ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-4/W4-2024
8th International Conference on Smart Data and Smart Cities (SDSC), 4–7 June 2024, Athens, Greece

Modelling multi-density urban expansion using Cellular Automata for Brussels Metropolitan Development Area

Anasua Chakraborty[1], Ahmed Mustafa[2], Jacques Teller[3]

[1] Department of ArGEnCO, Université de Liège, Liège, Belgium - a.chakraborty@uliege.be
[2] Urban Systems Lab, The New School, New York, USA - a.mustafa.eg@gmail.com
[3] Department of ArGEnCO, Université de Liège, Liège, Belgium - jacques.teller@uliege.be

Keywords: Multi-density, Urban Expansion, Cellular Automata, Land Use Studies, Urban Planning.

Abstract

Our studies aim at modelling and simulating urban expansion scenarios for Brussels capital region, Brabant of Flanders and Wallonia. Thereby we use a non-ordered multinomial logistic regression (MLR) coupled with cellular automata. Our model helps to study the probability for built-up development based on a) impact of different causative factors on expansion process and b) effect of the neighbouring cells on future built-up development. In our study, we have used 100×100 m raster data representing cadastral built-up of our study area. The model is then calibrated using the maps for years 2000-2010 and to simulate 2020. Thereto, simulated 2020 maps has been validated with observed 2020 built-up maps using fuzzy set theory. Our results show that all through our study, zoning or land use policies play an important role for expansion along all the built-up density classes. Besides, slope distance to highways and major cities encourages new urban development commonly known as 'Urban sprawls'. Distance to main roads, employment opportunities aids to further development from low density to medium density areas. While most of the factors impact negatively to further high density development showing that our area is highly governed by the zoning status in case of densification.

1. Introduction

Urban expansion, defined as land being changed into built-up areas, is a major driver of changes in land use that has very far reaching consequences on the environment, natural resources and human health (Zhang et al., 2011). Unchecked sprawl of urban development's poses challenges to sustainable urban planning and it calls for effective modelling techniques for understanding and managing its impacts. While existing models are mainly based on raster with low resolution, there is a growing recognition for higher resolutions so as to understand the complexity of urban landscapes (Han and Jia, 2017; Liao et al., 2014). In the quest for improved urban expansion modelling, scholars have ventured into innovative methodologies and diverse spatial scales. Larger grid dimensions, such as 100×100 m provide a balance between intricate details and processing load (Poelmans and Rompaey, 2010). However, they often blur the diversity of land use within each cell. On the other hand, smaller scales down to 10×10 m offer enhanced details but require abundant computational power, particularly when analysing large geographical areas (Mustafa et al., 2014; Chakraborty et al., 2022). Examining several built-up densities rather than using a binary categorization (i.e., non-built-up/built-up) is one way to handle the trade-off between heterogeneity and coarse regular cell spaces.

Cellular Automata (CA) are widely used to simulate patterns of urban growth; yet, precisely calibrating their transition rules is still a difficult task that depends on both spatial characteristics and causative factors. Traditional approaches depended on trial and error, but more recently, automated techniques like statistical analysis and machine learning have been adopted. Additional challenges arise in the validation of CA models, as traditional pixel-by-pixel techniques are unable to discriminate between various kinds of errors (Mustafa et al., 2018a). Spatial metrics may be misleading, prompting exploration of fuzzy set theory for more nuanced validation approaches (Ahmed et al., 2013). Comprehensive surveys emphasize the importance of robust validation methods for ensuring the reliability of CA models (Vliet et al., 2016).

The study presents a novel approach to model built-up expansion in the Brussels Development area using an integrated approach of multinomial logistic regression (MLR) and CA. The model accounts for transition of non-built-up areas to various built-up density class. Calibration and validation of the model are done using Belgian cadastral data for 2000, 2010 and 2020, with four built-up classes defined: non-built-up, low-density, medium-density, and high-density. Three maps can define one calibration interval (2000-2010) and one validation interval (2010–2020). The model considers a comprehensive set of 12 static causative factors encompassing accessibility, geophysical features, policies, socio-economic factors, and neighbourhood interactions, recognizing urbanization as a self-organizing system (Chakraborty et al., 2023).

