Geography and Sustainability 6 (2025) 100276

Contents lists available at ScienceDirect

Geography and Sustainability

journal homepage: www.elsevier.com/locate/geosus

Research Article

Model's parameter sensitivity assessment and their impact on Urban Densification using regression analysis

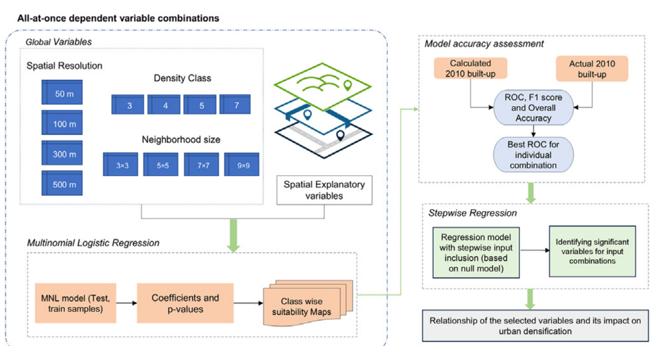
Anasua Chakraborty[a,*], Mitali Yeshwant Joshi[a,b], Ahmed Mustafa[c], Mario Cools[a], Jacques Teller[a]

[a] Local Environment Management and Analysis (LEMA), Urban and Environmental Engineering, University of Liege, Liege, Belgium
[b] Copernicus Institute of Sustainable Development, Utrecht University, Utrecht, the Netherlands
[c] Urban Systems Lab, The New School, New York, USA

HIGHLIGHTS

- A global sensitivity analysis of 48 input combinations impacting urban densification.
- Method combines multinomial logistic regression and stepwise model for the analysis.
- Zoning and population density identified as pivotal in shaping urban densification.
- 50 and 100 m spatial resolutions helps capture densification at a provincial scale.
- This study ecommends decentralization over conservative governance through local intervention.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 30 March 2024
Received in revised form 29 November 2024
Accepted 30 December 2024
Available online 8 February 2025

Keywords:
Urban densification
Sensitivity analysis
Multinomial logistic regression
Stepwise regression

ABSTRACT

The impact of different global and local variables in urban development processes requires a systematic study to fully comprehend the underlying complexities in them. The interplay between such variables is crucial for modelling urban growth to closely reflects reality. Despite extensive research, ambiguity remains about how variations in these input variables influence urban densification. In this study, we conduct a global sensitivity analysis (SA) using a multinomial logistic regression (MNL) model to assess the model's explanatory and predictive power. We examine the influence of global variables, including spatial resolution, neighborhood size, and density classes, under different input combinations at a provincial scale to understand their impact on densification. Additionally, we perform a stepwise regression to identify the significant explanatory variables that are important for understanding densification in the Brussels Metropolitan Area (BMA). Our results indicate that a finer spatial resolution of 50 m and 100 m, smaller neighborhood size of 5×5 and 3×3, and specific density classes—namely 3 (non-built-up, low and high built-up) and 4 (non-built-up, low, medium and high built-up)—optimally explain and predict urban densification. In line with the same, the stepwise regression reveals that models with a coarser resolution of 300 m lack significant variables, reflecting a lower explanatory power for densification. This approach aids in identifying optimal and significant global variables with higher explanatory power for understanding and predicting urban densification. Furthermore, these findings are reproducible in a global urban context, offering valuable insights for planners, modelers and geographers in managing future urban growth and minimizing modelling.

* Corresponding author.
E-mail addresses: a.chakraborty@uliege.be (A. Chakraborty), m.y.joshi@uu.nl (M.Y. Joshi), a.mustafa@newschool.edu (A. Mustafa), mario.cools@uliege.be (M. Cools), jacques.teller@uliege.be (J. Teller).

https://doi.org/10.1016/j.geosus.2025.100276

1. Introduction

The 21st-century wave of urbanization presents significant challenges and opportunities for sustainable development, with impacts markedly varying between the Global South, developing nations, and developed countries (UN-HABITAT, 2020). Rapid population growth in cities of the Global South and developing countries, such as Mumbai and São Paulo, contrasts with efforts in developed regions of Europe and North America to retrofit urban centers to meet sustainability standards without compromising growth (Waddell, 2002; Barredo et al., 2003; Seto et al., 2012; Wu et al., 2021). However, in developing countries, urbanization is a driver for economic growth, while also being the case for strain on the already existing infrastructure, necessitating innovative solutions for sustainable urban development (Todes, 2012). This global demographic shift towards urban living necessitates an in-depth understanding of urban densification, integrating socio-economic dynamics, environmental considerations and spatial planning policies to navigate the complexities of urban development across diverse contexts.

Urban densification, characterized, in our study, by the intensification of built-up for a given land unit, is a key focus in many urban modelling studies (Claassens et al., 2020; Chakraborty et al., 2022a). Despite the critical role of urban densification in shaping the future of sustainable cities, our knowledge of the interplay between socio-economic factors, planning policies and environmental sustainability across varied urban contexts remains minimal. A considerable gap exists in how models handle the variability and sensitivity of urban systems with varying input factors, which is critical for accurate predictions and successful policy interventions (Pijanowski et al., 2006; Franczyk, 2019).

Seminal works have advanced our understanding of urban modelling but revealed limitations. For instance, Dong et al. (2019) and Sun et al. (2020) highlighted the importance of model parameter selection and optimization but did not fully address dynamic urban growth or the interaction between local features and policies. Similarly, studies by Crooks et al. (2008), Shafizadeh-Moghadam et al. (2017) and Cuellar and Perez (2023) examine the role of neighbourhood sizes and spatial resolution in land transformation and segregation models but lack comprehensive consideration of multiple neighbourhood combinations and their suitable application for studying urban resilience. While Puente-Sotomayor et al. (2021) explore sensitivity analysis in disaster risk contexts, broader applications in urban planning remain underexplored. Despite these advances, gaps persist in comprehensively analysing the sensitivity of multi-input combinations in predictive urban modelling. To address this, our study conducts an all-at-a-time sensitivity analysis of global variables at a provincial scale, evaluating the strengths and weaknesses of different input combinations. Additionally, we include a multi-density approach, considering different density class combination to leverage its capability in capturing urban densification.

Multinomial logistic regression (MNL) is efficient at modelling categorical outcomes in urban areas including multi-level built-up density classes, enabling the estimation of outcome probabilities (Cao et al., 2020; Kantakumar et al., 2020; Chakraborty et al., 2022a). Our approach allows for a detailed examination of how various factors influence urban densification patterns (Hu and Lo, 2007; Yang et al., 2023). Complementing MNL, stepwise regression refines the model by selecting variables that significantly impact these outcomes, enhancing predictive accuracy by addressing multicollinearity and overfitting (Wang et al., 2016; Cortés-Borda et al., 2022). This combination of MNL and stepwise regression, in our study, systematically identifies and analyses key variables that drive urban land use changes. This strategy not only improves model reliability but also deepens our understanding of urban development mechanisms, distinguishing our work from previous studies that have primarily relied on traditional regression techniques or machine learning algorithms (Hu and Lo, 2007; Kantakumar et al., 2020; Cortés-Borda et al., 2022).

In contextualizing our study within a global framework, we scrutinize urban densification processes across a transboundary region in Belgium, showcasing varied strategies to manage multi-density expansion. Therefore, our study aims to equip policymakers, urban planners, and modellers with insights by elucidating the intricacies of complex urban densification.

These strategies range from enhancing accessibility and transport infrastructure to emphasizing socio-economic development. Our investigation, in line with the aforementioned aim, addresses the following research questions which include: (a) the evolution of the relative importance of explanatory variables across different input combinations of spatial resolution, (b) variations in the model's explanatory and predictive capabilities with changes in input variable combinations, and (c) differences in significant variables compared to those in the full model under various input criteria.

2. Materials and methods

2.1. Study area

Belgium, Europe's most urbanized nation with a population density of 375 inhabitants/km^2, has experienced significant urban growth since the 1820s (Giovanni et al., 2020). The country comprises three distinct regions—Flanders, Wallonia, and the Brussels Capital Region (BCR). Flanders and the BCR are characterized by dense urban centers, while Wallonia is dominated by rural and fragmented landscapes, creating a unique "rurban" mosaic. Our study focuses on the functional urban areas surrounding Brussels, the capital city, which encompass the Brussels Capital Region (161 km^2), the Flemish Brabant (2118 km^2), and the Walloon Brabant (1097 km^2). This area serves as a compelling case study due to its diverse mix of urban and rural features, reflecting the broader challenges of urbanization and land use planning (Deboosere et al., 2009; De Maesschalck et al., 2015; Guyot et al., 2021). Prominent cities within this region include Brussels, with approximately 1.2 million inhabitants; Leuven, with 103,009 inhabitants; and Nivelles, with 23,464 inhabitants (Fig. 1).

The region has witnessed a population shift from city centres to peripheral areas since the early 1960s, leading to urban sprawl and significant land take proportions (Seto et al., 2012). This phenomenon has implications for socio-economic development and environmental sustainability, highlighting the importance of strategies such as urban densification to mitigate adverse impact of expansion. The decentralization policies, characterized by significant autonomy at a regional and municipal level, have profoundly impacted urban development trajectories and are evidenced by rapid urban growth—for residential and commercial use. In the region of Flemish Brabant, for instance, devolving the planning authority led to fragmented suburbanization and low-density urban sprawl, as the regional government prioritized suburban development through regulations that encourage low-density housing and the expansion of road infrastructure (Boussauw and Boelens, 2015). On the other hand, the local authorities in the region of Walloon Brabant have opted for a different route that goes hand in hand with environmental conservation. They emphasize controlling urban sprawl by promoting higher-density development by utilizing existing public transportation, particularly around university towns such as Louvain-la-Neuve, where development is around a central urban core (Halleux et al., 2012). Coming to the Brussels-Capital Region, coordinated efforts between regional authorities and stakeholders have made it a success story for decentralization. Such diverse approaches by these regional governments showcase how regional development can be made to satisfy local needs (Boussauw et al., 2013). On the contrary, despite decentralization, parts of Belgium still suffer from sprawls requiring a more sophisticated planning strategy. Thus, our study contributes valuable insights to urban planning and policy-making efforts for sustainable urbanization (Poelmans and Van Rompaey, 2009).

