Urban development as a continuum:   
A multinomial logistic regression approach

Ahmed M Mustafa1,\*, Mario Cools1, Ismail Saadi1, Jacques Teller1

1 LEMA, University of Liège, Liège, Belgium

\*[a.mustafa@ulg.ac.be](mailto:a.mustafa@ulg.ac.be)

**Abstract.** Urban development is a complex process influenced by a number of driving forces, including spatial planning, topography and urban economics. Identifying these drivers is crucial for the regulation of urban development and the calibration of predictive models. Existing land-use models generally consider urban development as a binary process, through the identification of built versus non-built areas. This study considers urban development as a continuum, characterized by different level of densities, which can be related to different driving forces.

A multinomial logistic regression model was employed to investigate the effects of drivers on different urban densities during the past decade in Wallonia, Belgium. Sixteen drivers were selected from sets of driving forces including accessibility, geo-physical features, policies and socio-economic factors.

It appears that urban development in Wallonia is remarkably influenced by land-use policies and accessibility. Most importantly, our results highlight that the impact of different drivers varies along with urban density.

**Keywords:** urban development · driving forces · multinomial logistic regression model · cadastral data · urban densities.

1. Introduction

Urban development is a global issue with paramount socioeconomic and environmental implications, which may affect our well-being in terms of society, economy and/or culture. It may lead to a number of problems related to water quality degradation, air pollution, socio-economic disparities and social fragmentation [1, 2]. Several modelling approaches have been adopted in the analysis of urban growth. Those models can be aimed at predicting spatial location and/or the amount of change.

Prediction of the location of urban growth can be modeled through different approaches including statistical methods, cellular automata and agent-based models. For different modelling approaches, it is important to explore and analyze the main drivers of urban growth on space bases in order to better understand, control and model the future growth of urban settlements.

Urban growth models do usually not differentiate between high-density and low-density urban development [e.g. 3–5] . The main goal of this paper is to investigate the major drivers of different urban densities. This requires analyzing the relationship between urban growth and a number of forces related to people choices in terms of spatial location of new urban developments.

Empirical estimation models use statistical methods to model the relationship between urban growth and its drivers based on past observation. In this context, logistic regression models are commonly employed to model urban development potentials [e.g. 3, 5–8].

A multinomial logistic regression model (MNL) is a generalization of a logistic regression model for more than two discrete and unordered response categories. Thus, the MNL methodology allows for the consideration of several classes or urban densities as the dependent, responsible, variable (*Y*), using a set of independent, explanatory, variables (*X*), which were selected based on the following literature review.

* 1. Potential driving forces

The identification of dominant urban growth driving forces is the main objective of this paper. In literature [e.g. 3, 9–17], numerous explanatory urban development drivers are proposed, which can be grouped into four main sets: accessibility indicators, geo-physical features, land-use policies and socio-economic factors.

Accessibility indicators are often implemented in urban models by means of simple accessibility indicators, such as distance to cities, distance to the road network and distance to water bodies [10, 12, 13]. Road infrastructure consumes a high amount of urban land, around 25% of the total urban area in Europe and 30% in the USA [18]. In this paper, we considered Euclidean distances to different roads categories and to major 11 Belgian cities.

Geo-physical factors are reported as a fundamental driver of the spatial distribution and expansion of urban areas [3]. There is often a relationship between urban growth and a number of these factors, especially the topography of the study area [3, 19]. We considered elevation and slope as geo-physical factors in our study.

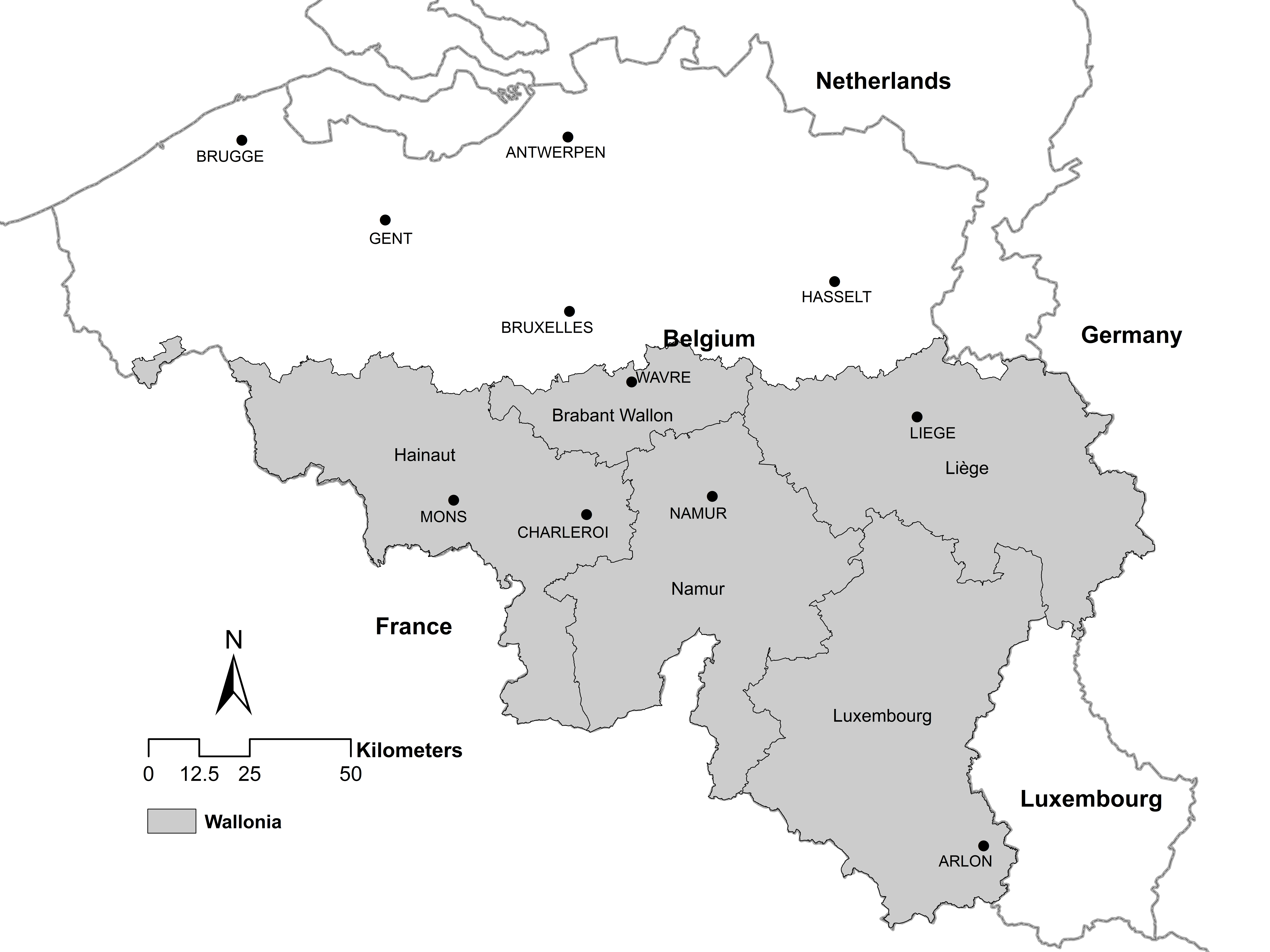
Zoning status (policies) is often considered as one of the major urban development drivers worldwide. It has been classified as the most pervasive driver in USA [20]. In Wallonia, land allocation is controlled by several regulations including the regional development plan, referred to as "plan de secteur (PDS)". In this paper, we consider this zoning plan, which defines the legally authorized land-use type for all the territory.