The model parameters undergo calibration through a combination of logistic regression and genetic algorithm techniques. More specifically, the neighbourhood interactions are dynamically adjusted by means of a multi-objective genetic algorithm (MOGA). Transitions between distinct built-up classes are represented by the dependant variable in the logistic regression model. The Relative Operating Characteristic (ROC) approach is used to validate the results of the logistic regression. Across all built-up classes, the MOGA's goal function seeks to optimize allocation accuracy rates (Mustafa et al., 2018a). A fuzzy membership function with an exponential decay that uses a neighbourhood window of four cells and a halving distance of two cells determines this accuracy rate. Furthermore, the accuracy rate function is used to validate the model.

https://doi.org/10.5194/isprs-annals-X-4-W4-2024-29-2024

2. Materials and Study Area

2.1 Study Area

In this section, we discuss our area of area of interest. Our model was carried at a provincial level of Belgium - one of the urbanised country of Europe. More specifically, it includes Brussels Capital Region (BRC) with a population of 1,241,175 inhabitants, Brabant of Flanders (FB) to the north having 1,187,483 inhabitants and Brabant of Wallonia (WB) to the south with 412,934 inhabitants as shown in Figure 1. The population shows a variation in them proving that this study represents an interesting cross border scenario. This is why, there are several densities that can be observed in the built-up where denser built-up can be noticed in the FB and BRC and sparse or low density is commonly found in WB.

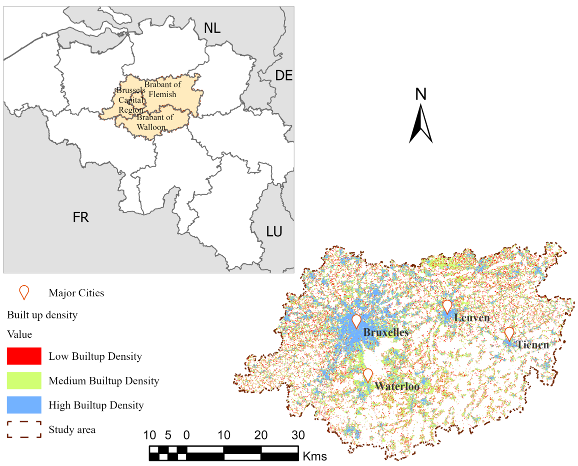

Figure 1. Study area.

Table 1 gives us the built-up transition for the year 2000-2010 which has been further used for calibrating our model. It shows in a top-down approach the class to class change between four density levels, where class-0 represents non built-up, class-1 represent low density built-up, class-2 represent medium density and class-3 shows high density built-up. It can be observed that in our study area, development from non built-up to low and medium density is predominant which is why this research the focus of calibrating and simulating urban expansion along multi-density classes.

2000-2010	Class-0	Class- 1	Class-2	Class-3
Class-0	-	-	-	-
Class-1	1469(0.70%)	-	-	-
Class-2	1147(0.54%)	2419(8.10%)	-	-
Class-3	416(0.20%)	170(0.50%)	3363(5.06%)	30260

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

2.2 Materials

This section of the paper discusses the different types and sources of data used for our modelling. In this study, the raster built-up maps for years 2000, 2010 and 2020 have been created using Belgian cadastral data provided by the Land registry administration of Belgium. These are vector data where each polygon represent a cadastral building. Hence, in order to minimise the heavy computation time, we rasterised the data in 2×2m cell resolution. This data was aggregated into 100×100 m (Mustafa et al., 2018b; Chakraborty et al., 2023; Tannier and Thomas, 2013) raster for creating density classes. All cell values $< 25m^2$ are considered non built-up because an average residential building of Belgium corresponds to values $\geq 25m^2$, represented by a cell size of $100m^2$ in our aggregated data. Since our data distribution was positively skewed, hence a geometric interval classification method was used to define the density classes.