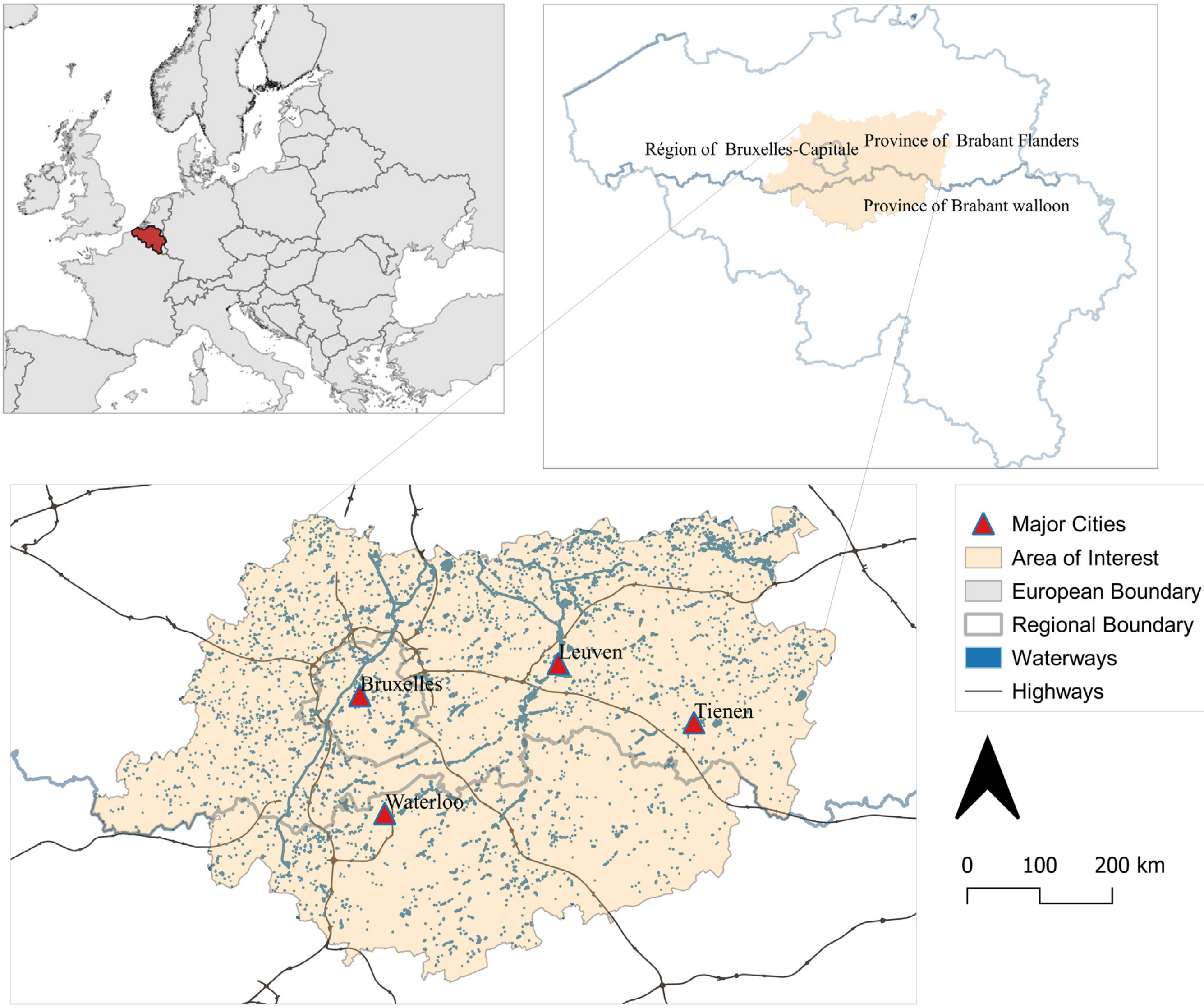


Fig. 1. Study area highlighting the three cross-border provinces of Belgium.

2.2. Built-up density mapping

In the pursuit of comprehensive urban modelling for Brabant, the integration of diverse datasets becomes imperative. Our approach involves crafting urban built-up maps through meticulous exploration of cadastral data sourced from the Belgian Land Registry. Specifically, we focus on the attribute of "construction year" to trace the evolution of the urban landscape across three years: 2000, 2010, and 2020. The data are then rasterized at a fine resolution of 2 m × 2 m per cell. This was adopted following the study previously done by Mustafa et al. (2018a). Each raster cell is assigned a density value, indicating the number of 2 m × 2 m cells enclosed within its boundary. This density value serves as the Built-Up Density (BDC) index for each cell, aggregated at different cell scaling factors of 25, 50, 150, 250, respectively. For instance, to define built-up areas at a 100 m × 100 m resolution, we establish a minimum density threshold of 25 cells—representing a building of 100 m^2 (Mustafa et al., 2018b; Chakraborty et al., 2023). This threshold has been determined, aligning with the average size of residential buildings in Belgium (Poelmans and Van Rompaey, 2010; Chakraborty et al., 2022b). However, the threshold changes while classifying for different resolutions. Hereto, we derive the different combinations of density classes using the geometric interval classification method. This method sets the ranges for density classes and works well with heavily skewed data that are not normally distributed (Arlinghaus and Kerski, 2013).

Table 1
Class (row) to class (column) changes (% of the reference class).

Years		Class-0	Class-1	Class-2	Class-3
2000–2010	Class-0	–	1469 (0.70 %)	1147 (0.54 %)	416 (0.20 %)
	Class-1	–	–	2419 (8.10 %)	170 (0.57 %)
	Class-2	–	–	–	3363 (5.06 %)
	Class-3	–	–	–	–
2010–2020	Class-0	–	1023 (0.49 %)	858 (0.41 %)	326 (0.16 %)
	Class-1	–	–	2544 (5.37 %)	148 (0.52 %)
	Class-2	–	–	–	2734 (5.41 %)
	Class-3	–	–	–	–

Table 1 provides insights into the actual built-up transitions observed over the modelled period, categorised into four density classes: class 0 (non-urban), class 1 (low density), class 2 (medium density), and class 3 (high density). These transitions underscore the dynamics of urban development—particularly the emergence of low-density and

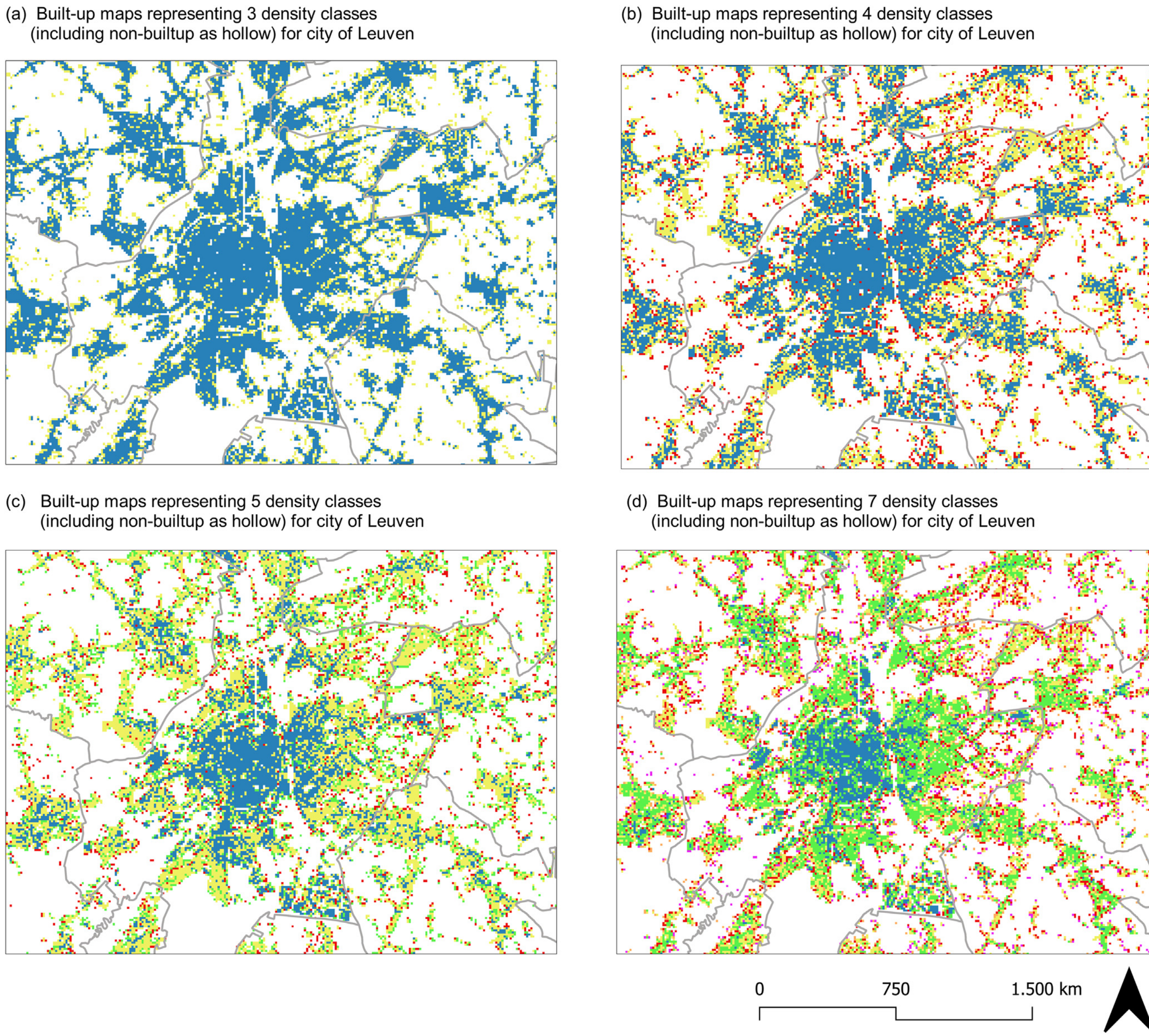


Fig. 2. Parts of Brussels classified under different urban built up density classes.

medium-density areas, which represent the predominant built-up patterns. Comparing our density classification with previous studies, such as Mustafa et al. (2018b) reveals the significance of transitions from class 1 to 2 and class 2 to 3 in understanding the built-up processes within the study area. We gather valuable insights into the evolving urban landscape by scrutinising transitions from non-built-up to high-density classes.

In this study, we keenly analyse what happens if we shift from traditionally used four density classes and further divide our datasets into different density classes with three, four, five and seven density classes, as shown in Fig. 2. This is because data classification techniques impact modelling outcomes (Shu et al., 2020; Chakraborty et al., 2023).

2.3. Potential explanatory variables

The explanatory variables influencing urban growth are presented in Table 2. The average number of built-up cells in the neighbourhood is included in our study to consider local interaction effects. We compute the statistical measure, i.e., the average for input cells within overlapping neighbourhood sizes. Socio-economic indicators such as number of households, job density, average household income and mean household size, are sourced from the Belgian Statistical Institute. Slope, derived from a 1 m × 1 m resolution Digital Elevation Model (DEM), is chosen for its relevance and collinearity with the DEM itself. The distance to different road categories, including motorways, primary roads, secondary roads, residential roads, and local roads, are extracted from OpenStreetMap (OSM) data (Chakraborty et al., 2023). Within the urban fabric, it is important to note that local roads typically connect neighbourhood areas, facilitating intra-community mobility. On the other hand, residential roads primarily provide access to residential properties. Furthermore, proximity to cities and urban green spaces, which are crucial for enhancing livability and environmental sustainability, are considered. These distance-based factors are calculated based on Euclidean distance, a commonly utilised metric in land-use change models (Mustafa et al., 2018b).

