, 199–212. doi - 10.1016/j.landusepol.2015.04.029
- Herold, M., Couclelis, H., & Clarke, K. C. (2005). The role of spatial metrics in the analysis and modeling of urban land use change. *Computers, Environment and Urban Systems*, 29(4), 369–399. doi - 10.1016/j.compenvurbsys.2003.12.001
- Herold, M., Liu, X., & Clarke, K. C. (2003). Spatial metrics and image texture for mapping urban land use. *Photogrammetric Engineering & Remote Sensing*, 69(9), 991–1001. doi - 10.14358/PERS.69.9.991

Hertogen, J. (2013). *Bevolkingsevolutie 1831-2012 in gemeenten, arr., prov. en gewesten (Evolution of population 1831-2012 in municipalities, arrondissements, provinces and regions)*. Non-profit data (npdata.be). Retrieved from <http://www.npdata.be/BuG/190-Bevolking-gemeente-1831-2012/>

Holland, J. H. (1975). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. Ann Arbor, MI, USA: University of Michigan Press.

Hoornaert, S. (Ed.) (2014). *Verkeersindicatoren hoofdwegennet Vlaanderen 2013*. Antwerp: Flemish Traffic Study Centre (Vlaams Verkeerscentrum). <http://www.verkeerscentrum.be/pdf/rapport-verkeersindicatoren-2013-v1.pdf>

Huang, J., Lu, X. X., & Sellers, J. M. (2007). A global comparative analysis of urban form: Applying spatial metrics and remote sensing. *Landscape and Urban Planning*, 82(4), 184–197. doi - 10.1016/j.landurbplan.2007.02.010

Huang, Q., Parker, D. C., Filatova, T., & Sun, S. (2014). A review of urban residential choice models using agent-based modeling. *Environment and Planning B: Planning and Design*, 41(4), 661–689. doi - 10.1068/b120043p

Iacono, M., Levinson, D., & El-Geneidy, A. (2008). Models of transportation and land use change: a guide to the territory. *Journal of Planning Literature*, 22(4), 323–340. doi - 10.1177/0885412207314010

Jackson, L. E. (2003). The relationship of urban design to human health and condition. *Landscape and Urban Planning*, 64(4), 191–200. doi - 10.1016/S0169-2046(02)00230-X

Jia, P., & Gaughan, A. E. (2016). Dasymeric modeling: A hybrid approach using land cover and tax parcel data for mapping population in Alachua County, Florida. *Applied Geography*, 66, 100–108. doi - 10.1016/j.apgeog.2015.11.006

Jia, P., Qiu, Y., & Gaughan, A. E. (2014). A fine-scale spatial population distribution on the High-resolution Gridded Population Surface and application in Alachua County, Florida. *Applied Geography*, 50, 99–107. doi - 10.1016/j.apgeog.2014.02.009

Kasanko, M., Barredo, J. I., Lavalle, C., McCormick, N., Demicheli, L., Sagris, V., et al. (2006). Are European cities becoming dispersed? A comparative analysis of 15 European urban areas. *Landscape and Urban Planning*, 77(1-2), 111-130. doi - 10.1016/j.landurbplan.2005.02.003

Kauffman, S. (1993). *The Origins of Order*. New York: Oxford University Press.

Keil, R. (Ed.) (2013). *Suburban constellations: governance, land and infrastructure in the 21st century*. Berlin: Jovis.

Kennedy, W. G. (2012). Modelling human behaviour in agent-based models. In A. J. Heppenstall, A. T. Crooks, L. M. See, & M. Batty (Eds.), *Agent-based models of geographical systems* (pp. 167–179). Dordrecht, Netherlands: Springer.

Koomen, E., & Stillwell, J. (2007). Modelling land-use change: Theories and methods. In E. Koomen, J. Stillwell, A. Bakema, & H. J. Scholten (Eds.), *Modelling Land-Use Change: Progress and Applications* (pp. 1–21). Dordrecht, Netherlands: Springer.

Krugman, P. (1998). What's new about the new economic geography? *Oxford Review of Economic Policy*, 14(2), 7–17. doi - 10.1093/oxrep/14.2.7

Langford, M., Maguire, D. J., & Unwin, D. J. (1991). The areal interpolation problem: estimating population using remote sensing in a GIS framework. In I. Masser, & M. Blakemore (Eds.), *Handling Geographical Information: Methodology and Potential Applications* (pp. 55–77). New York: Wiley.

Langton, C. (1992). Life at the Edge of Chaos. In C. Langton, C. Taylor, D. Farmer, & S. Rasmussen (Eds.), *Artificial Life II* (pp. 41–91). Redwood City, CA, USA: Addison-Wesley.

Lauf, S., Haase, D., Hostert, P., Lakes, T., & Kleinschmit, B. (2012). Uncovering land-use dynamics driven by human decision-making – A combined model approach using cellular automata and system dynamics. *Environmental Modelling & Software*, 27-28, 71–82. doi - 10.1016/j.envsoft.2011.09.005

- Li, X., Lin, J., Chen, Y., Liu, X., & Ai, B. (2013). Calibrating cellular automata based on landscape metrics by using genetic algorithms. *International Journal of Geographical Information Science*, 27(3), 594–613. doi - 10.1080/13658816.2012.698391
- Li, X., Yang, Q., & Liu, X. (2008). Discovering and evaluating urban signatures for simulating compact development using cellular automata. *Landscape and Urban Planning*, 86(2), 177–186. doi - 10.1016/j.landurbplan.2008.02.005
- Li, X., & Yeh, A. G.-O. (2002). Neural-network based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science*, 16(4), 323–343. doi - 10.1080/13658810210137004
- Li, X., & Yeh, A. G.-O. (2004). Analyzing spatial restructuring of land use patterns in a fast growing region using remote sensing and GIS. *Landscape and Urban Planning*, 69(4), 335–354. doi - 10.1016/j.landurbplan.2003.10.033
- Liao, J., Tang, L., Shao, G., Qiu, Q., Wang, C., Zheng, S., et al. (2014). A neighbor decay cellular automata approach for simulating urban expansion based on particle swarm intelligence. *International Journal of Geographical Information Science*, 28(4), 720–738. doi - 10.1080/13658816.2013.869820
- Liao, J., Tang, L., Shao, G., Su, X., Chen, D., & Xu, T. (2016). Incorporation of extended neighborhood mechanisms and its impact on urban land-use cellular automata simulations. *Environmental Modelling & Software*, 75, 163–175. doi - 10.1016/j.envsoft.2015.10.014
- Lin, Y.-P., Chu, H.-J., Wu, C.-F., & Verburg, P. H. (2011). Predictive ability of logistic regression, auto-logistic regression and neural network models in empirical land-use change modeling – a case study. *International Journal of Geographical Information Science*, 25(1), 65–87. doi - 10.1080/13658811003752332
- Litman, T. (2015). Analysis of public policies that unintentionally encourage and subsidize urban sprawl. *The New Climate Economy*, online publication, March 2015. Supporting paper on behalf of the Global Commission on the Economy and Climate for the New Climate Economy