This built-up maps are used as our response variable for MLR model. As shown in Table 1, the causative factors of our model includes digital elevation model (DEM) of 1m resolution procured from geospatial data repository of regional development authorities. Elevation and Slope were both calculated using the DEM data. We then created Euclidean distance maps for Road1 (Highways), Roads2 (main roads), Roads3 (secondary roads) and Roads4 (local roads) from open street maps (OSM), using geospatial software. The major railways stations have been obtained from the OSM database as well, and a sophisticated cost distance method has been obtained for the same, instead of Euclidean distance. Distance to main cities which comes under our study area was taken from Atlas of Belgium. This comprises of the geophysical and accessibility factors respectively. However, our model includes dynamic socio-economic factors like number of jobs - which was used to calculate jobs density per $100m^2$. Another important socio-economic factor taken into account was the average household income which has a causal relationship with built-up development. Along with these two, the total number of population was also considered and sourced from National Statistical authority of Belgium (Statbel).

Factor	Name	Unit	Source
X1	Slope	percent	Calculated from Digital Elevation Model from ALOS DEM
X2	Euclidean distance to Motorways	meters	Open Street Maps
X3	Euclidean distance to Primary Roads	meters	Open Street Maps
X4	Euclidean distance to Secondary Roads	meters	Open Street Maps
X5	Euclidean distance to Local Roads	meters	Open Street Maps
X6	Euclidean distance to Residential Roads	meters	Open Street Maps
X7	Cost Distance to Railways Stations	meters	Atlas of Belgium
X8	Euclidean distance to Major Cities	meters	Open Street Maps
X9	Jobs density	Num/$100m^2$	Self Calculation based on Belgian Statistical Institute
X10	Average household income	€	Self Calculation based on Belgian Statistical Institute
X11	Total number of Population	Inhabitant/km^2	Self Calculation based on Belgian Statistical Institute
X12	Zoning status	Binary	Self Calculation based on Belgian Statistical Institute

Table 2. List of selected causative factors.

https://doi.org/10.5194/isprs-annals-X-4-W4-2024-29-2024 | © Author(s) 2024. CC BY 4.0 License.

3. Methods

In this section, we discuss an integrated approach of MLR and CA model. The model considers the expansion process that is transition from class-0 as the reference class to class-1, class-2 and class-3. The map of 2000-2010 has been used for calibration while validation has been done using the 2010-2020 map to simulate the actual quantity of new built-up divided for a span of 10 years.

3.1 Defining transition rules

The transition rules have been set using two main components. The first component concerns coefficients derived from our MLR model, while the second involves neighbourhood interactions. Equation 1, below, gives the transition of a cell state from non built-up to any of the built-up class in a specific time step.

$$P_{ij} = \sqrt{(P_c)_{ij} \times (P_n)_{ij}^{\sigma}} \tag{1}$$

Where $(P_c)_{ij}$ is the built-up probability based on the causative factors derived from the MLR model and $(P_n)_{ij}^{\sigma}$ is the neighbourhood cell effect of ij, where σ expresses the relative importance of the same.

3.2 Defining Causative factor calibration

In the MLR model, we study the empirical relationship between a multi density dependent variable and its causative factors. Equation 2 given below is the generalized form of a non-ordered logistic regression:

$$\begin{aligned} log(k_1) &= \alpha k_1 + \beta_{k_1 1} X_1 + \beta_{k_1 2} X_2 + \cdots + \beta_{k_1 \nu} X_\nu \\ &\vdots \\ log(k_n) &= \alpha k_n + \beta_{k_n 1} X_1 + \beta_{k_n 2} X_2 + \cdots + \beta_{k_n \nu} X_\nu \end{aligned} \tag{2}$$

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 $(X_1, X_2, \ldots, X_\nu)$, 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, Equation 3 gives the probabilities for each class.

$$\begin{aligned} &(P_c)_{ij}, Y = k_0 \\ &= \frac{1}{1 + exp(log(k_1)) + exp(log(k_2)) + \cdots + exp(log(k_n))} \\ &(P_c)_{ij}, Y = k_1 \\ &= \frac{exp(log(k_1))}{1 + exp(log(k_1)) + exp(log(k_2)) + \cdots + exp(log(k_n))} \\ &\vdots \\ &(P_c)_{ij}, Y = k_n \\ &= \frac{exp(log(k_n))}{1 + exp(log(k_1)) + exp(log(k_2)) + \cdots + exp(log(k_n))} \end{aligned} \tag{3}$$

where $(P_c)_{ij}, Y = k_n$ is the probability of change from the reference class to class k_n occurring in cell ij (Mustafa et al., 2015; Chakraborty et al., 2022). Our MLR model uses a maximum likelihood estimation method to achieve the best set of coefficients for each X variable. The outcomes or the set of coefficients that describe the impact of each causative factors on the expansion process and thereby generate a probability map using the same coefficients.