Unlike most studies on land-use modelling (Achmad et al., 2015; Cao et al., 2020; Karimi et al., 2021), we include cost path distance instead of Euclidean distance for the point vector datatype of bus and train stops. While Euclidean distance is simpler and computationally effective, it fails to capture the various factors associated with accessibility

Table 2
List of selected explanatory variables.

Variable	Name	Type[x]	Unit	Resolution[y]	Source[z]
D1	Slope	1	%	a	DEM
D2	Euclidean distance to Motorways	1	m	a	OSM
D3	Euclidean distance to Primary Roads	1	m	a	OSM
D4	Euclidean distance to Secondary Roads	1	m	a	OSM
D5	Euclidean distance to Residential Roads	1	m	a	OSM
D6	Euclidean distance to Local Roads	1	m	a	OSM
D7	Euclidean distance to Major Cities	1	m	a	OSM
D8	Euclidean distance to Urban Green Space	1	m	a	OSM
D9	Cost Distance to Railways Stations	1	m	a	OSM
D10	Cost Distance to Bus Stops	1	m	a	OSM
D11	Total Number of Households	1	Number	c	SI
D12	Average size of housing	1	Number	c	SI
D13	Population Density	1	Inh/km^2	d	SI
D14	Jobs density	1	Num/100m^2	b	SI
D15	Net taxable household income	1	€	c	SI
D16	Zoning status	2	Binary	a	SI
D17	Average number of urban built-up cells in the neighbourhood	1	Number	a	CD

[x] 1, Continuous; 2, Categorical.
[y] a, Cell level; b, Municipal level; c, Statistical level; d, Grid level = 1 km × 1 km.
[z] DEM, Calculated from Digital Elevation Model; OSM, Open Street Maps; SI, Self-calculation based on Belgian Statistical Institute; CD, Self-calculation based on built-up Cadastral data.

from one point to another. On the other hand, cost path distance considers the heterogeneity of the surface between two points representing a near-realistic scenario. Though computationally complex, cost path distance fulfils our study requirement—where the surrounding travel behaviour influences densification.

Zoning areas, obtained from the regional zoning plan for all respective regions, serve as a binary variable distinguishing between urbanisable (1) and non-urbanisable areas (0), offering insights into land use regulations and urban growth boundaries.

2.4. Preliminary exploratory data analysis

Our methodology begins with thorough data preparation and preliminary analysis to ensure the reliability and validity of our subsequent regression analysis. As seen from Table 2, all the explanatory variables are measured in different units. Thus, we initiate the process of standardising all the variables by employing the z-score normalisation/standardisation technique. This ensures that each variable is transformed to have a mean (μ) of 0 and a standard deviation (σ) of 1, facilitating comparability and reducing the influence of scale variations. This technique is used when the data that needs to be compared have different scales or units.

For each element i, in a given data, the z-score standardisation follows the below equation:

$$z_i = \frac{x_i - \mu}{\sigma} \tag{1}$$

With this transformation to x_i, it can be easily proven that the required data is now in the standardised form. In our context, having non-comparable data produces biased models, as each data does not contribute to the model equally. Moreover, since logistic regression is a likelihood optimisation problem, it uses iterative algorithms that, without normalised data, can undergo numerical instability.

Furthermore, we conducted a Moran's I analysis (Wang et al., 2023) to examine the effects of aggregation in the context of Modifiable Areal Unit Problem (MAUP) across multiple spatial resolutions: 50 m, 100 m, 300 m and 500 m. In essence, Moran's I compares the value of a variable at one location to the values at neighbouring locations. Nearby areas with similar values indicate spatial clustering of urban patches, while dissimilar values suggest a scattered distribution (Wang et al., 2021a). Moran's I can be defined as:

$$\text{Moran's } I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left(y_i - \bar{y}\right)\left(y_j - \bar{y}\right)}{W \sum_{i+1}^{n} \left(y_i - \bar{y}\right)^2} \tag{2}$$

where n is the total number of elements in the calculation, w_{ij} is the spatial weight between elements i and j, $(y_i - \bar{y})$ is the deviation of the attributes of element i from the average value, and W is the sum of all spatial weights.

A result of the sensitivity analysis indicates a strong autocorrelation at higher resolutions as it is more predominant in the proximity of already urbanised cells (Kantakumar et al., 2020). Moran's I values are lower for 500 m and 300 m, showing the lowest value at 500 m of 0.70, as seen in Fig. 3. To ensure the robustness of our model, the 500 m cell size is discarded from the scope of our analysis as it shows a steep fall from Moran's I of the highest resolution of 50 m. It is also a common practice for land use studies to use spatial resolution between 30 m and 300 m (Hu and Lo, 2007; Liu et al., 2008; Feng et al., 2011; Vermeiren et al., 2012; Hao et al., 2013). A subset of our study area under different spatial resolution ranging from 2 m to aggregated 300 m is demonstrated in Fig. 4.

2.5. Conceptual framework of sensitivity analysis

Sensitivity analysis (SA) is a fundamental pillar enabling a systematic approach to determine the influence of explanatory variables on the model's outcome (Cuellar and Perez, 2023). SA uses two common approaches—one-at-a-time (OAT) and all-at-a-time (AAT). The OAT approach involves varying one model parameter while leaving the others constant, allowing researchers to isolate the effects of individual parameters on model outcomes. Conversely, AAT considers simultaneous variations in multiple parameters, providing a systematic perspective on their combined influence. Both approaches offer unique insights into the sensitivity of urban models and are often used in tandem to comprehensively assess model behaviour (Saltelli et al., 2019; Razavi et al., 2021; Gamboa et al., 2022). By using this analytical framework, researchers may better understand the relative importance of various characteristics and how they affect the complex dynamics of urban growth (Chakraborty et al., 2023). By adopting the AAT approach, we seek to capture the interdependent nature of model parameters. This methodological framework enables an all-encompassing investigation of the interactions between variables. It offers a thorough comprehension of their combined influence on the process of urban densification. A total of 48 permutations, including various combinations of spatial resolutions, density classes, and neighbourhood sizes, were created (Fig. 5).

The selection of spatial resolutions, including 50 m, 100 m and 300 m, allows us to examine the variation of urban built-up density at different levels of detail, from fine-grained local variations to broader regional trends (Fig. 4). Our multi-resolution approach enhances the robustness of our sensitivity study, ensuring thorough and applicable conclusions across varying spatial scales (Kocabas and Dragicevic, 2006; Wang et al., 2021b). Including different density classes is essential to comprehend the effect of multi-level densities on urban modelling. Drawing inspiration from influential studies such as Wang et al. (2021b), our research explores the complex interactions between varying degrees of urban density and the resulting spatial resolutions. Spanning classes 3 to 7, our research encapsulates a broad spectrum of urban landscapes, ranging from sprawling suburbs to compact urban cores. While

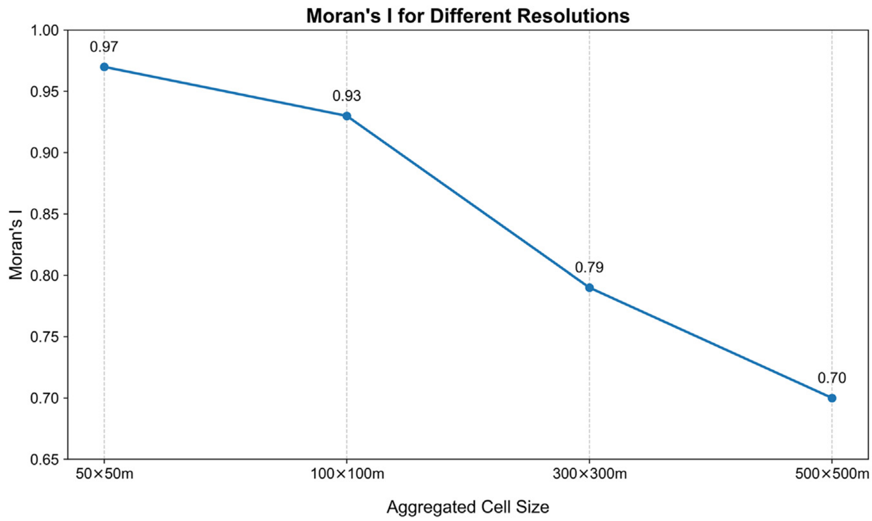


Fig. 3. Effects of spatial autocorrelation on data aggregation based on Moran's *I*.

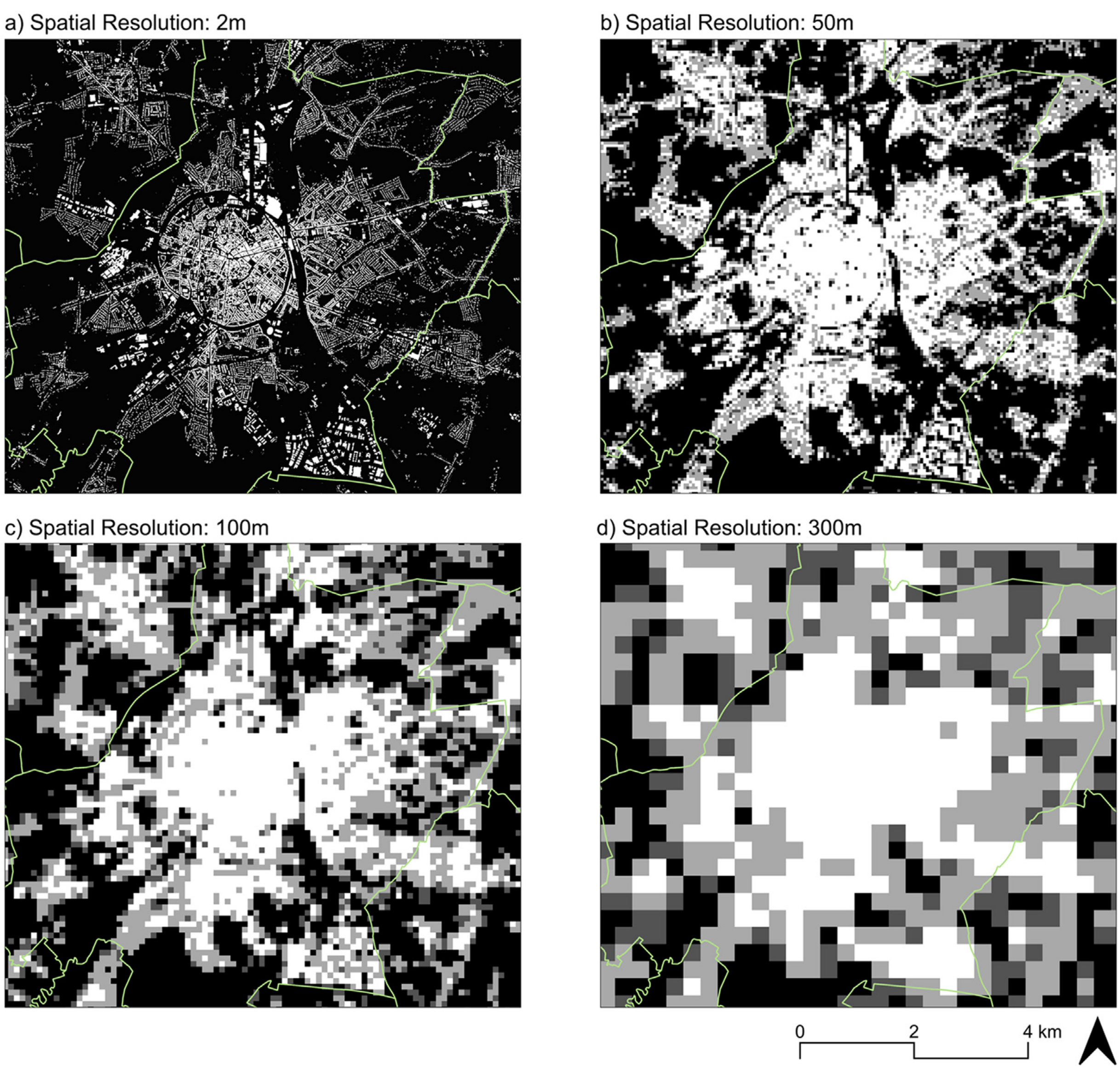


Fig. 4. Area of Brussels Capital Region at four spatial resolutions: (a) 2 m, (b) 50 m, (c) 100 m and (d) 300 m.