Cities Program. Retrieved from <http://static.newclimateeconomy.report/wp-content/uploads/2015/03/public-policies-encourage-sprawl-nce-report.pdf>

Liu, X., Li, X., Chen, Y., Tan, Z., Li, S., & Ai, B. (2010a). A new landscape index for quantifying urban expansion using multi-temporal remotely sensed data. *Landscape Ecology*, 25(5), 671–682.
doi - 10.1007/s10980-010-9454-5

Liu, X., Li, X., Shi, X., Wu, S., & Liu, T. (2008). Simulating complex urban development using kernel-based non-linear cellular automata. *Ecological Modelling*, 211, 169–181. doi - 10.1016/j.ecolmodel.2007.08.024

Liu, X., Li, X., Shi, X., Zhang, X., & Chen, Y. (2010b). Simulating land-use dynamics under planning policies by integrating artificial immune systems with cellular automata. *International Journal of Geographical Information Science*, 24(5), 783–802. doi - 10.1080/13658810903270551

Liu, X., Li, X., Yeh, A. G.-O., He, J., & Tao, J. (2007). Discovery of transition rules for geographical cellular automata by using ant colony optimization. *Science in China Series D: Earth Sciences*, 50(10), 1578–1588. doi - 10.1007/s11430-007-0083-z

Liu, Y., & Phinn, S. R. (2003). Modelling urban development with cellular automata incorporating fuzzy-set approaches. *Computers, Environment and Urban Systems*, 27(6), 637–658. doi - 10.1016/S0198-9715(02)00069-8

Lösch, A. (1954 [1940]). *The Economics of Location*. Translated by W. Woglom, & W. Stolper. New Haven: Yale University Press.

Lowry, I. S. (1964). *A Model of Metropolis*. RM-4035-RC. Santa Monica, CA, USA: Rand Corporation.

Lu, D., Weng, Q., & Li, G. (2006). Residential population estimation using a remote sensing derived impervious surface approach. *International Journal of Remote Sensing*, 27(16), 3553–3570.
doi - 10.1080/01431160600617202

Mandelbrot, B. (1982). *The Fractal Geometry of Nature*. San Francisco: W. H. Freeman.

Martínez, L. M., & Viegas, J. M. (2013). A new approach to modelling distance-decay functions for accessibility assessment in transport studies. *Journal of Transport Geography*, 26, 87-96.
doi - 10.1016/j.jtrangeo.2012.08.018

Mather, P. M. (2004). *Computer processing of remotely-sensed images: An introduction*. Chichester: Wiley.

Matthews, R. B., Gilbert, N. G., Roach, A., Polhill, J. G., & Gotts, N. M. (2007). Agent-based land-use models: a review of applications. *Landscape Ecology*, 22(10), 1447–1459. doi - 10.1007/s10980-007-9135-1

McGarigal, K., Cushman, S. A., Neel, M. C., & Ene, E. (2002). *FRAGSTATS v3: Spatial pattern analysis program for categorical maps*. Computer software program produced at Amherst, MA, USA: University of Massachusetts. Retrieved from
<http://www.umass.edu/landeco/research/fragstats/fragstats.html>

Meeus, B., & De Decker, P. (2015). Staying put! A housing pathway analysis of residential stability in Belgium. *Housing Studies*, 30(7), 1116–1134. doi - 10.1080/02673037.2015.1008424

Mennis, J. (2003). Generating surface models of population using dasymetric mapping. *The Professional Geographer*, 55(1), 31–42.
doi - 10.1111/0033-0124.10042

Monserud, R. A., & Leemans, R. (1992). Comparing global vegetation maps with the Kappa statistic. *Ecological Modelling*, 62(4), 275–293.
doi - 10.1016/0304-3800(92)90003-W

Moreno, N., Ménard, A., & Marceau, D. J. (2008). VecGCA: a vector-based geographic cellular automata model allowing geometric transformations of objects. *Environment and Planning B: Planning and Design*, 35(4), 647–665. doi - 10.1068/b33093

Moreno, N., Wang, F., & Marceau, D. J. (2009). Implementation of a dynamic neighborhood in a land-use vector-based cellular automata model. *Computers, Environment and Urban Systems*, 33(1), 44-54.
doi - 10.1016/j.compenvurbsys.2008.09.008

- Muth, R. (1969). *Cities and Housing: The Spatial Pattern of Urban Residential Land Use*. Chicago: University of Chicago Press.
- Næss, P. (2010). Residential location, travel, and energy use in the Hangzhou Metropolitan Area. *Journal of Transport and Land Use*, 3(3), 27–59. doi - 10.5198/jtlu.v3i3.98
- Newman, P., & Kenworthy, J. (1989). *Cities and automobile dependence. A sourcebook*. Aldershot: Gower.
- Ngo, T. A., & See, L. (2012). Calibration and validation of agent-based models of land-cover change. In A. J. Heppenstall, A. T. Crooks, L. M. See, & M. Batty (Eds.), *Agent-based models of geographical systems* (pp. 181–197). Dordrecht, Netherlands: Springer.
- Omrani, H., Judge, V., Antoni, J. P., & Klein, O. (2016). Modelling challenge of small amounts of change: a new sampling strategy in land use science. In *Proceedings of the 2nd International Conference on Cellular Automata Modeling for Urban and Spatial Systems (CAMUSS 2016)*, September 21 – 23, 2016, Quebec, Canada, (pp. 51 – 54).
<http://www.camuss2016.ulaval.ca/>
- O'Sullivan, D. (2001). Graph-cellular automata: a generalised discrete urban and regional model. *Environment and Planning B: Planning and Design*, 28(5), 687–705. doi - 10.1068/b2707
- Oxley, T., McIntosh, B. S., Winder, N., Mulligan, M., & Engelen, G. (2004). Integrated modelling and decision-support tools: a Mediterranean example. *Environmental Modelling & Software*, 19(11), 999–1010. doi - 10.1016/j.envsoft.2003.11.003
- Parker, D. C., Brown, D. G., Filatova, T., Riolo, R., Robinson, D. T., & Sun, S. (2012). Do land markets matter? A modeling ontology and experimental design to test the effects of land markets for an agent-based model of ex-urban residential land-use change. In A. J. Heppenstall, A. T. Crooks, L. M. See, & M. Batty (Eds.), *Agent-based models of geographical systems* (pp. 525–542). Dordrecht, Netherlands: Springer.
- 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. doi - 10.1111/1467-8306.9302004