This study also selects a number of socio-economic factors. Population is one of the most active driver of urban development [15]. In this respect, the evolution of net and gross population densities was considered. Economic development could also be considered as a driver of urban development; there is a relation between economic increase and urban development [15] and furthermore economic development has an important influence on people's location choices. In this respect, employment potential, richness level, housing and land prices are considered. The number of households is another factor to be considered in this paper. This number may rely on population lifecycle, migration, societal values, gender relationships and the relationships between parents and children.

1. Methodology
   1. Study area

The study area is Wallonia, occupying the southern part of Belgium (Fig. 1). Wallonia is the predominantly French-speaking region of Belgium. It has a territory of 16,844 km² that makes up 55% of the territory of Belgium but with only a third of its population. The population volume in 2010 was 3,498,384 inhabitants [21]. Administratively, it comprises five provinces: Hainaut, Liège, Luxembourg, Namur, and Walloon Brabant. It has 20 administrative arrondissements and 262 municipalities. The geography of the area goes from flat to hilly with altitude ranges from 0 to 693 m above see-level. Over the last decades, land-use change in Wallonia was related to urban growth and a subsequent loss of agricultural land.

Major cities in Wallonia are characterized by a strong center–periphery structure with well-off households located in the peripheries [22]. The main metropolitan areas are Charleroi, Liège, Mons and Namur. They are all characterized by a historical city-center around which the urban development was spread. Urban sprawl has affected Wallonia for decades leading to fragmented and isolated landscapes that were developed in space and time [23]. Wallonia is highly affected by its neighbors especially by the trans-border workers number evolution.



**Fig. .** Study area

* 1. Multinomial logistic regression model

An MNL model was applied to investigate the contribution of the selected driving forces on the probability of urban development along different densities. MNL consists of three components: multi-temporal urban maps, a multivariate function of the hypothesized drivers of change, and the resulting prediction map of urban development potential [24].

MNL analysis yields coefficients for each *X*. These coefficients are then interpreted as weights in a formula that generates a map for each urban density class depicting the probability of each cell in the landscape to be converted into this class. If the *Y* variable is a categorical map with *k* classes, taking on values *0, 1,..., k-1* and *X* is a set of explanatory variables *X1, X2,..., Xn* then the logit for each non-reference class *k1,..., kn* against the reference class *k0* model is calculated through:

 ()

where log(*k*) is a logit function of class *k* against the reference class, *α* is the intercept, *β* is the regression coefficients of class *k*.

The conditional probabilities of each class can be calculated with the following formula:

 ()

The goodness-of-fit, in terms of predictive ability and the interpretability, of the MNL outcomes can be evaluated using McFadden pseudo R-square (MFR2) and Relative Operating Characteristic (ROC) statistic respectively [e.g. 5, 13, 17, 25, 26].

MFR2 tries to mimic the R-squared analysis of linear regression. An MFR2 of 1 indicates a perfect fit, where MFR2 of 0 indicates no relationship. Clark & Hosking [25] stated that a MFR2 greater than 0.2 can be considered a good fit and it is calculated according to the following formula:

 ()

where *Lm* is the value of the likelihood function for the full model as fitted with *X* and *L0* is the value of the likelihood function if all *β* except *α* are 0.

The ROC statistic compares the probability map, produced by MNL, to a map with the observed changes of the urban cells for each class between two time-steps. It first divides the probability outcomes into percentile groups from high to low probability and then calculates the proportion of true-positives and false-positives for a range of specified threshold values and relates them to each other in a graph. The ROC measures the area under the curve and its value could range between 0.5 (no relationship) and 1 (perfect fit). ROC values higher than 0.70 are considered as a reasonable fit [19, 27, 28].

Prior to performing the MNL model, we have to consider three aspects that may exist in *X*: disparity in units, autocorrelation and multicollinearity. These aspects potentially affect any regression analysis results. Due to disparity in units and scale of *X* (table.1), the logit coefficients cannot be used directly to measure the relative contribution of the *X* in urban development process. Consequently, all continuous *X* were standardized before performing MNL analysis according to the following formula:

 (4)

where *zi* is the standardized score of cell *i*, *xi*is the original value, *μ* is the mean value of *X* variable and *σ* is the standard deviation of the *X* variable. *z* score is negative when the raw value is below the mean and positive when above. Categorical *X* were not standardized to keep the meaning of the dummy variable.

Spatial autocorrelation in one or more *X* will bias the results of the regression analysis. Autocorrelation is the propensity for cell value to be similar to surrounding cells. Moran's I statistic was processed to detect spatial autocorrelation for each *X.*It is given as:

 (5)

where *M\_I* is the Moran's I statistic for each *X*, *n* is the number of neighbor cells to be taken into account, *w* spatial weights and *xi/j* cells values at location *i/j*. Locations depend on cell neighbors which can be considered as only shared-border neighbors to the cell under evaluation (*xi*) or also diagonal neighbors (*xj*). We considered only *xj* neighbors. Moran's I value ranges between -1 and +1, where +1 means absolute autocorrelation and -1 none autocorrelation. All *X* show strong degree of spatial autocorrelation with Moran's I value between 0.746 for zoning and 0.999 for distance to cities. A number of scholars [e.g. 3, 19, 29] suggested that this problem could be addressed through a data sampling approach. A random sample of 15,675 cells, around 1.15% of the study area, distributed throughout the study area were used in the MNL model.