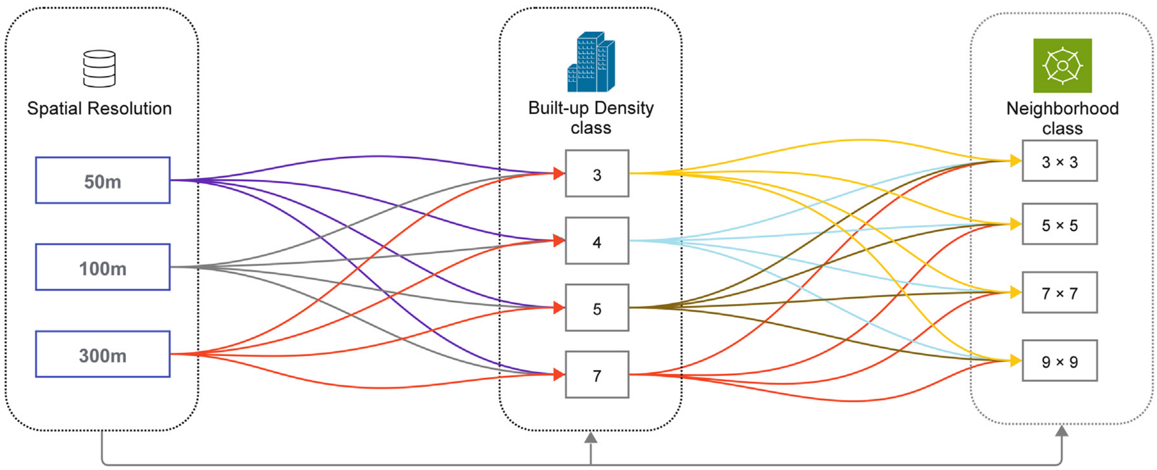


Fig. 5. Illustration of the input combinations used for Sensitivity Analysis.

high-density metropolitan regions struggle with difficulties, including overcrowded public services and infrastructure, low-density places frequently represent unrestrained expansion and resource-intensive development (Shu et al., 2020). This empirical investigation clarifies the complex interactions between density gradients and environmental, socio-economic, and behavioural dynamics. It thus provides important information for the development of evidence-based policies, which are needed for resilient and liveable urban futures.

The neighbourhood parameter refers to the spatial window size used to analyse urban growth patterns. It is key in capturing spatial relationships and dependencies within a given area. In urban growth models, the choice of neighbourhood size significantly impacts the detection of spatial patterns and variations in urban densities. Our model calculates the neighbourhood as the average of the total density classes present in the given window size as a model variable at a *meso* scale. The influence of neighbourhood size on urban growth models has been extensively explored in the literature (Wu, 2002; Poelmans and Van Rompaey, 2009; Chen et al., 2014). Smaller neighbourhood sizes are adept at capturing localised variations in urban densities, while larger sizes offer a broader perspective on regional dynamics (Roodposhti et al., 2020). While many studies debate using weighted regression for studying neighbourhood effects at a local scale (Harris et al., 2010; Lu et al., 2014), understanding neighbourhood dynamics at a global level can be a vantage point to gain a broader perspective of meso-scale trends of urban development. Drawing from the insights from prior research by (Poelmans and Van Rompaey, 2010; Mustafa et al., 2018c; Chakraborty et al., 2022a), it is observed that among various window sizes analysed, the model utilising a 3 × 3 neighbourhood size yielded a land-use pattern that closely resembled the actual pattern. Thus, our study employs a neighbourhood size centred around the central cell incorporating varied sizes, specifically 3 × 3, 5 × 5, 7 × 7, and 9 × 9, to explore their influence on urban growth and densification patterns. Unlike previous studies, which often focus on a single neighbourhood size, our approach encompasses a range of spatial scales to comprehensively capture local interactions and spatial dependencies within a compact spatial extent. The weighted sum value for a given cell is calculated using:

$$S_{\mathrm{w}} = \sum_{i=1}^{N} w_i x_i \tag{3}$$

where, S_{w} is the weighted sum value for the processing block, N is the number of cells in the neighbourhood, w_i represents the weight parameter assigned to cells with different classes at distinct positions and distances, and x_i is the input cell value.

Fig. 6 illustrates the application of a 3 × 3 neighbourhood size centred around the highlighted central cell. This approach enables the examination of the effect of the neighbourhood on the model outcome, i.e., indicating the extent of the model's capability to analyse the impact of neighbourhood dynamics on the probability of new urban development. Through this analysis, we aim to elucidate the spatial relationships between neighbourhood size and urban growth, thereby contributing to a deeper understanding of urban development processes.

2.6. Statement of model

Our study employs an MNL model to examine the relationship between explanatory variables and dependent variables of urban density class with the different input combinations. Even though our dependent variable exhibits natural ordering, we choose an MNL model because the proportional odds assumption test was violated. Prior to modelling, we perform a Variance Inflation Factor (VIF) test for multicollinearity, as using collinear variables in MNL can distort the relationships between predictors and the outcome variable (Getu and Gangadhara Bhat, 2024). Thus far, the model has been evaluated using multiple input combinations, incorporating spatial resolution, density class, and neighbourhood size, to assess its sensitivity to the model's output (Abolhasani et al., 2016; Anputhas et al., 2016; Calka et al., 2017). The general form of MNL can be represented as:

$$\begin{aligned}
\log\left(k_1\right) &= \alpha k_1 + \beta_{k_{11}} X_1 + \beta_{k_{12}} X_2 + \cdots + \beta_{k_{1v}} X_v \\
&\vdots \\
\log\left(k_n\right) &= \alpha k_n + \beta_{k_{n1}} X_1 + \beta_{k_{n2}} X_2 + \cdots + \beta_{k_{nv}} X_v
\end{aligned} \tag{4}$$

where $\log(k_n)$ is the natural logarithm of class k_n versus the reference class k_0, X is a set of explanatory variables $(X_1, X_2, \ldots, X_v)$, αk_n is the intercept term for class k_n versus the reference class, and β is the slopes for the classes (the coefficient vector). Thus, the probabilities of each class are obtained using the following formula:

$$\begin{aligned}
\left(P_c\right)_{ij,Y=k_0} &= \frac{1}{1+\exp\left(\log\left(k_1\right)\right)+\exp\left(\log\left(k_2\right)\right)+\cdots+\exp\left(\log\left(k_n\right)\right)} \\
\left(P_c\right)_{ij,Y=k_1} &= \frac{\exp\left(\log\left(k_1\right)\right)}{1+\exp\left(\log\left(k_1\right)\right)+\exp\left(\log\left(k_2\right)\right)+\cdots+\exp\left(\log\left(k_n\right)\right)} \\
&\vdots \\
\left(P_c\right)_{ij,Y=k_n} &= \frac{\exp\left(\log\left(k_n\right)\right)}{1+\exp\left(\log\left(k_1\right)\right)+\exp\left(\log\left(k_2\right)\right)+\cdots+\exp\left(\log\left(k_n\right)\right)}
\end{aligned} \tag{5}$$

where $\left(P_c\right)_{ij,Y=k_n}$ is the probability of change from the reference class to class k_n occurring in cell *ij*.

MNL outcomes consist of coefficients indicating the relative contribution of each factor to the built-up process (Cao et al., 2020). These coefficients quantify the influence of explanatory variables on transitioning between density classes. Additionally, MNL probability maps

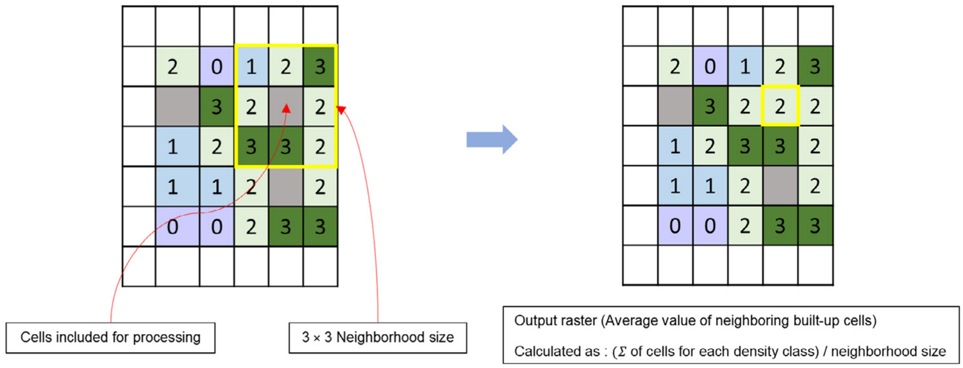


Fig. 6. Example of 3 × 3 neighbourhood size where the highlighted central cell is a result of the average of its neighbourhood cells.

for each density class offer spatial representations of built-up likelihood (Mustafa et al., 2018a).

2.7. Sampling strategy for MNL

While effective in many contexts, the logistic regression method does not inherently account for spatial dependence, crucial in analysing phenomena such as urban growth. Urban growth often exhibits strong spatial autocorrelation, with new urban areas tending to develop within proximity to existing ones. This underscores the importance of carefully designing the sampling scheme when repeatedly applying our model in spatial contexts (Cheng and Masser, 2003; Li et al., 2013). To mitigate potential issues arising from spatial autocorrelation among model residuals, we adopt a hybrid sampling approach, drawing inspiration from the systematic-random sampling technique similar to the one used by Mustafa et al. (2018a).

We employ a random sampling strategy to address potential biases resulting from an imbalanced sample size, where transitions from non-urban to urban categories are relatively infrequent compared to the persistence of non-urban pixels. Initially, we select systematic samples, ensuring an equal number of samples of class 0 with the rest of the density classes. This is done to ensure that the sample accurately reflects land use distribution across the study area. The selection of samples is based on 200 runs of MNL with different random samples for varied density classes.