An ROC method commonly used for understanding the goodness of fit of the model has been used to check the accuracy. It is considered to be a high quality method that can predict the occurrence of an event by comparing the probability map with the actual changes (Hu and Lo, 2007). ROC closer to the value of 1 is considered to be a perfect fit. All causative factors data were resampled to the same resolution of 100×100 m using nearest neighborhood. Since, our data was available at different units it was imperative to standardize all the continuous variables. However, zoning was kept as binary (0,1) showing allowance of built-up development.

Henceforth, variation inflation factor (VIF), a commonly used method was employed to examine the multicollinearity between the causative factors, to ensure there is no bias in model's output. It is recommended by Montgomery and Runger, 2010 that VIF values exceeding 4 should not be taken into consideration for modelling. The selection of samples excludes the other existing class than the reference class, example – in our case, class-0 sampling process considers new transitions from class-0 to classes 1, 2, and 3.

3.3 Defining neighborhood calibration

Since MLR models are not temporally explicit, and as a result of which cannot exhibit a trend dependent self organized development scenario typical for expansion (Hu and Lo, 2007; Verburg et al., 2004). This is why it is important to calibrate neighborhood interaction independently as a dynamic phenomenon using CA modelling. In our study, we have considered a 3×3 neighborhood window for understanding this interaction. This is typically based on the study by Chen et al., 2014 and Poelmans and Rompaey, 2009, where they extensively worked on several neighborhood window sizes and concluded 3×3 windows as the best fit. This can also differ based on spatial resolution and area of study. However, currently that does not involve our scope of study. The $(P_n)_{ij}^{\sigma}$ is calculated based on the proposed method by White et al., 2012:

$$(P_n)_{ij} = \sum_k \sum_x \sum_d w_{kxd} \cdot I_{kxd} \tag{4}$$

where w_{kxd} is the weighing parameter assigned to one of the built-up class for position x at a distance d, and I_{kxd} will be 1 if a cell in distance d is allocated by class k or else 0.

The aim of our study is also to define an accurate parameter for CA to achieve the best allocation accuracy rate for the expansion process. In order to calibrate this, we used a MOGA method. This method is a trade off among multiple conflicting objectives and all of them at once (Al-Ahmadi et al., 2009). It is to be noted that MOGA is one of the most effective algorithms for solving both constrained and unconstrained scenarios of optimization.

The MOGA uses a stochastic operator to generate new genes from random initial population based on a fitness function. This fitness function is used to evaluate next generations based on a relative fitness score. After multiple iterations it produces the optimum fitness value, where we propose that a new generation is obtained depending on the best combination of crossover operator. Initially, the MOGA started with a random population and a large number of generations evolved to obtain the best possible solution. However, each individual solution is highly computationally draining (on an Intel core i-9 processor), and takes a timespan of 6 hours for one optimization run. In order to minimize that iteratively, we have reached a set of values where 100 generations with 30 parameters were selected to calibrate expansion process. This calibration function was validated based on a fuzzy membership function as discussed below.

3.4 Validation

We validated the accuracy of our optimization process by comparing the simulated map of 2020 with the observed map of 2020. This only considered the new built-up between 2010 and 2020. As discussed earlier, the fuzziness of a cell depends on the cell itself and its neighboring cell. As multiple authors proposed (Ahmed et al., 2013; Hagen, 2003; Loibl and Tötzer, 2003), we have considered an exponential decay function with a halving distance of two cells and a neighborhood of four cell radius to understand the influence of neighborhood cells on fuzzy index (Equation 5).

$$A_k = \frac{\sum_{x_k \in X_{k,sim}} |I_{x_k 0} \cdot \left(\frac{1}{2}\right)^{\frac{0}{2}}, I_{x_k 1} \cdot \left(\frac{1}{2}\right)^{\frac{1}{2}}, \ldots, I_{x_k d} \cdot \left(\frac{1}{2}\right)^{\frac{d}{2}}|_{max}}{X_{k,actual}} \tag{5}$$

Where we compute the average fuzziness index depending on the cell in simulated map in a neighborhood is identical to a cell in actual map within halving distance. We thus employ recurring weightage to them depending on its location where perfectly allocated is 1 and allocated away from four neighboring cells is 0.