Multicollinearity represents a high degree of dependency among a number of *X*. It commonly occurs when a large number of *X* are introduced in a regression model. It is because some of *X* may measure the same phenomena. Strong collinearities cause the erroneous estimation of parameters and further affect the MNL results [26]. In this context, a number of procedures is proposed to detect multicollinearity among *X* such as tolerance value, variance inflation factor and Belsley diagnostics [30–33]. We used Belsley diagnostics, one of the most common procedures, to detect multicollinearity. Although Belsley diagnostics are normally applied to linear regression models, it is still valid to apply it for MNL since multicollinearity is a problem among *X* [34]. The outcomes of Belsley diagnostics are condition indices and variance-decomposition proportions for each *X*. A condition index greater than 30 represents strong multicollinearity [33]. In this case, it is highly recommended to omit all *X* with variance-decomposition proportions exceeding the tolerance of 0.5 [33]. All *X* show low degree of multicollinearity with condition indices between 1 and 9.15 for all *X*maps and 1 to 9.86 for the selected samples. Thus, all *X*will be used in the MNL model.

* 1. Data

### Dependent variable**.**

The *Y* is constituted by cells whose status not changed from non-urban to urban and changed from non-urban to one of different urban classes between 2000 and 2010.

The cadastral dataset (CAD) was used to develop *Y*. CAD is a vector map representing buildings in two dimensions as polygons provided by the Land Registry Administration of Belgium. Each building comes with different attributes from which the construction date is the most important attribute for our study. Using construction date, two urban land-use maps were developed for 2000 and 2010 years.

#### Preparing the data*.*

CAD vector data were rasterized at a very fine cell dimension (2x2 m2). The rasterized cells were aggregated with a factor of 50 by which to multiply the cell size to obtain 100x100 m2 raster grid. Each aggregated cell has a density value that represents the number of rasterized 2x2 cells. This value will be used to introduce dwellings density in the aggregated CAD maps.

Scholars always define a minimum map unit (MMU) to avoid overestimation of one class in land cover data [35, 36]. The CORINE Land-Cover (CLC) dataset, as an example, was provided with MMU of 25 ha. Due to the nature of the study area, it is possible to find several scattered buildings in a hectare. Consequently, to produce more accurate data we set MMU at one hectare. In order to avoid overestimation of urban lands, two procedures were applied to the aggregated data: minimum building density per cell (MBDC) and minimum building density per neighbor (MBDN).

#### Minimum building density per cell*.*

The average size of residential building in Belgium is about 10x10 m² [37]. This figure somehow corresponds to the average size of households in Wallonia [38] multiplied by the single family house floor space per capita in north & west Europe, which is about 41 m² [39]. The multiplication gives an average size of 2.3\*41= 94.3 m². Thus, we set MBDC at 25 (100 m²).

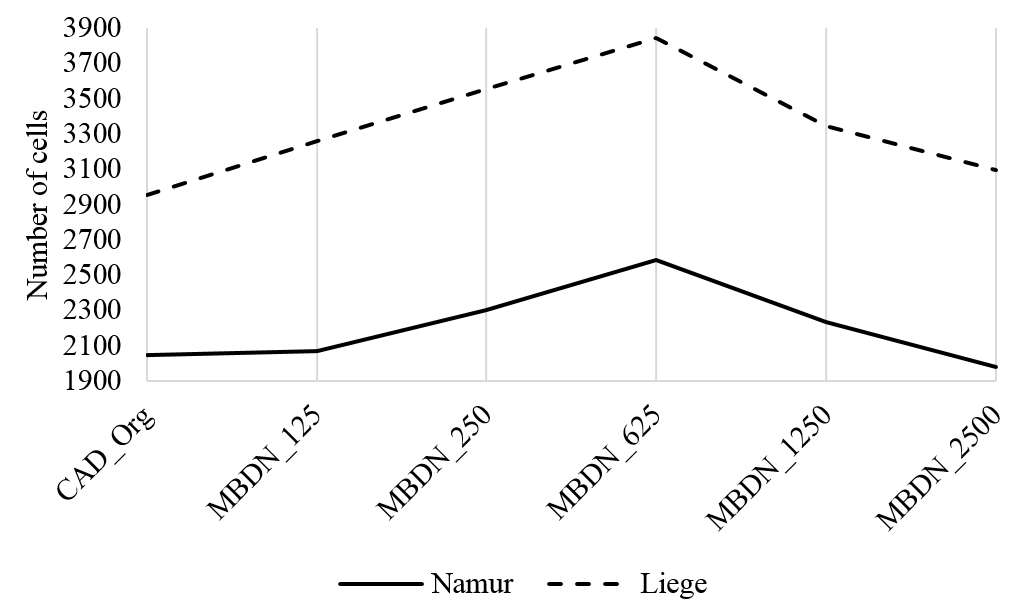
#### Minimum building density per neighbor.

A threshold of five dwellings per ha was fixed for considering that a cell was urbanized. Neighborhoods with such a density are indeed observable in Wallonia. We then performed an analysis using different thresholds of MBDN using a search window of 3x3 cells for each MBDN cell less than 125 (5x25). These thresholds are 125, 250, 625, 1250 and 2500. Table 1 lists a comparison between CLC data, CAD original aggregated data and different MBDN thresholds.

**Table 1.** Comparison of area (km²) between CLC, CAD\_Org original aggregated CAD data, MBDN\_125, MBDN\_250, MBDN\_625, MBDN\_1250 and MBDN\_2500.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Year | CLC | CAD\_Org | MBDN125 | MBDN  250 | MBDN  625 | MBDN  1250 | MBDN  2500 |
| 2000 | 2506 | 3229 | 2599 | 2468 | 2093 | 1744 | 1579 |
| 2006 | 2513 | - | - | - | - | - | - |
| 2010 | - | 3339 | 2716 | 2594 | 2230 | 1868 | 1693 |

We assumed that the number of changed cells between two time-steps would increase until a specific value of MBDN and then start declining along with the increase of MBDN. Actually, those cells that are under urban development at time-step 1 and reach the threshold of MBDN at time-step 2 are then considered as urban. If the MBDN threshold is very high, this condition will not be reached because this threshold exceeds the observed number of built cells at time-step 1 and 2. The number of changed cells calculated in two provinces of Wallonia confirmed our assumption (Fig. 2). The result showed that the most appropriate threshold for MBDN is 625.