Parry, H. R., & Bithell, M. (2012). Large scale agent-based modelling: a review and guidelines for model scaling. In A. J. Heppenstall, A. T. Crooks, L. M. See, & M. Batty (Eds.), *Agent-based models of geographical systems* (pp. 271–308). Dordrecht, Netherlands: Springer.

Pauwels, C., & Andries, P. (2016). *Diagnostiek Woon-Werkverkeer 2014*. Brussels: Belgian Federal Department of Mobility and Transport. Retrieved from https://mobilit.belgium.be/sites/default/files/final_report_nl_5.0.pdf [Accessed 24 April 2017].

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(6), 553–575. doi - 10.1016/S0198-9715(01)00015-1

Pinto, N. N., & Antunes, A. P. (2010). A cellular automata model based on irregular cells: application to small urban areas. *Environment and Planning B: Planning and Design*, 37(6), 1095–1114. doi - 10.1068/b36033

Poelmans, L., & Engelen, G. (2017). De bebouwde ruimte in Vlaanderen – waar komen we van en waar gaan we naartoe? Ruimtelijke toekomstscenario's ter ondersteuning van het Beleidsplan Ruimte Vlaanderen (The built-up environment in Flanders – where from, where to? Spatial scenarios in support of the spatial policy plan). *Ruimte & Maatschappij*, 8(3), 70–77.

Poelmans, L., Van Esch, L., Janssen, L., & Engelen, G. (2016). *Landgebruiksbestand voor Vlaanderen, 2013 (Land-use dataset of Flanders, 2013)*. Report 2016/RMA/R/0846. Mol, Belgium: Flemish Institute for Technological Research.

Poelmans, L., & Van Rompaey, A. (2009). Detecting and modelling spatial patterns of urban sprawl in highly fragmented areas: a case study in the Flanders-Brussels region. *Landscape and Urban Planning*, 93(1), 10–19. doi - 10.1016/j.landurbplan.2009.05.018

- Poelmans, L., & Van Rompaey, A. (2010). Complexity and performance of urban expansion models. *Computers, Environment and Urban Systems*, 34(1), 17–27. doi - 10.1016/j.compenvurbsys.2009.06.001
- Pontius, R. G. Jr., Huffaker, D., & Denman, K. (2004). Useful techniques of validation for spatially explicit land-change models. *Ecological Modelling*, 179(4), 445–461. doi - 10.1016/j.ecolmodel.2004.05.010
- Pontius, R. G. Jr., Thontteh, O., & Chen, H. (2008). Components of information for multiple resolution comparison between maps that share a real variable. *Environmental And Ecological Statistics*, 15(2), 111–142. doi - 10.1007/s10651-007-0043-y
- Portugali, J., Benenson, I., & Omer, I. (1997). Spatial cognitive dissonance and sociospatial emergence in a self-organizing city. *Environment and Planning B: Planning and Design*, 24(2), 263–285. doi - 10.1068/b240263
- Poundstone, W. (1985). *The Recursive Universe: Cosmic Complexity and the Limits of Scientific Knowledge*. Oxford: Oxford University Press.
- Prigogine, I. (1980). *From Being to Becoming*. San Francisco: W. H. Freeman.
- Prigogine, I., & Stengers, I. (1984). *Order out of Chaos: Man's New Dialogue with Nature*. New York: Bantam Books.
- Prokop, G., Jobstmann, H., & Schönbauer, A. (2011). *Overview on best practices for limiting soil sealing and mitigating its effects in EU-27*. Technical Report 2011-050, Environment Agency Austria. Retrieved from <http://ec.europa.eu/environment/soil/sealing.htm> [Accessed 25 January 2017].
- Pumain, D. (2012). Multi-agent system modelling for urban systems: the series of SIMPOP models. In A. J. Heppenstall, A. T. Crooks, L. M. See, & M. Batty (Eds.), *Agent-based models of geographical systems* (pp. 721–738). Dordrecht, Netherlands: Springer.
- Rand, W. (2012). Business Applications and Research Questions Using Spatial Agent-Based Models. In A. J. Heppenstall, A. T. Crooks, L. M. See,

& M. Batty (Eds.), *Agent-based models of geographical systems* (pp. 463–480). Dordrecht, Netherlands: Springer.

Rand, W., Zellner, M., Page, S. E., Riolo, R., Brown, D. G., & Fernandez, L. E. (2002). The complex interaction of agents and environments: An example in urban sprawl. In C. Macal & D. Sallach, (Eds.), *Proceedings of Agent 2002: Social Agents: Ecology, Exchange, and Evolution* (pp. 149–161). Chicago: University of Chicago and Argonne National Lab.

Ravetz, J., Fertner, C., & Nielsen, T. S. (2013). The dynamics of peri-urbanization. In K. Nilsson, S. Pauleit, S. Bell, C. Aalbers, & T. A. S. Nielsen (Eds.), *Peri-urban futures: scenarios and models for land use change in Europe* (pp. 13–44). Berlin/Heidelberg: Springer-Verlag.
doi - 10.1007/978-3-642-30529-0_2

Renders, N., Aernouts, K., Cornelis, E., Moorkens, I., Uljee, I., Van Esch, L., et al. (2015). *Warmte in Vlaanderen. Studie uitgevoerd in opdracht van Vlaams Energie Agentschap, VEA (Heating in Flanders. Study commissioned by the Flemish Energy Agency, VEA)*. Mol, Belgium: Flemish Institute for Technological Research. Retrieved from <http://www.energiesparen.be/warmtekaart> [Accessed 12 April 2017].