4. Results and Discussion

Our study area is an amalgamation of two main urban cores of Brussels in BCR and Leuven in FB. Though there are certain urban cores in WB, they are predominantly scattered in nature. This pattern is familiar to many built-up development in the world. We, in our study, employed an MLR model where all the causative factors were less than VIF value 4, with highest value of 2.86. Hence, they were all considered for modelling inputs. In Figure 2, we demonstrated the MLR parameters used to calibrate for year 2000-2010.

The result shows a steady positive trend for zoning impact across all the density classes. This means that high density development is permitted following the legal plan to avoid any administrative risk. However, for new built-up in suburbs, it does not strictly abide by the same pattern. This is in line with Mustafa et al., 2018a. It is also interesting to see that for higher density classes, slope has a negative impact as it appears to be a hindrance for its adverse effects on land use. Furthermore, it should be noted that new development is slope dependent for the ease of construction. Distance to road also shows a negative impact for expansion except for the fact that medium density can get encouraged by existence of local roads connecting to highways. The job density has a positive impact on new development and expansion to medium density that leads to suburbanization. This could have a relationship with income ability. This implies that people with middle to high income can well settle in low to medium density. This is also be backed by the positive impact of average household income which is positive for all the three density classes. Population however has a negative impact showing that inhabitants are avoiding staying at the urban core with already crowded areas and preferring to stay at the peripheral of cores. This brings us to the point where we can see a similar kind of negative impact of distance to major cities. This is in line with study by Poelmans and Rompaey, 2010, who reported that development occurs around the city center but not exactly in it.

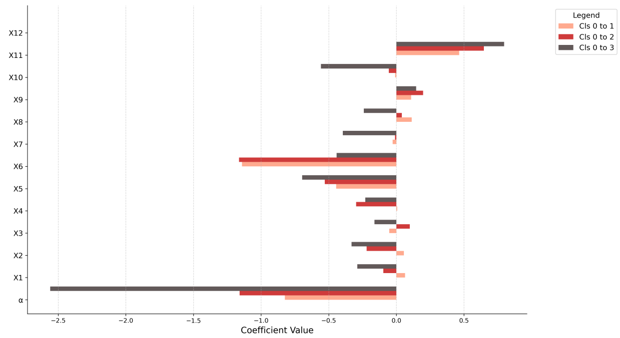

Figure 2. MLR coefficient for expansion for 2000-2010.

The ROC values for our MLR model are 0.84, 0.86 and 0.91 for class-0 to class-1, class-0 to class-2, and class-0 to class-3, respectively. As seen in the study by Cammerer et al., 2013, ROC values greater than 0.70 are a reasonable fit and can be used for further analysis. This is also seen below in Figure 3.

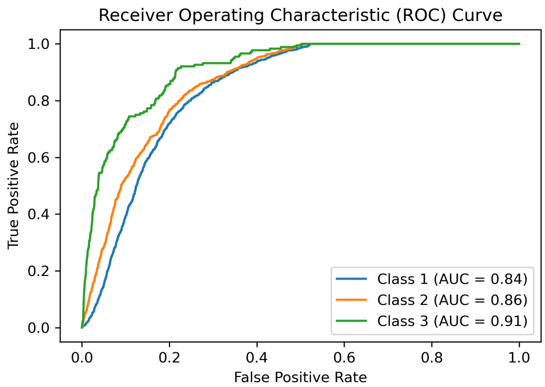

Figure 3. Density class wise ROC value.

After 256 iterations, the average change in the Pareto solution spread for MOGA optimization was less than 0.000001. The calibration shows that the likelihood of low-density expansion highly increases with increase in the number of existing low density and medium density lands, and decreases with the decrement in high density lands. This is also correlated to the similar kind of neighborhood pattern. This says that a similar kind of density pattern can be seen depending on the neighborhood expansion type. A low to medium density transition is highly probable where there is non built-up, low and medium density built-up. Whereas over the time of calibration and validation, we can see that there is a tendency of high density built-up towards the centrality creating the urban cores denser and the peripheral as it is. This finding suggests that in a cross-border situation like our study a few existing high density cores might have a possibility to cover the leftover places showing a trend of densification, while places experiencing low built-up pattern will restrict densification. This heterogeneity produces a rural-urban (rurban) scenarios.