**Fig. .** Number of changed cells between 2000 and 2010

#### Measuring density*.*

Different definitions of urban density can be found in the literature, according to the application it is used for. Density can refer to either dwellings or inhabitants per unit area [40]. In this paper, we defined urban density as a number of built-up units per cell of one hectare. We performed different MNL models for 4, 6, 8 and 10 urban densities quantile classes and measured the goodness-of-fit. An MFR2 statistic is not appropriate to compare the goodness-of-fit in this case because MFR2 depends on *Y* which is changeable, in terms of number of classes. Consequently, we measured misclassification rate instead. It equaled 24.23%, 23.60%, 22.70% and 26.07% respectively. As a result, for the final MNL, we used eight classes for urban densities, from class0 (non-urban) to class7 (highest density), each class has almost the same number of cells except for class0 (Fig. 3). Table 2 lists the density range for each class.

Urban class7 can be considered as urban cores and urban classes 4, 5 and 6 may be considered as urban peripheries and suburbs. Urban classes 2 and 3 may be considered as rural areas whereas urban class1 may be considered as remote locations. This can be further assessed by measuring how land-use policies control urban development within each class. Land-use policies can highly control urban developments within urban cores whereas urban development in peripheries and suburbs is not strictly following policies. Rural areas are expected to follow policies whereas very remote sites are normally not controlled by policies.

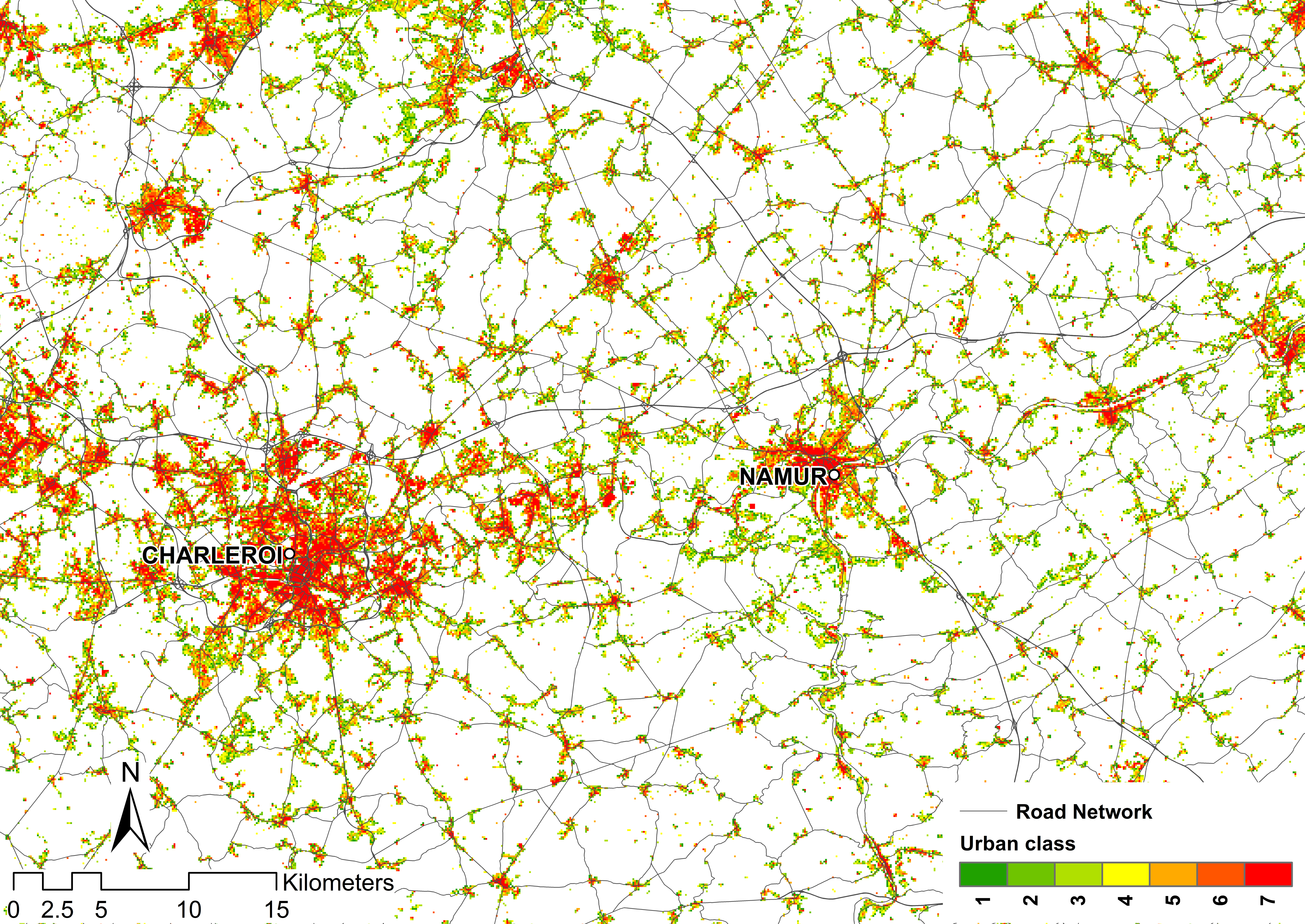
**Table 2.** Urban classes density ranges in number of 2x2 pixels (% of 100x100 cell area)   
covered by building footprints.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Min | Max | *μ* | Mode |
| Class0a | - | - | - | - |
| Class1 | 25 (1.0%) | 78 (3.2%) | 51.5 | 32 |
| Class2 | 79 (3.2%) | 132 (5.3%) | 105.5 | 127 |
| Class3 | 133 (5.3%) | 180 (7.2%) | 156.5 | 138 |
| Class4 | 181 (7.2%) | 243 (9.8%) | 212.0 | 182 |
| Class5 | 244 (9.8%) | 330 (13.2%) | 287.0 | 254 |
| Class6 | 331 (13.2%) | 491 (19.7%) | 411.0 | 333 |
| Class7 | 492 (19.7%) | 2500 (100.0%) | 1365.9 | 504 |

a class0 represents all non-urban cells and affected cells by MBDC and MBDN procedures

### Independent variables**.**

Statistical data related to population volume, households, employment rate, richness index and mean land/housing price were acquired from the official Belgian statistics [38, 41] and mapped with a resolution of 100x100 m² raster grid at municipality level. Gross population density was calculated for each municipality as the number of inhabitants divided by the area of municipality in km² whereas net population density was calculated as the number of inhabitants divided by the area of built-up zones of the municipality in km².