Riahi, K., Kriegler, E., Johnson, N., Bertram, C., den Elzen, M., Eom, J., et al. (2015). Locked into Copenhagen pledges—implications of short-term emission targets for the cost and feasibility of long-term climate goals. *Technological Forecasting and Social Change*, 90, 8–23.
doi - 10.1016/j.techfore.2013.09.016

Ricardo, D. (2001 [1821]). *On the principles of political economy and taxation*, 3rd Edition. Ontario, Canada: Batoche Books.

Richardson, H. W. (1995). Economies and diseconomies of agglomeration. In H. Giersch (Ed.), *Urban agglomeration and economic growth* (pp. 123–155). Berlin/Heidelberg: Springer-Verlag.
doi - 10.1007/978-3-642-79397-4_6

Rogelj, J., den Elzen, M., Höhne, N., Fransen, T., Fekete, H., Winkler, H., et al. (2016). Paris Agreement climate proposals need a boost to keep warming well below 2 C. *Nature*, 534, 631–639. doi - 10.1038/nature18307

Ruimte Vlaanderen (2016). *Witboek BRV. Samen aan de slag om Vlaanderen te transformeren – een opstap naar een volwaardig omgevingsbeleid (White Paper BRV. Transforming Flanders together – a stepping stone towards a comprehensive environment policy)*. Brussels: Ruimte Vlaanderen. Retrieved from <https://www.ruimtelijkeordening.be/BRV> [Accessed 24 January 2017].

Ruimte Vlaanderen (2017). *Working together on the space of tomorrow, a brochure to the White Paper on the Spatial Policy Plan for Flanders*. Brussels: Ruimte Vlaanderen. Retrieved from <https://www.ruimtevlaanderen.be/NL/Diensten/Publicaties/articleType/ArticleView/articleId/9079> [Accessed 14 April 2017].

Saey, P., Kesteloot, C., & Vandermotten, C. (1998). Unequal economic development at the origin of the federalization process. In K. Deprez, & L. Vos (Eds.), *Nationalism in Belgium. Shifting Identities, 1780–1995* (pp. 165 – 176). London: Palgrave Macmillan UK.

Santé, I., García, A. M., Miranda, D., & Crecente, R. (2010). Cellular automata models for the simulation of real-world urban processes: A review and analysis. *Landscape and Urban Planning*, 96(2), 108–122. doi - 10.1016/j.landurbplan.2010.03.001

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, 1(2), 143–186. doi - 10.1080/0022250X.1971.9989794

Schillebeeckx, E., De Decker, P., & Oosterlynck, S. (2013). *Rapport WP2. Internationale migratie en vergrijzing in Vlaanderen: data en kaarten (Report WP2. International migration and ageing in Flanders: data and maps)*. Heverlee: Steunpunt Ruimte. Retrieved from http://www.steunpuntruimte.be/index.php?option=com_joomdoc&task=doc_download&gid=1733&Itemid=112&lang=nl

Schmitz, O., de Kok, J.-L., & Karssenbergh, D. (2016). A software framework for process flow execution of stochastic multi-scale integrated models. *Ecological Informatics*, 32, 124–133. doi - 10.1016/j.ecoinf.2016.01.009

Schuermans, N., Meeus, B., & De Decker, P. (2015). Geographies of whiteness and wealth: white, middle class discourses on segregation and social mix in Flanders, Belgium. *Journal of Urban Affairs*, 37(4), 478–495. doi - 10.1111/juaf.12155

Schwanen, T., & Mokhtarian, P. L. (2005). What affects commute mode choice: neighborhood physical structure or preferences toward neighborhoods? *Journal of Transport Geography*, 13(1), 83–99. doi - 10.1016/j.jtrangeo.2004.11.001

Sembolini, F. (2000). The dynamic of an urban cellular automata in a 3-D spatial pattern. In *Proceedings of the XXI National Conference Aisre*, September 20-22, 2000, Palermo, Italy.

Shan, J., Alkheder, S., & Wang, J. (2008). Genetic algorithms for the calibration of cellular automata urban growth modeling. *Photogrammetric Engineering & Remote Sensing*, 74(10), 1267–1277. doi - 10.14358/PERS.74.10.1267

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(6), 525–552. doi - 10.1016/S0198-9715(01)00014-X

Slegers, K., Kesteloot, C., Van Criekingen, M., & Decroly, J.-M. (2013). Fordist housing behaviour in a post-fordist context. *Journal of Settlements and Spatial Planning*, 3(2), 77–91.

Sohn, J. (2005). Are commuting patterns a good indicator of urban spatial structure? *Journal of Transport Geography*, 13(4), 306–317. doi - 10.1016/j.jtrangeo.2004.07.005

Stanilov, K. (2012). Space in agent-based models. In A. J. Heppenstall, A. T. Crooks, L. M. See, & M. Batty (Eds.), *Agent-based models of geographical systems* (pp. 253–269). Dordrecht, Netherlands: Springer.

Stanilov, K., & Batty, M. (2011). Exploring the historical determinants of urban growth patterns through cellular automata. *Transactions in GIS*, 15(3), 253–271. doi - 10.1111/j.1467-9671.2011.01254.x

Stevens, D., Dragičević, S., & Rothley, K. (2007). iCity: A GIS–CA modelling tool for urban planning and decision making. *Environmental Modelling & Software*, 22(6), 761–773. doi - 10.1016/j.envsoft.2006.02.004

Stevens, F. R., Gaughan, A. E., Linard, C., & Tatem, A. J. (2015). Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. *PloS ONE*, 10(2), e0107042. doi - 10.1371/journal.pone.0107042

Straatman, B., White, R., & Engelen, G. (2004). Towards an automatic calibration procedure for constrained cellular automata. *Computers, Environment and Urban Systems*, 28(1-2), 149–170. doi - 10.1016/S0198-9715(02)00068-6

Tapp, A. F. (2010). Areal Interpolation and Dasymetric Mapping Methods Using Local Ancillary Data Sources. *Cartography and Geographic Information Science*, 37(3), 215–228. doi - 10.1559/152304010792194976