2.8. Assessment of model performance

In assessing the effectiveness and reliability of our model for predicting urban growth patterns, we utilise a rigorous validation process encompassing multiple metrics (Overmars et al., 2003; Pontius et al., 2004). We randomly select 200 samples from the test data for each density class for the periods of 2000–2010 and 2010–2020. We avoid placing class 0 test samples in order to avoid biases and over-fitting with the samples used in the training process. Overall, the testing sample accounts for 30 % of the total dataset, whereas 70 % is considered the training sample size. The model's Overall Accuracy (OA) is calculated to assess its performance in correctly classifying instances across different density classes. The OA is calculated as the ratio of the sum of True Positives (TP) and True Negatives (TN) to the total number of instances in the test dataset:

$$\mathrm{OA} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{6}$$

where FP and FN are False Positives and False Negatives, respectively.

While overall accuracy provides a comprehensive measure of the model's performance across all classes, it may not adequately capture the performance of the model for each individual class, especially in situations where one class significantly outweighs the others. Hence, we employ the F1 score, which combines precision and recall into a single metric, offering a balanced assessment of the model's performance across different classes. A higher F1 score indicates a better balance between precision and recall, suggesting the model performs well in correctly classifying instances across different classes while minimising false positives and false negatives. The F1 score is calculated as the harmonic mean of precision and recall:

$$\mathrm{F1} = \frac{2 \times \mathrm{TP}}{2 \times \mathrm{TP} + \mathrm{FP} + \mathrm{FN}} \tag{7}$$

2.9. Model validation using ROC

Relative Operating Characteristic (ROC) is used to validate the MNL model. By plotting the true positive rate against the false positive rate at various threshold values, the ROC curve allows planners to assess the model's ability to accurately classify urban and non-urban areas across different settings. This analysis aids planners in understanding the trade-offs between sensitivity and specificity, guiding decisions on the optimal threshold for classifying urban growth patterns. In our study, the implementation of the ROC method serves as a statistical tool to address a key question: "How effectively does urban growth align with locations characterised by elevated suitability for urban development?".

To conduct model validation, we generate urban growth probability maps by plugging in the coefficients of the MNL model. This is then compared with the observed map of 2020 to validate the model's ability to specify the location of change, independently from the predicted amount of change (Fawcett, 2006). The ROC value above 0.5 can refer to a better fit than random, while an ROC value equal to 0.5 represents random fit and below 0.5 is worse (Hu and Lo, 2007). Amongst the multiple runs of MNL, the model with the highest ROC for each spatial resolution comes out as the optimal choice for further analysis. This selection criterion ensures that the chosen model demonstrates the highest accuracy in delineating density class and neighbourhood size (Mustafa et al., 2018c).

2.10. Sequential variable selection for MNL

Stepwise regression serves as a robust methodology for variable selection, especially in scenarios where the number of explanatory variables is large, and their relative importance is uncertain. However, not all variables may significantly contribute to the predictive probability of MNL (Liu et al., 2021). The rationale for choosing Forward Stepwise Regression (FSR) over Backward Stepwise Regression in our methodology lies in its systematic approach to variable selection, especially when dealing with many explanatory variables. FSR starts with an empty model, which is unbiased in nature and incrementally adds variables based on their statistical significance, as indicated by their *p*-values. This method ensures that only variables contributing significantly to

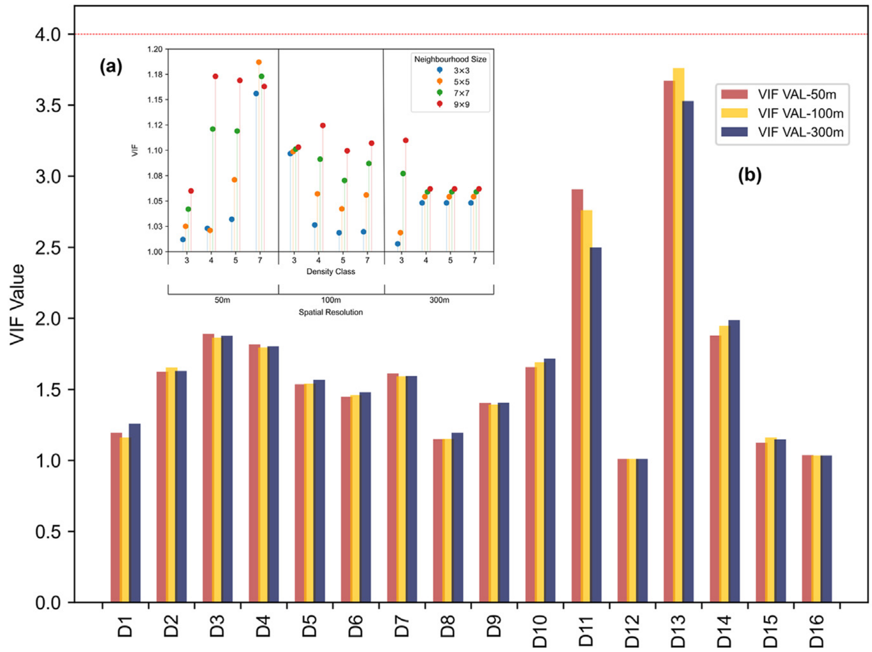


Fig. 7. Variation Inflation Factor (VIF) values are plotted for (a) different density class for spatial resolution of 50 m, 100 m and 300 m, and (b) and for the sixteen model's parameters selected for modelling.

the predictive probability of the model are retained, enhancing its explanatory power. In contrast, Backward Stepwise Regression begins with all variables included and removes them iteratively, which may lead to the exclusion of potentially important variables early in the process. Therefore, we opt for FSR to systematically identify the subset of variables that best explain the variability across different density classes, spatial resolutions, and neighbourhoods for urban growth prediction/ modelling.

3. Results

3.1. Multicollinearity analysis

The results of the VIF test for different spatial resolutions are outlined in Fig. 7. As mentioned earlier, this is conducted to understand the existence of multicollinearity between the independent variables in Table 2. While the conventional threshold of VIF is debatable (James et al., 2021), we opt for a more stringent limit—setting VIF < 4 based on studies for better efficiency (Saganeiti et al., 2021; Chakraborty et al., 2022a). It is important to note that variables such as total number of households (D11) and population density (D13) exhibit a high VIF of approximately 2.8 and 3.5, respectively (Fig. 7(a)). Moreover, we also observe the VIF values across different combinations of spatial resolution, density class and neighbourhood sizes, considered as input variables in our model. Furthermore, it is important to note that there is a steep increase in the VIF value for 50 m resolution, reaching its peak at approximately 1.2 for seven density classes and a 5 × 5 neighbourhood size. On the contrary, for 100 m, a decline is observed in VIF values with an increase in density class and neighbourhood size, while the VIF for 300 m shows a constant value across the density class and neighbourhood combinations.

3.2. Performance metrics assessment for density class and spatial resolution

Accuracy metrices comprehensively evaluate our model's performance across a spectrum of spatial resolutions and density classes. Overall Accuracy (OA) and F1 score serve as indicators of the model's predictive accuracy and ability to capture built-up patterns accurately for the calibration period of 2000–2010, as seen in Table 3.

Table 3
Performance metrics indicated by overall accuracy (OA) and F1 scores.

Spatial resolution	50 m		100 m		300 m	
Accuracy metric	OA	F1	OA	F1	OA	F1
3 density classes	0.880	0.663	0.868	0.652	0.669	0.594
4 density classes	0.875	0.671	0.827	0.627	0.675	0.565
5 density classes	0.871	0.658	0.819	0.602	0.643	0.532
7 density classes	0.811	0.597	0.805	0.594	0.629	0.519

Our model exhibits notable performance at a spatial resolution of 50 m, for 3 and 4 density classes, where it achieves an OA of 0.880 and 0.875, respectively. Although marginal, as the density class increases to 5 and 7 classes, a decline in OA is observed, with values of 0.871 and 0.811. On the other hand, the F1 score for 3 and 4 density class is 0.663 and 0.671, which subsequently drops to 0.597 at 7 density classes. Akin to that, upon transitioning to coarser spatial resolutions of 100 m and 300 m, there is a decrease in OA and F1 scores across all density classes. For instance, at 100 m, the OA for 3 density classes decreases to 0.868, while for 7 classes it falls to 0. 805. Similarly, the F1 scores decline from 0.652 to 0.594 with increased density classes from 3 to 7. At 300 m resolution with 7 density classes, both OA and F1 score deteriorate to a value of 0.629 and 0.519, respectively.

3.3. Relative importance of spatial parameters on model performance

To evaluate the efficacy and stability of our model, we carry out an extensive analysis that made use of all explanatory variables. We estimate urban development probabilities for varying neighbourhood sizes by using calibrated models for the 2000–2010 period. As a part of our validation process, we then compare the observed 2020 urban built-up map with the projected probability maps of 2020 to estimate ROC.

The result, elucidated in Fig. 8, shows that a spatial resolution of 50 m produces a high ROC value across different density classes and neighbourhood sizes. As seen in Fig. 8(a), in density class 3, the ROC values ranged from 0.8823 to 0.8890 across different neighbourhood sizes, with the highest value observed for a 5 × 5 neighbourhood size.