**Fig. .** Urban density classes of 2010 (7 highest density, 1 lowest density)

Digital Elevation Model (DEM) provided by the Belgian National Geographic Institute was used to calculate elevation and slope in percentage for each cell.

Accessibility was measured in terms of Euclidean distance of a cell from the nearest road and city. Road networks for 2002 were provided by Navteq. Four functional classes of roads were introduced in MNL (R\_class1: high speed and volume controlled access roads, R\_class2: quick travel between and through cities, R\_class3: moderate speed travel within cities and R\_class4: moderate speed travel between neighborhoods). Major cities of Belgium (Antwerp, Brussels, Wavre, Brugge, Gent, Charleroi, Mons, Liege, Hasselt, Arlon and Namur) were used to develop a map of distances to cities.

According to the most recent zoning plan of Wallonia, urban development is only allowed in those zones that are designated for residential, economic or leisure development. In other zones, such as agricultural and forest areas, urban development is not permitted unless specific conditions. A zoning map was developed by discerning zones where urban development is not permitted (code 0) and zones that are designated for urban development (code 1).

All maps were created as raster grids with a resolution of 100x100 m² (table 1). The spatial resolution is defined based on the availability of data. The statistical data are available at municipality level whereas other variables could be calculated at cell level. That is common to coupling both resolution levels [e.g. 19, 42].

**Table 3.** List of the selected drivers of urban development.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Driver | Name | Typea | Unit | Resolutionb | *μ* | *σ* |
| *X1* | Elevation | 1 | m | 1 | 257.14 | 183.4 |
| *X2* | Slope | 1 | % | 1 | 5.51 | 57.02 |
| *X3* | Dist to city | 1 | m | 1 | 29028.16 | 15479.34 |
| *X4* | Dist to R\_class1 | 1 | m | 1 | 7936.12 | 8282.57 |
| *X5* | Dist to R\_class2 | 1 | m | 1 | 4174.5 | 3757.23 |
| *X6* | Dist to R\_class3 | 1 | m | 1 | 1668.27 | 1425.25 |
| *X7* | Dist to R\_class4 | 1 | m | 1 | 818.63 | 850.46 |
| *X8* | Dist to rail stations | 1 | m | 1 | 6962.07 | 5710.64 |
| *X9* | Num households | 1 | number | 2 | 6421.52 | 12040.71 |
| *X10* | Mean housing price | 1 | € | 2 | 139487 | 31965.1 |
| *X11* | Mean land price | 1 | €/m² | 2 | 51.1 | 99.73 |
| *X12* | Employment potential | 1 | % | 2 | 48.39 | 98.42 |
| *X13* | Richness index | 1 | % | 2 | 95.71 | 61.62 |
| *X14* | Growth population density | 1 | inh/km² | 2 | 206.95 | 354.33 |
| *X15* | Net population density | 1 | inh/urban km² | 2 | 819.52 | 522.97 |
| *X16* | Zoning status | 2 | binary | 1 |  |  |

a 1. Continuous, 2. Categorical. b 1. Cell level, 2. Municipality level.

1. Results and discussions

In the last decade, Wallonia experienced urban growth distributed between different urban classes. The total area of urban land in 2000 was 2093 km², accounting for 12.4% of the total area, and in 2010, urban land increased to 2230 km², accounting for 13.2% of the total area. More than 137.5 km² of land was converted between 2000 and 2010 from non-urban to different urban densities as: 50.4 km² to urban class1, 34.7 km² to class2, 23 km² to class3, 10.5 km² to class4, 7.12 km² to class5, 5.3 km² to class6 and 6.68 km² to class7. Of course, one km² of urban class1 has not the same weight as one km² of urban class7, as densities are very different.

The goodness-of-fit for predictive ability was evaluated using MFR2 and it equals 0.244. The model reveals a very good correspondence with ROC values: 0.775, 0.819, 0.829, 0.805, 0.793, 0.813 and 0.914 for classes 1, 2, 3, 4, 5, 6 and 7 respectively. This indicates that the MNL performs well and the MNL’s outcomes could effectively interpret the process of urban development in Wallonia.

The results of the MNL model are illustrated in table 4. Scholars and decision-makers are not only interested in the identification of potential driving forces of urban development process but also interested in measuring the relative contribution of these drivers to urban development process [3, 10]. To relatively measure the contribution of each *X* to urban development process, Odds Ratio (OR) that equals *exp(β)* is calculated. It is difficult to directly interpret an MNL coefficients and they become useful when they are converted into an OR. An OR greater than 1 indicates a positive effect, the probability of urban development increases by increasing the OR of the variable, whereas less than 1 indicates a negative effect, the probability of urban development decreases by increasing the OR of the variable. An OR of 1 means neutral contribution to the urban development process [13]. Table 5 lists OR values for all *X*.

For interpreting continuous *X*, multiply a one unit increase in *X* variable by OR. For instance, five units increase in elevation increase the class7 portability by *exp(5x0.110) ~* 1.7 whereas decrease the class1 portability by *exp(5x-0.078) ~* 0.7. For categorical *X*, zoning status, the OR value for class1 is about 15.4. This means that it is around 15.4 *(exp(2.735))* times more likely to find a new urban cell related to class1 in zones designated to urban development than in zones designated to other uses in zoning plan.

Generally, the impact of different drivers varies with different urban densities. These drivers can be grouped into common drivers with impacts on different urban classes and special drivers with impacts on individual classes.