Thys, B., & Andries, P. (2011). *Diagnostiek Woon-Werkverkeer van 30 juni 2011*. Brussels: Belgian Federal Department of Mobility and Transport. http://mobilit.belgium.be/sites/default/files/downloads/RapportWWV_2011_NL_bijlagen_cover.pdf

Tobler, W. R. (1979). Cellular geography. In S. Gale, & G. Olsson (Eds.), *Philosophy in Geography* (pp. 379–386). Dordrecht: Springer Netherlands. doi - 10.1007/978-94-009-9394-5_18

Torrens, P. M., & O'Sullivan, D. (2001). Cellular automata and urban simulation: where do we go from here? *Environment and Planning B: Planning and Design*, 28(2), 163–168. doi - 10.1068/b2802ed

UNFCCC, 2015. *Adoption of the Paris Agreement*. Report No. FCCC/CP/2015/L.9/Rev.1. Retrieved from <http://unfccc.int/resource/docs/2015/cop21/eng/109r01.pdf> [Accessed 26 January 2017].

van de Coevering, P., & Schwanen, T. (2006). Re-evaluating the impact of urban form on travel patterns in Europe and North-America. *Transport Policy*, 13(3), 229–239. doi - 10.1016/j.tranpol.2005.10.001

Vandenbulcke, G., Steenberghen, T., & Thomas, I. (2009). Mapping accessibility in Belgium: a tool for land-use and transport planning? *Journal of Transport Geography*, 17(1), 39–53. doi - 10.1016/j.jtrangeo.2008.04.008

Vanderhaegen, S., & Canters, F. (2016). Use of earth observation for monitoring soil sealing trends in Flanders and Brussels between 1976 and 2013. *Belgeo* [Online], 2 | 2016. Retrieved from <http://belgeo.revues.org/18025>

Vanderhaegen, S., De Munter, K., & Canters, F. (2015). High resolution modelling and forecasting of soil sealing density at the regional scale. *Landscape and Urban Planning*, 133, 133–142. doi - 10.1016/j.landurbplan.2014.09.016

van der Kwast, J., Poelmans, L., Van de Voorde, T., de Jong, K., Uljee, I., Karssenbergh, D., et al. (2012). Uncertainty analysis and data-assimilation of remote sensing data for the calibration of cellular automata based land-use models. In R. Seppelt, A. A. Voinov, S. Lange, & D. Bankamp (Eds.), *Managing Resources of a Limited Planet: Pathways and Visions under Uncertainty, Proceedings of the 6th International Congress on Environmental Modelling and Software (iEMSs 2012)*, July 1-5, 2012, Leipzig, Germany, (pp. 997–1004). http://www.iemss.org/sites/iemss2012//proceedings/D1_2_0452_vanderKwast_et_al.pdf

van der Kwast, J., Uljee, I., Engelen, G., Van de Voorde, T., Canters, F., & Lavallo, C. (2009). Using remote sensing derived spatial metrics for the calibration of land-use change models. *IEEE Proceedings of the 7th International Urban Remote Sensing Conference (URS 2009)*, May 20-22, 2009, Shanghai, China. doi - 10.1109/URS.2009.5137596

Vandermotten, C. (1998). Dynamiques spatiales de l'industrialisation et devenir de la Belgique. *Le Mouvement social*, 185, 75–100. doi - 10.2307/3779048

Vandermotten, C., & Vandewattyne, P. (1985). Groei en vorming van het stadsstramien in België. *Tijdschrift van het Gemeentekrediet van België*, 39(154), 41–62.

Van de Voorde, T., De Roeck, T., & Canters, F. (2009). A comparison of two spectral mixture modelling approaches for impervious surface mapping in urban areas. *International Journal of Remote Sensing*, 30(18), 4785–4806. doi - 10.1080/01431160802665918

Van de Voorde, T., Jacquet, W., & Canters, F. (2011). Mapping form and function in urban areas: An approach based on urban metrics and continuous impervious surface data. *Landscape and Urban Planning*, 102(3), 143–155. doi - 10.1016/j.landurbplan.2011.03.017

Van de Voorde, T., van der Kwast, J., Poelmans, L., Canters, F., Binard, M., Cornet, Y., et al. (2016). Projecting alternative urban growth patterns: The development and application of a remote sensing assisted calibration framework for the Greater Dublin Area. *Ecological Indicators*, 60, 1056–1069. doi - 10.1016/j.ecolind.2015.08.035

Van Esch, L., Poelmans, L., Engelen, G., & Uljee, I. (2011). *Landgebruiksk kaart voor Vlaanderen en Brussel, studie uitgevoerd in opdracht van MIRA, Milieurapport Vlaanderen (Land-use map for Flanders and Brussels, study commissioned by MIRA, the Environmental Report of Flanders)*. MIRA/2011/09. Mechelen: Vlaamse Milieumaatschappij. <http://www.milieurapport.be/nl/publicaties/mira-onderzoeksrapporten/>

Van Herzele, A., & Wiedemann, T. (2003). A monitoring tool for the provision of accessible and attractive urban green spaces. *Landscape and Urban Planning*, 63(2), 109–126. doi - 10.1016/S0169-2046(02)00192-5

van Meeteren, M., Boussauw, K., Derudder, B., & Witlox, F. (2016). Flemish Diamond or ABC-Axis? The spatial structure of the Belgian metropolitan area. *European Planning Studies*, 24(5), 974–995. doi - 10.1080/09654313.2016.1139058

Van Nuffel, N., & Saey, P. (2005). Commuting, hierarchy and networking: the case of Flanders. *Tijdschrift voor Economische en Sociale Geografie*, 96(3), 313–327. doi - 10.1111/j.1467-9663.2005.00462.x

Van Reeth, W., & Van Daele, T. (2009). *Budgetten en beleidsprestaties onder alternatieve landgebruiksscenario's. Wetenschappelijk rapport – NARA 2009 (Budget and policy performance of alternative land-use scenarios. Scientific report – NARA 2009)*. Report INBO.R.2009.19.

Brussels: Instituut voor Natuur- en Bosonderzoek. Retrieved from https://data.inbo.be/purews/files/711082/VanReeth_VanDaele_2009_BudgettenBeleidsprestatiesOnderAlternatieveLandgebruiksscenarios.pdf [Accessed 17 April 2017].