Considering the fuzzy accuracy index, the calibration and validation for allocation accuracy have been done. The simulated map has been compared with actual map for 2020 as given in Figure 4 for Nivelles, a city in Wallon Brabant. For our study, the σ for MOGA optimization converges after iteration 26. The fuzzy accuracy index shows that the calibration has a higher accuracy rate than the validation. This is possible because of underlying uncertainty which was also the case for the study by Mustafa et al., 2018a. Stochastic perturbation techniques can be used as a plausible solution in the further research for this. It also highly depends on the data sources, parameters differentiation and needs a more exhaustive framework to address.

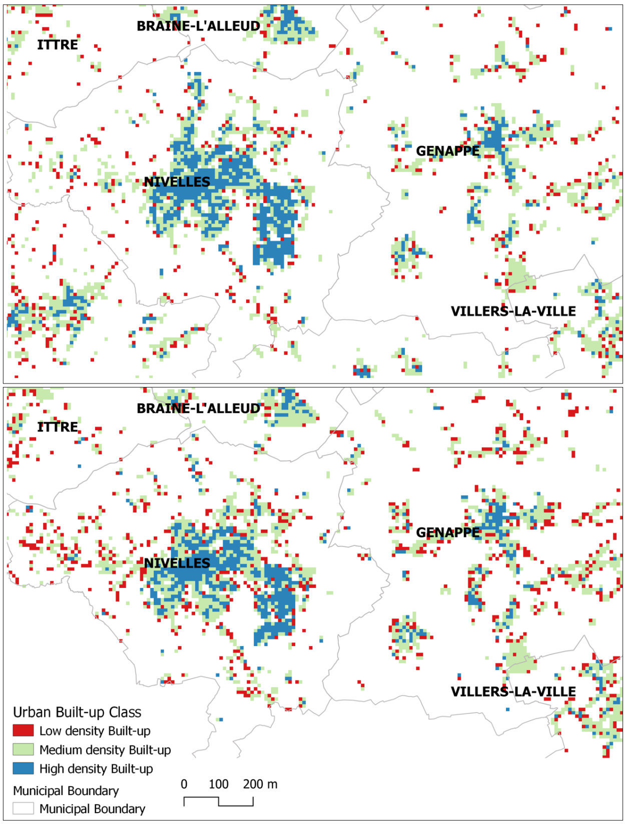

Figure 4. 2020 actual and simulated built-up expansion pattern of Nivelles.

5. Conclusions

Our study addresses some limitations of existing studies where urban expansion models consider the process as simple as a new development over a vacate land parcel. It is imperative - for having a global applicability to see this process through the multi-label lens. Our findings produce rare and interesting scenarios of cross border planning where the expansion process is not only heterogeneous but is also governed by different administrative bodies. While the urban cores show a slow expansion process in existing built-up areas, the rural areas still have a long way to cover and hinder such dense growth. In order to validate the model, we calibrated our model for 2000-2010, which shows a fuzzy accuracy of 0.26, 0.55, 0.48 for class-0 to class-1, class-0 to class-2, and class-0 to class-3, respectively. We then simulated it for 2020 and observed it with respect to an actual 2020 map. Here we derived the fuzzy accuracy rate of 0.20, 0.50 and 0.41 for same class transitions, which can be seen in Figure 5.

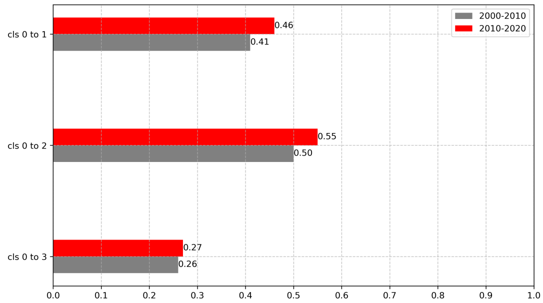

Figure 5. The average fuzzy similarity rates for calibration and validation.