**Table 4.** The coefficients (*β*) of MNL model (class0 is the reference class).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Class1 | Class2 | Class3 | Class4 | Class5 | Class6 | Class7 |
| *α* | -4.158 | -4.603 | -4.700 | -4.583 | -4.409 | -4.500 | -5.969 |
| Elevation | -0.078 | -0.032 | 0.043 | 0.105 | 0.023 | 0.204 | 0.110 |
| Slope | -0.237 | -0.078 | -0.210 | -0.644 | -0.693 | -0.840 | -1.185 |
| Dist to city | 0.072 | 0.051 | -0.099 | -0.036 | 0.130 | 0.008 | 0.070 |
| Dist to R\_class1 | -0.129 | 0.014 | -0.004 | -0.165 | -0.146 | -0.306 | -0.917 |
| Dist to R\_class2 | -0.113 | -0.036 | -0.021 | -0.084 | -0.244 | -0.215 | -0.587 |
| Dist to R\_class3 | -0.265 | -0.257 | -0.141 | -0.283 | -0.214 | -0.197 | -0.278 |
| Dist to R\_class4 | -0.651 | -0.536 | -0.587 | -0.552 | -0.427 | -0.394 | -0.228 |
| Dist to rail stations | 0.002 | -0.020 | 0.069 | 0.024 | -0.136 | -0.141 | -0.289 |
| Num households | 0.001 | -0.016 | -0.079 | -0.098 | -0.131 | -0.082 | -0.086 |
| Mean housing price | 0.027 | 0.029 | 0.074 | 0.037 | 0.096 | 0.038 | -0.170 |
| Mean land price | 0.079 | 0.127 | 0.083 | 0.191 | -0.054 | 0.250 | 0.198 |
| Employment potential | -0.174 | -0.098 | -0.027 | 0.001 | 0.105 | 0.205 | 0.236 |
| Richness index | 0.159 | 0.032 | 0.057 | -0.076 | -0.043 | -0.371 | -0.305 |
| Gross population density | 0.311 | 0.161 | 0.210 | 0.213 | 0.257 | -0.076 | 0.045 |
| Net population density | -0.362 | -0.364 | -0.451 | -0.233 | -0.078 | 0.120 | -0.070 |
| Zoning status | 2.735 | 3.639 | 3.745 | 3.278 | 2.942 | 2.807 | 3.775 |

**Table 5.** The OR value for *X* (class0 is the reference class).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Class1 | Class2 | Class3 | Class4 | Class5 | Class6 | Class7 |
| Elevation | 0.925 | 0.969 | 1.043 | 1.111 | 1.023 | 1.226 | 1.117 |
| Slope | 0.789 | 0.925 | 0.811 | 0.525 | 0.500 | 0.432 | 0.306 |
| Dist to city | 1.074 | 1.053 | 0.906 | 0.965 | 1.139 | 1.008 | 1.072 |
| Dist to R\_class1 | 0.879 | 1.014 | 0.996 | 0.848 | 0.864 | 0.736 | 0.400 |
| Dist to R\_class2 | 0.893 | 0.964 | 0.980 | 0.920 | 0.783 | 0.807 | 0.556 |
| Dist to R\_class3 | 0.767 | 0.774 | 0.869 | 0.754 | 0.807 | 0.822 | 0.758 |
| Dist to R\_class4 | 0.521 | 0.585 | 0.556 | 0.576 | 0.653 | 0.674 | 0.796 |
| Dist to rail stations | 1.002 | 0.980 | 1.071 | 1.024 | 0.873 | 0.868 | 0.749 |
| Num households | 1.001 | 0.984 | 0.924 | 0.907 | 0.877 | 0.922 | 0.918 |
| Mean housing price | 1.027 | 1.029 | 1.076 | 1.038 | 1.101 | 1.039 | 0.843 |
| Mean land price | 1.082 | 1.136 | 1.087 | 1.210 | 0.947 | 1.284 | 1.219 |
| Employment potential | 0.840 | 0.907 | 0.974 | 1.001 | 1.111 | 1.228 | 1.266 |
| Richness index | 1.172 | 1.033 | 1.059 | 0.927 | 0.958 | 0.690 | 0.737 |
| Gross population density | 1.365 | 1.175 | 1.233 | 1.237 | 1.293 | 0.927 | 1.046 |
| Net population density | 0.696 | 0.695 | 0.637 | 0.792 | 0.925 | 1.128 | 0.932 |
| Zoning status | 15.405 | 38.050 | 42.317 | 26.523 | 18.952 | 16.555 | 43.598 |