Van Steertegem, M., Bossuyt, M., Brouwers, J., De Geest, C., Maene, S., Maes, F., et al. (Eds.) (2009). *Milieuverkenning 2030. Milieurapport Vlaanderen, MIRA 2009*. Erembodegem: Vlaamse Milieumaatschappij. <http://www.milieurapport.be/nl/publicaties/toekomstverkenningen/milieuverkenning-2030/>

van Vliet, J., Bregt, A. K., Brown, D., van Delden, H., Heckbert, S., & Verburg, P. (2016). A review of current calibration and validation practices in land-change modeling. *Environmental Modelling & Software*, 82, 174–182. doi - 10.1016/j.envsoft.2016.04.017

van Vliet, J., Bregt, A. K., & Hagen-Zanker, A. (2011). Revisiting Kappa to account for change in the accuracy assessment of land-use change models. *Ecological Modelling*, 222(8), 1367–1375. doi - 10.1016/j.ecolmodel.2011.01.017

van Vliet, J., Hagen-Zanker, A., Hurkens, J., & van Delden, H. (2013a). A fuzzy set approach to assess the predictive accuracy of land use simulations. *Ecological Modelling*, 261-262, 32–42. doi - 10.1016/j.ecolmodel.2013.03.019

van Vliet, J., Hagen-Zanker, A., Hurkens, J., & Vanhout, R. (2010). *Map Comparison Kit 3. User Manual*. Software version 3.2.1. Maastricht: Research Institute for Knowledge Systems. Retrieved from <http://mck.riks.nl/downloads>

van Vliet, J., Hurkens, J., White, R., & van Delden, H. (2012). An activity-based cellular automaton model to simulate land-use dynamics. *Environment and Planning B: Planning and Design*, 39(2), 198–212. doi - 10.1068/b36015

van Vliet, J., Naus, N., van Lammeren, R. J., Bregt, A. K., Hurkens, J., & van Delden, H. (2013b). Measuring the neighbourhood effect to calibrate land use models. *Computers, Environment and Urban Systems*, 41, 55–64. doi - 10.1016/j.compenvurbsys.2013.03.006

van Vliet, J., White, R., & Dragičević, S. (2009). Modeling urban growth using a variable grid cellular automaton. *Computers, Environment and Urban Systems*, 33(1), 35–43. doi - 10.1016/j.compenvurbsys.2008.06.006

Verachtert, E., Mayeres, I., Poelmans, L., van der Meulen, M., Vanhulsel, M., & Engelen, G. (2016). *Ontwikkelingskansen op basis van knooppuntwaarde en nabijheid voorzieningen, studie uitgevoerd in opdracht van Ruimte Vlaanderen (Development probabilities based on node values and proximity of services, study commissioned by Ruimte Vlaanderen)*. Brussels, Belgium: Ruimte Vlaanderen. Retrieved from <https://www.ruimtevlaanderen.be/NL/Diensten/Onderzoek/Studies/articleType/ArticleView/articleId/8954/Ontwikkelingskansen-obv-knooppuntwaarde-en-voorzieningen> [Accessed 24 January 2017].

Verachtert, E., Mayeres, I., Poelmans, L., van der Meulen, M., Vanhulsel, M., Engelen, G., et al. (in preparation). Spatial development potential in Flanders based on transit node value and service level. Manuscript in preparation.

Verbeek, T., Boussauw, K., & Pisman, A. (2014). Presence and trends of linear sprawl: Explaining ribbon development in the north of Belgium. *Landscape and Urban Planning*, 128, 48–59. doi - 10.1016/j.landurbplan.2014.04.022

Verburg, P. H. (2006). Simulating feedbacks in land use and land cover change models. *Landscape Ecology*, 21(8), 1171–1183. doi - 10.1007/s10980-006-0029-4

Verburg, P. H., Eickhout, B., & van Meijl, H. (2008). A multi-scale, multi-model approach for analyzing the future dynamics of European land use. *The Annals of Regional Science*, 42(1), 57–77. doi - 10.1007/s00168-007-0136-4

Verburg, P. H., & Overmars, K. P. (2009). Combining top-down and bottom-up dynamics in land use modeling: exploring the future of abandoned farmlands in Europe with the Dyna-CLUE model. *Landscape Ecology*, 24(9), 1167–1181. doi - 10.1007/s10980-009-9355-7

Verburg, P. H., Soepboer, W., Veldkamp, A., Limpiada, R., Espaldon, V., & Mastura, S. S. A. (2002). Modeling the spatial dynamics of regional land

use: the CLUE-S model. *Environmental Management*, 30(3), 391–405.
doi - 10.1007/s00267-002-2630-x

Vlaamse Overheid (2008). *Ruimtelijk Structuurplan Vlaanderen. Addendum 12 december 2008 (Spatial Structure Plan for Flanders. Addendum 12 December 2008)*. Retrieved from http://www2.vlaanderen.be/ruimtelijk/rsv/herziening/doc/voorontwerp_addendum_richtinggevend_bindend.pdf

Voets, J., De Peuter, B., Vandekerckhove, B., Broeckaert, D., Le Roy, M., Maes, P., et al. (2010). *Evaluerend onderzoek naar de effectiviteit van de uitvoering van het ruimtelijk beleid in Vlaanderen. Voorbereidend onderzoek voor het Beleidsplan Ruimte (Evaluative research on the effectiveness of the implementation of spatial policy in Flanders. Preparatory research for the Spatial Policy Plan)*. Report commissioned by the Flemish Department of Spatial Planning, Housing Policy and Immovable Heritage. Brussels: Flemish Government.

Volkery, A., Ribeiro, T., Henrichs, T., & Hoogeveen, Y. (2008). Your vision or my model? Lessons from participatory land use scenario development on a European scale. *Systemic Practice and Action Research*, 21(6), 459–477. doi - 10.1007/s11213-008-9104-x

von Neumann, J. (1966). *Theory of Self-Reproducing Automata*. Edited by A. W. Burks. Champaign, IL, USA: University of Illinois Press.

von Thünen, J. F. (1966 [1826]). *Von Thünen's "Isolated State"*. Translated by C. Wartenberg. Oxford: Pergamon.

Vorel, J. (2015). *Urban Simulation Modeling. An Introduction and Experimental Applications in the Czech Republic*. Prague: Czech Technical University Publishing House.