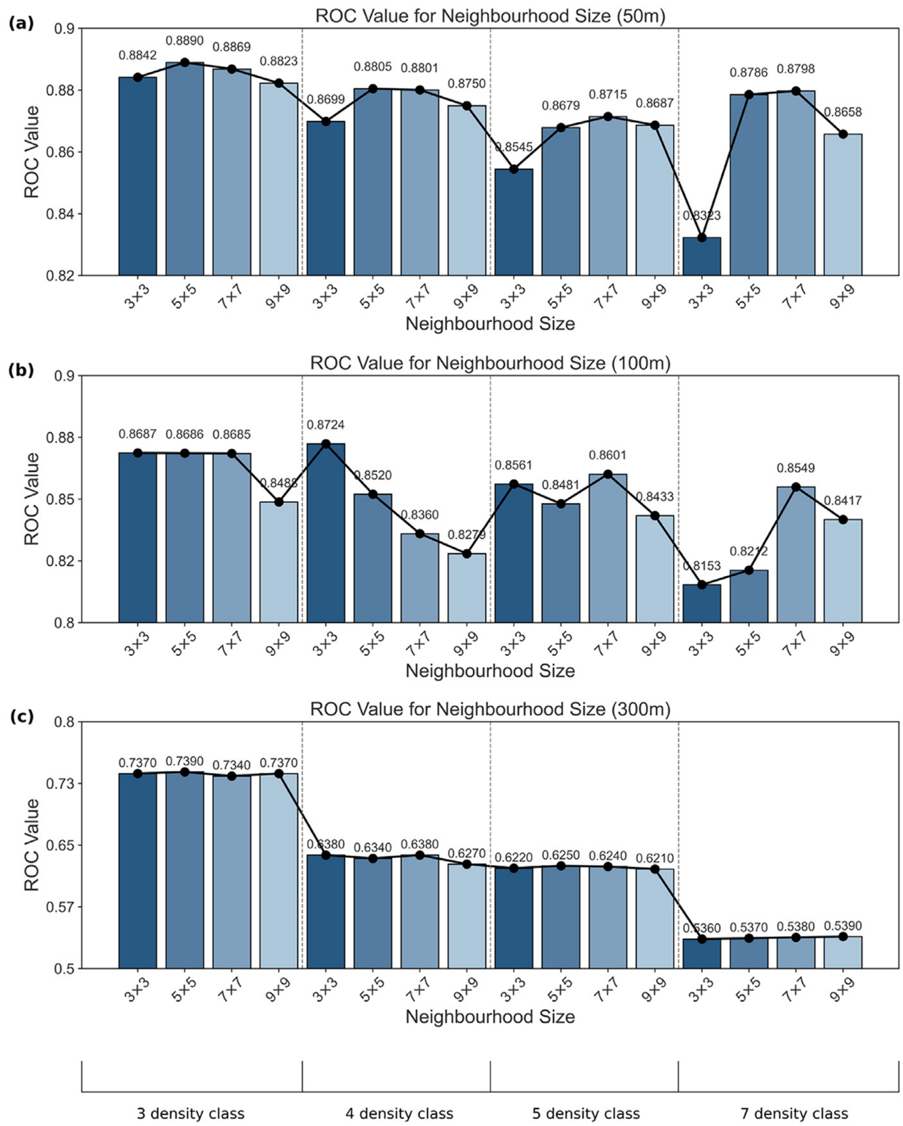


Fig. 8. Accuracy metrics for the model, using multiple inputs to find the best combination. Plotted are the Relative Operating Characteristic (ROC) values for different resolutions of (a) 50 m, (b) 100 m, and (c) 300 m.

Conversely, as the spatial resolution increases to 100 m, we observe a slight decline in overall predictive performance, particularly in areas with lower urban density. Our model produces the highest ROC of 0.8724 with 4 density classes and 3 × 3 neighbourhood size. Furthermore, the coarser resolution of 300 m exhibits pronounced limitations across all density classes and neighbourhood sizes (Fig. 8(c)). A resolution of 300 m poorly captures the densification scenario, which is highly dependent on existing urban development in its vicinity. With three density classes, for instance, the ROC values range from 0.734 to 0.737, indicating a moderate level of predictive accuracy. However, when compared to the ROC values obtained at 50 m and 100 m, the decrease in performance is obvious.

3.4. *Refinement of model complexity based on* p*-values*

In the previous subsection, we observed the result for the MNL fitted with all explanatory variables for different input combinations. These variables are further used to fit the model with a stepwise selection to assess AAT sensitivity. For the stepwise selection of explanatory variables, we adopt a sequential addition of variables having p-values < 0.05 at each step. Examining the outcomes of our stepwise regression, we find that non-significant predictor variables are most yielded at a spatial resolution of 300 m.

Multiple explanatory variables emerge as common variables in both 50 m and 100 m, as seen in Table 4. The presence of these variables underscores their relevance in shaping urban growth patterns for high resolutions. Our results show a correlation between residential roads (D5) and bus stops (D10) at finer spatial resolutions (50 m and 100 m) reflecting the significance of transportation infrastructure in urban development. D11 and D13 indicate that areas with a dense population will increase the demand for housing, driving increased urban growth. D16 and D17 reflect the significance of zoning policies for urban development along all spatial scale and density classes.

Some explanatory variables appear to be scale dependent. For instance, in 100 m calculations, D5 (like cost distance to railway stations (D9)), plays a more dominant role for lower density classes to further develop to medium density. On the other hand, D12 and D14 becomes noticeable especially in 50 m, however, remains not so significant in 100 m. This suggests that the provision of large average size housing will stimulate further densification. The inclusion of D14 in 50 m in-

Table 4
Summary of selected significant explanatory variables based on p-values < 0.05.

Variable	Name	50 m, 3 density classes, 5 × 5 neighbourhood		100 m, 4 density classes, 3 × 3 neighbourhood		
		Class 1: Low density	Class 2: High density	Class 1: Low density	Class 2: Medium density	Class 3: High density
D1	Slope					
D2	Euclidean distance to Motorways					
D3	Euclidean distance to Primary Roads					
D4	Euclidean distance to Secondary Roads					
D5	Euclidean distance to Residential Roads	×	×	×	×	
D6	Euclidean distance to Local Roads	×	×			
D7	Euclidean distance to Major Cities					
D8	Euclidean distance to Urban Green Space					
D9	Cost Distance to Railways Stations				×	
D10	Cost Distance to Bus Stops	×	×	×		
D11	Total Number of Households	×	×	×	×	×
D12	Average size of housing	×				
D13	Population Density	×	×			×
D14	Jobs density	×	×			
D15	Net taxable household income					
D16	Zoning status	×	×	×	×	×
D17	Average number of urban built-up cells in the neighbourhood	×	×	×	×	×

dicated the importance of employment opportunities in driving urban growth at a fine spatial scale.

4. Discussion

4.1. Contribution of variables in modelling

The VIF values (presented in Section 3.1) of the independent variables indicate a moderate level of multicollinearity. However, inclusion of these variables is important for their contribution to study urban densification. These variables reflect the interplay between population distribution and its impact on household formation, pivotal for determining the built-up density of an urban core like Brussels Metropolitan Area. Additionally, for neighborhood size, the values remain lower than the threshold, across all the scenarios investigated, with the majority falling under 4 (Fig. 7(b)). It demonstrates that the model captures more localized effects in higher resolutions, leading to increased dependencies amongst the variables. This also suggests the importance of considering density class and neighborhood size in analyzing input interactions during urban densification modelling. The decline observed in VIF values at a coarser resolutions reflects the model's capability to effectively capture urban built-up variability without having significant multicollinearity. This could be attributed to the smoothening effect at large spatial resolutions, where localized variations are attenuated, resulting in weaker correlations between the variables.

4.2. Performance accuracy of model under multiple input combinations

The overall accuracy and F1 score calculated for model's validation underscores its superior performance in delineating built-up patterns, especially in lower-density regions. This aligns with previous research emphasizing the challenges posed by high-density classes (Mustafa et al., 2018a). Though every model's accuracy is dependent on various parameters like sample sizes or the training versus testing ratio, an F1 score of 0.6 is widely considered to be precise. Since the F1 score shows model's ability to perform a multiclass prediction, a poor F1 score could be due to unavailability of enough samples, resulting in imbalanced data distribution, common for increase in density classes. Interestingly, as observed from Table 3, a constant declining trend emerges across all spatial resolutions, where high-density classes exhibit inferior model performance compared to their lower-density counterparts. This disparity stresses the formidable challenges associated with accurately classifying built-up areas in densely populated urban environments. Moreover, our study area is an amalgamation of rural and urban regions, making it difficult to capture fragmented growth at a coarser resolution. The decline in the model accuracy can also result from model overfitting, where the testing set tends to reflect a pattern too similar to that of the training set. This results in poor prediction between observed maps and simulated maps.

4.3. Delineating model's predictive ability

A model's sensitivity to the change in its variables can be well captured through its performance accuracy. This study employs the ROC value, a metric widely used to ascertain model's ability to capture spatial dependencies and interactions between explanatory variables in studying urban land use patterns. Our results, in Fig. 8, indicate that the model's predictive accuracy depends largely on resolution, density classes and neighborhood size. A high resolution (50 m) complimented with lower density class (3) and small neighborhood size (5 × 5) effectively identifies areas undergoing urban growth, including expanding urban centers and new residential or commercial areas. The model maintains a relatively strong predictive accuracy, suggesting that it can still effectively capture growth patterns, albeit with slightly reduced precision at 100 m. Previous studies by Poelmans and Van Rompaey (2009), Mustafa et al. (2018a), Chakraborty et al. (2022a) also concluded that a spatial resolution of 100 m proves to be efficient at capturing fine-scale urban features and reducing computational complexity. Moreover, in the context of Belgium, 100 m has been a standard resolution for land use and planning studies.

Besides, both 50 m and 100 m resolutions outperform 300 m spatial resolution. The ROC values obtained at 300 m resolution indicate a decrease in predictive accuracy compared to higher resolutions, suggesting that the model struggles to capture urban growth dynamics effectively at this resolution in our study area. Interestingly, the class size is proportional to neighborhood size, i.e., lower-class density captures urban densification better at smaller neighborhood sizes and vice versa, as given in Fig. 8(b). This evinces an observation that a high-resolution image helps to understand different multiclass densification scenarios where the development depends on existing built-up. Furthermore, a neighborhood distance of around 100 m radius is well adept to studying urban densification. However, it is also possible that these models work well at coarser resolutions like 300 m while capturing urban sprawl in rural areas with built-up situated far from one another. This section thus helps us conclude that in our study, considering a resolution of 50 m and 5 neighborhood classes (which creates a 250 m radius for capturing densification) is the best selection for studying a process like urban multi-density expansion or densification. Similarly, 100 m works best with 3 neighborhood classes, i.e., with a 300 m radius. Since densification is highly dependent on the existing built-up's proximity, a drop in model accuracy at 300 m shows that a broader search area is not helpful for studying compact phenomena like densification at a provincial scale.

4.4. Impact of model parameters on urban densification

This stepwise selection strategy ensures that our model offers insightful information into the multifaceted factors driving urbanization process. Consequently, as indicated by its decreased discriminatory power in terms of ROC values, the 300 m spatial resolution does not provide any meaningful contribution to the prediction performance of our model. A previous study by Hu and Lo (2007) considered a spatial resolution of 225 m with an ROC of 0.850, but, in contrast, our goal is to check both the predicted accuracy and explanatory robustness. Thus, we do not include this resolution in our final model definition. The necessity to prioritize spatial resolutions that demonstrate statistically significant correlations with the dynamics of urban expansion drove this conclusion.

Our result is consistent with the concept of accessibility-driven development, where areas with better transportation links are more attractive for residential and commercial activities, leading to increased urban development. Medium-density regions typically exhibit a balanced mix of urban and suburban characteristics, where access to transportation networks plays a critical role in facilitating connectivity and accessibility. In higher resolutions, the influence of these factors becomes apparent, capturing their influence more intensely. A coarser resolution dilutes the relation as it broadens to capture larger areas and explain their effect on initial growth and further urban development. Bus stops have connected roads from dense residential areas with main city centers and other connectivity points like railway stations. They also tie the exteriors of a region to its central urban parts. Hence, traffic variables, like general road congestion or average traffic flow, might be less directly tied to the initial phases of urban expansion and more relevant to the broader functioning of already developed areas. It is also derived that multi-density expansion or densification, is largely impacted by zoning policies and that urban built-up development is largely governed by policies. Areas with a high density of urban built-up cells in the neighbourhood encourage further development due to the presence of robust urban infrastructure and amenities (Poelmans and Van Rompaey, 2009; Mustafa et al., 2018a, 2018b). In regions with low population density, housing attributes such as size and type play a crucial role in attracting new residents and driving development activities. Larger housing units are often associated with suburban or peri-urban developments, where land availability allows for the construction of spacious dwellings and amenities. It is also a representation of the resident's preferences against high density development seeking spacious living environments forming close knit community interactions. However, higher job density suggests areas with thriving economic activity, likely attracting development and population influx can influence existing urban expansion patterns. This study of variables underscores the need for an AAT approach of sensitivity study emphasizing the importance of land use planning and regulation in managing urban growth and ensuring sustainable development.