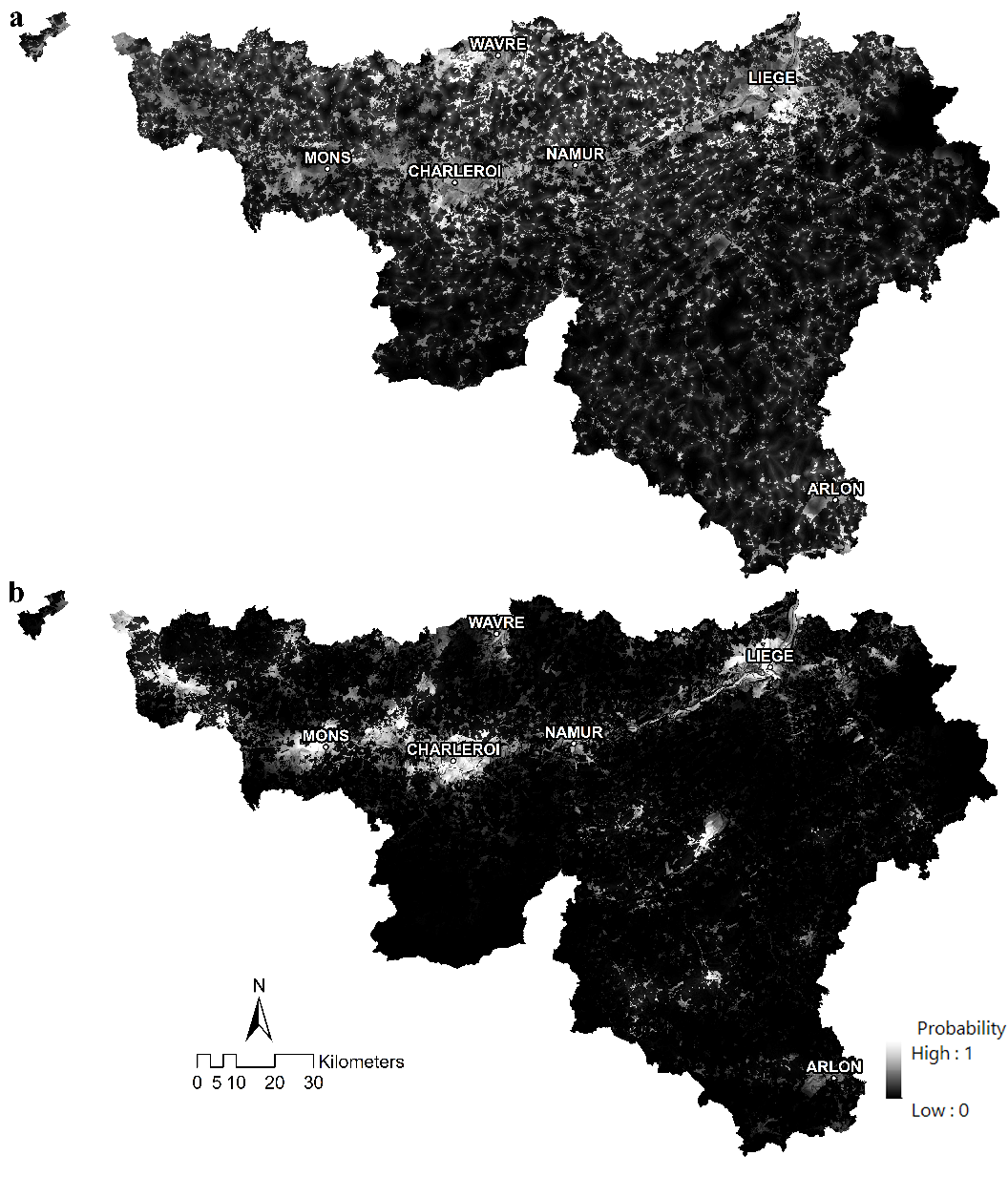
We found that the likelihood of urban development is markedly influenced by policies (zoning status). Zoning status has the strongest impact on urban development within all urban classes. Slope, distance to R\_class4, distance to R\_class3, net/gross population densities and mean land price respectively also demonstrate an impact on all urban classes, but far less important than zoning status. The result shows that the impact of slope is generally increasing with built-up densities. Quite logically dense urban projects are rather developed in flat areas, in many cases floodplains in Wallonia. In contrast, the impact of distance to R\_class4, namely intra-urban or inter-villages roads, is generally decreasing with built-up densities. Distances to R\_class1 and R\_class2 have a noticeable impact on the development of high density projects (class7) with OR of 0.40 and 0.56 respectively. That indicates a strong relationship between highest urban density class and proximity to high speed roads. It should be stressed at this respect that a number of urban cores are directly accessible via high-speed roads in Wallonia. Employment potential has a significant attraction impact on urban class7. It is generally increasing with density, which is what can be expected. Richness index and elevation have moderate impacts on urban class6. Distance to rail stations has a moderate positive influence on urban class7. Still this influence is much lower than the proximity of high-speed roads, suggesting that urban areas located nearby train stations are not yet sufficiently attractive for new dense urban developments in Wallonia. This should be a major source of concern for urban policy makers. Quite significantly, mean housing price represents a low influence on urban growth. This influence is negative for high density developments, which is another source of concern given the shortage of available housing, especially apartments, in areas characterized by a strong pressure on the real estate market.

Our assumption regarding general identification of different urban classes can be assessed by measuring the influence of land-use policies on urban development. OR values for zoning show that policy has a very strong impact on the highest density developments (class7). Those high density developments will most naturally be developed in areas where the legally-binding plan allows such developments, in order to minimize the administrative and financial risks of such operations. Further on they are often located in dense urban areas where the existing plan already allocates a significant share of the land to urban development and where non-urban zones (parks, green areas, etc) are strongly protected for environmental, social and/or heritage reasons. Policies impact is taken a downward trend with classes 4, 5 and 6 respectively. We consider those classes as suburbs. Quite understandably that urban developments in suburbs do not strictly follow policies. The impact of policy on class1 is very low compared to other classes. This class can be considered as remote developments, consisting in scattered constructions (1 to 3 buildings per ha), which can sometimes deviate from existing zoning plans especially in agricultural zones. Land-use policies also show a noticeable impact on classes 2 and 3. We considered those both classes as low density developments in rural areas. It is not surprising that new developments are mainly directed to urbanizable zones, where there is an excess supply of such land. Fig. 4 presents urban development probability maps for classes 1 and 7 as examples of MNL model outcomes.

1. Conclusions

Urban development process in Wallonia is dynamic and diverse. The prediction of spatial distribution of such development can be modelled based on a set of geo-physical and socioeconomic attributes that represent proximate causes of urban growth. Considering urban development as a continuum allows us to better understand the interactions between different drivers and different urban densities.

In this paper, we examined the driving forces of urban development process in Wallonia over a period of 10 years (2000 to 2010). Multinomial logistic regression model was employed to relatively measure the impact of different drivers on probability of urban development. Sixteen drivers were selected from four sets of driving forces including geo-physical features, land-use policies, socio-economic and accessibility. Generally, result reveals that policies and accessibility are the most important determinants of urban growth process. Most importantly, our results highlight that the impact of different drivers varies along with urban density. This is especially the case of land planning driver, whose effects are much more significant for smaller densities than for higher ones, except in the case of urban cores. This study findings can support urban growth modeling, urban planning, and decision-making process in identifying urban development likelihood for each location in the present and medium term.



**Fig. 4.** MNL probability maps for: **a.** class1 (scattered), **b.** class7 (generally located in urban zones)

**Acknowledgments.** The research was funded through the ARC grant for Concerted Research Actions, financed by the Wallonia-Brussels Federation.

**References**

1. Arnfield, A.J.: Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. Int. J. Climatol. 23, 1–26 (2003).

2. Xian, G., Crane, M.: Assessments of urban growth in the Tampa Bay watershed using remote sensing data. Remote Sens. Environ. 97, 203–215 (2005).

3. Li, X., Zhou, W., Ouyang, Z.: Forty years of urban expansion in Beijing: What is the relative importance of physical, socioeconomic, and neighborhood factors? Appl. Geogr. 38, 1–10 (2013).

4. Maimaitijiang, M., Ghulam, A., Sandoval, J.S.O., Maimaitiyiming, M.: Drivers of land cover and land use changes in St. Louis metropolitan area over the past 40 years characterized by remote sensing and census population data. Int. J. Appl. Earth Obs. Geoinformation. 35, Part B, 161–174 (2015).

5. Mustafa, A., Saadi, I., Cools, M., Teller, J.: Measuring the Effect of Stochastic Perturbation Component in Cellular Automata Urban Growth Model. Procedia Environ. Sci. 22, 156–168 (2014).