Waddell, P. (2002). UrbanSim: modeling urban development for land use, transportation and environmental planning. *Journal of the American Planning Association*, 68(3), 297–314. doi - 10.1080/01944360208976274

Warntz, W. (1965). *Macrogeography and Income Fronts*. Philadelphia: Regional Science Research Institute.

- Weber, A. (1957 [1909]). *Alfred Weber's "Theory of the Location of Industries"*. Translated by C. Friedrich. Chicago: University of Chicago Press.
- Wegener, M. (2004). Overview of land-use transport models. In D.A. Hensher, et al. (Eds.), *Handbook of transport geography and spatial systems* (pp. 127–146). Kidlington, UK: Pergamon/Elsevier Science.
- Weng, Q. (2012). Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. *Remote Sensing of Environment*, 117, 34–49. doi - 10.1016/j.rse.2011.02.030
- White, R. (1977). Dynamic central place theory: Results of a simulation approach. *Geographical Analysis*, 9(3), 226–243. doi - 10.1111/j.1538-4632.1977.tb00576.x
- White, R. (1989). The artificial intelligence of urban dynamics: Neural network modeling of urban structure. *Papers of the Regional Science Association*, 67(1), 43–53. Springer-Verlag. doi - 10.1007/BF01934666
- White, R. (2006a). Modelling multi-scale processes in a cellular automata framework. In J. Portugali (Ed.), *Complex artificial environments* (pp. 165–178). Berlin: Springer-Verlag. doi - 10.1007/3-540-29710-3_11
- White, R. (2006b). Pattern based map comparisons. *Journal of Geographical Systems*, 8, 145–164. doi - 10.1007/s10109-006-0026-9
- White, R., & Engelen, G. (1993). Cellular Automata and Fractal Urban Form: A Cellular Modelling Approach to the Evolution of Urban Land Use Patterns. *Environment and Planning A*, 25(8), 1175–1199. doi - 10.1068/a251175
- White, R., & Engelen, G. (1997). Cellular Automata as the Basis of Integrated Dynamic Regional Modelling. *Environment and Planning B*, 24(2), 235–246. doi - 10.1068/b240235
- White, R., & Engelen, G. (2000). High resolution integrated modelling of the spatial dynamics of urban and regional systems. *Computers, Environment, and Urban Systems*, 24(5), 383–400. doi - 10.1016/S0198-9715(00)00012-0

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 and Design*, 24(3), 323-343. doi - 10.1068/b240323

White, R., Engelen, G., & Uljee, I. (2000). Modelling land-use change with linked cellular automata and socio-economic models: a tool for exploring the impact of climate change on the island of St. Lucia. In M. Hill, & R. Aspinall (Eds.), *Spatial Information for Land Use Management* (pp. 189–204). Amsterdam: Gordon and Breach.

White, R., Engelen, G., & Uljee, I. (2015). *Modeling cities and regions as complex systems. From theory to planning applications*. Cambridge, MA, USA, & London, UK: The MIT Press.

White, R., Shahumyan, H., & Uljee, I. (2011). Activity based variable grid cellular automata for urban and regional modelling. In D. Marceau, & I. Benenson (Eds.), *Advanced geosimulation models* (pp. 14–29). Hilversum: Bentham.

White, R., Uljee, I., & Engelen, G. (2012). Integrated modelling of population, employment and land-use change with a multiple activity-based variable grid cellular automaton. *International Journal of Geographical Information Science*, 26(7), 1251–1280. doi - 10.1080/13658816.2011.635146

Williams, B., & Shiels, P. (2002). The expansion of Dublin and the policy implications of dispersal. *Journal of Irish Urban Studies*, 1(1), 1–20.

Winters, S., Ceulemans, W., Heylen, K., Pannecoucke, I., Vanderstraeten, L., Van den Broeck, K., et al. (2015). *Wonen in Vlaanderen anno 2013. De bevindingen uit het Grote Woononderzoek 2013 gebundeld (Housing in Flanders anno 2013. The compiled findings of the Great Housing Research 2013)*. Leuven: Steunpunt Wonen. Retrieved from https://steunpuntwonen.be/Documenten/Onderzoek_Werkpakketten/gwo-volume-1-eind.pdf

Wolfram, S. (1984). Universality and complexity in cellular automata. *Physica D: Nonlinear Phenomena*, 10(1), 1–35. doi - 10.1016/0167-2789(84)90245-8

Wolfram, S. (2002). *A New Kind of Science*. Champaign, IL, USA: Wolfram Media.

Wu, C. (2004). Normalized spectral mixture analysis for monitoring urban composition using ETM+ imagery. *Remote Sensing of Environment*, 93(4), 480–492. doi - 10.1016/j.rse.2004.08.003

Wu, C., & Murray, A. T. (2007). Population estimation using Landsat Enhanced Thematic Mapper imagery. *Geographical Analysis*, 39(1), 26–43. doi - 10.1111/j.1538-4632.2006.00694.x

Wu, F. (1998). Simulating urban encroachment on rural land with fuzzy-logic-controlled cellular automata in a geographical information system. *Journal of Environmental Management*, 53(4), 293–308. doi - 10.1006/jema.1998.0195

Wu, S., Qiu, X., & Wang, L. (2005). Population estimation methods in GIS and remote sensing: a review. *GIScience and Remote Sensing*, 42(1), 80–96. doi - 10.2747/1548-1603.42.1.80

Xie, Z. (2006). A framework for interpolating the population surface at the residential-housing-unit level. *GIScience & Remote Sensing*, 43(3), 233–251. doi - 10.2747/1548-1603.43.3.233

Yang, Q., Li, X., & Shi, X. (2008). Cellular automata for simulating land use changes based on support vector machines. *Computers & Geosciences*, 34(6), 592–602. doi - 10.1016/j.cageo.2007.08.003

Yang, X., Chen, R., & Zheng, X. Q. (2016). Simulating land use change by integrating ANN-CA model and landscape pattern indices. *Geomatics, Natural Hazards and Risk*, 7(3), 918–932. doi - 10.1080/19475705.2014.1001797

Yang, X., Zheng, X. Q., & Chen, R. (2014). A land use change model: Integrating landscape pattern indexes and Markov-CA. *Ecological Modelling*, 283, 1–7. doi - 10.1016/j.ecolmodel.2014.03.011

Zandbergen, P. A. (2011). Dasymetric mapping using high resolution address point datasets. *Transactions in GIS*, 15(s1), 5–27. doi - 10.1111/j.1467-9671.2011.01270.x

Zandbergen, P. A., & Ignizio, D. A. (2010). Comparison of dasymetric mapping techniques for small-area population estimates. *Cartography and Geographic Information Science*, 37(3), 199–214.
doi - 10.1559/152304010792194985

Zipf, G. (1949). *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Cambridge, MA, USA: Addison-Wesley.