However, in further studies we can imbibe uncertainty factors to address such differences in accuracy rates. Thus, our study identifies the important causative factors amongst many of which are applicable to global south scenarios. The amalgamation of a sophisticated optimization process like MOGA brings out a temporally effective way to map the expansion process in the study. We can conclude the work with the fact that expansion in different density level is an outcome of different factors and can vary with space and time. Having said that and working upon the limitations on its way, this study can be a pioneer towards mapping land use, understanding planning intricacies and a step to zero net land take.

Acknowledgement

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, 2020–2023).

References

Ahmed, B., Ahmed, R., Zhu, X., 2013. Evaluation of Model Validation Techniques in Land Cover Dynamics. *ISPRS Int. J. Geo Inf.*, 2, 577–577.

Al-Ahmadi, K., See, L., Heppenstall, A., Hogg, J., 2009. Calibration of a fuzzy cellular automata model of urban dynamics in Saudi Arabia. *Ecological Complexity*, 6, 80–101.

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

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, Belgium. Springer, Urban Book Series, Cham, Switzerland.

Chakraborty, A., Omrani, H., Teller, J., 2022. A Comparative Analysis of Drivers Impacting Urban Densification for Cross Regional Scenarios in Brussels Metropolitan Area. *Land 2022*, 11, 2291.

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. *International Journal of Geographical Information Science*, 28, 234–255.

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

Han, Y., Jia, H., 2017. Simulating the spatial dynamics of urban growth with an integrated modeling approach: A case study of Foshan, China. *Ecological Modelling*, 353, 107–116.

Hu, Z., Lo, C., 2007. Modeling urban growth in Atlanta using logistic regression. *Comput. Environ. Urban Syst.*, 31, 667–688.

Liao, J., Tang, L., Shao, G., Qiu, Q., Wang, C., Zheng, S., Su, X., 2014. A neighbor decay cellular automata approach for simulating urban expansion based on particle swarm intelligence. *International Journal of Geographical Information Science*, 28, 720–738.

Loibl, W., Tötzer, T., 2003. Modeling growth and densification processes in suburban regions-simulation of landscape transition with spatial agents. *Environ. Model. Softw.*, 18, 553–563.

Montgomery, D. C., Runger, G. C., 2010. *Applied statistics and probability for engineers*. John Wiley & Sons.

Mustafa, A., Cools, M., Saadi, I., Teller, J., 2015. *Urban Development as a Continuum: A Multinomial Logistic Regression Approach*. Springer, Computational Science and Its Applications – ICCSA 2015, Cham, Switzerland.

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

Mustafa, A., Saadi, I., Cools, M., Teller, J., 2014. Measuring the Effect of Stochastic Perturbation Component in Cellular Automata Urban Growth Model. *Procedia environmental sciences*, 22, 156–168.

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 Studies*, 55, 3279–3298.

Poelmans, L., Rompaey, A., 2009. Detecting and modelling spatial patterns of urban sprawl in highly fragmented areas: A case study in the Flanders–Brussels region. *Landscape and Urban Planning*, 93, 10–19.

Poelmans, L., Rompaey, A., 2010. Complexity and performance of urban expansion models. *Comput. Environ. Urban Syst.*, 34, 17–27.

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

Verburg, P., van Eck, J., de Nijs, T., Dijst, M., Schot, P., 2004. A review of current calibration and validation practices in land-change modeling. *Environ. Model. Softw.*, 82, 174–182.

Vliet, J., Bregt, A., Brown, D., Delden, H., Heckbert, S., Verburg, P., 2016. A review of current calibration and validation practices in land-change modeling. *Environ. Model. Softw.*, 82, 174–182.

White, R., Uljee, I., Engelen, G., 2012. Integrated modelling of population, employment and land-use change with a multiple activity-based variable grid cellular automaton. *International Journal of Geographical Information Science*, 26, 1251–1280.

Zhang, Q., Ban, Y., Liu, J., Hu, Y., 2011. Simulation and analysis of urban growth scenarios for the Greater Shanghai Area, China. *Comput. Environ. Urban Syst.*, 35, 126–139.