4.5. Global perspective for policy implications

Our study goes beyond its current applicability to our specific region. It can provide insightful information for similar cross-border urban planning scenarios and for large scale studies. Since cadastral data is not widely available for its innate administrative confidentiality and data privacy (Zevenbergen, 2002; Bennett et al., 2008), several countries across the globe do not have access to such data. Therefore, to reproduce similar studies, use of widely available remote sensing images like Sentinel, Landsat or Aerial orthophotos would be beneficial. They are also available for multiple dates and can be resampled and incorporated using the adapted methodology of 100 m × 100 m. Additionally, this study can lay a foundation for planners and modelers, enabling them to apply these methods in similar cross-border case studies as can be seen for areas like the Greater Toronto-Hamilton region in the United States and Canada, or the Greater Mekong region of Southeast Asia (Krongkaew, 2004; Moos and Mendez, 2015). Hence, we conclude that greater emphasis should be placed on research like this to better understand a model's performance in response to changes in input and the variation in its ability to produce efficient results. This will be helpful in informed urban planning and modelling in a global context - as a step towards sustainable urban development.

5. Conclusions

Our study conducts a global sensitivity analysis in three cross-border provinces of Belgium, investigating the effect of various input combinations on the model's performance. We examine 48 combinations of varied spatial resolutions, urban density classes, and neighbourhood sizes. The aim is to study urban densification by highlighting the sensitivity of a range of spatial parameters in an AAT approach, which were missing in earlier research. Our results clearly show that high-resolution data is crucial for understanding the densification dynamics of our study area, while coarser resolutions limit explanatory power. Changes in neighborhood size affect urban densification due to the influence of surrounding local agents. However, considering neighborhood as an implicit variable in our MNL models can limit its full potential. Explicit and dynamic modelling approaches like cellular automata can be an efficient alternative to capture neighborhood effects on densification. Results show that the model's performance enhances at high resolutions of 50 m and 100 m and deteriorates with coarser resolution of 300 m. The combinations which best work for capturing urban densification are (i) 50 m, 3 density classes with 5 × 5 neighborhoods, and (ii) 100 m, 4 density classes with 3 × 3 neighborhoods. These findings inform tailored policies required for managing urban development, emphasizing that studying densification or existing built-up development in Belgium requires spatial resolution with a coverage of a maximum of 300 m (i.e., 50 m × 5 or 100 m × 3 neighbourhoods). However, it can be contextually varying and dependent on built-up density classes and areas.

On the other hand, stepwise feature selection helps assess the variable significance in modelling. Our findings reveal similar observations in line with MNL models, where 50 m and 100 m resolutions demonstrate superior explanatory ability compared to 300 m resolution. A 3 × 3 neighborhood captures densification better. It can be concluded, that even though spatial resolution has proven to be a key parameter throughout the study, variation in density classes hold an equal sensitivity especially for densification or multi-density expansion. The result of stepwise regression shows that along with residential roads and population density, zoning policies largely impact the process of densification in Belgium across all the density classes. However, policy intervention is one of our study's bottlenecks, especially because it involves three different regions with separate federal governances. Since these authorities behave according to their respective rules and regulations, it is recommended that a decentralized manner of management is implemented to encourage urban densification, which is not yet very conspicuous. Thus, we strongly recommend a transition from conservative zoning practices to decentralisation as a new instrument for sustainable urbanization in Belgium, preventing adverse expansion.

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Anasua Chakraborty: Writing – original draft, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mitali Yeshwant Joshi:** Writing – review & editing, Methodology, Investigation, Formal analysis, Conceptualization. **Ahmed Mustafa:** Writing – review & editing, Supervision, Conceptualization. **Mario Cools:** Supervision, Conceptualization. **Jacques Teller:** Writing – review & editing, Supervision, Investigation, Funding acquisition, Formal analysis.

Acknowledgements

This research was funded by the INTER program and co-funded by the Fond National de la Recherche, Luxembourg (FNR) and the Fund for Scientific Research-FNRS, Belgium (F.R.S—FNRS), T.0233.20,—'Sustainable Residential Densification' project (SusDens, 2020–2024).

References

Abolhasani, S., Taleai, M., Karimi, M., Rezaee Node, A., 2016. Simulating urban growth under planning policies through parcel-based cellular automata (ParCA) model. Int. J. Geogr. Inf. Sci. 30, 2276–2301. doi:10.1080/13658816.2016.1184271.

Achmad, A., Hasyim, S., Dahlan, B., Aulia, D.N., 2015. Modeling of urban growth in tsunami-prone city using logistic regression: analysis of Banda Aceh, Indonesia. Appl. Geogr. 62, 237–246. doi:10.1016/j.apgeog.2015.05.001.

Giovanni, A., Simon, A., Andrew, R., 2020. Eurostat 2020. European Union, Luxembourg.

Anputhas, M., Janmaat, J.J.A., Nichol, C.F., Wei, X., 2016. Modelling spatial association in pattern based land use simulation models. J. Environ. Manage. 181, 465–476. doi:10.1016/j.jenvman.2016.06.034.

Arlinghaus, S.L., Kerski, J.J., 2013. Spatial Mathematics: Theory and Practice through Mapping. CRC Press, Boca Raton doi:10.1201/B15049.

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

Bennett, R., Wallace, J., Williamson, I., 2008. Organising land information for sustainable land administration. Land Use Policy 25, 126–138. doi:10.1016/J.LANDUSEPOL.2007.03.006.

Boussauw, K., Allaert, G., Witlox, F., 2013. Colouring inside what lines? Interference of the urban growth boundary and the political–administrative border of Brussels. Eur. Plan. Stud. 21, 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. Environ. Plann. C Gov. Policy 33, 1376–1393. doi:10.1068/c12327.

Calka, B., Nowak Da Costa, J., Bielecka, E., 2017. Fine scale population density data and its application in risk assessment. Geomat. Nat. Hazards Risk 8, 1440–1455. doi:10.1080/19475705.2017.1345792.

Cao, Y., Zhang, X., Fu, Y., Lu, Z., Shen, X., 2020. Urban spatial growth modeling using logistic regression and cellular automata: a case study of Hangzhou. Ecol. Indic. 113, 106200. doi:10.1016/J.ECOLIND.2020.106200.

Chakraborty, A., Mustafa, A., Omrani, H., Teller, J., 2023. A framework to probe uncertainties in urban cellular automata modelling using a novel framework of multilevel density approach: a case study for Wallonia Region. In: Goodspeed, R., Sengupta, R., Kyttä, M., Pettit, C. (Eds.), Intelligence for Future Cities: Planning Through Big Data and Urban Analytics (The Urban Book Series). Springer, Belgium, pp. 325–341.

Chakraborty, A., Omrani, H., Teller, J., 2022a. A comparative analysis of drivers impacting urban densification for cross regional scenarios in Brussels Metropolitan Area. Land 11, 2291. doi:10.3390/LAND11122291.

Chakraborty, A., Omrani, H., Teller, J., 2022b. Modelling the drivers of urban densification to evaluate built-up areas extension: a data-modelling solution towards zero net land take. In: Gervasi, O., Murgante, B., Hendrix, E.M.T., Taniar, D., Apduhan, B.O. (Eds.), Computational Science and Its Applications – ICCSA 2022. ICCSA 2022. Lecture Notes in Computer Science, 13376. Springer, Cham doi:10.1007/978-3-031-10450-3_21.

Chen, Y., Li, X., Liu, X., Ai, B., 2014. Modeling urban land-use dynamics in a fast developing city using the modified logistic cellular automaton with a patch-based simulation strategy. Int. J. Geogr. Inf. Sci. 28, 234–255. doi:10.1080/13658816.2013.831868.

Cheng, J., Masser, I., 2003. Urban growth pattern modeling: a case study of Wuhan city, PR China. Landsc. Urban Plan. 62, 199–217. doi:10.1016/S0169-2046(02)00150-0.

Claassens, J., Koomen, E., Rouwendal, J., 2020. Urban density and spatial planning: the unforeseen impacts of Dutch devolution. PLoS One 15, e0240738. doi:10.1371/JOURNAL.PONE.0240738.

Cortés-Borda, D., Polanco, J.andrés, Escobar-Sierra, M., 2022. Social perception assessment of hydropower sustainability: a stepwise logistic regression modeling. Environ. Sci. Policy 134, 108–118. doi:10.1016/J.ENVSCI.2022.03.026.

Crooks, A., Castle, C., Batty, M., 2008. Key challenges in agent-based modelling for geo-spatial simulation. Comput. Environ. Urban Syst. 32, 417–430. doi:10.1016/j.compenvurbsys.2008.09.004.

Cuellar, Y., Perez, L., 2023. Assessing the accuracy of sensitivity analysis: an application for a cellular automata model of Bogota's urban wetland changes. Geocarto Int. 38 (1), 2186491. doi:10.1080/10106049.2023.2186491.

De Maesschalck, F., De Rijck, T., Heylen, V., Bruel, P., 2015. Crossing borders: social-spatial relations between Brussels and Flemish Brabant. Bruss. Stud. doi:10.4000/BRUSSELS.1260.

Deboosere, P., Eggerickx, T., Van Hecke, E., Wayens, B., 2009. The population of Brussels: a demographic overview. Bruss. Stud. doi:10.4000/BRUSSELS.891.

Dong, F., Li, J., Wang, Y., Zhang, X., Zhang, S., Zhang, S., 2019. Drivers of the decoupling indicator between the economic growth and energy-related CO_2 in China: a revisit from the perspectives of decomposition and spatiotemporal heterogeneity. Sci. Total Environ. 685, 631–658. doi:10.1016/J.SCITOTENV.2019.05.269.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recognit. Lett. 27, 861–874. doi:10.1016/J.PATREC.2005.10.010.

Feng, Y., Liu, Y., Tong, X., Liu, M., Deng, S., 2011. Modeling dynamic urban growth using cellular automata and particle swarm optimization rules. Landsc. Urban Plan. 102, 188–196. doi:10.1016/J.LANDURBPLAN.2011.04.004.

Franczyk, A., 2019. Using the Morris sensitivity analysis method to assess the importance of input variables on time-reversal imaging of seismic sources. Acta Geophys. 67, 1525–1533. doi:10.1007/S11600-019-00356-5.

Gamboa, F., Gremaud, P., Klein, T., Lagnoux, A., 2022. Global sensitivity analysis: a novel generation of mighty estimators based on rank statistics. Bernoulli 28, 2345–2374. doi:10.3150/21-BEJ1421.