Publications

Peer-reviewed articles related to the PhD

1. Crols, T., van der Meulen, M., White, R., Poelmans, L., Uljee, I., Engelen, G., & Canters, F. (in preparation). Semi-automated calibration of an activity-based cellular automata model. *Environment and Planning B*, special issue on Cellular Automata Modeling for Urban and Spatial Systems. Abstract submitted.
2. Crols, T., Vanderhaegen, S., Canters, F., Engelen, G., Poelmans, L., Uljee, I., & White, R. (2017). Dated high-resolution population density maps using sealed surface cover time series. *Landscape and Urban Planning*, 160, 96-106.
Doi: 10.1016/j.landurbplan.2016.12.009
3. Crols, T., White, R., Uljee, I., Engelen, G., Poelmans, L., & Canters, F. (2015). A travel time-based variable grid approach for an activity-based cellular automata model. *International Journal of Geographical Information Science*, 29(10), 1757-1781.
Doi: 10.1080/13658816.2015.1047838

Chapters

1. White, R., Crols, T., & Engelen, G. (in review). Modelling the spatial dynamics of population, economic activity and land use for policy and planning. In: Y. Jin (Ed.), *Applied Urban Modelling*, Chapter 7. Proceedings of the British Academy.
2. Le Xuan, Q., De Munter, K., Crols, T., Khan Mahsud, A. Z., & Corijn, E. (2013). Sustainability in Coastal Urban Environment: Identifying Resources and Users in Belgian Case Study Areas. In A. Z. Khan, L. X. Quynh, E. Corijn, & F. Canters (Eds.), *Sustainability in the Coastal Urban Environment: Thematic Profiles of Resources and their Users*. (pp. 109-146). Sapienza Universita Editrice.

Conference papers

1. Crols, T., White, R., Uljee, I., Poelmans, L., Engelen, G., & Canters, F. (2017, in press). High-resolution simulations of population-density change with an activity-based cellular automata land-use model. In *Proceedings of the 14th International Conference on GeoComputation (GeoComputation 2017), Leeds, UK, 3-7 September 2017*. Accepted.
2. Crols, T., van der Meulen, M., White, R., Uljee, I., Engelen, G., Poelmans, L., & Canters, F. (2016). Semi-automated calibration of an activity-based CA model. In *Proceedings of the 2nd International Symposium on Cellular Automata Modeling for Urban and Spatial Systems (CAMUSS 2016), Québec, Canada, 21-23 September 2016*. (pp. 12–17).
3. Crols, T., White, R., Uljee, I., Engelen, G., Canters, F., & Poelmans, L. (2012). An activity-based CA model for Flanders, Belgium. Robustness analysis and improved distance computation in a variable grid representation. In N. Pinto, J. Dourado, & A. Natálio (Eds.), *Proceedings of CAMUSS, the International Symposium on Cellular Automata Modeling for Urban and Spatial Systems, Oporto, November 8 to 10, 2012*. (pp. 195-207). Coimbra: Department of Civil Engineering of the University of Coimbra.
4. Crols, T., White, R., Uljee, I., Engelen, G., Canters, F., & Poelmans, L. (2012). Development of an activity-based cellular automata land-use model: the case of Flanders, Belgium. In R. Seppelt, A. Voinov, S. Lange, & D. Bankamp (Eds.), *Proceedings of the 6th International Congress on Environmental Modelling and Software (iEMSs 2012), Leipzig, Germany, 1-5 July 2012*. (Vol. F1, pp. 2000-2007). International Environmental Modelling and Software Society.
5. Canters, F., & Crols, T. (2011). Low-error equal-area map projections for global mapping of the terrestrial environment. In *Proceedings of the 25th International Cartographic Conference, Paris, 3-8 July 2011*. (Vol. D1). International Cartographic Association.

Meeting abstracts

1. Crols, T., Poelmans, L., Verachtert, E., Uljee, I., White, R., Canters, F., & Engelen, G. (2017, in press). *Evaluating a land-take neutral policy scenario in the region of Flanders, Belgium with an activity-based cellular automata model*. Abstract from Social Simulation Conference 2017, Dublin, Republic of Ireland. Abstract accepted.
2. Crols, T., Vanderhaegen, S., White, R., Canters, F., Engelen, G., Uljee, I., & Poelmans, L. (2016). *A historical population and densification analysis of Flanders and Brussels (1986-2013) to calibrate an activity-based land-use change model*. Abstract from Doctoral Seminar On Sustainability Research in the Built Environment 2016 (DS²BE 2016), Vaalbeek, Belgium.
3. Crols, T., White, R., Uljee, I., Engelen, G., Vanderhaegen, S., Canters, F., & Poelmans, L. (2015). *Calibrating an activity-based cellular automata land-use model for Flanders and Brussels*. Abstract from 6th Belgian Geography Days, Brussels, Belgium.
4. Crols, T., White, R., Uljee, I., Vanderhaegen, S., Engelen, G., Canters, F., & Poelmans, L. (2015). *Initial conditions analysis and historical calibration of an activity based cellular automata land-use model for Flanders and Brussels*. Abstract from 19th European Colloquium of Theoretical and Quantitative Geography (ECTQG 2015), Bari, Italy.
5. Crols, T., Uljee, I., Van Esch, L., Canters, F., White, R., Engelen, G., & Poelmans, L. (2013). *Disaggregation of initial employment data in an Activity-based Cellular Automata (ACA) land-use model for Flanders, Belgium*. Abstract from 18th European Colloquium of Theoretical and Quantitative Geography (ECTQG 2013), Dourdan, France.
6. Crols, T., White, R., Uljee, I., Canters, F., Engelen, G., & Poelmans, L. (2013). *Network-based time-distance computations for an Activity-based Cellular Automata (ACA) land-use model for Flanders*. Abstract from 5th Belgian Geographical Day, Louvain-la-Neuve, Belgium.