6. Puertas, O.L., Henríquez, C., Meza, F.J.: Assessing spatial dynamics of urban growth using an integrated land use model. Application in Santiago Metropolitan Area, 2010–2045. Land Use Policy. 38, 415–425 (2014).

7. Wu, F.: Calibration of stochastic cellular automata: the application to rural-urban land conversions. Int. J. Geogr. Inf. Sci. 16, 795–818 (2002).

8. Munshi, T., Zuidgeest, M., Brussel, M., van Maarseveen, M.: Logistic regression and cellular automata-based modelling of retail, commercial and residential development in the city of Ahmedabad, India. Cities. 39, 68–86 (2014).

9. Bičı́k, I., Jeleček, L., Štěpánek, V.: Land-use changes and their social driving forces in Czechia in the 19th and 20th centuries. Land Use Policy. 18, 65–73 (2001).

10. Serneels, S., Lambin, E.F.: Proximate causes of land-use change in Narok District, Kenya: a spatial statistical model. Agric. Ecosyst. Environ. 85, 65–81 (2001).

11. Verburg, P.H., Schot, P.P., Dijst, M.J., Veldkamp, A.: Land use change modelling: current practice and research priorities. GeoJournal. 61, 309–324 (2004).

12. QUAN, B., CHEN, J.-F., QIU, H.-L., RÖMKENS, M.J.M., YANG, X.-Q., JIANG, S.-F., LI, B.-C.: Spatial-Temporal Pattern and Driving Forces of Land Use Changes in Xiamen. Pedosphere. 16, 477–488 (2006).

13. Braimoh, A.K., Onishi, T.: Spatial determinants of urban land use change in Lagos, Nigeria. Land Use Policy. 24, 502–515 (2007).

14. Poelmans, L., Van Rompaey, A.: Detecting and modelling spatial patterns of urban sprawl in highly fragmented areas: A case study in the Flanders–Brussels region. Landsc. Urban Plan. 93, 10–19 (2009).

15. Liu, C., Ma, X.: Analysis to driving forces of land use change in Lu’an mining area. Trans. Nonferrous Met. Soc. China. 21, Supplement 3, s727–s732 (2011).

16. Hallowell, G.D., Baran, P.K.: Suburban change: A time series approach to measuring form and spatial configuration. J. Space Syntax. 4, 74–91 (2013).

17. Shu, B., Zhang, H., Li, Y., Qu, Y., Chen, L.: Spatiotemporal variation analysis of driving forces of urban land spatial expansion using logistic regression: A case study of port towns in Taicang City, China. Habitat Int. 43, 181–190 (2014).

18. Camagni, R., Gibelli, M.C., Rigamonti, P.: Urban mobility and urban form: the social and environmental costs of different patterns of urban expansion. Ecol. Econ. 40, 199–216 (2002).

19. Cammerer, H., Thieken, A.H., Verburg, P.H.: Spatio-temporal dynamics in the flood exposure due to land use changes in the Alpine Lech Valley in Tyrol (Austria). Nat. Hazards. 68, 1243–1270 (2013).

20. Brueckner, J.K.: Lectures on Urban Economics. MIT Press (2011).

21. Belgian Federal Government: Population, http://statbel.fgov.be/fr/modules/publications/statistiques/population/population\_-\_chiffres\_population\_1990-2010.jsp.

22. Verhetsel, A., Thomas, I., Beelen, M.: Commuting in Belgian metropolitan areas: The power of the Alonso-Muth model. J. Transp. Land Use. 2, (2010).

23. Antrop, M.: Landscape change and the urbanization process in Europe. Landsc. Urban Plan. 67, 9–26 (2004).

24. Lambin, E.F.: Modelling Deforestation Process - A review - Trees Tropical Ecosystem Environment Observations by Satellites -. European Commission Luxembourg (1994).

25. Clark, W.A.V., Hosking, P.L.: Statistical Methods for Geographers. Wiley, New York (1986).

26. Lin, Y., Deng, X., Li, X., Ma, E.: Comparison of multinomial logistic regression and logistic regression: which is more efficient in allocating land use? Front. Earth Sci. 1–12 (2014).

27. Jr, D.W.H., Lemeshow, S.: Applied Logistic Regression. John Wiley & Sons (2004).

28. Poelmans, L.: Modelling urban expansion and its hydrological impacts, (2010).

29. Poelmans, L., Van Rompaey, A.: Complexity and performance of urban expansion models. Comput. Environ. Urban Syst. 34, 17–27 (2010).

30. Belsley, D.A., Kuh, E., Welsh, R.E.: Regression Diagnostics. John Wiley and Sons, New York (1980).

31. Judge, G.G., Griffiths, W.E., Hill, R.C., Lütkepohl, H., Lee, T.-C.: The Theory and Practice of Econometrics. Wiley, New York (1985).

32. Belsley, D.A.: Conditioning diagnostics. Wiley Online Library (1991).

33. Kennedy, P.: A Guide to Econometrics. MIT Press (2003).

34. Flom, P.L.: Multinomial and ordinal logistic regression using PROC LOGISTIC. NESUG. NESUG, Baltimore (2010).

35. Knight, J.F., Lunetta, R.S.: An experimental assessment of minimum mapping unit size. IEEE Trans. Geosci. Remote Sens. 41, 2132–2134 (2003).

36. Saura, S.: Effects of minimum mapping unit on land cover data spatial configuration and composition. Int. J. Remote Sens. 23, 4853–4880 (2002).

37. Tannier, C., Thomas, I.: Defining and characterizing urban boundaries: A fractal analysis of theoretical cities and Belgian cities. Comput. Environ. Urban Syst. 41, 234–248 (2013).

38. Belgian Federal Government: Statistics Belgium, http://statbel.fgov.be/fr/statistiques/chiffres/.

39. Economidou, M., Atanasiu, B., Despret, C., Maio, J., Nolte, I., Rapf, O.: Europe’s buildings under the microscope. Buildings Performance Institute Europe (BPIE), Brussels (2011).

40. Jenks, M., Dempsey, N.: Future Forms and Design for Sustainable Cities. Routledge (2005).

41. Institut wallon de l’évaluation, de la prospective et de la statistique: Statistiques, http://www.iweps.be/themes-page.

42. Roy Chowdhury, P.K., Maithani, S.: Modelling urban growth in the Indo-Gangetic plain using nighttime OLS data and cellular automata. Int. J. Appl. Earth Obs. Geoinformation. 33, 155–165 (2014).