Getu, K., Gangadhara Bhat, H., 2024. Application of geospatial techniques and binary logistic regression model for analyzing driving factors of urban growth in Bahir Dar city, Ethiopia. Heliyon 10, e25137. doi:10.1016/J.HELIYON.2024.E25137.

Guyot, M., Araldi, A., Fusco, G., Thomas, I., 2021. The urban form of Brussels from the street perspective: the role of vegetation in the definition of the urban fabric. Landsc. Urban Plan. 205, 103947. doi:10.1016/J.LANDURBPLAN.2020.103947.

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, 887–898. doi:10.1016/j.landusepol.2012.01.008.

Hao, P., Hooimeijer, P., Sliuzas, R., Geertman, S., 2013. What drives the spatial development of urban villages in China? Urban Stud. 50, 3394–3411. doi:10.1177/0042098013484534.

Harris, P., Charlton, M., Fotheringham, A.S., 2010. Moving window kriging with geographically weighted variograms. Stoch. Environ. Res. Risk Assess. 24, 1193–1209. doi:10.1007/S00477-010-0391-2.

Hu, Z., Lo, C.P., 2007. Modeling urban growth in Atlanta using logistic regression. Comput. Environ. Urban Syst. 31, 667–688. doi:10.1016/J.COMPENVURBSYS.2006.11.001.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2021. An Introduction to Statistical Learning: With Applications in R, 2nd Ed. Springer, New York, NY.

Kantakumar, L.N., Kumar, S., Schneider, K., 2020. What drives urban growth in Pune? A logistic regression and relative importance analysis perspective. Sust. Cities Soc. 60, 102269. doi:10.1016/J.SCS.2020.102269.

Karimi, F., Sultana, S., Babakan, A.S., Suthaharan, S., 2021. Urban expansion modeling using an enhanced decision tree algorithm. Geoinformatica 25, 715–731. doi:10.1007/s10707-019-00377-8.

Kocabas, V., Dragicevic, S., 2006. Assessing cellular automata model behaviour using a sensitivity analysis approach. Comput. Environ. Urban Syst. 30, 921–953. doi:10.1016/J.COMPENVURBSYS.2006.01.001.

Krongkaew, M., 2004. The development of the Greater Mekong Subregion (GMS): real promise or false hope? J. Asian Econ. 15, 977–998. doi:10.1016/J.ASIECO.2004.09.006.

Li, X., Zhou, W., Ouyang, Z., 2013. Forty years of urban expansion in Beijing: what is the relative importance of physical, socioeconomic, and neighborhood factors? Appl. Geogr. 38, 1–10. doi:10.1016/J.APGEOG.2012.11.004.

Liu, B., Zhao, Q., Jin, Y., Shen, J., Li, C., 2021. Application of combined model of stepwise regression analysis and artificial neural network in data calibration of miniature air quality detector. Sci. Rep. 11, 3247. doi:10.1038/s41598-021-82871-4.

Liu, X., Li, X., Shi, X., Wu, S., Liu, T., 2008. Simulating complex urban development using kernel-based non-linear cellular automata. Ecol. Modell. 211, 169–181. doi:10.1016/J.ECOLMODEL.2007.08.024.

Lu, B., Charlton, M., Harris, P., Fotheringham, A.S., 2014. Geographically weighted regression with a non-Euclidean distance metric: a case study using hedonic house price data. Int. J. Geogr. Inf. Sci. 28, 660–681. doi:10.1080/13658816.2013.865739.

Moos, M., Mendez, P., 2015. Suburban ways of living and the geography of income: how homeownership, single-family dwellings and automobile use define the metropolitan social space. Urban Stud. 52, 1864–1882. doi:10.1177/0042098014538679.

Mustafa, A., Saadi, I., Cools, M., Teller, J., 2018a. Understanding urban development types and drivers in Wallonia: a multi-density approach. Int. J. Bus. Intell. Data Min. 13, 309–330. doi:10.1504/IJBIDM.2018.088434.

Mustafa, A., Van Rompaey, A., Cools, M., Saadi, I., Teller, J., 2018b. Addressing the determinants of built-up expansion and densification processes at the regional scale. Urban Stud. 55 (15), 3279–3298. doi:10.1177/0042098017749176.

Mustafa, A., Heppenstall, A., Omrani, H., Saadi, I., Cools, M., Teller, J., 2018c. Modelling built-up expansion and densification with multinomial logistic regression, cellular automata and genetic algorithm. Comput. Environ. Urban Syst. 67, 147–156. doi:10.1016/J.COMPENVURBSYS.2017.09.009.

Overmars, K.P., De Koning, G.H.J., Veldkamp, A., 2003. Spatial autocorrelation in multi-scale land use models. Ecol. Modell. 164, 257–270. doi:10.1016/S0304-3800(03)00070-X.

Pijanowski, B.C., Alexandridis, K.T., Müller, D., 2006. Modelling urbanization patterns in two diverse regions of the world. J. Land Use Sci. 1, 83–108. doi:10.1080/17474230601058310.

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. Landsc. Urban Plan. 93, 10–19. doi:10.1016/J.LANDURBPLAN.2009.05.018.

Poelmans, L., Van Rompaey, A., 2010. Complexity and performance of urban expansion models. Comput. Environ. Urban Syst. 34, 17–27. doi:10.1016/J.COMPENVURBSYS.2009.06.001.

Pontius, R.G., Huffaker, D., Denman, K., 2004. Useful techniques of validation for spatially explicit land-change models. Ecol. Modell. 179, 445–461. doi:10.1016/J.ECOLMODEL.2004.05.010.

Puente-Sotomayor, F., Mustafa, A., Teller, J., 2021. Landslide susceptibility mapping of urban areas: logistic regression and sensitivity analysis applied to Quito, Ecuador. Geoenviron. Disasters 8, 19. doi:10.1186/S40677-021-00184-0.

Razavi, S., Jakeman, A., Saltelli, A., Prieur, C., Iooss, B., Borgonovo, E., Plischke, E., Lo Piano, S., Iwanaga, T., Becker, W., Tarantola, S., Guillaume, J.H.A., Jakeman, J., Gupta, H., Melillo, N., Rabitti, G., Chabridon, V., Duan, Q., Sun, X., Smith, S., Sheikholeslami, R., Hosseini, N., Asadzadeh, M., Puy, A., Kucherenko, S., Maier, H.R., 2021. The future of sensitivity analysis: an essential discipline for systems modeling and policy support. Environ. Modell. Softw. 137, 104954. doi:10.1016/J.ENVSOFT.2020.104954.

Roodposhti, M.S., Hewitt, R.J., Bryan, B.A., 2020. Towards automatic calibration of neighbourhood influence in cellular automata land-use models. Comput. Environ. Urban Syst. 79, 101416. doi:10.1016/J.COMPENVURBSYS.2019.101416.

Saganeiti, L., Mustafa, A., Teller, J., Murgante, B., 2021. Modeling urban sprinkling with cellular automata. Sust. Cities Soc. 65, 102586. doi:10.1016/J.SCS.2020.102586.

Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., Wu, Q., 2019. Why so many published sensitivity analyses are false: a systematic review of sensitivity analysis practices. Environ. Modell. Softw. 114, 29–39. doi:10.1016/J.ENVSOFT.2019.01.012.

Seto, K.C., Güneralp, B., Hutyra, L.R., 2012. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. Proc. Natl. Acad. Sci. U.S.A. 109, 16083–16088. doi:10.1073/PNAS.1211658109.

Shafizadeh-Moghadam, H., Asghari, A., Taleai, M., Helbich, M., Tayyebi, A., 2017. Sensitivity analysis and accuracy assessment of the land transformation model using cellular automata. GISci. Remote Sens. 54, 639–656. doi:10.1080/15481603.2017.1309125.

Shu, B., Zhu, S., Qu, Y., Zhang, H., Li, X., Carsjens, G.J., 2020. Modelling multi-regional urban growth with multilevel logistic cellular automata. Comput. Environ. Urban Syst. 80, 101457. doi:10.1016/J.COMPENVURBSYS.2019.101457.

Sun, L., Chen, J., Li, Q., Huang, D., 2020. Dramatic uneven urbanization of large cities throughout the world in recent decades. Nat. Commun. 11, 5366. doi:10.1038/s41467-020-19158-1.

Todes, A., 2012. Urban growth and strategic spatial planning in Johannesburg, South Africa. Cities 29, 158–165. doi:10.1016/J.CITIES.2011.08.004.

Vermeiren, K., Van Rompaey, A., Loopmans, M., Serwajja, E., Mukwaya, P., 2012. Urban growth of Kampala, Uganda: pattern analysis and scenario development. Landsc. Urban Plan. 106, 199–206. doi:10.1016/J.LANDURBPLAN.2012.03.006.

Waddell, P., 2002. UrbanSim: modeling urban development for land use, transportation, and environmental planning. J. Am. Plan. Assoc. 68, 297–314. doi:10.1080/01944360208976274.

Wang, M., Wright, J., Brownlee, A., Buswell, R., 2016. A comparison of approaches to stepwise regression on variables sensitivities in building simulation and analysis. Energy Build. 127, 313–326. doi:10.1016/J.ENBUILD.2016.05.065.

Wang, R., Feng, Y., Tong, X., Zhao, J., Zhai, S., 2021a. Impacts of spatial scale on the delineation of spatiotemporal urban expansion. Ecol. Indic. 129, 107896. doi:10.1016/J.ECOLIND.2021.107896.

Wang, H., Guo, J., Zhang, B., Zeng, H., 2021b. Simulating urban land growth by incorporating historical information into a cellular automata model. Landsc. Urban Plan. 214, 104168. doi:10.1016/J.LANDURBPLAN.2021.104168.

Wang, Y., Lv, W., Wang, M., Chen, X., Li, Y., 2023. Application of improved Moran's *I* in the evaluation of urban spatial development. Spat. Stat. 54, 100736. doi:10.1016/J.SPASTA.2023.100736.

UN-HABITAT, 2020. World Cities Report 2020: The Value of Sustainable Urbanization. UN-HABITAT, Nairobi, Kenya.

Wu, F., 2002. Calibration of stochastic cellular automata: the application to rural-urban land conversions. Int. J. Geogr. Inf. Sci. 16, 795–818. doi:10.1080/13658810210157769.

Wu, H., Lin, A., Xing, X., Song, D., Li, Y., 2021. Identifying core driving factors of urban land use change from global land cover products and POI data using the random forest method. Int. J. Appl. Earth Obs. Geoinf. 103, 102475. doi:10.1016/j.jag.2021.102475.

Yang, W., Zou, X., Wang, M., Liu, P., 2023. A multinomial logistic regression model-based seismic risk assessment method for museum exhibition halls. J. Build. Eng. 69, 106312. doi:10.1016/J.JOBE.2023.106312.

Zevenbergen, J., 2002. A Systems Approach to Land Registration and Cadastre. International Federation of Surveyors (FIG) XXII International Congress, Washington, D.C. USA April 19–26 2002.